Janine E. Deakin · Paul D. Waters
Jennifer A. Marshall Graves *Editors*

Marsupial Genetics and Genomics

Marsupial Genetics and Genomics

Janine E. Deakin · Paul D. Waters ·
Jennifer A. Marshall Graves
Editors

Marsupial Genetics and Genomics

Editors
Janine E. Deakin
ARC Centre of Excellence for Kangaroo Genomics
Evolution, Ecology and Genetics
The Australian National University
Research School of Biology
Building 46
0200 Canberra ACT
Australia
janine.deakin@anu.edu.au

Paul D. Waters
ARC Centre of Excellence for Kangaroo Genomics
Evolution, Ecology and Genetics
The Australian National University
Research School of Biology
Building 46
0200 Canberra ACT
Australia

Jennifer A. Marshall Graves
ARC Centre of Excellence for Kangaroo Genomics
Evolution, Ecology and Genetics
The Australian National University
Research School of Biology
Building 46
0200 Canberra ACT
Australia

ISBN 978-90-481-9022-5 e-ISBN 978-90-481-9023-2
DOI 10.1007/978-90-481-9023-2
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010928910

© Springer Science+Business Media B.V. 2010
No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Marsupials are appealing subjects for genetics and genomics studies because of their unique phylogenetic position, and their unique biology, as well as their tractable genome.

In the vertebrate phylogeny they are placed in Class Mammalia, sharing features common to all mammals such as possessing fur, feeding their young with milk produced by mammary glands, and being warm blooded homeotherms. Marsupials last shared a common ancestor with eutherian mammals ∼150 million years ago (MYA), thereby bridging the gap between the divergence of Sauropsida (birds/reptiles) from the mammalian ancestor (350 MYA), and the eutherian radiation (105 MYA). As the most distantly related therian mammals to humans, they have special value for comparative genetic studies, and their inclusion in comparative genomic studies has provided many surprising findings in genome and gene evolution.

Marsupials are characterized especially by their reproductive system, which evolved to fit them for a life in a harsh continent with uncertain food supplies. Unlike eutherian ("placental") mammals, they give birth to altricial young, and most of their development occurs outside the protective environment of the uterus. This is usually in the confines of a pouch (hence their name, from "marsupium", meaning pouch); however, in many species the teats are arrayed in folds of the skin.

Marsupial genetics is not new. The large size of their chromosomes and their low diploid number has made them easy to study at the cytogenetic level and, in fact, they were amongst the first mammalian chromosomes to be examined early last century.

Much has happened since the landmark publication of *Mammals from Pouches and Eggs* (edited by J. Graves, R. Hope and D. Cooper) almost 20 years ago. This volume featured a collection of papers summarising research in marsupial breeding and genetics that has guided and shaped the field. But despite many major advances in our understanding of the marsupial genome, this is the only book to be dedicated to marsupial genetics and genomics. During the last decade, particularly, technological advances in genome sequencing and analysis have had a huge impact on the pace at which marsupial genetic and genomic research has, and is, progressing. We felt that it is time, once again, to compile a comprehensive collection of new work in marsupial genetics and genomics for those working in the field, their collaborators,

advanced students and those with a general interest in the genomes of these amazing creatures.

For many years, those working on marsupial genomics were not taken seriously by those studying more traditional model mammals, mouse and human. However, after sequencing of the human genome was completed, NHGRI recognised the need to sequence genomes of other mammals, particularly those of divergent species, to gain a better understanding of the organization, function and evolution of the human genome. Among mammal species chosen to be sequenced was a marsupial. The sequencing of the grey short-tailed opossum (published in 2007), and the genome of a second marsupial, the tammar wallaby, to be published shortly, has attracted worldwide attention, and enabled marsupials to be included in analyses of many genes and genome features of mammals, as illustrated in many chapters of this book. Of course most contributors to this book recognised the value of marsupials long before the sequencing revolution, and have been using the current marsupial genome sequence data to address long standing questions about the unique features of these fascinating animals.

Research into marsupial genetics and genomics informs many fields, as reflected by the content of this book, from animal breeding to genome sequencing projects, the development of bioinformatic tools and programs to deal with some of the unique features of marsupial genomes, to gene and genome evolution, the role of genes in reproduction and development, and of course population genetics and conservation. Often these areas overlap, and there is a high degree of collaboration between different groups working in this field, as is evident by the extent of cross-referencing between chapters. Through the compilation of work featured here, we hope this book will highlight more areas of potential collaboration and result in even more rapid progress in this field.

Marsupial genetics and genomics studies are critical for the conservation of Australia's native fauna, particularly as so many Australian marsupials are on the threatened or endangered species lists, including the iconic Tasmanian devil. The devil has recently been the subject of worldwide concern due to the devastating effects that devil facial tumour disease is having on the population. This unique transmissible tumour, with chromosomes and gene markers attesting to its clonal nature, has mobilized the genetic resources in Australia to genetically characterize the tumour and the population of normal animals, in order to develop a strategy to save the devil by enlightened management and captive breeding, or even by developing a vaccine. In order to achieve such a goal, we need to know much more about the normal and tumour genomes, as well as marsupial immunology. In the meantime, the knowledge of conservation geneticists is called upon to maximise the potential of captive breeding colonies. We hope that the collection of reviews we have compiled will be a useful resource to those taking up the fight to protect the Tasmanian devil, and other vulnerable Australian marsupials, from extinction.

We wish to thank Des Cooper, who inspired two of the editors to work on these interesting creatures, and has been an inspiration, friend and mentor to many people working in this field. We must acknowledge Max Haring, our publishing editor at Springer, who approached us with the idea for putting together this volume. We

are also extremely grateful to all those who contributed chapters, and hope this volume will encourage them to continue their excellent work in marsupial genetics and genomics.

Canberra, ACT Janine E. Deakin
Canberra, ACT Paul D. Waters
Canberra, ACT Jennifer A. Marshall Graves
December 2009

Contents

Part I Marsupial Breeding
Janine E. Deakin

1 Breeding and Genetic Management of Captive Marsupial Populations 5
Emily J. Miller and Catherine A. Herbert

Part II Marsupial Chromosomes and Gene Maps
Janine E. Deakin

2 The Conserved Marsupial Karyotype: Chromosome Painting and Evolution 37
Willem Rens and Malcolm Ferguson-Smith

3 Marsupial Centomeres and Telomeres: Dynamic Chromosome Domains 55
Dawn M. Carone and Rachel J. O'Neill

4 Marsupial Linkage Maps 75
Paul B. Samollow

5 Physical and Comparative Gene Maps in Marsupials 101
Janine E. Deakin

Part III Marsupial Genome Sequencing
Jennifer A. Marshall Graves

6 Marsupial Sequencing Projects and Bioinformatics Challenges . . 121
Anthony T. Papenfuss, Arthur Hsu, and Matthew Wakefield

7 Insight into Evolution of Gene Regulation Networks from the Opossum Genome 133
Matthew J. Wakefield and Anthony T. Papenfuss

Part IV Marsupial Sex Chromosomes
Paul D. Waters

8 Organization and Evolution of the Marsupial X Chromosome 151
Hardip R. Patel, Margaret L. Delbridge, and Jennifer A.M. Graves

9 Gene Content of the Mammalian X Chromosome 173
Margaret L. Delbridge

10 Marsupial Sex Chromosome Behaviour During Male Meiosis . . . 187
Raúl Fernández-Donoso, Soledad Berríos, Julio S. Rufas, and Jesús Page

11 Compact but Complex – The Marsupial Y Chromosome 207
Veronica J. Murtagh, Paul D. Waters, and Jennifer A. Marshall Graves

Part V Marsupial Epigenetics
Paul D. Waters

12 The Evolution of Genomic Imprinting – A Marsupial Perspective . 233
Timothy A. Hore, Marilyn B. Renfree, Andrew J. Pask, and Jennifer A. Marshall Graves

13 Marsupial Genetics Reveals Insights into Evolution of Mammalian X Chromosome Inactivation 259
Shafagh Al Nadaf, Paul D. Waters, Janine E. Deakin, and Jennifer A. Marshall Graves

Part VI Marsupial Reproductive and Developmental Genetics
Jennifer A. Marshall Graves

14 Molecular Regulation of Marsupial Reproduction and Development 285
Andrew J. Pask and Marilyn B. Renfree

15 Marsupial Milk – Identifying Signals for Regulating Mammary Function and Development of the Young 317
Julie A. Sharp, Christophe Lefevre, Joly Kwek, Denijal Topcic, Laurine Buscara, Cate Pooley, Robb de Iongh, Mary Familari, and Kevin R. Nicholas

Part VII Marsupial Immune Genes
Janine E. Deakin

16 The Marsupial Major Histocompatibility Complex 339
Yuanyuan Cheng, Hannah V. Siddle, and Katherine Belov

17 Marsupial Immunoglobulin and T Cell Receptor Genomics 357
Michelle L. Baker, Xinxin Wang, and Robert D. Miller

18 Use of Genomic Information to Gain Insights into Immune Function in Marsupials: A Review of Divergent Immune Genes . . 381
Katrina Morris, Emily S.W. Wong, and Katherine Belov

Part VIII Marsupial Genes and Gene Evolution
Paul D. Waters and Janine E. Deakin

19 Visual Pigments and Colour Vision in Marsupials and Monotremes 403
Samir S. Deeb

20 The Evolutionary History of Globin Genes: Insights from Marsupials and Monotremes 415
Vidushi S. Patel and Janine E. Deakin

21 The Olfactory Receptor Gene Family of Marsupials 435
Margaret L. Delbridge, Amir Mohammadi, and Jennifer A. Marshall Graves

Part IX Marsupial Conservation Genetics
Jennifer A. Marshall Graves

22 Marsupial Population and Conservation Genetics 461
Mark D.B. Eldridge

23 Devil Facial Tumour Disease (DFTD): Using Genetics and Genomics to Investigate Infectious Disease in an Endangered Marsupial 499
Hannah S. Bender

Index 517

Contributors

Shafagh Al Nadaf Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, shafagh.alnadaf@anu.edu.au

Michelle L. Baker Australian Animal Health Laboratory, CSIRO Livestock Industries, Geelong, VIC 3220, Australia, michelle.baker@csiro.au

Katherine Belov Faculty of Veterinary Science, The University of Sydney, Sydney, NSW 2006, Australia, kbelov@usyd.edu.au

Hannah S. Bender Department of Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, hannah.bender@anu.edu.au

Soledad Berríos Programa de Genética Humana, Facultad de Medicina, Instituto de Ciencias Biomédicas, Universidad de Chile, Santiago, Chile, sberrios@med.uchile.cl

Laurine Buscara Institute of Technology Research and Innovation, Deakin University, Geelong, VIC 3214, Australia, lbusc@deakin.edu.au

Dawn M. Carone Department of Molecular and Cell Biology, Center for Applied Genetics and Technology, University of Connecticut, Storrs CT 06269, USA, dawn.carone@uconn.edu

Yuanyuan Cheng Faculty of Veterinary Science, The University of Sydney, Sydney, NSW 2006, Australia, yuanyuan.cheng@sydney.edu.au

Robb de Iongh Department of Anatomy and Cell Biology, The University of Melbourne, Melbourne, VIC 3010, Australia, r.deiongh@unimelb.edu.au

Janine E. Deakin ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, janine.deakin@anu.edu.au

Samir S. Deeb Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA, USA, sdeebA@u.washington.edu

Margaret L. Delbridge ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, margaret.delbridge@anu.edu.au

Mark D.B. Eldridge Australian Museum, Sydney, NSW 2010, Australia, mark.eldridge@austmus.gov.au

Mary Familari Department of Zoology, The University of Melbourne, Melbourne,VIC 3010, Australia, m.familari@zoology.unimelb.edu.au

Malcolm Ferguson-Smith Cambridge Resource Centre for Comparative Genomics, Department of Veterinary Medicine, University of Cambridge, Cambridge, CB3 0ES, UK, maf12@cam.ac.uk

Raul Fernández-Donoso Programa de Genética Humana, Facultad de Medicina, Instituto de Ciencias Biomédicas, Universidad de Chile, Santiago, Chile, rfernand@med.uchile.cl

Catherine A. Herbert Faculty of Veterinary Science, The University of Sydney, Camperdown, NSW 2006, Australia, catherine.herbert@sydney.edu.au

Timothy A. Hore ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia; Laboratory of Developmental Genetics and Imprinting, The Babraham Institute, Cambridge, CB22 3AT, UK, tim.hore@bbsrc.ac.uk

Arthur Hsu Bioinformatics, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia, hsu@wehi.edu.au

Joly Kwek Department of Zoology, The University of Melbourne, Melbourne, VIC 3010, Australia, joly.kwek@alumni.unimelb.edu.au

Christophe Lefevre Institute of Technology Research and Innovation, Deakin University, Geelong, VIC 3214, Australia, christophe.lefevre@deakin.edu.au

Jennifer A. Marshall Graves ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, jenny.graves@anu.edu.au

Robert D. Miller Department of Biology, Center for Evolutionary and Theoretical Immunology, The University of New Mexico, Albuquerque, NM 87131, USA, rdmiller@unm.edu

Emily J. Miller Faculty of Veterinary Science, The University of Sydney, Camperdown, NSW 2006, Australia, emily.miller@sydney.edu.au

Amir Mohammadi ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, amir.mohammadi@anu.edu.au

Katrina Morris Faculty of Veterinary Science, The University of Sydney, Sydney, NSW 2006, Australia, kmor9570@usyd.edu.au

Veronica J. Murtagh Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, Veronica.Murtagh@anu.edu.au

Kevin R. Nicholas Institute of Technology Research and Innovation, Deakin University, Geelong, VIC 3214, Australia, kevin.nicholas@deakin.edu.au

Rachel J. O'Neill Department of Molecular and Cell Biology, Center for Applied Genetics and Technology, University of Connecticut, Storrs CT 06269, USA, rachel.oneill@uconn.edu

Jesús Page Unidad de Biología Celular, Departamento de Biología, Universidad Autónoma de Madrid, Madrid, Spain, jesus.page@uam.es

Anthony T. Papenfuss Bioinformatics, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia, papenfuss@wehi.edu.au

Andrew J. Pask Department of Molecular and Cell Biology, University of Connecticiut Storrs, CT 06269, USA; Department of Zoology, ARC Centre of Excellence for Kangaroo Genomics, The University of Melbourne, Melbourne, VIC 3010, Australia, andrew.pask@uconn.edu

Hardip R. Patel ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, hardip.patel@anu.edu.au

Vidushi S. Patel ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, vidushi.patel@anu.edu.au

Cate Pooley Institute of Technology Research and Innovation, Deakin University, Geelong, VIC 3214, Australia, cate.pooley@deakin.edu.au

Marilyn B. Renfree Department of Zoology, ARC Centre of Excellence for Kangaroo Genomics, The University of Melbourne, Melbourne, VIC 3010, Australia, m.renfree@unimelb.edu.au

Willem Rens Department of Veterinary Medicine, Cambridge Resource Centre for Comparative Genomics, University of Cambridge, Cambridge, CB3 0ES, UK, wr203@mole.bio.cam.ac.uk

Julio S. Rufas Unidad de Biología Celular, Departamento de Biología, Universidad Autónoma de Madrid, Madrid, Spain, julio.s.rufas@uam.es

Paul B. Samollow Department of Veterinary Integrative Biosciences, Faculty of Genetics, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843-4458, USA, psamollow@cvm.tamu.edu

Julie A. Sharp Institute of Technology Research and Innovation, Deakin University, Geelong, VIC 3214, Australia, julie.sharp@deakin.edu.au

Hannah V. Siddle Faculty of Veterinary Science, The University of Sydney, Sydney, NSW 2006, Australia, h.siddle@usyd.edu.au

Denijal Topcic Department of Zoology, The University of Melbourne, Melbourne, VIC 3010, Australia, denijal.topcic@bakeridi.edu.au

Matthew J. Wakefield Bioinformatics, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia, wakefield@wehi.edu.au

Xinxin Wang Department of Biology, Center for Evolutionary and Theoretical Immunology, The University of New Mexico, Albuquerque, NM 87131, USA, wxx1981@unm.edu

Paul D. Waters Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia, paul.waters@anu.edu.au

Emily S.W. Wong Faculty of Veterinary Science, The University of Sydney, Sydney, NSW 2006, Australia, emily.wong@sydney.edu.au

Abbreviations

AE	Axial elements
AFLPs	Amplified fragment length polymorphisms
ALT	Alternative lengthening of telomeres
AS	Angelman's syndrome
BAC	Bacterial artificial chromosome
BCR	B cell receptor
BLAST	Basic local alignment search tool
BLAT	BLAST-like alignment tool
BLG	β-lactoglobulin
CE	Central element
CEN	Centromere
cM	centiMorgans
CML	Chronic myelogenous leukaemia
crasiRNA	Centromere repeat associated small interacting RNA
CT	Cancer-testis
CTVT	Canine transmissible venereal tumour
DFTD	Devil facial tumour disease
DP	Dense plate
DSB	Double strand breaks
dsRNA	Double-stranded RNA
DTH	Delayed type hypersensitivity
EB	Evolutionary breakpoint
ERV	Endogenous retrovirus
ESD	Environmental sex determination
ESU	Evolutionary significant units
EU	Euchromatic region
FIL	Feedback inhibitor of lactation
FISH	Fluorescence in situ hybridisation
GSD	Genetic sex determination
HLA	Human leukocyte antigen
HMG	High mobility group proteins
HMM	Hidden Markov models
HS	Hypersensitive site

HSP	High-scoring segment pairs
ICR	Imprint control region
Ig	Immunoglobulin
KERV	Kangaroo endogenous retrovirus
kLTR	KERV long terminal repeat
LCR	Locus control region
LDL-C	Low-density-lipoprotein cholesterol
LE	Lateral elements
LG	Linkage group
LINE	Long interspersed nucleotide element
LTR	Long terminal repeat
LWS	Long-wave-sensitive
Mb	Megabases
MGR	Multiple genome rearrangement
MHC	Major histocompatibility complex
miRNA	Micro-RNA
MLR	Mixed lymphocyte reactions
MRE	Major regulatory region
mRNA	Messenger RNA
MSCI	Meiotic sex chromosome inactivation
MSP	Microspectrophotometric
MSUC	Meiotic silencing of unsynapsed chromatin
MSY	Male specific region of the Y
mtDNA	Mitochondrial DNA
MWS	Middle-wave-sensitive
MYA	Millions of years ago
NOR	Nucleolar organising regions
NT	Northern territory
OR	Olfactory receptor
ORF	Open reading frame
PAR	Pseudoautosomal region
PBR	Peptide binding region
piRNA	Piwi-interacting RNA
PMSC	Post meiotic sex chromatin
PWS	Prader-Willi syndrome
PY	Pouch young
QTL	Quantitative trait loci
RAG	recombination activating gene
RAPD	Random-amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RNAi	RNA interference
RSS	Recombination signal sequences
SA	South Australia
SC	Synaptonemal complex
siRNA	Short interfering RNA

snoRNA	Small nucleolar RNA
SNP	Single nucleotide polymorphism
SpA	Staphylococcal protein A
SRR	Sex and reproduction related
SSCP	Single-strand conformation polymorphism
STR	Simple tandem repeat
SWS	Short-wave-sensitive
TCR	T cell receptor
TDF	Testis determining factor
TF	Transverse filaments
TSD	Temperature-dependent sex determination
V(D)J	Variable (diversity) Joining
VH	IgH V domains
Vic	Victoria
VNTR	Variable number of tandem repeat
WA	Western Australia
WAP	Whey acidic protein
Xa	Active X chromosome
XAR	X added region
XCE	X controlling element
XCI	X chromosome inactivation
XCR	X conserved region
Xi	Inactive X chromosome
XIC	X inactivation centre
XIST	X inactive specific transcript
XLMR	X-linked mental retardation

Part I
Marsupial Breeding

Janine E. Deakin

Well established and managed captive breeding colonies of marsupials have been extremely important resources for research into marsupial genetics and genomics. In fact, much of the work described in this book would have been impossible without captive bred animals. Initially, many different species where used for marsupial genetics research but three distantly related species were chosen to be "model" marsupial species 20 years ago at the Boden Conference from which "Mammals from Pouches and Eggs" was derived (Hope and Cooper, 1990). The tammar wallaby (*Macopus eugenii*) was chosen as a model kangaroo, representing the large Family Macropodidae, and the fat-tailed dunnart (*Sminthopsis crassicaudata*) as a representative of the speciose Family Dasyuridae. The grey short-tailed opossum (*Monodelphis domestica*) was chosen to represent South American marsupials.

Most work on marsupial genetics and genomics since then has been done on two of these three species, the tammar wallaby and the grey opossum. The fat-tailed dunnart was used to derive some of the first linkage maps, but this research has now ceased and the colony is no longer genetically managed (Graves et al., 2003) A closely related species, the stripe-faced dunnart (*S. macroura*) is now bred for research into the marsupial egg (Selwood and Hickford, 1999), but it is likely that the Tasmanian devil (*Sarcophilus harissii*) will soon replace *S. crassicaudata* as the model dasyurid species for genetics and genomics research, with transcriptome and genome projects for this species currently underway (see Papenfuss et al., Chapter 6 and Bender, Chapter 23 for more details), and a physical map of the genome being constructed (Deakin, Chapter 5).

Both the tammar wallaby and the grey short-tailed opossum continue to be important model species. Their genomes have been sequenced (http://www.genome.gov/12512299; Mikkelsen et al., 2007), cementing their status as model marsupials for genetic and genomics studies. The combination of genome sequence data with the physiological data obtained from research on these animals will lead to a greater understanding of the unique features of marsupials.

The tammar wallaby (*Macropus eugenii*), a species now extinct on mainland Australia but still thriving on islands off the coast of South Australia (Kangaroo Island) and Western Australia (e.g. Abrolhos Island), was chosen as the model macropod species (Hinds et al., 1990). Its small size makes handling of these

animals easy and its reproduction is easily manipulated, although limited to one offspring per year. The many genetic polymorphisms detected between the two geographically isolated populations were recognised as a further asset, with animals bred from crosses between these two populations being particularly valuable for linkage mapping studies (McKenzie and Cooper, 1997; see Samollow Chapter 4, in Part II of this book) and research for which information on the parent of origin for allele expression is important (e.g. imprinting and X chromosome inactivation; see Al Nadaf et al., Chapter 13).

The grey short-tailed opossum (*M. domestica*), a South American species, is a true "laboratory marsupial" that has been used for biomedical research. It is a small marsupial weighing between 70 and 160 g in its adult form, can be bred in cages similar to rodents, and is a prolific breeder, with litter sizes on average of eight young and about three litters per year (Samollow, 2008). This species has been raised in pedigreed colonies for over 30 years (Samollow, 2008). These well-defined pedigrees were essential for the construction of genetic linkage maps and mapping of quantitative traits (described in Samollow, Part II of this book).

In Chapter 1, Miller and Herbert discuss the many factors that need to be considered when establishing and maintaining captive breeding colonies of marsupials, such as the reproductive strategies of the species. They highlight the importance of genetic factors in maintaining captive breeding colonies and the procedures which should be carried out to maintain a healthy level of genetic diversity, using their own experience with a captive research colony of tammar wallabies as an example.

Some marsupial species are bred in captivity for conservation, rather than research purposes. Miller and Herbert highlight the increasing need for captive breeding colonies for the conservation of the ever growing list of threatened and endangered marsupial species. A very prominent example is the Tasmanian devil, a species recently added to the endangered species list as a result of devil facial tumour disease (DFTD) (see Bender, Chapter 23), which has killed more than half the world's devils. Captive colonies of devils have been established on mainland Australia to insure against the extinction of this species. In this case, the consideration of genetic factors is critical. It appears that the prevalence of the disease in wild devils is due to low levels of genetic diversity in key immune genes of the Major Histocompatibility Complex (MHC) (Siddle et al., 2007). Each animal entering an insurance colony is typed for its MHC alleles with the intent of maximising genetic diversity for these loci (TDSC, 2007).

Obviously marsupial captive breeding colonies are vital for research and conservation. Unfortunately, maintaining such colonies is an extensive and expensive endeavour, but the loss of these valuable resources would be devastating for marsupial research and the conservation of Australia's native fauna. These colonies are particularly important now that we have genome sequence data. Genome sequences provide the basic information required to now begin exploring, at the genetic and genomic level, the unique biology of marsupials.

References

Graves JAM, Wakefield MJ, Renfree MB, et al. (2003) Proposal to sequence the genome of the model marsupial *Macropus eugenii* (Tammar Wallaby). http://wwwgenomegov/Pages/Research/Sequencing/SeqProposals/WallabySEQpdf.

Hinds LA, Poole WE, Tyndale-Biscoe CH, van Oorschot RAH, Cooper DW (1990) Reproductive biology and the potential for genetic studies in the tammar wallaby *Macropus eugenii*. In: Graves JAM, Hope RM, Cooper DW (eds) Mammals from Pouches and Eggs: Genetics, Breeding and Evolution of Marsupials and Monotremes. CSIRO Press, Melbourne.

Hope RM, Cooper DW (1990) Marsupial and monotreme breeding in wild and captive populations: towards a laboratory marsupial. In: Graves JAM, Hope RM, Cooper DW (eds) Mammals from Pouches and Eggs: Genetics, Breeding and Evolution of Marsupials and Monotremes. CSIRO Press, Melbourne.

McKenzie LM, Cooper DW (1997) Hybridisation between tammar wallaby (*Macropus eugenii*) populations from Western and South Australia. J Hered 88:398–400.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Samollow PB (2008) The opossum genome: insights from an alternative mammal. Genome Res 18:1199–1215.

Selwood L, Hickford D (1999) Early cell lineages in marsupial embryos. In: Moody SA (ed) Cell Fate and Cell Determination. Academic Press, New York.

Siddle HV, Kreiss A, Eldridge MDB, et al. (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA 104:16221–16226.

TDSC (2007) Insurance Population Strategy. Save the Tasmanian Devil Steering Committee.

Chapter 1
Breeding and Genetic Management of Captive Marsupial Populations

Emily J. Miller and Catherine A. Herbert

Abstract Numerous marsupial species are held in captive breeding colonies throughout the world for the purposes of research, conservation and education. Research colonies have facilitated the development of model marsupial species, including the tammar wallaby (*Macropus eugenii*), grey short-tailed opossum (*Monodelphis domestica*) and fat-tailed dunnart (*Sminthopsis crassicaudata*), the former two of which have been the subject of extensive genome sequencing projects. These species have played a fundamental role in improving our knowledge of marsupial biology, and much of this knowledge now contributes to the conservation and management of other species. Captive breeding programs have become an integral component of species recovery plans as an increasing number of marsupial species become threatened with extinction in the wild, and display colonies also play an important role in conservation education. Genetics has become an intrinsic component of conservation biology, and an instrumental tool in the management of captive populations. All captive colonies need to be actively managed to maximise the retention of extant genetic diversity, minimise levels of inbreeding and avoid genetic bottlenecks, adaptation to captivity and unwanted introgression, all of which can reduce the fitness of individuals and the evolutionary potential of the population. Successful breeding and genetic management of captive marsupial populations requires a thorough understanding of the reproductive and life history traits of the species. This chapter explores the extent to which these traits influence the retention of genetic diversity in relation to extrinsic factors associated with colony management.

Keywords Captive breeding · Conservation · Genetic diversity · Inbreeding · Mating system

---

C.A. Herbert (✉)
Faculty of Veterinary Science, The University of Sydney, Camperdown, NSW 2006, Australia
e-mail: catherine.herbert@sydney.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_1, © Springer Science+Business Media B.V. 2010

1.1 Introduction

Marsupials have been bred in captivity in Australia and internationally for several decades (George, 1990). Historically, captive populations were established to satisfy curiosity or provide entertainment. Today, there has been a philosophical shift towards conservation. Many wild animal populations currently require some form of active management to ensure their survival. Human modifications to the environment, including habitat destruction and fragmentation, modification of faunal assemblages through the removal or introduction of predators and competitors, modification of floral communities, alteration to fire regimes and water sources, and the process of urbanisation have resulted in dramatic changes to the distribution and abundance of endemic species (Caughley, 1987). The current losses of biodiversity and rates of extinction are unprecedented (Myers and Knoll, 2001) and the role of conservation programs, including captive breeding, has become crucial for the survival of many species in the twenty-first century.

The rationale for maintaining marsupial captive breeding colonies generally falls within one of the following three broad categories, all of which contribute to conservation in some way:

Research colonies, housing one or more model marsupial species which provide a source of individual animals for fundamental biological or applied research;

Conservation colonies, which generally aim to breed threatened or vulnerable species in captivity for long-term conservation or reintroduction into their natural habitat once threatening processes have been ameliorated in the wild; and

Display colonies, which are open to the public and play an important role in education.

For the purpose of this review, a captive breeding colony is defined as the breeding of a species in a controlled environment such as a zoo, conservation or research facility. This includes individual populations that are housed at a single facility, as well as the management of numerous colonies as a self-sustaining metapopulation. A list of the marsupial species commonly kept in captivity is provided in Table 1.1.

Research colonies provide an important contribution to the understanding of many facets of biology, from behaviour to genomics. While there are innumerable marsupial species that have been the subject of research, there are only a few species that could truly be considered model species for marsupial research. These include the tammar wallaby (*Macropus eugenii*) (Fig. 1.1a), grey short-tailed opossum (*Monodelphis domestica*), fat-tailed dunnart (*Sminthopsis crassicaudata*) and the brushtail possum (*Trichosurus vulpecula*) (Fig. 1.1b). The focus of much of this research has been within the fields of genetics and genomics (Hope, 1993; Graves and Hesterman, 2002); reproductive biology (Tyndale-Biscoe and Renfree,

Table 1.1 List of marsupial species commonly held in captive breeding colonies throughout the world

Common name	Scientific name	Purpose
Didelphidae		
Grey short-tailed opossum	*Monodelphis domestica*	Research
Phascolarctidae		
Koala	*Phascolarctos cinereus*	Education
Dasyuridae		
Tasmanian devil	*Sarcophilus harrisi*	Education, conservation
Tiger quoll	*Dasyurus maculates*	Education, conservation
Fat-tailed dunnart	*Sminthopsis crassicaudatus*	Research
Potoroidae		
Burrowing bettong	*Bettongia lesueur*	Conservation
Macropodidae		
Tammar wallaby	*Macropus eugenii*	Research, education
Red kangaroo	*M. rufus*	Education
E. Grey kangaroo	*M. gigantueus*	Education
Bennett's wallaby	*M. rufogriseus rufogriseus*	Education
Parma wallaby	*M. parma*	Education
Yellow-footed rock-wallaby	*Petrogale xanthopus*	Education/conservation
Matschie's tree kangaroo	*Dendrolagus matschiei*	Education
Goodfellow's tree kangaroo	*D. goodfellowi*	Education
Petauridae		
Sugar glider	*Petaurus breviceps*	Education
Acrobatidae		
Feathertail glider	*Acrobates pygmaues*	Education, research[a]
Peramelidae		
Eastern-barred bandicoot	*Perameles gunnii*	Education, conservation
Greater bilby	*Macrotis lagotis*	Conservation
Vombatidae		
Common wombat	*Vombatus ursinus*	Education
Phalangeridae		
Brushtail possum	*Trichosurus vulpecula*	Research
Spotted cuscus	*Spilocuscus maculates*	Education, research[a]

[a]Research on husbandry techniques

1987) and immunology (Ashman, 1977; Belov et al., 2006). The precocial nature of marsupial offspring at birth also makes them excellent candidates for studying developmental biology and they are increasingly being used as model species for biomedical research (Tyndale-Biscoe and Janssens, 1988; Mikkelsen et al., 2007). While many research colonies do not set out to address questions related

Fig. 1.1 Model marsupial species: (**a**) tammar wallaby (*Macropus eugenii*) and (**b**) brushtail possum (*Trichosurus vulpecula*) © E.J. Miller

to marsupial conservation, the core scientific knowledge gained from these research activities often facilitate marsupial conservation programs. A good example of this is the characterisation of major histocompatability complex (MHC) genes in marsupials and the subsequent development of associated genetic markers. These studies initially focused on the evolution of immune genes in vertebrates (e.g. Deakin et al., 2006) but the subsequent development of MHC-linked markers enabled identification of a link between devil facial tumour disease (DFTD) and low MHC diversity in Tasmanian devils (*Sarcophilus harrisii*) (Siddle et al., 2007). This has resulted in

a coordinated approach to Tasmanian devil captive breeding programs to encourage pairings that will enhance the maintenance of genetic diversity, especially at MHC loci (TDSC, 2007; Bender, Chapter 23).

Conservation colonies have become an important component of threatened species recovery programs and often form a necessary component of conserving threatened taxa. These colonies vary from intensive one-on-one pairings of individual animals to large, semi-free range enclosures of breeding individuals. The ultimate goal of these populations is usually to produce progeny for release into the wild, but meeting this goal is dependent on addressing the initial cause(s) for their decline in the wild. If the threatening processes have not been removed, then release into the wild may not be possible, and long-term maintenance in captivity will be required. The marsupial species that fall within this category vary temporally and geographically. The species list in Table 1.1 merely illustrates some well known examples and is by no means exhaustive.

Display colonies of many marsupial species are held in zoos and wildlife parks throughout the world. The motivation for keeping these animals is primarily conservation education rather than a direct role in conservation per se. For example, the Regional Collection Plan (RCP) for marsupials within the member institutions of the American Zoo and Aquarium Association (AZA) focuses on keeping collections of "common [marsupial] species with interesting stories" that can play a role in educational programs by demonstrating important concepts in wildlife conservation and management (Hutchins et al., 1998). Most zoos, particularly in the Australasian region, maintain a diversity of species that is reflected in Table 1.1. Many zoos also try to place an emphasis on research activities in specific taxa, e.g. the common spotted cuscus (*Spilocuscus maculatus*) was included in the RCP for North American Zoos because it can serve as the focus for basic biological and husbandry research. Many species of cuscus are considered threatened or endangered and little is known about their basic biology or appropriate methods of care in captivity (Hutchins et al., 1998). Relatively common species, such as the common spotted cuscus, can therefore serve as a model species for other less common species within this and other closely related genera.

All captive colonies need to be actively managed to maximise the retention of extant genetic diversity, minimise levels of inbreeding and avoid genetic bottlenecks, adaptation to captivity and unwanted introgression, all of which affect population growth, survival and evolutionary potential (Ralls and Ballou, 1986; Reed and Frankham, 2003). Factors that influence the maintenance of genetic diversity in captivity include intrinsic factors, such as the species mating system and reproductive biology, and extrinsic factors, such as enclosure design, stocking density and colony management strategies. The aim of this chapter is to discuss the interplay of reproductive and genetic factors that influence successful breeding and genetic management of marsupial populations in captivity. As increasing numbers of marsupial species decline in the wild and become the subject of intensive (or extensive) captive breeding programs, understanding these factors will become increasingly important for long term species survival.

1.2 Aspects of Species Biology Influencing Captive Breeding

1.2.1 Reproductive Strategies

Most marsupials are seasonal breeders and produce relatively small, immature young. The maternal investment in pregnancy is small, with a correspondingly longer period of lactation. Marsupial reproductive strategies have been divided into four basic reproductive patterns (Table 1.2) by Tyndale-Biscoe and Renfree (1987), following on from the initial attempt to classify marsupial reproductive patterns by Sharman et al. (1966), who first recognised the four groups. The division into four groups is primarily based on the duration of the oestrous cycle as a function of gestation, the type of placentation, whether they are polyoestrous or monoestrous, polyovular or monovular, and other specialised features, such as the occurrence of embryonic diapause. It is important to be aware of these fundamental aspects of their reproductive biology to facilitate effective management of population size and paternity of offspring.

Group 1 is thought to represent the basic or ancestral pattern from which the other three patterns are derived (Tyndale-Biscoe and Renfree, 1987). Species in this group are polyoestrous and polyovular (with the exception of the brushtail possum, which is monovular, and *Antechinus* sp. and *Phascogale* sp. which are monoestrous). The gestation period is considerably shorter than the oestrous cycle and coincides with the luteal phase, so that post-partum oestrus and ovulation are suppressed during lactation. Group 2 species have similar characteristics (see Table 1.2), but are unique in possessing a well formed, short-lived chorioallantoic placenta (Tyndale-Biscoe and Renfree, 1987). Species in this group are also renowned for having some of the shortest gestation periods of any mammal (Tyndale-Biscoe, 2005).

Group 3 is the only monovular group and comprises the macropodids and potoroids. Gestation is almost the same length as the oestrous cycle and extends into the follicular phase so that post-partum oestrus and ovulation occur. If fertilisation occurs, subsequent development of the corpus luteum is inhibited and the resultant conceptus remains in embryonic diapause, whereby development does not proceed past the blastocyst stage and the embryo is maintained in a state of dormancy (see Section 1.2.2 below). Within the Macropodidae family there are a number of species in which gestation occupies only 80–88% of the oestrous cycle and post-partum ovulation does not occur, such as the eastern (Fig. 1.2) and western grey kangaroos (*M. giganteus* and *M. fuliginosus* respectively), whiptail (prettyface) wallaby (*M. parryi*), parma wallaby (*M. parma*) and Matschie's tree kangaroo (*Dendrolagus matschiei*). These species are classified in an intermediate group (Group 3i), which is described as essentially a Group 1 pattern (Renfree, 1993). Females of these species will often mate during the later stages of lactation and the resultant embryo will enter diapause. The exception is the western grey kangaroo (*M. fuliginosus*), musky rat kangaroo (*Hypsiprymnodon moschatus*) and Lumholtz's tree kangaroo (*D. lumholtzi*), in which embryonic diapause does not occur at any time (Tyndale-Biscoe, 2005).

Table 1.2 The characteristics of the main reproductive patterns observed in marsupials

Reproductive pattern	Families	Oestrous characteristics	Oestrous cycle: gestation[a]	Specialised features	Exceptions
Group 1	Didelphidae Dasyuridae Petauridae Phalangeridae Vombatidae	Polyoestrous Polyovular	0.41–0.68		*Antechinus, Phascogale* (both monoestrous) *Trichosurus* and *Vombatus* (both monovular)
Group 2	Peramelidae Thylacomyidae	Polyoestrous Polyovular	0.60–0.69	Well-developed chorioallantoic placenta	
Group 3	Macropodidae Potoroidae	Polyoestrous Monovular	0.94–1.09	Post partum mating Embryonic diapause	*M. fuliginosus, M. giganteus, M. parma, M. parryi, Dendrolagus matschiei, D. lumholtzi, Hypsiprymnodon moschatus* (Group 3i)[b]
Group 3i	Some Macropodidae	Polyoestrous Monovular	0.8–0.88	No post partum mating Embryonic diapause	*M. fuliginosus D. lumholtzi, H. moschatus* [b] (no embryonic diapause)
Group 4	Burramyidae Acrobatidae Tarsipedidae	Polyoestrous Polyovular		Prolonged gestation with embryonic diapause	Burramys (Group 1 pattern)
Group 5	Phascolarctidae	Polyoestrous Monovular	0.70[c]	Induced ovulator	

[a]Proportional length of gestation relative to the length of oestrous cycle
References: Tyndale-Biscoe and Renfree, 1987 unless otherwise stated: [b]Tyndale-Biscoe, 2005; [c]Johnston et al., 2000 (NB. See this reference for a detailed account on variations in the oestrous cycle duration in koala mated versus non-mated cycles).

Fig. 1.2 Female eastern grey kangaroo (*Macropus giganteus*) with a joey © E.J. Miller

Much less is known about the Group 4 species. They are polyoestrous and polyovular, with a very prolonged luteal phase and gestation with a post-partum oestrus. The most striking aspect of their reproduction is the long period of embryonic diapause, the control of which is not clearly understood (Renfree and Shaw, 2000). Interestingly, blastocyst reactivation in the honey possum does not appear to be controlled by lactation, and birth has been observed to coincide with peaks in the production of pollen and nectar of flowering plants (Renfree and Shaw, 2000).

The majority of marsupial species are spontaneous ovulators. One notable exception to this is the koala (*Phascolarctos cinereus*), which is an induced ovulator. As such, this species do not fit clearly within any of the four generalised modes of reproduction reported by Tyndale-Biscoe and Renfree (1987), and it has been suggested that they should probably be placed within their own reproductive group (Handasyde et al., 1990; Johnson et al., 2000). For these reasons they have been included as a separate group (Group 5) in Table 1.2.

There are a number of marsupial species in which presence of the male appears to be essential for the induction of oestrus, including the grey short-tailed opossum,

brush-tailed bettong (*B. penicillata*) and greater bilby (Ballantyne et al., 2009). In the opossum, females are believed to respond to olfactory stimulus from a male, which induces oestrus and ovulation (Fadem, 1985). In the case of the bilby, it has not yet been determined if the female is an induced ovulator (like the koala) or if they respond to other factors associated with the presence of the male (Ballantyne et al., 2009).

1.2.2 Embryonic Diapause

Embryonic diapause is a common phenomenon that has been extensively studied in macropodids (Group 3). Little is known about embryonic diapause in Group 4 species, so they have been excluded from discussions in this section. If fertilisation occurs as a result of a post partum oestrus in macropodids, the embryo will develop to the stage of a unilaminar blastocyst of approximately 100 cells (taking approximately 8 days in the tammar wallaby), before entering a dormant state (diapause). Neural stimulation from the suckling of the newborn young in the pouch causes the pituitary to inhibit further corpus luteum and embryonic development (Renfree, 1979) until the first young vacates the pouch (or dies). The quiescent state of the ovaries as a result of lactation driven embryonic diapause is termed lactational quiescence (Tyndale-Biscoe et al., 1974). Two macropodid species, the tammar and Bennett's wallaby (*M. eugenii* and *M. rufogriseus* respectively) also undergo a period of seasonal quiescence. In these two species, the corpus luteum and blastocyst are inhibited by the suckling stimulus of the pouch young during the period of decreasing daylength (from birth in late January to the winter solstice in June in the Southern Hemisphere). Loss of the young during this period will result in corpus luteum and embryo reactivation, with birth occurring approximately 26 days later in the case of the tammar. However, if the young is lost during the period of increasing daylength, the blastocyst and corpus luteum will not reactivate (due to a photoperiod inhibition) until daylength begins to decrease again after the summer solstice (Sadlier and Tyndale-Biscoe, 1977).

1.2.3 Monitoring Reproductive Status

Monitoring the reproductive status of colony animals, combined with knowledge of the reproductive strategy of the species, is essential to effectively manage population size and breeding opportunities. It is necessary to be aware when specific individuals reach sexual maturity and when females are likely to return to oestrus, so that the appropriate pairings between males and females can be facilitated, or animals can be isolated if no further breeding is desired. Consideration must also be given to the occurrence of post-partum mating and embryonic diapause in the macropodids. This will influence population management and paternity planning, with colony managers and curators needing to pair the appropriate male with females at the time of oestrus, up to 11 months before the birth of the neonate.

Given the short duration of gestation in marsupials, there is little need for the use of hormone assays to determine the reproductive status of a female. Faecal hormone concentrations are primarily used to determine oestrous cycle length, timing of oestrus and gestation length, which can then be correlated with breeding success and behavioural observations. Such methods are often employed in the initial stages of a captive breeding program or in species that have experienced poor breeding success in captivity, e.g. Gilbert's potoroo (Stead-Richardson et al., 2010). The hope is that this information can be used to optimise husbandry techniques to increase reproductive success.

1.2.4 Reproductive Behaviour and Mating Systems

Large groups of breeding animals in captivity will display varying levels of competition for access to mates. This will be influenced by the species mating system, but also conditions imposed by the nature of the captive environment, e.g. density and sex ratio. Mating systems influence individual fitness and reproductive strategies, and are strongly linked with the genetic characteristics of a population. These are important to consider, as the genetic properties of a population will influence the evolutionary potential of a species. Furthermore, reproductive strategies are strongly linked to effective population size (N_e) (Ralls and Ballou, 1986) which is defined as "the number of individuals that would give rise to the calculated inbreeding coefficient, loss of heterozygosity or variance in allele frequency if they behaved in the manner of an idealised population" (Frankham et al., 2002). For example, in mating systems where there is high reproductive skew and some individuals are excluded from breeding (Nunney, 1991; 1993), N_e can become very low (Nunney and Elam, 1994; Frankham, 1995b). In populations of conservation concern, this is of particular importance as N_e determines the rate of loss of genetic diversity and inbreeding which contributes to the risk of extinction (Frankham, 1995b; Frankham et al., 2002).

There is limited empirical data available on the genetic mating system of many marsupial species. Table 1.3 summaries the broad patterns of the social organisation, degree of sexual dimorphism and mating system for each marsupial family. The classic mating system classifications are monogamy, polygyny, polyandry and promiscuity, though the precise definition of each can differ even within the field of evolutionary biology (Andersson, 1994).

Typically, monogamy is defined as the continuing bond and exclusive relationship between a male and a female. The only marsupial known to show any indication of monogamy is the allied rock-wallaby, *Petrogale assimilis* (Spencer et al., 1998). Using genetic techniques, Spencer et al. (1998) found evidence to suggest the allied rock-wallaby forms long-term socially monogamous pair bonds, although there was evidence of extra-pair fertilisations, which could lead to the assertion that they are effectively promiscuous (Dobson and Zinner, 2003). In a promiscuous mating system, both males and females mate with more than one partner and there is no long term relationship (Wittenberger, 1979). Polygyny occurs when a single male mates

Table 1.3 Broad patterns of reproductive groups, social organisation, degree of male-biased sexual size dimorphism and mating system in Marsupial families

Family	Genera	Common name	Reproductive group	Social organisation	Sexual size dimorphism[a]	Mating system	References
Dasyuridae	*Antechinus*	Antechinus	Group 1	Nocturnal, generally solitary or small groups	Strong	Promiscuous	Holleley et al. (2006), Kraaijeveld-Smit et al. (2003), (2002) and Strahan (2002)
	Dasyurus	Quoll		Usually nocturnal	Strong	Promiscuous	Glen et al. (2009) and Strahan (2002)
	Sarcophilus	Tasmanian devil		Nocturnal, generally solitary	Strong	Promiscuous	Fisher et al. (2001), Strahan (2002) and Taggart et al. (1998)
	Sminthopsis	Dunnart		Nocturnal	Strong	Polygynous	Strahan (2002) and Taggart et al. (1998)
Peramelidae	*Isoodon*	Short-nosed bandicoot	Group 2	Nocturnal, solitary	Moderate	Promiscuous or polygynous	Taggart et al. (1998)
	Perameles	Long-nosed bandicoot					
Thylacomyidae	*Macrotis*	Greater bilby	Group 2	Nocturnal, solitary, burrowing	Moderate	Overlap promiscuity	Lee and Cockburn (1985), Johnson and Johnson (1983) and Miller et al. (unpublished data)

Table 1.3 (continued)

Family	Genera	Common name	Reproductive group	Social organisation	Sexual size dimorphism[a]	Mating system	References
Vombatidae	*Lasiorhinus*	Hairy-nosed wombat	Group 1	Nocturnal, hierarchical, burrowing territorial	Mild	Unknown; polygyny proposed	Strahan (2002), Taggart et al. (1998) and Tyrell (2001)
	Vombatus	Common wombat					
Phascolarctidae	*Phascolarctos*	Koala	Group 5	Solitary, overlapping territories	Strong	Territorial polygyny proposed	Bercovitch et al. (2006) and Strahan (2002)
Potoroidae	*Bettongia*	Bettong	Group 3	Usually solitary, nocturnal, burrowing	None	Described as "promiscuous with a hint of monogamy"	Jarman (1991) and Pope et al. (2005)
	Potorous	Potoroo	Group 3		Moderate	Unknown	Shaw and Rose (1978)
Macropodidae	*Dendrolagus*	Tree-kangaroo	Group 3	Nocturnal, Predomiantly solitary or unimale groups, territorial	Strong	Polygynous	Strahan (2002)

Table 1.3 (continued)

Family	Genera	Common name	Reproductive group	Social organisation	Sexual size dimorphism[a]	Mating system	References
	Macropus	Kangaroo/ wallaby		Crepuscular, solitary to gregarious	Strong	Polygynous	Croft (1981), Fisher and Lara (1999), Hynes et al. (2005), Jarman (1991), Johnson (1989), Miller et al. (2010a), Rudd (1994) and Sigg et al. (2005)
	Petrogale	Rock-wallaby		Nocturnal, solitary or unimale groups, territorial	Mild	Socially monogamous (*P. assimilis*) to poylgynous (*P. pencillata*)	Hazlitt et al. (2006) and Spencer et al. (1998)
	Setonix	Quokka		Nocturnal, Hierarchical	Strong	Polygynous or Promiscuous	McLean et al. (2009) and McLean and Schmitt (1999)
	Thylogale	Pademelon			Moderate	Unknown	
Phalangeridae	*Trichosurus*	Possum	Group 1	Nocturnal, solitary with overlapping home ranges, territorial	Moderate	Polygynous	Strahan (2002), Taggart et al. (1998) and Taylor et al. (2000)

[a]Russell (1984)

with several females, and polyandry is the reverse, whereby a single female mates with several males (Wittenberger, 1981). Polygyny is common among marsupials, and is often linked to sexual size dimorphism (see Table 1.3). Male-biased sexual size dimorphism is exhibited to varying degrees in all marsupial families except Bettongia.

The influence that mating system and male reproductive skew has on various measures of genetic diversity can be predicted using Chesser's breeding-group model (Cheeser, 1991a, b; Chesser et al., 1993; Sugg and Chesser, 1994), which uses behavioural and ecological data to calculate the gene dynamics of social groups. Under any mating system, as the number of fathers increases, the levels of kinship and genetic differentiation among groups decreases, and N_e is increased (Dobson and Zinner, 2003). Species with a polygynous mating system often have male hierarchies to varying degrees (Clutton-Brock,1989) and reproductive success can therefore be highly skewed (e.g. eastern grey kangaroos, Miller et al. (2010b)). In such cases, a potential strategy to maximise the N_e in captive populations would be to manipulate mating patterns by using a specific number of sires (Oyama et al., 2007) to maximise founder representation. When more males participate in breeding, such as in socially monogamous and promiscuous mating systems, the lower variance in male reproductive success slows the rates of inbreeding, lowers the level of relatedness within the population and increases the N_e (Frankham et al., 2002). A good example of this is the greater bilby, in which the promiscuous mating system is believed to have contributed to the maintenance of genetic diversity in free-ranging captive breeding programs for this species (Miller et al., unpublished data).

This discussion then leads to the question: should animals be artificially paired based on genetic considerations or allowed to mate randomly? Mate choice plays an important evolutionary role, providing individuals the opportunity to choose the fittest individual to mate with, thereby passing on good genes to their offspring and increasing its chance of survival and successful reproduction. If a captive population is genetically depauperate, then every effort should be made to retain the maximum amount of genetic variation, particularly if the species is conservation dependent and has limited genetic variation across all populations. This will likely involve selective breeding of individual animals, unless the mating system encourages the retention of genetic diversity through limited male reproductive skew. If a population is genetically diverse, there is less need for specific pairings. But, given that captive conditions will likely alter the duration of a male's reign of dominance through improved health and nutrition, artificial male turnover may be necessary to prevent inbreeding.

1.3 Captive Breeding Strategies to Maintain Genetic Diversity

The maintenance of genetic diversity is a central focus of captive breeding programs, especially for conservation dependent species, as it provides the raw material for adaptive change, which is fundamental for the long-term sustainability of

populations (Darwin, 1859; Frankham et al., 2002). Management strategies for captive colonies should aim to minimise the genetic changes to the captive population over time in order to be representative of wild populations if, and when they are reintroduced back into their natural habitat (Ballou and Lacy, 1995; Frankham, 1995a). There are four main undesirable genetic changes that occur in captivity that can jeopardise the ability of populations to be self-sustaining (i) loss of genetic diversity, (ii) inbreeding depression, (iii) accumulation of deleterious mutations, and (iv) genetic adaptation to captivity (Frankham et al., 2002; Frankham, 2008). Such genetic changes can reduce reproductive fitness and threaten reintroduction programs (Frankham and Loebel, 1992; Frankham, 2008).

Genetic variation can be measured on several levels: the individual level, between individuals in a population, between populations, and between species. Allelic diversity and heterozygosity are the most common measures of diversity used. Allelic diversity is a measure of diversity based on the average number of alleles per locus in a population (Allendorf and Luikart, 2007), which is important for a population's long-term ability to adapt (Allendorf, 1986). Heterozygosity refers to the variation that accounts for either the observed or expected proportion of individuals in a population that are heterozygotes, that is an individual who carries two different alleles at a locus (Allendorf and Luikart, 2007). Maintaining heterozygosity is important for a population's short-term adaptive potential (Allendorf, 1986). The latter tends to be used as an indicator of a population's genetic health as it provides a proportional indication of the amount of genetic diversity at a locus (Allendorf, 1986). Although heterozygosity provides a useful insight, it can be deceptive. For instance, during a founding event or genetic bottleneck, allelic diversity is lost more rapidly than heterozygosity (Allendorf, 1986) and in such circumstances, using heterozygosity as a single indicator for genetic diversity would provide an inaccurate representation of the population (Ballou and Foose, 1996).

Molecular markers provide information about allelic diversity and heterozygosity. The increasing availability of genetic markers allows an in depth analysis and evaluation of genetic diversity. The most popular markers used today are microsatellites (simple tandem repeats, STRs) and amplified fragment length polymorphisms (AFLPs). In marsupials, there is an abundance of such markers for Macropodidae and Dasyuridae, but much less in known about most other families (Eldridge, Chapter 22). Other types of markers available include single strand polymorphisms (SSCP), random amplified polymorphic DNA (RAPD), restriction length polymorphisms (RFLP) and variable number of tandem repeats (VNTRs). Additionally mitochondrial DNA (mtDNA) and Y chromosome markers are used to trace maternal and paternal lineages, respectively. The recent advances in marsupial genome characterisation have lead to the development of microsatellite markers linked to functional regions of the genome, namely the MHC; Cheng et al., Chapter 16. The MHC contains genes that are important for immune response and can be used to measure a population's immunological fitness. As such, these markers will become increasingly popular as population genetics tools.

Most captive populations are founded with a small number of individuals, and small populations are prone to extinction events due to stochastic and genetic

problems (Ralls and Ballou, 1986). Population size is intrinsically linked to genetic diversity (Frankham, 1996), and small closed populations are likely to lose diversity more rapidly due to random genetic drift (Allendorf, 1986; Lacy, 1989; Frankham et al., 2002). Therefore it is expected that genetic diversity in captive populations will erode over time, and this is determined not so much by the number of individuals present in a population (*N*, census size), but by the genetically effective population size (N_e) (Wright, 1969). In general, the rate of loss of diversity is faster in smaller populations and the longer the time in captivity the larger the overall loss (Ballou and Foose, 1996). This raises two fundamental questions when devising management plans, (i) how long should the population be maintained in captivity; and (ii) how much diversity is "enough"?

Management plans for species being bred in captivity generally aim to achieve long-term self-sustainability where possible and follow the principles outlined by Frankel and Soulé (1981). Soulé et al. (1986) proposed that captive breeding programs should aim to retain 90% of the source population genetic variation over 200 years. However this is often unattainable due to the small number of founder individuals available for establishing populations and/or space limitations. Consequently, many captive breeding programs aim to retain 90% of diversity over 100 years, but sometimes this is relaxed to 80% for 100 years, or 90% for 50 years (Frankham et al., 2002).

Captive breeding programs for threatened species are often established with few founders because of a lack of animal availability. Many of these populations provide a source of individuals for reintroductions and as such should be capable of establishing a self-sustaining wild population with high reproductive fitness and ample genetic diversity (Frankham et al., 2002). For these reasons it is recommended that captive populations are established before wild populations drop below 1,000 individuals (IUCN, 2006). The advantages of establishing a captive population at this stage include allowing sufficient time to develop suitable husbandry techniques, using wild individuals with low levels of inbreeding, and reducing the impact of removing individuals from the wild (IUCN, 2006).

When founding a captive population there are some important factors that need to be considered. First, resolving taxonomic uncertainty. Often the founders of a population come from different sources or are of unknown origin so it is important to resolve such issues prior to the commencement of captive breeding to avoid unwanted hybridisation and/or outbreeding depression (Frankham et al., 2002). Second, to determine how many individuals are required to capture sufficient genetic diversity that is representative of wild populations. A greater number of animals are required to sample rare alleles. This is a crucial phase of establishing captive populations as it impacts the N_e, the extent of the initial genetic bottleneck, and the demographic security of the population (Ballou and Foose, 1996; Rudnick and Lacy, 2008).

Several management strategies have been recommended to retain maximum levels of genetic diversity and minimise levels of inbreeding in captive populations. Table 1.4 provides a summary of breeding strategies that can be applied to minimise the loss of genetic diversity and potential risks associated with captive breeding,

Table 1.4 Strategies applied to minimise the loss of genetic diversity and potential risks associated with captive breeding

Strategy	Advantages	Disadvantages	References
Minimising kinship	Chooses parents with lowest kinship using a pedigree Equalises founder contributions Equalises family sizes Retains reproductive fitness	Complete pedigrees rarely available Not suitable for species in large groups with unknown paternity	Frankham et al. (2002), Montgomery et al. (1997) and Rudnick and Lacy (2008)
Maximum avoidance of inbreeding	Equalises family sizes Doubles N_e Delays inbreeding Retains reproductive fitness	Must be applied from first (founder) generation Must be followed precisely No practical allowance for mortality, reduced fertility, incompatibility of breeders	Frankham et al. (2002) and Montgomery et al. (1997)
Random choice of parents	Reduces unequal founder representation Retains reproductive fitness	Can result in sib mating if not controlled Retains lower levels of diversity than minimising kinship Elimination of mate choice may reduce offspring viability	Montgomery et al. (1997) and Wedekind (2003)
Population supplementation with wild individuals	Reduces local adaptation to captive environment Prevents inbreeding depression Maintains genetic diversity Increases N_e Can be used to simulate gene flow between different populations	Difficult to monitor the genetic effects of supplementation Requires accurate knowledge of species taxonomy and origin	Eldridge and Killebrew (2008), Ryman and Laikre (1991) and Miller et al. (unpublished data)
Minimise generations in captivity	Reduces selection for domestication in captivity Reduces potential loss of genetic diversity in captivity		Frankham (2008) and Frankham et al. (2002)

Table 1.4 (continued)

Strategy	Advantages	Disadvantages	References
Free mate choice (rather than random selection of parents)	Potentially improves offspring fitness	Often larger space requirements	Wedekind (2003)
Maintaining high N_e			Ballou and Cooper (1992), Ballou and Lacy (1995), Borlase et al. (1993), Frankham (1995b), Lacy (1989) and Loebel et al. (1992)
Equalising family sizes	Approximately doubles N_e Minimises inbreeding and loss of genetic diversity		
Equalising sex ratio	Increases N_e		
Maximising generation length	Reduces amount of diversity lost in every generation	Risk of stochastic losses of animals before they breed Risk of reduction in fertility due to age-dependent factors	
Equalising founder contribution	Limits the reproduction and genetic representation of the founder individuals Thought to diminish genetic drift by enlarging effective population size resulting in a higher retention of allelic diversity	Requires equalisation of founding population to avoid irreversible loss of diversity Requires adequate number of founders	

Table 1.4 (continued)

Strategy	Advantages	Disadvantages	References
Cryopreservation and artificial insemination	Facilitates breeding between individuals unable to breed naturally Can be used to equalise or increase founder contribution Transport frozen semen between colonies, rather than animals (more cost effective) Can be used to equalise sex ratio of breeders Provides genetic material for future use Protected from deleterious environmental and genetic changes Slows inbreeding and loss of diversity	Technology needs to be customised for each species Often requires semen from animals that are no longer available or accessible	Ballou (1984), Ballou and Cooper (1992) and Johnston and Lacy (1995)
Supportive breeding from local populations	Assists populations that are not self-sustaining Reduces loss of adaptations to local environment Prevents outbreeding depression	Reduces N_e in long-term Reduces reproductive fitness Increases risk of genetic adaptation to captivity Risk of inbreeding depression Risk of unwanted hybridsation Risk of artificial selection for adaption to captivity	Brannon et al. (2004), Ryman et al. (1995), Ryman and Laikre (1991), Waples (1991) and Wedekind (2003)

such as adaptation to captivity. The most widely applied method is minimising kinship (Rudnick and Lacy, 2008). This strategy aims to reduce the overall kinship of a population by managing pedigreed populations so that the matings are made between individuals that are the most distantly related in the population. Simulations (Ballou and Lacy, 1995; Fernandez et al., 2004) and empirical data (Montgomery et al., 1997) provide evidence that this is the most effective strategy for maintaining genetic diversity (Rudnick and Lacy, 2008).

Captive breeding programs are often managed through pedigrees that are maintained in studbooks (Hedrick and Millar, 1992). Studbooks assist in minimising kinship among individuals, maintaining the representation of wild founder genetic diversity as well as providing a tool for assessing the genetic health of a population based on demographic data. The basic data that should be recorded for every individual includes animal identification, sex, parentage, birth and death dates, as well as animal movements (Ballou and Foose, 1996). The calculation of kinship values for captive breeding programs depends on a population's pedigree. These calculations are only effective when the relationships among individuals in a population are accurately known (Hedrick and Kalinowski, 2000; Nielsen et al., 2007; Rudnick and Lacy, 2008). Unfortunately all pedigrees have some level of uncertainty, especially since many were established during the early days of captive breeding when records were not always consistently maintained (Rudnick and Lacy, 2008). Errors in studbook estimates of the "genetic health" of a population can arise due to missing or erroneous pedigree data. For example, estimates of genetic diversity and inbreeding were calculated from 5 years of studbook (pedigree) data and compared to those calculated from microsatellite data to assess studbook reliability for a captive breeding colony of the greater bilby (*Macrotis lagotis*). This study found that studbook genetic estimates did not accurately reflect the microsatellite estimates of diversity, suggesting they should not be solely relied upon to evaluate the genetic health of the captive populations (Miller et al., unpublished data).

To improve the concordance between studbook and genetic estimates of diversity, institutions need to maintain more complete pedigree records. Validation of the "founder assumption" of individuals being unrelated and not inbred (Ballou, 1984), should be prioritised when there are limited resources available. Genetic analyses of parentage can help resolve questions regarding individual relationships within a population, and verify pedigrees. Accurate parentage data is essential for wildlife managers when making decisions regarding the pairing of individuals for mating and the transfer of individuals among institutions. The widespread application of genetic techniques has revealed inconsistencies in the presumed relationship between social organisation, parentage and mating system (Ambs et al., 1999; Coltman et al., 1999; Worthington Wilmer et al., 1999; Issac, 2005), transforming how mating systems are understood.

A significant biological problem associated with captive breeding is that each colony represents a small, fragmented population (Ballou and Foose, 1996) and therefore has a reduced probability of long-term survival (Frankham, 1995a). Gene flow needs to be resurrected between populations, by allowing for the introduction of "immigrants" and mutation (Ballou, 1984). Often animals are translocated

between institutions to simulate gene flow. Introducing new genotypes into a population in this way can also help alleviate the effects of inbreeding and loss of heterozygosity (Spielman and Frankham, 1992; Bryant et al., 1999) in a form of "genetic rescue" (Hedrick, 1995; 2000; Hedrick, 2005; Hedrick and Fredrickson, 2008). However this is costly, there is a level of stress involved for the animals, potential risk of disease transfer and sometimes problems of acceptance into populations, particularly if the species is highly territorial. The genetic advantages of translocating animals between institutions must therefore be weighed up against these other potentially negative consequences (Tallmon et al., 2004).

1.4 Case Study: Management of a Research Colony of a Model Marsupial (*Macropus eugenii*)

The management of a tammar wallaby research colony, housed at the University of New South Wales field station in the north of Sydney, Australia, will be used to illustrate the complex interplay between intrinsic and extrinsic factors and their influence on the maintenance of genetic diversity in a captive marsupial colony. As the tammar wallaby is a model macropodid marsupial that has been extensively studied over the last 50 years, their reproductive biology has been well-characterised, which provides valuable baseline information for the management of captive colonies of this species.

The principal strategies employed to maintain genetic diversity include an initial large founder size (> 100 animals) with occasional supplementation of wild animals ($n = 30$–80 every 4 years) into the population. This degree of supplementation is feasible given the large population present on Kangaroo Island (SA), which is currently managed by annual population harvesting (Wright and Stott, 1999), and the ease with which this (sub)species adapts to captive conditions. Sexually mature males are rotated every breeding season to maximise the potential for females to breed with numerous males throughout their lifetime. Females are, however, given some degree of mate-choice, with most enclosures housing multiple males, usually at a ratio of one male to four–eight females. Based on our knowledge of the mating system of the tammar wallaby, the dominant male on average will sire approximately 50% of young (Hynes et al., 2005; Miller et al., 2010a), so subordinate males will have some capacity for breeding.

Prevention of inbreeding is enhanced by knowledge of the reproductive biology of this species and careful attention to the timing of reproductive events. Female tammar wallabies will undergo a post-partum mating, usually within 1 h of giving birth (Rudd, 1994), and the resultant blastocyst will remain in embryonic diapause for up to 11 months (Berger, 1966). Hence, any young born as a result of blastocyst reactivation triggered by the summer solstice will have been conceived in the previous breeding season. Paternity records need to reflect this to prevent subsequent pairing of related individuals. If paternity is unknown, the large suite of microsatellilte markers available for this species (see Eldridge, Chapter 22) should be utilised to confirm paternity. Maternity should also be confirmed by

Fig. 1.3 Tammar wallaby (*Macropus eugenii*) from North Island in the Abrohlos Archipelago, Western Australia © E.J. Miller

microchipping pouch young when they are still confined to their mother's pouch whenever possible.

The recent clarification of taxonomic relationships between the Western Australian and South Australian tammar wallaby populations (Eldridge et al., unpublished data), and the acquisition of wild-caught Abrolhos Island (WA) tammars (Fig. 1.3), will likely result in the maintenance of two independently managed tammar wallaby lineages within captive research colonies. Historically, the divergence between these two lineages, combined with their ability to hybridise in captivity (McKenzie and Cooper, 1997), has been exploited to generate a linkage map the tammar wallaby genome (Zenger et al., 2002; see also Samollow, Chapter 4).

1.5 Future Research and Conclusions

Marsupials have been successfully maintained and bred in captivity for decades, but the increasing occurrence of captive breeding for species conservation necessitates greater attention to the genetic consequences of captive management strategies. Marsupial captive breeding programs have rarely been evaluated in terms of the extent to which genetic diversity is maintained over time and there is little information on the genetic mating system of many marsupial species. This dramatically

limits our ability to effectively manage these populations for long term species conservation and reintroduction goals. If captive breeding is to contribute significantly to both in situ and ex situ marsupial conservation then genetic considerations are paramount. The recent dramatic decline of the Tasmanian devil is a prime example of the potentially disastrous consequences that loss of genetic diversity can have at both a population and species level. More research is urgently needed to characterise marsupial mating systems and to evaluate the genetic consequences of captive breeding strategies in the context of specific mating systems.

Glossary

Chorioallantoic placenta an enveloping vascular fetal membrane formed by the fusion of chorion and allantois.

Corpus luteum is a temporary structure in mammals that develops from the ruptured ovarian follicle after ovulation and is involved in the production of progesterone.

Follicular phase phase of the oestrous cycle during which follicles in the ovary mature, secreting increasing amounts of oestradiol as maturation progresses.

Luteal phase the latter phase of the oestrous cycle when the corpus luteum is actively producing progesterone.

Monoestrous animals display one oestrous cycle within a breeding season.

Monovular ovulation of a single ovum.

Polyoestrous animals display multiple oestrous cycles within a breeding season.

Polyovular ovulation of more than one ovum during oestrus.

References

Allendorf FW (1986) Genetic drift and the loss of alleles versus heterozygosity. Zoo Biol 5: 181–190.

Allendorf FW, Luikart G (2007) Conservation and the Genetics of Populations. Blackwell Publishing, Malden, MA.

Ambs SM, Boness DJ, Bowen WD, Perry EA, Fleischer RC (1999) Proximate factors associated with high levels of extraconsort fertilisation in polygynous grey seals. Anim Behav 58:527–535.

Andersson M (1994) Sexual Selection. Princeton University Press, Princeton, NJ.

Ashman RB (1977) Marsupial immunology: a brighter future? Dev Comp Immunol 1:283–284.

Ballantyne K, Matson P, Noakes N, Nicolson V, Johnson SD (2009) Reproductive physiology of the female greater bilby (*Macrotis lagotis* Thylacomyidae): evidence for a male-induced luteal phase. Reprod Fert Dev 21:274–282.

Ballou JD (1984) Strategies for maintaining genetic diversity through reproductive technology. Zoo Biol 3:311–323.

Ballou JD, Cooper KA (1992) Application of biotechniology to captive breeding of endangered species. Sym Zool S 64:183–296.

Ballou JD, Foose TJ (1996) Demographic and genetic management of captive populations. In: Kleiman DG (ed) Wild Mammals in Captivity. Chicago University Press, Chicago, IL.
Ballou JD, Lacy RC (1995) Identifying genetically important individuals for management of genetic variation in pedigreed populations. In: Ballou JD, Gilpin M, Foose T (eds) Population Management for Survival and Recovery. Columbia University Press, New York, NY.
Belov K, Deakin JE, Papenfuss AT, et al. (2006) Reconstructing an ancestral mammalian immune supercomplex from a marsupial major histocompatibility complex. PLoS Biol 4:317–328.
Bercovitch FB, Tobey JR, Andrus CH, Doyle L (2006) Mating patterns and reproductive success in captive koalas (*Phascolarctos cinereus*). J Zool 270:512–516.
Berger PJ (1966) Eleven-month “embryonic diapause” in a marsupial. Nature 211:435–436.
Borlase SC, Loebel DA, Frankham R, Nurthen RK, Briscoe DA, Daggard GE (1993) Modeling problems in conservation genetics using captive *Drosophila* populations: effects of equalising family sizes. Conserv Biol 7:122–131.
Brannon EL, Amend DF, Cronin MA, et al. (2004) The controversy about salmon hatcheries. Fisheries 29:12–31.
Bryant EH, Backus VL, Clark ME, Reed DH (1999) Experimental tests of captive breeding for endangered species. Conserv Biol 13:1487–1496.
Caughley G (1987) Introduction to the sheep rangelands. In: Caughley G, Shepherd N, Short J (eds) Kangaroos: Their Ecology and Management in the Sheep Rangelands of Australia. Cambridge University Press, Cambridge, MA.
Cheeser RK (1991a) Gene diversity and female philopatry. Genetics 127:437–447.
Cheeser RK (1991b) Influence of gene flowand breeding tactics on gene diveristy within populations. Genetics 129:573–583.
Chesser RK, Rhoades OE, Sugg DW, Schnabel A (1993) Effective sizes for subdivided populations. Genetics 135:1221–1232.
Clutton-Brock TH (1989) Mammalian mating systems. P Roy Soc Lond B Bio 236:339–372.
Coltman DW, Bancroft DR, Robertson A, Smith JA, Clutton-Brock TH (1999) Male reproductive success in a promiscuous mammal: behavioural estimates compared with genetic paternity. Mol Ecol 8:1199–1209.
Croft DB (1981) Behaviour of the red kangaroo, *Macropus rufus* (Desmarest, 1822) in northwestern New South Wales. Aust Mammal 5:5–13.
Darwin C (1859) On the Origin of Species by Means of Natural Selection. John Murray, London.
Deakin JE, Papenfuss AT, Belov K, et al. (2006) Evolution and comparative analysis of the MHC Class III inflammatory region. BMC Genomics 7:281–295.
Dobson FS, Zinner B (2003) Social groups, genetic structure and conservation. In: Festa-Bianchet M, Apollonio M (eds) Animal Behaviour and Wildlife Conservation. Island Press, Washington.
Eldridge W, Killebrew K (2008) Genetic diversity over multiple generations of supplementation: an example from Chinook salmon using microsatellite and demographic data. Conserv Genet 9:13–28.
Fadem BH (1985) Evidence for the activation of female reproduction by males in a marsupial, the gray short-tailed opossum (*Monodelphis domestica*). Biol Reprod 33:112–116.
Fernandez J, Toro MA, Caballero A (2004) Managing individuals’ contributions to maximize the allelic diversity maintained in small, conserved populations. Conserv Biol 18:1358–1367.
Fisher DO, Lara MC (1999) Effects of body size and home range on access to mates and paternity in male bridled nailtail wallabies. Anim Behav 58:121–130.
Fisher DO, Owens IPF, Johnson CN (2001) The ecological basis of life history variation in marsupials. Ecology 82:3531–3540.
Frankel OH, Soulé ME (1981) Conservation and Evolution. Cambridge University Press, Cambridge, MA.
Frankham R (1995a) Conservation genetics. Annu Rev Genet 29:305–327.
Frankham R (1995b) Effective population size/adult population size ratios in wildlife: a review. Genet Res 66:95–107.

Frankham R (1996) Relationship of genetic variation to population size in wildlife. Conserv Biol 10:1500–1508.

Frankham R (2008) Genetic adaptation to captivity in species conservation programs. Mol Ecol 17:325–333.

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to Conservation Genetics. Cambridge University Press, Cambridge, MA.

Frankham R, Loebel DA (1992) Modeling problems in conservation genetics using captive *Drosophila* populations: rapid genetic adaptation to captivity. Zoo Biol 11:333–342.

George GG (1990) Monotreme and marsupial breeding programs in Australian zoos. Aust J Zool 37:181–205.

Glen AS, Cardoso MJ, Dickman CR, Firestone KB (2009) Who's your daddy? Paternity testing reveals promiscuity and multiple paternity in the carnivorous marsupial *Dasyurus maculatus* (Marsupialia: Dasyuridae). Biol J Linn Soc 96:1–7.

Graves JA, Hesterman M (2002) Marsupial genetics and gemonics. Trends Genet 18:517–521.

Handasyde KA, McDonald R, Than KA, Michaelides J, Martin RW (1990) Reproductive hormones and reproduction in the koala. In: Lee AK, Handasyde KA, Sanson GD (eds) Biology of the Koala. Surrey Beatty and Sons, Sydney.

Hazlitt SL, Sigg DP, Eldridge MDB, Goldizen AW (2006) Restricted mating dispersal and strong breeding group structure in a mid-sized marsupial mammal (*Petrogale penicillata*). Mol Ecol 15:2997–3007.

Hedrick PW (1995) Gene flow and genetic restoration: the Florida panther as a case study. Conserv Biol 9:996–1007.

Hedrick PW (2000) Genetics of Populations, 2nd edn. Jones and Bartlett, Sudbury.

Hedrick P (2005) 'Genetic restoration': a more comprehensive perspective than 'genetic rescue'. Trends Ecol Evol 20:489–496.

Hedrick PW, Fredrickson RJ (2008) Captive breeding and the reintroduction of Mexican and red wolves. Mol Ecol 17:344–350.

Hedrick PW, Kalinowski ST (2000) Inbreeding depression in conservation biology. Annu Rev Ecol Syst 31:139–162.

Hedrick PW, Millar JS (1992) Conservation genetics techniques and fundamentals. Ecol Appl 2:30–46.

Holleley CE, Dickman CR, Crowther MS, Oldroyd BP (2006) Size breeds success: multiple paternity, multivariate selection and male semelparity in a small marsupial, *Antechinus stuartii*. Mol Ecol 15:3439–3448.

Hope RM (1993) Selected features of marsupial genetics. Genetica 90:165–180.

Hutchins M, Roberts M, Cox C, Crotty MJ (1998) Marsupials and monotremes: a case study in regional planning collection. Zoo Biol 17:433–451.

Hynes EF, Rudd CD, Temple-Smith PD, Sofronidis G, Paris D, Shaw G, Renfree MB (2005) Mating sequence, dominance and paternity success in captive male tammar wallabies. Reproduction 130:123–130.

IUCN (2006) 2006 IUCN Red List of Threatened Species. Retrieved 26 November 2009, from www.iucnredlist.org

Issac JL (2005) Potential causes and life-history consequences of sexual size dimorphism in mammals. Mammal Rev 35:101–115.

Jarman PJ (1991) Social behaviour and organisation in the Macropodoidea. Adv Study Behav 20:1–37.

Johnson CN (1989) Social interactions and reproductive tactics in red-necked wallabies (*Macropus rufogriseus banksianus*). J Zool 217:267–280.

Johnson CN, Johnson KA (1983) Behaviour of the bilby, *Macrotis lagotis* (Reid), (Marsupialia, Thylacomyidae) in captivity. Aust Wildlife Res 10:77–87.

Johnson SD, McGowan MR, O'Callaghan PO, Cox R, Nicolson V (2000) Studies of the oestrous cycle, oestrus and pregnancy in the koala (*Phascolarctos cinereus*). J Reprod Fert 120: 49–57.

Johnston LA, Lacy RC (1995) Genome resource banking for species conservation:selection of sperm donors. Cryobiology 32:68–77.

Kraaijeveld-Smit FJL, Ward SJ, Temple-Smith PD (2003) Paternity success and the direction of sexual selection in a field population of a semelparous marsupial, *Antechinus agilis*. Mol Ecol 12:475–484.

Kraaijeveld-Smit FJL, Ward SJ, Temple-Smith PD, Paetkau D (2002) Factors influencing paternity success in *Antechinus agilis*: last-male sperm precedence, timing of mating and genetic compatibility. J Evol Biol 15:100–107.

Lacy RC (1989) Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. Zoo Biol 8:111–123.

Lee AK, Cockburn A (1985) The Evolutionary Ecology of Marsupials. Cambridge University Press, Cambridge, MA.

Loebel DA, Nurthen RK, Frankham R, Briscoe DA, Craven D (1992) Modeling problems in conservation genetics using *Drosophila* populations: consequences of equalising founder representation. Zoo Biol 11:319–332.

McKenzie LM, Cooper DW (1997) Hybridisation between tammar wallaby (*Macropus eugenii*) populations from Western and South Australia. Heredity 88:398–400.

McLean IG, Cameron EZ, Linklater WL, Schmitt NT, Pulskamp KSM (2009) Partnerships in the social system of a small macropod marsupial, the quokkas (*Setonix brachyurus*). Behaviour 146:89–112.

McLean IG, Schmitt NT (1999) Copulatory behaviour in the quokka, *Setonix brachyurus*. Aust Mammal 21:139–142.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial Monodelphis domestica reveals innovation in non-coding sequences. Nature 447:167–177.

Miller EJ, Eldridge MDB, Herbert CA (2010a) Dominance and paternity in the tammar wallaby (*Macropus eugenii*). In: Coulson G, Eldridge MDB (eds) Macropods: The Biology of Kangaroos, Wallabies and Rat-Kangaroos. CSIRO Publishing, Melbourne, pp. 77–86.

Miller EJ, Eldridge MDB, Cooper DW, Herbert CA (2010b) Dominance, body size and internal relatedness influence male reproductive success in eastern grey kangaroos (*Macropus giganteus*). Reprod Fert Dev 22(3):539–549.

Montgomery ME, Ballou JD, Nurthen RK, England PR, Briscoe DA, Frankham R (1997) Minimizing kinship in captive breeding programs. Zoo Biol 16:377–389.

Myers N, Knoll AH (2001) The biotic crisis and the future of evolution. Proc Natl Acad Sci USA 98:5389–5392.

Nielsen RK, Pertoldi C, Loeschcke V (2007) Genetic evaluation of the captive breeding program of the Persian wild ass. J Zool 272:349–357.

Nunney L (1991) The influence of age structure and fecundity on effective population size. Proc Roy Soc Lond B Biol Sci 246:71–76.

Nunney L (1993) The influence of mating system and overlapping generations on effective population size. Evolution 47:1329–1341.

Nunney L, Elam DR (1994) Estimating the effective population size of conserved populations. Conserv Biol 8:175–184.

Oyama K, Nojima M, Shojo M, Fukushima M, Anada K, Mukai F (2007) Effect of sire mating patterns on future genetic merit and inbreeding in a closed beef cattle population. J Anim Breed Genet 124:73–80.

Pope LC, Blair D, Johnson CN (2005) Dispersal and population structure of the rufous bettong, Aepyprymnus rufescens (Marsupialia: Potoroidae). Austral Ecol 30:572–580.

Ralls K, Ballou J (1986) Captive breeding programs for populations with a small number of founders. Trends Ecol Evol 1:19–22.

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. Conserv Biol 17:230–237.

Renfree MB (1979) Initiation of development of diapausing embryo by mammary denervation during lactation in a marsupial. Nature 278:549–551.

Renfree MB (1993) Endocrinology of pregnancy, parturition and lactation in marsupials. In: Lamming GE (ed) Marshall's Physiology of Reproduction. Chapman and Hall, London.

Renfree MB, Shaw G (2000) Diapause. Annu Rev Physiol 62:353–375.

Rudd CD (1994) Sexual behaviour of male and female tammar wallabies (*Macropus eugenii*) at post-partum oestrus. J Zool 232:151–162.

Rudnick JA, Lacy RC (2008) The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. Conserv Genet 9:1439–1450.

Russell EM (1984) Social behaviour and social organization of marsupials. Mammal Rev 14: 101–154.

Ryman N, Jorde PE, Laikre L (1995) Supportive breeding and variance in effective population size. Conserv Biol 9:1619–1628.

Ryman N, Laikre L (1991) Effects of supportive breeding on the genetically effective population size. Conserv Biol 5:325–329.

Sadlier RM, Tyndale-Biscoe CH (1977) Photoperiod and the termination of embryonic diapause in the marsupial *Macropus eugenii*. Biol Reprod 16:605–608.

Sharman GB, Calaby JH, Poole WE (1966) Patterns of reproduction in female diprotodont marsupials. In: Rowlands IW (ed) Comparative Biology of Reproduction in Mammals. Academic Press, London.

Shaw G, Rose RW (1978) Delayed gestation in the Potoroo *Potorous tridactylus* (Kerr). Aust J Zool 27:901–912.

Siddle HV, Kreiss A, Eldridge MDB, et al. (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA 104:16221–16226.

Sigg DP, Goldizen AW, Pople AR (2005) The importance of mating system in translocation programs: reproductive success of released male bridled nailtail wallabies. Biol Conserv 123:289–300.

Soulé ME, Gilpin M, Conway W, Foose T (1986) The millennium ark: how long a voyage, how many staterooms, how many passengers? Zoo Biol 5:101–113.

Spencer PBS, Horsup AB, Marsh HD (1998) Enhancement of reproductive success through mate choice in a social rock-wallaby, *Petrogale assimilis* (Macropodidiae) as revealed by microsatellite markers. Behav Ecol Sociobiol 43:1–9.

Spielman D, Frankham R (1992) Modeling problems in conservation genetics using captive *Drosophila* popualations: improvement of reproductive fitness due to immigration of one individuals into small partially inbred populations. Zoo Biol 11:343–351.

Stead-Richardson E, Bradshaw D, Friend T, Fletcher T (2010) Monitoring reproduction in the critically endangered marsupial, Gilbert's potoroo (*Potorous gilbertii*): Preliminary analysis of faecal oestradiol-17b, cortisol and progestagens. Gen Comp Endocr 165(1): 155–162.

Strahan R (2002) The Mammals of Australia. Reed New Holland, Sydney

Sugg DW, Chesser RK (1994) Effective population sizes with multiple paternity. Genetics 137:1147–1155.

Taggart DA, Breed WA, Temple-Smith PD, Purvis A, Shimmin G (1998) Reproduction, mating strategies and sperm competition marsupials and monotremes. In: Birkhead TR, Møller AP (eds) Sperm Competition and Sexual Selection. Academic Press, London.

Tallmon DA, Luikart G, Waples RS (2004) The alluring simplicity and complex reality of genetic rescue. Trends Ecol Evol 19:489–496.

Taylor AC, Cowan PE, Fricke BL, Cooper DW (2000) Genetic analysis of the mating system of the common brushtail possum (*Trichosurus vulpecula*) in New Zealand farmland. Mol Ecol 9:869–879.

TDSC (2007) Insurance Population Strategy. Save the Tasmanian Devil Steering Committee.

Tyndale-Biscoe CH (2005) Life of Marsupials. CSIRO Publishing, Collingwood.

Tyndale-Biscoe CH, Hearn JP, Renfree MB (1974) Control of reproduction in macropodid marsupials. J Endocrinol 63:589–614.

Tyndale-Biscoe CH, Janssens PA (1988) The Developing Marsupial: Models for Biomedical Research. Springer, Heidelberg.

Tyndale-Biscoe CH, Renfree MB (1987) Reproductive Physiology of Marsupials. Cambridge University Press, Cambridge, MA.

Tyrell J (2001) The reproductive biology of the female Southern Hairy-nosed Wombat, *Lasiorhinus latifrons*. Melbourne University, Victoria.

Waples RS (1991) Pacific salmon, oncorhynchus spp., and the definition of “species” under the endangered species act. Mar Fish Rev 53:11.

Wedekind C (2003) Pathogen-driven sexual selection for “good genes” versus genetic variability in small populations. In: Bianchet-Festa M, Apollonio M (eds) Animal Behaviour and Wildlife Conservation. Island Press, Washington.

Wittenberger JF (1979) The evolution of mating systems in birds and mammals. In: Marler P, Vandenbergh JG (eds) Handbook of Behavioral Neurobiology. Plenum, New York, NY.

Wittenberger JF (1981) Animal Social Behaviour. Duxbury Press, Boston, MA.

Worthington Wilmer J, Allen PJ, Pomeroy PP, Twiss SD, Amos W (1999) Where have all the fathers gone? An extensive microsatellite analysis of paternity in the grey seal (*Halichoerus grypus*). Mol Ecol 8:1417–1429.

Wright S (1969) Evolution and the Genetics of Populations: The Theory of Gene Frequencies. University of Chicago Press, Chicago, IL.

Wright M, Stott P (1999) The Kangaroo Island Tammar Wallaby: Assessing Ecologically Sustainable Commercial Harvesting. Rural Industries Research and Development Corporation, Kingston, Canberra.

Zenger KR, McKenzie LM, Cooper DW (2002) The first comprehensive genetic linkage map of a marsupial: the tammar wallaby (Macropus eugenii). Genetics 162:321–330.

Part II
Marsupial Chromosomes and Gene Maps

Janine E. Deakin

The earliest work on marsupial genetics focussed on characterising marsupial chromosomes. Their large size and low diploid numbers have made them ideal to study by cytogenetic methods. The remarkable conservation of chromosome number and chromosome morphology observed by the pioneers in marsupial chromosome research (David Hayman and Peter Martin) showed that marsupial chromosomes have changed little since divergence from a common ancestor (Hayman and Martin, 1969); much less than have eutherian mammals. In this current era of marsupial genomics, we have the ability to delve much deeper into the characteristics of marsupial chromosomes, permitting a more thorough investigation of karyotypic and sequence conservation among divergent species, and allowing a much more penetrating analysis of chromosome evolution in this branch of mammals.

In the first chapter of this part (Chapter 2), Rens and Ferguson-Smith recount the history of early marsupial chromosome studies, and review how the development of chromosome painting confirmed the findings of G-banding studies. In this beautiful technique, DNA isolated from a particular chromosome from one species can be labelled and hybridised to the chromosome of another species to visualise directly regions of homology. Remarkably, chromosome painting demonstrated that karyotypic diversity among marsupials can be attributed to the rearrangement of just 19 conserved chromosomal segments. With the release of marsupial genome data, "e-painting", a method combining information from chromosome painting of some species with whole genome sequence data from others, has emerged as an *in silico* approach for determining chromosome evolution and reconstructing ancestral karyotypes.

Within marsupials, the macropods (kangaroos and wallabies) display the most karyotypic diversity, largely as the result of repeated Robertsonian fusions between chromosome arms. In Chapter 3, Carone and O'Neill discuss how centromeres have played an important role in marsupial, particularly macropod, chromosome evolution by centric fusion. An endogenous retrovirus known as KERV (kangaroo endogenous retrovirus) appears to play a major role in chromosome evolution, being particularly prevalent at centromeres, telomeres and evolutionary breakpoints. Certain repeat elements associated with KERV appear to be important for transcription of centromeric sequences. Carone and O'Neill also review how

invaluable macropod interspecies hybrids have been for gaining insight into hybrid dysgenesis. Continued sequencing efforts and analysis should lead to a greater understanding of marsupial centromeres and the role they play in genome evolution.

Chromosome painting, described in Chapter 2, provides information on the global homology between chromosomes from different marsupial species, but provides no information on the genic content or order within these chromosomes. For many years now, it has been of great interest to construct gene maps in order to determine gene content of marsupial chromosomes and compare the arrangement of these genes between marsupials and other vertebrates. Among the first genes to be mapped in any marsupial were several on the kangaroo X chromosome, identified by their X-linked mode of inheritance in pedigrees. Autosomal linkage groups were established later in the dunnart, then opossum species. Methods have progressed since these X-linked genes were mapped in marsupials. Microsatellites (anonymous sequence markers), account for most markers on marsupial linkage maps. Mapping these markers provides valuable information on recombination events, but does not provide information on gene content of marsupial chromosomes. In Chapter 4, Samollow reviews the history and current status of marsupial linkage mapping and explains the peculiarities of these recombination events in marsupials and the potential applications of linkage mapping data.

Linkage maps are vital for integrating phenotypic information with genetic markers. However, most gene mapping in marsupials has been accomplished using physical mapping techniques such as fluorescence in situ hybridisation (FISH). Even the rudimentary maps produced a decade ago have been important in deducing how parts of the human genome, particularly the sex chromosomes, evolved (Graves, 1995).

In Chapter 5, Deakin reviews the essential role physical mapping has played in marsupial genome projects, and describes the immense impact the availability of marsupial genome sequence data has on the ease and speed at which marsupial gene maps can be constructed. By mapping genes found at the ends of evolutionary conserved gene blocks, a detailed picture of the makeup of the tammar wallaby genome has been rapidly generated. The comparison of gene order between species has made it possible to construct comparative gene maps and determine ancestral gene arrangements. Both Samollow (Chapter 4) and Deakin (Chapter 5) explain how linkage and physical maps are being integrated for the opossum and tammar wallaby, and the benefit gained from this integration during the assembly of the genomes of these species.

The chapters in this part detail how technological advances have allowed research on marsupial chromosomes to proceed at a rapid pace. Armed with data on the gross chromosome homologies between species obtained from chromosome painting, genome sequence data and detailed maps, we can now begin to piece together the ancestral marsupial karyotype and gain a greater understanding of the events which have shaped marsupial – and eutherian (including human) – genomes.

References

Graves JAM (1995) The origin and function of the mammalian Y chromosome and Y-borne genes – an evolving understanding. Bioessays 17:311–320.

Hayman DL, Martin PG (1969) Cytogenetics of marsupials. In: Benirschke K (ed) Comparative Mammalian Cytogenetics. Springer Verlag, New York.

Chapter 2
The Conserved Marsupial Karyotype: Chromosome Painting and Evolution

Willem Rens and Malcolm Ferguson-Smith

Abstract Marsupials are an ancient group of mammals that fill diverse niches in Australia, New Guinea, the East Indies and America. These species separated from eutherians around 150 million years ago and the American species diverged from the Australian species around 70 million years ago. Compared to other mammalian species, the karyotypes of marsupials are highly conserved; their diploid numbers range from 2n = 10–32 but with a predominance of 2n = 14 or 2n = 22. The first chromosome comparative studies were performed mainly by searching for similarities in G-banding patterns leading to hypotheses on the ancestral marsupial karyotype and the chromosome rearrangement mechanisms that resulted in the karyotypes seen in the extant species. The advent of chromosome painting allowed chromosome comparisons to be based on chromosome-wide sequence similarities, which is a more accurate method than the indirect method of banding analysis. This chapter is divided into six sections. The first section describes early marsupial karyotype studies performed by G-banding and introduces hypotheses on marsupial chromosome evolution. The second explains chromosome painting techniques including flow karyotyping and flow sorting, and presents results in the form of chromosome paint images and chromosome homology maps. The third section describes marsupial chromosome evolution in terms of phylogeny, ancestral karyotypes, chromosome conserved regions, and mechanisms of chromosome rearrangements. The fourth section explains the role of centromere dynamics in marsupial chromosome evolution. The fifth section focuses on recent work on the sequenced genome of the opossum. This section is followed by concluding remarks.

Keywords Chromosome painting · Chromosome evolution · Centromeres · G-banding · Marsupial

W. Rens (✉)
Department of Veterinary Medicine, Cambridge Resource Centre for Comparative Genomics, University of Cambridge, Cambridge, UK
e-mail: wr203@mole.bio.cam.ac.uk

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_2, © Springer Science+Business Media B.V. 2010

2.1 Marsupial Karyotype Comparison by G-Banding

More than 330 marsupial species are found in America, Australia, New Guinea and islands of eastern Indonesia. Examples are the Virginia opossum in North America, the shrew opossums in the South American Andes, the Mexican mouse opossum seen in Central America, the gliders at the northern and eastern coastal regions of Australia, and the bilbies in central Australia. Marsupial species are classified in seven extant orders as shown in Fig. 2.1. These orders are further categorised into families. For instance, Diprodontia is divided into Vombatiformes (wombats and koala), Phalangeroidea (cuscuses, brushtail possums and scaly-tailed possums), Macropodiformes (kangaroos and wallabies), and Petauroidea (gliders). The marsupial phylogeny shown in Fig. 2.1 is based on studies using sequence data from five nuclear genes (Meredith et al., 2009). This figure shows the gliders on top and American opossums at the base. The root of the marsupial tree is between Paucituberculata (shrew opossums) and the other marsupials. However, the location of this root is not certain (Meredith et al., 2009). Amrine-Madsen et al. (2003) prefer the root at the base of Didelphimorphia (opossums), however, this is only weakly supported by molecular data on nuclear genes with fewer taxa. Larger and other types of genomic and molecular data sets are required to determine the root of Marsupialia (Meredith et al., 2009). The 2n chromosome numbers are given to the right of representative species. Some species are not presented with a chromosome number to illustrate that several species still need to be investigated. Three species shown have different male and female chromosome numbers, which is due to an X-autosome or Y-autosome fusion in these species. In two species, *Petauroides volans*, and *Echymipera kalubu*, B-chromosomes have been observed. B-chromosomes are supernumerary chromosomes that differ from the normal chromosome complement by being dispensable, often heterochromatic and exhibiting unusual meiotic behaviour.

A short glance at the chromosome numbers is enough to observe that chromosome evolution has been relatively limited in marsupials when compared to other mammalian species. Two numbers, 2n = 14 and 2n = 22, dominate this list and a debate started between proponents of an ancestral marsupial karyotype with 2n = 14 chromosomes or 2n = 22 chromosomes. The first party base their argument on similarities between the 2n = 14 karyotypes and their occurrence in many families (Rofe and Hayman, 1985), the second uses data on interstitial telomeres to support their hypothesis (Svartman and Vianna-Morgante, 1998).

The first reports on marsupial chromosome studies appeared in the 1920s (see for instance Altmann and Ellery, 1925), but more extensive studies were published from the 1970s and onwards (Hayman and Martin, 1969 and see references in Hayman, 1990) and established most of the diploid numbers presented in Fig. 2.1.

An early study (Murray et al., 1980) compared the karyotype (un-banded or C-banded) of two species of gliders in the Petauroidea order. One species has a diploid number of 2n = 20 (*Pseudocheirus peregrinus*) and one has 2n = 16 (*Pseudochirops dahlia*). The latter was believed to have evolved by Robertsonian fusions (centromere-fusion between two acrocentric chromosomes). This report was

14
Petauroidea
Macropodiformes
Diprotodontia
Phalangeroidea
Vombatiformes
Australasia
Americas
14 Dasyuromorphia
Peramelemorphia
Notoryctemorphia
Microbiotheria
Didelphimorphia
14 Paucituberculata

Acrobates pygmaeus 14
Dactylopsila palpator 18
Gymnobelideus leadbeateri 22
Petaurus breviceps 22
Petaurus norfolcencis 22
Hemibelideus lemuroides 22
Petauroides volans 22 + B
Pseudocheirus peregrinus 20
Pseudochirulus herbertensis 12
Petroseudes dahli 16
Pseudochirops archeri
Tarsipes rostratus 24
Aepyprymnus rufescenc 32 *
Bettongia galmardi 22
Bettongia penicillata 22
Potorous longipes 24
Potorous tridactylous 13m,12f *
Dendrolagus dorianus 12
Dendrolagus matchiei 14 *
Peradorcas concinna 16
Petrogale xanthopus 22 *
Thylogale thetis 22 *
Dorcopsis veterum
Dorcopsis vanheurni 18 *
Lagorchestes conspicillata 15m,16f
Lagorchestes hirsutus 20
Macropus aglilis 16
Macropus irma 16
Macropus bicolor 11m,10f *
Macropus rufus 20 *
Macropus fulginosus 16
Macropus giganteus 16
Setonix brachyurus 22
Onychogalea fraenata 18
Onychogalea unguifera 20
Lagostrophus fasciatus 24
Hypsiprymnodon moschatus 22
Allurops ursinus
Phalanger gymnotis 14
Phalanger orientalis 14
Strigocuscus pelengensis
Spilocuscus maculatus
Trichosurus vulpecula 20 *
Wyulda squamicaudata
Burramys parvus 14
Cercartetus lepidus 14
Vombatus ursinus 14 *
Phascolarctos cinereus 16
Antechinus 14
Phascogale tapoatafa 14
Sminthopsis crassicaudata 14 *
Planigale 14
Myrmecobius fasciatus 14
Echymipera kalubu 14+B
Microperoryctes longicauda 14
Peroryctes raffrayana
Isoodon macrourus 14
Macrotis lagotis 19m,18f
Notoryctes typhlops 20
Dromiciops gliroides 14
Caluromys 14
Didelphis marsupialis 22
Monodelphis domestica 18 *
Caenolestes fuliginosus 14
Rhycholestes raphanurus 14

Fig. 2.1 Marsupial phylogeny. Representative species are shown on the *right*. The root of the marsupial tree is between Paucituberculata and the other marsupials and is shown on the *left*. The number “14” refers to the 2n = 14 ancestral karyotype explained in the text. “*” refers to species used in chromosome painting studies

followed by karyotype studies on six species of the Petauroidea order (Murray et al., 1990).

Direct (no banding) karyotype comparison was performed for 19 Didelphimorphia species (Carvalho et al., 2002). Eight species have a diploid number of 2n = 14, four species have 2n = 18, and seven species have a 2n = 22 diploid number. These findings together with studies performed by Reig et al. (1977) show that 2n = 14 is predominant (58%) in these south American marsupials species followed by 2n = 22 (27%); the remaining species (15%) have a diploid number of 2n = 18. These are the only diploid numbers found in American marsupials. Reig et al. (1977) hypothesized that the 2n = 18 and 2n = 22 karyotypes evolved from the 2n = 14 through a process of Robertsonian fissions.

One of the first papers presenting and comparing marsupial G-banded chromosomes was from Rofe and Hayman (1985). Eight Dasyuromorphia species were compared with two Peramelemorphia, four Vombatiformes, and one Didelphimorphia species, all with 2n = 14. This comparison showed that the karyotypes of these species were very similar with almost identical banding and observed differences could be explained by pericentric inversions. This similarity indicates that a 2n = 14 karyotype was present at least around 70 million years ago when Didelphimorphia and Australidelphia diverged. Alternatively, a similar 2n = 14 karyotype must have evolved independently in the lineages leading to the species studied (Rofe and Hayman, 1985).

South American marsupials with a diploid number of 2n = 14 also have similar G-banded karyotypes. The four species *Caluromys philander* (bare-tailed woolly opossum), *Caluromys lanatus* (brown-eared woolly opossum), *Marmosa cinerea* (woolly mouse opossum), and *Marmosa murina* (murine mouse opossum) have a high degree of chromosome conservation (Souza de et al., 1990). This similarity strengthened the 2n = 14 pleisiomorphic karyotype hypothesis.

More recently, karyotypes of seven species from the Didelphimorphia order were compared by G- and C-banding. Four of those species have a 2n = 14 karyotype, one (*Monodelphis domestica*, gray short-tailed opossum, Fig. 2.1) has a 2n = 18 karyotype and two have a 2n = 22 karyotype (*Philander Opossum*, gray four-eyed opossum, and *Didelphis marsupialis*, common opossum, Fig. 2.1). Only one mechanism, namely Robertsonian rearrangements, seem to be the cause of change in chromosome numbers in these Didelphimorphia species. However the direction, increase in number by fission or decrease by fusions, of chromosome evolution could not be determined (Svartman and Vianna-Morgante, 1998, 1999).

Thus a 2n = 14 karyotype is favoured by most of the above reports as ancestral as firstly the observed 2n = 14 karyotypes are very similar and secondly, they are found in different marsupial families from both South America and Australia. Other karyotypes evolved through fissions. The only objection to this hypothesis comes from observations of interstitial telomeres, with their existence indicating telomeric fusions. This points to 2n = 22 as the representative of the ancestral karyotype, with other karyotypes evolving through fusions (Svartman and Vianna-Morgante, 1998; Carvalho and Mattevi, 2000).

G-banding comparison in rock wallabies was important to shed light on chromosome evolution in the superfamily Macropodiformes (Fig. 2.1) with high chromosome variability (Eldridge et al., 1992a, b). Comparison between six rock wallaby species demonstrated a 2n = 22 karyotype that is plesiomorphic for this superfamily and is retained in *Thylogale thetis* (red-necked pademelon). It is important to note that this 2n = 22 karyotype is not similar to the 2n = 22 karyotype seen in other marsupial superfamilies.

Figure 2.1 shows that *Potorous tridactylus* (long-nosed potoroo) contains an X-autosome fusion causing two different diploid numbers in male and female (Johnston et al., 1984). Sinclair et al. show by G-banding that this karyotype is almost identical to *Potorous gilbertii* (Gilbert's potoroo) indicating that this fusion was not species specific, which had implications for the Potorous phylogeny (Johnston et al., 1984; Sinclair et al., 2000). The 2n = 24 karyotype of *Potorous longipes* (long-footed potoroo) is seen as ancestral for Potorous species.

A beautiful collection of 55 G-banded marsupial chromosomes can be found in the Atlas of Mammalian Chromosomes (Eldridge and Metcalfe, 2006). Karyotypes of ten Didelphimorphia species can be compared with ten Dasyuromorphia, twenty-nine Diprotondia, three Peramelemorphia and one species each of Paucituberculata, Microbiotheria and Notoryctemorphia. The set of G-banded marsupial karyotypes produced by Rofe for his PhD investigations deserves special mention for its high quality.

G-banded karyotypes have been very instrumental in gaining insights into chromosome evolution but this method is based on chromosome morphology and appearance. A more unambiguous method for whole genome comparison based on direct DNA comparison is cross-species chromosome painting discussed in the next section.

2.2 Marsupial Karyotype Comparison by Cross-Species Chromosome Painting

2.2.1 History

Cross-species chromosome painting became possible after a technique was developed for labelling specific chromosomes with a fluorescence tag. Two requirements had to be fulfilled in order to accomplish this technique. First, each chromosome type of a species of interest needed to be collected in a separate tube, i.e. for human there would be 24 different tubes. The first flow cytometer that could separate and sort human chromosomes into 21 fractions was presented in 1982 by Lebo and Bastian (1982). This was a major step forwards in cytology. Since then improvements have been made concerning accuracy and speed but these are only adaptations of the original design. The second requirement for chromosome painting was the method for DNA labelling with equal coverage along the length of the chromosome. This method was developed by Hakan Telenius (Telenius et al., 1992) 10 years after

the development of chromosome sorting. General amplification was achieved by using a degenerate primer together with a specific PCR protocol in which initial primer annealing occurs at a low temperature. Labelling of DNA occurs in a second round of PCR amplification by incorporating a hapten. Immediately it was realised that this process is species independent and that chromosome specific DNA and paints can be produced for a large variety of vertebrate species.

2.2.2 Marsupial Chromosome Painting

Studies into marsupial chromosome evolution started to benefit from chromosome painting after 1997. Several reports were published focussing on different parts of the marsupial phylogenetic tree. The detected chromosome rearrangements were subsequently placed on branches of this tree. In chronological order these reports are the following.

In 1997 chromosome paints were produced from tammar wallaby (*Macropus eugenii*, 2n = 16) sorted chromosomes and these were hybridized onto metaphases of the swamp wallaby (*Wallabia bicolor*, 2n = 11♂,10♀). Chromosome painting revealed that the reduction in chromosome number in the swamp wallaby was caused by fusions of "tammar" chromosomes. For instance, swamp wallaby chromosome 1 consists of a region homologous to tammar chromosome 4, a region on tammar chromosome 1 and a region on tammar chromosome 6 (Toder et al., 1997).

Tammar wallaby chromosome paints were also painted on three rock wallaby species metaphases (all three 2n = 22) to shed light on chromosome evolution within *Petrogale* species and detect conserved chromosome regions between members of *Macropus* and *Petrogale*. The karyotypes of *Petrogale lateralis lateralis*, *P. l. pearsoni* and *P. xanthopus xanthopus* are identical: they have identical paint distributions after hybridization with tammar chromosome paints. Each *Petrogale* chromosome corresponds to only one tammar chromosome (region). Tammar chromosomes are thus a result of three fusions of six chromosomes of the 2n = 22 karyotype, if the 2n = 22 Petrogale karyotype is regarded as pleisiomorphic for *Petrogale* and *Macropus* (O'Neill et al., 1999).

Cross-species chromosome painting becomes more informative when hybridization is performed bidirectionally (reciprocally): chromosomes of species A are hybridized on metaphases of species B and vice versa. A report in 1999 (Rens et al., 1999) shows results of bidirectional chromosome painting between four distantly related marsupial species: two Macropodiformes, one Phalangeroidea and one Dasyuromorphia species. The latter species is *Sminthopsis crassicaudata* (fat-tailed dunnart), which has a karyotype that, apart from inversions, represents the putative 2n = 14 ancestral karyotype. Five different chromosome homology maps were produced, each important for comparative gene mapping and for showing the limited number of chromosome rearrangements which have occurred between these species. A karyotype phylogeny was constructed from these data. The karyotype of *Trichosurus vulpecula* (brushtail possum; Phalangeridea, Fig. 2.1) can be derived from the 2n = 14 ancestral karyotype by four fissions and one fusion. The

Macropodiformes plesiomorphic 2n = 20 karyotype is related to the 2n = 14 ancestral karyotype by four fissions and one fusion. The chromosome complement of *Potorous tridactylus* evolved after one fission and five fusions. The 2n = 22 karyotype that is ancestral to Macropods and Petrogale (see above) is derived after one fission. The *Macropus eugenii* complement evolved after three fusions. The *S. crassicaudata* karyotype is related to the 2n = 14 marsupial ancestral karyotype by six inversions. Fifteen chromosome segments are conserved in the karyotypes of these Australian marsupials (Rens et al., 1999), but this number has been increased to nineteen based on studies on more marsupial species, see below (Rens et al., 2003).

Tammar wallaby chromosome specific paints were used to investigate the relatively large karyotypic diversity within Macropodiformes (Glas et al., 1999). These paints were hybridized onto metaphases of *Macropus rufus*, *Dendrolagus matschiei*, *Thylogale thetis* and an unknown *Dorcopsis* species. This study showed that the *Thylogale thetis* and *Dorcopsis* karyotypes are identical to the *Petrogale* 2n = 22 karyotype. All three karyotypes can be considered as representatives of the Macropodid plesiomorphic karyotype. Although the *Dendrolagus matschiei* karyotype has a diploid number of 2n = 14, it is not similar to the 2n = 14 marsupial ancestral, but derived through independent fusions from the 2n = 22 Macropodid plesiomorphic karyotype. Only one fusion separates *Macropus rufus* from this 2n = 22 karyotype.

Sminthopsis macroura and *Vombatus latifrons* are marsupials species from different orders, Diprodontia and Dasyuromorphia, with both 2n = 14 karyotypes. Painting metaphases of these species with tammar wallaby chromosome paints revealed that these karyotypes are closely related and that each differs from the ancestral 2n = 14 karyotypes by a different set of two inversions (De Leo et al., 1999).

In 2001 the American species *Monodelphis domestica* (2n = 18) was added to the list of marsupial species studied by cross-species chromosome painting (Rens et al., 2001). Chromosome specific paints of this species were hybridized to *Sminthopsis crassicaudata* (2n = 14), *Macropus eugenii* (2n = 16) and *Trichosurus vulpecula* (2n = 22) and vice versa. Fourteen of the fifteen chromosome segments that are conserved in *S. crassicaudata*, *M. eugenii*, *P. tridactylus*, and *T. vulpecula* (see above) are also conserved in the karyotype of *Monodelphis domestica* although 70 million years separate the American *Didelphis* species from Australian marsupial species. *M. domestica* chromosome 1, 2, 5 and 8 are homologous to *S. crassicaudata* chromosome 2, 4, 6, and 5 respectively. Chromosome 4 and 6 of *M. domestica* together form *S. crassicaudata* chromosome 1. *S. crassicaudata* chromosome 3 is homologous to *M. domestica* chromosome 4 and 7. Apart from inversions, the karyotype of *Sminthopsis* represents the 2n = 14 marsupial ancestral karyotype.

One species *Aepyprymnus rufuscens* (2n = 32) within the Macropodiformes family has a noticeable higher diploid number. Chromosome evolution studies using *A. rufuscens* specific paints delivered higher resolution chromosome wide homology maps (Rens et al., 2003). Figure 2.2 shows an example of cross species chromosome painting: a paint of *A. rufuscens* chromosome 9 was hybridized on *P. tridactylus* chromosome 2, the region highlighted is conserved region C16,

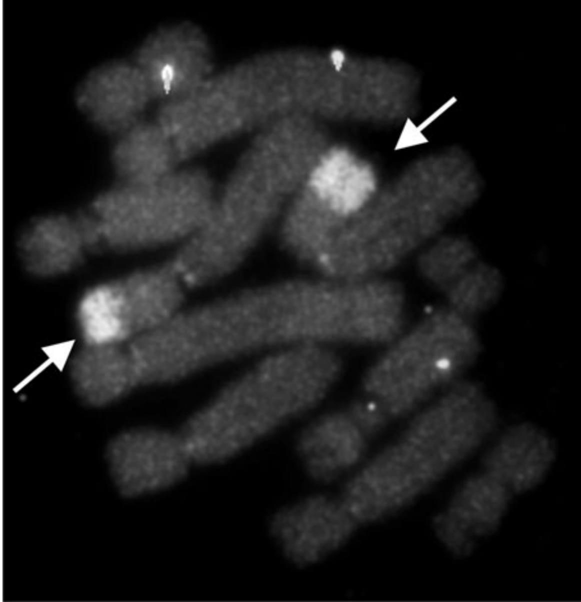


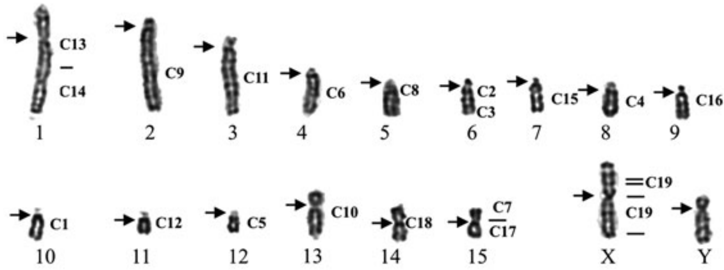
Fig. 2.2 Cross-species chromosome painting: a paint of *Aepyprymnu rufuscens* chromosome 9 was hybridized on *Potorous tridactylus* chromosome 2p

see below. Figure 2.3 presents a chromosome homology map of six marsupial species: *M. domestica*, *S. crassicaudata*, *T. vulpecula*, *M. eugenii*, *P. tridactylus*, *A. rufuscens*. The chromosome painting results and the homology maps resulted in a karyotype phylogeny discussed further in the next section.

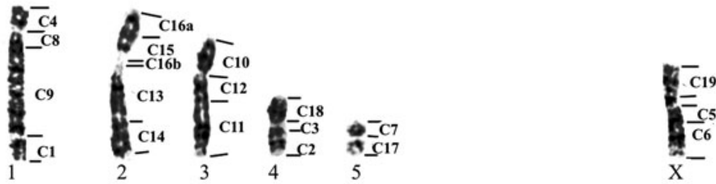
2.3 Marsupial Karyotype Phylogeny and Chromosome Evolution

Whole genome comparisons by G-banding and by chromosome painting as described in the previous two sections (Sections 2.1 and 2.2) can now be combined to describe the chromosome rearrangements that took place during marsupial chromosome evolution and to establish ancestral karyotypes. All the diversity of karyotypes can be accounted for by the rearrangement of just 19 evolutionary conserved regions (Rens et al., 2003). A karyotype of six pairs of autosomes together

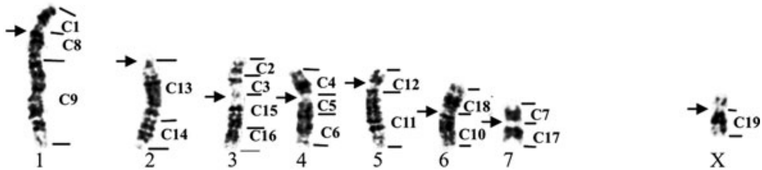
Fig. 2.3 Chromosome homology map of six marsupial species: *Monodelphis domestica*, *Sminthopsis crassicaudata*, *Trichosurus vulpecula*, *Macropus eugenii*, *P. tridactylus*, *A. rufuscens*. The *numbers* to the *right* of each chromosome refer to the 19 conserved regions, see text. The *arrows* point to centromeres



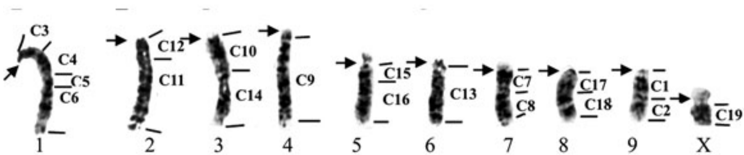
Aepyprymnus rufescens, rufous bettong



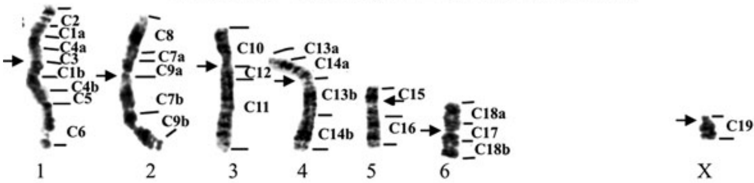
Potorous tridactylus, long-nosed potoroo



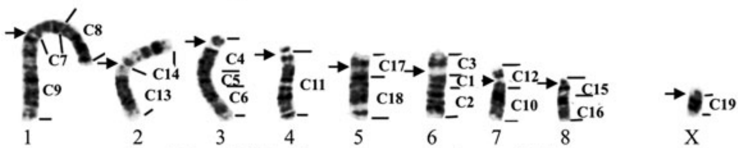
Macropus eugenii, tammar wallaby



Trichosurus vulpecula, common brushtail possum



Sminthopsis crassicaudata, fat-tailed dunnart



Monodelphis domestica, gray short-tailed opossum

with the sex chromosomes (2n = 14) can be regarded as ancestral for all marsupials: each marsupial order, including Paucituberculata that first branched from the marsupial root, has representatives with a 2n = 14 karyotype. Svartman and Vianne-Morgante (1998) together with Carvalho and Mattevi (2000) disputed the 2n = 14 ancestral karyotype based on their findings of interstitial telomeres in both a 2n = 14 and a 2n = 18 South-American marsupial, which indicate a set of fusions up in the evolution tree of these species (i.e. diverged and not ancestral). These authors favoured the 2n = 22 didelphid karyotype to be ancestral to all marsupial karyotypes. However, as this view is an exception (but see Section 2.6), chromosome evolution starting from a 2n = 14 ancestral karyotype is described here below.

Figure 2.4 shows the 2n = 14 karyotype with a list of marsupial species from both continents which, apart from inversions, have retained this karyotype. The chromosomes contain regions numbered 1–19 that are conserved in all karyotypes shown. Chromosome evolution is discussed starting with the South American species and then upwards in the list of species at the right in Fig. 2.1.

The chromosomes of the South American opossum *Monodelphis domestica* 2n = 18 evolved from the 2n = 14 ancestral karyotype through two fissions. One fission occurred between region C3 and C4 and one fission separated region C11 from region C10 and C12. All other chromosomes are conserved in *Monodelphis domestica*. The chromosomes of *Didelphis marsupialis* have evolved further by an additional set of two fissions, one between region C8 and C9 and one between region C13 and C14. No fusions have occurred and thus *Monodelphis domestica* chromosome number has increased from 2n = 14 to 2n = 18 and *Didelphis marsupialis* to 2n = 22. Both species are members of the Didelphimorphia order.

Chromosomes of both orders Microbiotheria and Notoryctemorphia have not been studied in detail. *Isoodon* (Peramelemorphia), *Sminthopsis* (Dasyuromorphia), and *Vombatus* (Vombatiformes) have retained a 2n = 14 karyotype (see previous paragraphs).

A set of four fissions and one fusion created the 2n = 20 karyotype of *Trichosurus vulpecula* (Phalangeroidea). As can be seen in Fig. 2.4 these rearrangements are not all specific for *Trichosurus vulpecula*. The fission between region C8 and C9 creating a chromosome containing region C9 and a chromosome containing regions C7 and C8 has occurred in the *Trichosurus vulpecula* lineage, the *Didelphis marsupialis* lineage, and *Aepyprymnus rufuscens* lineage. The fission between region C13 and C14 is also observed in the *D. marsupialis* lineage. The separation of region C10 and region C11.12 is shared between *T. vulpecula* and the 2n = 22 macropodid pleisiomorphic karyotype, see below. The other two rearrangements are specific for *T. vulpecula*.

Six rearrangements, five fissions and one fusion, formed the macropodid plesiomorphic 2n = 22 karyotype. The fission between region C3 and region C4 is the fifth example of a convergent event, it occurred in both the *Monodelphis domestica* lineage and the 2n = 22 plesiomorphic karyotype. The fission separating region C10 is also seen in *T. vulpecula*, see above. The fissions between region C1 and C2, between C7 and C8.9, and between C17 and C18 are specific

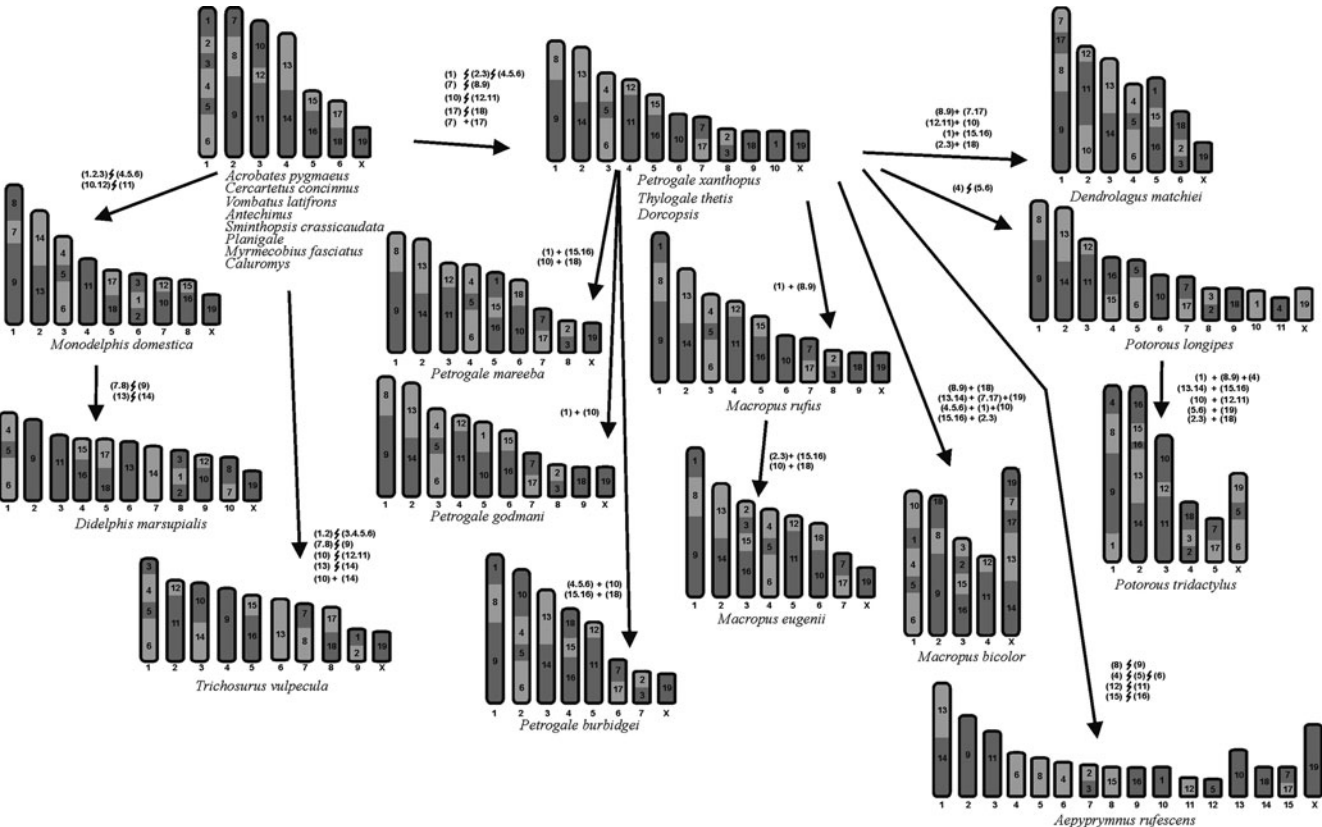


Fig. 2.4 Marsupial chromosome evolution. The 2n = 14 ancestral karyotype is shown at the *top* on the *left* with a list of species that retained this karyotype. The *numbers* refer to the 19 conserved regions. Chromosome rearrangements are indicated next to the *arrow*. A "+" sign indicates fusions, the "*lightning symbol*" indicates fissions

for this lineage. The fusion between region C7 and C17 formed a chromosome that is conserved in all macropodid marsupials studied. The macropodid plesiomorphic is retained in at least the three marsupials species listed under this karyotype in Fig. 2.4. Note that this 2n = 22 karyotype is not similar to the 2n = 22 *D. marsupialis* karyotype. A pre-dominance of 2n = 22 should not be an argument for favouring this karyotype to be ancestral for all marsupials.

The matschie's tree-kangaroo, also known as the huon tree-kangaroo (*Dendrolagus matschiei*) has a 2n = 14 karyotype that should not be confused with the 2n = 14 ancestral karyotype. Figure 2.4 shows that these two karyotypes are different. *D. matchiei* karyotype evolved from the 2n = 22 macropodid pleisiomorphic karyotype through four fusions. Two of those may be reversal events. The fusion between the C8.9 region and C7.17 region may directly connect region C7 with region C8 as it was in the 2n = 14 ancestral karyotype. However the orientation of the C7.17 region on *D. matchiei* chromosome 1 is yet unknown and region C17 may be connected to region C8 as presented in Fig. 2.4. A similar argument regards the fusion between region C10 and C12 as this also may be a reversal event. Region C10 and C12 are connected in the 2n = 14 ancestral karyotype (chromosome 3) and possible in *D. matschiei* chromosome 2. The other two fusions are specific for this tree kangaroo.

Potorous longipes chromosome 5 and 11 were formed after fission between region C4 and C5 in chromosome 3 of the macropodid pleisiomorphic karyotype. All other *P. longipes* chromosomes are conserved. *Potorous tridactylus* has a reduced number of chromosomes due to a set of six fusions, including one that created the relatively large X chromosome. One of these fusions is a reversal: the fusion between region C10 and C12 reconnects these two regions after the fission between the 2n = 14 and 2n = 22 karyotypes.

Five fissions in the chromosomes of the macropodid ancestor created the 2n = 32 karyotype of the rufous bettong, *Aepyprymnus rufuscens*. The separation of region C8 and C9 was a convergent event as it was also observed in the *D. marsupialis* and *T. vulpecula* lineage. The two fissions that separated regions C4, C5 and C6 were exceptional as the C4.5.6 chromosome region is conserved in all marsupials except in the potoroo. The fission between region C4 and C5 happened both in the *A. rufuscens* and *P. longipes* lineages and is thus a convergent event. The fission of region C12 from C11 is an additional convergent event as it occurred also in the *M. domestica* lineage. The fission between region C15 and C16 is specific for *A. rufuscens*.

The karyotype of *Macropus rufus* evolved by one fusion from the 2n = 22 macropodid pleisiomorphic karyotype. The tammar wallaby (*Macropus eugenii*) karyotype was formed after two fusions, the fusion between the region C2.3 and C15.16 is also seen in the Macropus bicolor lineage. The fusion between C10 and C18 is also observed in *Petrogale mareeba*, see below.

The rock wallabies have a relative high rapidity of chromosome rearrangements when compared to other marsupial families. The macropodid pleisiomorphic karyotype is retained in *Petrogale xanthopus*, *P. lateralis*, and *P. rothschildi*, but other rock wallaby karyotypes evolved after a set of fusions and centromere shifts

(Eldridge and Close, 1993). Three examples are given in Fig. 2.4. Two fusions, one between region C1 and C15.16, and one between C10 and C18 formed the *P. mareeba* karyotype. The latter is a convergent event as the same is seen in the *Macropus eugenii* lineage. A fusion between region C1 and C10 is observed in the karyotype of *P. godmani*. The karyotype of *P. burbidgei* is formed after two fusions, between region C10 and C4.5.6, and between region C15.16 and C18.

In summary, nine recurrent (convergence) and three reversals are observed during chromosome evolution of the marsupials in this review. These events may be a reason to reconsider the order of the sequential rearrangements in the phylogenetic tree. However, Rens et al. (2003) have shown that reordering will produce different reversals or convergences.

The Petauroidea remain to be studied by chromosome painting. *Acrobates pygmaeus* represent the 2n = 14 ancestral karyotype (as shown by G-banding). It will be of interest to investigate the extent of convergence and reversals in Petauroidea chromosome evolution.

2.4 Centromere Dynamics in Chromosome Evolution

A relatively high number of chromosome rearrangements that shaped the marsupial karyotypes can be classified as convergent or reversal events, as mentioned in Section 2.3. Marsupial chromosome rearrangements are therefore not ideal to be used as parameters for construction of phylogenetic trees. An explanation for reuse of breakpoints may be found by investigating the centromere sequence composition. That centromeres can influence chromosome stability was shown by characterising the centromeres of *Macropus* hybrids (Metcalfe et al., 2007). It appeared that some centromeres were elongated and segmented in the hybrid and that specific sequences were amplified at the centromeres, one of them is a kangaroo endogenous retrovirus, KERV-1 (Ferreri et al., 2004; see also Carone and O'Neill, Chapter 3). A correlation between centromere sequences and convergent breakpoint reuse can be determined by mapping both parameters in the marsupial phylogenetic tree. It appears that expansion or diminishment of centromere sequences matches convergent breakpoint reuse at centromeres, i.e. centromere-satellite sequence evolution follows chromosomal evolution (Bulazel et al., 2007). Thus the observation of convergence and reversal in marsupial chromosome evolution actually has a biological explanation; chromosome rearrangements are not necessarily random events. The reuse of marsupial breakpoints extends beyond marsupials and even into human karyotype evolution/instability. Longo et al. (2009) shows that both centromeres and evolutionary breakpoints are enriched for specific interspersed repeats, endogenous retroviruses and long interspersed nuclear elements, linking centromeres with chromosome instability. An evolutionary breakpoint on tammar wallaby chromosome 1q is orthologous to human 14q32.33, which itself is an evolutionary breakpoint; both sites are enriched for the above types of repeats. Activity of these repeat elements determines the instability observed at these locations.

2.5 Marsupial Whole Genome Sequencing

In the last few years chromosome painting has been joined by "E-painting" to discover chromosome homologies and rearrangements. The latter technique became possible with the availability of whole genome sequences of many vertebrate species. The tammar wallaby is currently sequenced with 2x coverage and the opossum (*Monodelphis domestica*) has been sequenced with 6x coverage, and assigned to chromosomes (Mikkelsen et al., 2007). These sequencing endeavours provide us with a wealth of comparative genomic data, which is covered elsewhere in this book. E-painting uses the sequence data to discover regions of conserved synteny between species. Although it relies on "computer" data, the resolution is much greater than that obtained by cross-species chromosome painting. Between marsupial comparisons cannot be done yet as other marsupials await sequencing, but E-painting on eutherian species has been performed (Kemkemer et al., 2009). Six different mammalian species (including opossum) and chicken were compared revealing a large set of 526 evolutionary breakpoints. The eutherian ancestral karyotype (Yang et al., 2003) is supported by the conserved regions detected in opossum as ancestral syntenic segment associations were present in *Monodelphis domestica*.

2.6 Conclusion

Since the 1970s a considerable amount of whole genome comparative data, either by G-banding or chromosome painting, has been collected and used to follow chromosome evolution in different marsupial lineages and to reconstruct ancestral karyotypes. Pioneers in this field had to rely on chromosome banding patterns to establish homology between chromosomes from different marsupial species, which lead to the first hypotheses on marsupial chromosome evolution. Cross-species chromosome painting made these studies more accurate and reliable which enabled a relatively precise description of chromosome rearrangements that shaped the karyotypes of extant marsupial species.

These studies are not yet completed. There are four major marsupial lineages missing in the chromosome painting analysis. Paucituberculata species all have $2n = 14$ karyotypes that presumably correspond to the proposed marsupial ancestral karyotype. Likewise, Microbiotheria have not been studied yet. The missing Notoryctemorphia contain species with karyotypes with higher numbers and could be necessary to support or refute the $2n = 14$ hypothetical ancestral karyotype. The Petauroidea lineage contains species that differ in chromosome number and are interesting to study with respect to convergent and reversal events as discussed in this chapter and with respect to the $2n = 14$ hypothetical ancestral karyotype.

At the current status quo there is still an apparent conflict between the assumption that the ancestral karyotype contained 14 pairs of chromosomes and the observation of interstitial telomeres in the American Didelphimorphia species. The location of these telomeres indicate Robertsonian fusions and thus a reduction in chromosome number from $2n = 22$ to $2n = 14$. It is therefore unfortunate that the root

of marsupials is uncertain. The marsupial phylogeny proposed by Meredith et al. (2009)and used in this chapter presents the Paucituberculata as outgroup species to Didelphimorphia and thus favours the 2n = 14 as ancestral. In contrast, Amrine-Madsen et al. (2003) prefers the root of marsupials at the base of Didelphimorphia. In that case, a 2n = 22 ancestral karyotype implies fusions that are convergent in the Didelphimorphia lineage and the lineage leading to all other marsupials. Studies using non-marsupial species, such as monotremes, as outgroup will resolve this conflict. These studies can only be performed by comparative gene mapping, because these species are too distantly related (150 million years ago) to permit chromosome painting. Precise structural studies at evolutionary breakpoints such as performed by O'Neill's group will also aid in the determination of the direction of chromosome change, i.e. fusion or fission.

Whole genome comparison as discussed in this chapter will be very valuable for and is followed up by gene evolution and comparative functional genomic research. Important examples are the studies of MHC (Siddle et al., 2009; Cheng et al., Chapter 16) and other gene clusters, and studies on the evolution of X-inactivation (Al Nadaf et al., Chapter 13) and genomic imprinting (Edwards et al., 2007, 2008; Renfree et al., 2009; Hore et al., Chapter 12). There is no doubt that evolutionary and developmental biologists are also very interested in marsupial genomics.

References

Altmann SCA, Ellery MEW (1925) The chromosomes of four species of marsupials. Q J Microsc Sci 69:463–469.

Amrine-Madsen H, Scally M, Westerman M, Stanhope MJ, Krajewski C, Springer MS (2003) Nuclear gene sequences provide evidence for the monophyly of australidelphian marsupials. Mol Phylogenet Evol 28:186–196.

Bulazel KV, Ferreri GC, Eldridge MD, O'Neill RJ (2007) Species-specific shifts in centromere sequence composition are coincident with breakpoint reuse in karyotypically divergent lineages. Genome Biol 8:R170.

Carvalho BD, Mattevi MS (2000) (T2AG3)n telomeric sequence hybridization suggestive of centric fusion in karyotype marsupials evolution. Genetica 108:205–210.

Carvalho BD, Oliveira CLFB, Nunes AP, Mattevi MS (2002) Karyotypes of nineteen marsupial species from Brazil. J Mammal 83:58–70.

De Leo AA, Guedelha N, Toder R, et al. (1999) Comparative chromosome painting between marsupial orders: relationships with a 2n = 14 ancestral marsupial karyotype. Chromosome Res 7:509–517.

Edwards CA, Mungall AJ, Matthews L, et al. (2008) The evolution of the DLK1-DIO3 imprinted domain in mammals. PLoS Biol 6:e135.

Edwards CA, Rens W, Clarke O, et al. (2007) The evolution of imprinting: chromosomal mapping of orthologues of mammalian imprinted domains in monotreme and marsupial mammals. BMC Evol Biol 7:157.

Eldridge MD, Close RL (1993) Radiation of chromosome shuffles. Curr Opin Genet Dev 3: 915–922.

Eldridge MD, Johnston PG, Close RL (1992a) Chromosomal rearrangements in rock wallabies, *Petrogale* (Marsupialia: Macropodidae). VI. Determination of the plesiomorphic karyotype: G-banding comparison of *Thylogale* with *Petrogale persephone, P. xanthopus*, and *P. l. lateralis*. Cytogenet Cell Genet 61:29–33.

Eldridge MD, Johnston PG, Lowry PS (1992b) Chromosomal rearrangements in rock wallabies, *Petrogale* (Marsupialia: Macropodidae). VII. G-banding analysis of *Petrogale brachyotis* and *P. concinna*: species with dramatically altered karyotypes. Cytogenet Cell Genet 61:34–39.
Eldridge MD, Metcalfe CJ (2006) Marsupialia. In: O'Brien SJ, Menniger J, Nash W (eds) Atlas of Mammalian Chromosomes. Wiley, New York, NY.
Ferreri GC, Marzelli M, Rens W, O'Neill RJ (2004) A centromere-specific retroviral element associated with breaks of synteny in macropodine marsupials. Cytogenet Genome Res 107:115–118.
Glas R, De Leo AA, Delbridge ML, et al. (1999) Chromosome painting in marsupials: genome conservation in the kangaroo family. Chromosome Res 7:167–176.
Hayman DL (1990) Marsupial cytogenetics. Aust J Zool 37:331–349.
Hayman DL, Martin PG (1969) Cytogenetics of marsupials. In: Benirschke K (ed) Comparative Mammalian Cytogenetics. Springer, New York, NY.
Johnston PG, Davey RJ, Seebeck JH (1984) Chromosome homologies in *Potorous tridactylus* and *P. longipes* (Marsupialia:Macropodidae) based on G-banding patterns. Aust J Zool 32: 319–324.
Kemkemer C, Kohn M, Cooper DN, et al. (2009) Gene synteny comparisons between different vertebrates provide new insights into breakage and fusion events during mammalian karyotype evolution. BMC Evol Biol 9:84.
Lebo RV, Bastian AM (1982) Design and operation of a dual laser chromosome sorter. Cytometry 3:213–219.
Longo MS, Carone DM, Program NC, Green ED, O'Neill MJ, O'Neill RJ (2009) Distinct retroelement classes define evolutionary breakpoints demarcating sites of evolutionary novelty. BMC Genom 10:334.
Meredith RW, Westerman M, Springer MS (2009) A phylogeny of Diprotodontia (Marsupialia) based on sequences for five nuclear genes. Mol Phylogenet Evol 51:554–571.
Metcalfe CJ, Bulazel KV, Ferreri GC, et al. (2007) Genomic instability within centromeres of interspecific marsupial hybrids. Genetics 177:2507–2517.
Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.
Murray JD, Donnellan S, McKay GM, et al. (1990) The chromosomes of four genera of possums from the family Petauridae (Marsupialia;Diprotodonta). Aust J Zool 38:33–39.
Murray JD, Sharman GB, McKay GM, Calaby JH (1980) Karyotypes, constitutive heterochromatin and taxonomy of ringtail opossums of the genus *Pseudocheirus* (Marsupialia: Petauridae). Cytogenet Cell Genet 27:73–81.
O'Neill RJ, Eldridge MD, Toder R, Ferguson-Smith MA, O'Brien PC, Graves JA (1999) Chromosome evolution in kangaroos (Marsupialia: Macropodidae): cross species chromosome painting between the tammar wallaby and rock wallaby spp. with the 2n = 22 ancestral macropodid karyotype. Genome 42:525–530.
Reig OA, Gardner AL, Bianchi NO, Patton JL (1977) The chromosomes of the Didelphidae (Marsupialia) and their evolutionary significance. Biol J Linn Soc 9:191–216
Renfree MB, Hore TA, Shaw G, Graves JA, Pask AJ (2009) Evolution of genomic imprinting: insights from marsupials and monotremes. Annu Rev Genomics Hum Genet 10:241–262.
Rens W, O'Brien PC, Fairclough H, Harman L, Graves JA, Ferguson-Smith MA (2003) Reversal and convergence in marsupial chromosome evolution. Cytogenet Genome Res 102:282–290.
Rens W, O'Brien PC, Yang F, et al. (2001) Karyotype relationships between distantly related marsupials from South America and Australia. Chromosome Res 9:301–308.
Rens W, O'Brien PC, Yang F, Graves JA, Ferguson-Smith MA (1999) Karyotype relationships between four distantly related marsupials revealed by reciprocal chromosome painting. Chromosome Res 7:461–474.
Rofe R, Hayman D. (1985) G-banding evidence for a conserved complement in the Marsupialia. Cytogenet Cell Genet 39:40–50.

Siddle HV, Deakin JE, Coggill P, et al. (2009) MHC-linked and un-linked class I genes in the wallaby. BMC Genom 10:310.

Sinclair EA, Murch AR, Renzo MD, Palermo M (2000) Chromosome morphology in Gilbert's potoroo, *Potorous gilbertii* (Marsupialia:Potoroidae). Aust J Zool 48:281–287.

Souza de MJ, Maia V, Santos dos JF (1990) Nucleolar organizer regions, G- and C-bands in some Brazilian species of Didelphidae. Braz J Genet 13:767–775.

Svartman M, Vianna-Morgante AM (1998) Karyotype evolution of marsupials: from higher to lower diploid numbers. Cytogenet Cell Genet 82:263–266.

Svartman M, Vianna-Morgante AM (1999) Comparative genome analysis in American marsupials: chromosome banding and in-situ hybridization. Chromosome Res 7:267–275.

Telenius H, Pelmear AH, Tunnacliffe A, et al. (1992) Cytogenetic analysis by chromosome painting using DOP-PCR amplified flow-sorted chromosomes. Genes Chromosomes Cancer 4:257–263.

Toder R, O'Neill RJ, Wienberg J, O'Brien PC, Voullaire L, Marshall-Graves JA (1997) Comparative chromosome painting between two marsupials: origins of an XX/XY1Y2 sex chromosome system. Mamm Genome 8:418–422.

Yang F, Alkalaeva EZ, Perelman PL, et al. (2003) Reciprocal chromosome painting among human, aardvark, and elephant (superorder Afrotheria) reveals the likely eutherian ancestral karyotype. Proc Natl Acad Sci USA 100:1062–1066.

Chapter 3
Marsupial Centomeres and Telomeres: Dynamic Chromosome Domains

Dawn M. Carone and Rachel J. O'Neill

Abstract What has become clear from a synthesis of work on marsupial chromosomes over the last 100 years is that the centromere is more than simply an architectural feature of the chromosome. Rather, it has been participant, either directly or indirectly, in the evolution of the diversity of karyotypes observed in marsupials. Across marsupial lineages, a family of model species stands out as an ideal system in which to study centromere function and evolution: macropodines (kangaroos and wallabies). This chapter focuses on the study of centromeres in marsupials, as both a region critical to ensuring the distribution of sister chromatids to daughter cells during cell division and a chromosomal domain involved in karyotypic stability and evolution. We will explore the role played by elements found at centromeres and telomeres in cell division and karyotypic change as supported by both historic and current experimental evidence.

Keywords Centromere · Telomere · Small RNA · Retroelement · KERV

3.1 Centromere Evolution in Marsupials

3.1.1 Satellite Sequence Convergence in Macropus

"Satellite DNA" is a phrase that originally applied to satellite bands observed in ultracentrifuge density gradients, yet now is commonly used to describe any tandemly repetitive sequence (John, 1988). It has been known for quite some time that satellite DNA families can be species and/or group specific (for example, the α satellites of primates) (Singer, 1982). However, their seemingly ubiquitous presence

R.J. O'Neill (✉)
Department of Molecular and Cell Biology, Center for Applied Genetics and Technology, University of Connecticut, Storrs, CT 06269, USA
e-mail: rachel.oneill@uconn.edu

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_3, © Springer Science+Business Media B.V. 2010

at or near centromeric domains across a wide variety of organisms implicates them in centromere function (see Willard, 1990; Csink and Henikoff, 1998; Eichler, 1999; Henikoff et al., 2001; for reviews). Contrarily, studies have shown that, while satellites are capable of centromere function, they are neither necessary or required for functioning centromere identity. For instance, neocentromeres, aberrant ectopic centromeres, are devoid of satellites. These observations have led to the hypothesis that kinetochore establishment on chromatin is an epigenetic phenomenon independent of underlying DNA sequence (reviewed in Allshire and Karpen, 2008). The only commonality among centromere-associated sequences that have been identified to date is that most comprise large arrays of repeated DNAs. One notable exception to this includes *Saccharomyces cerevisiae* within which the centromere is defined as a "point" centromere that consists of only ∼125 bp of DNA (Clarke, 1990). This simple structure appears to be an evolutionarily derived characteristic of *S. cerevisiae* as other fungal species carry arrays of repeats within their centromeres (see Cambareri et al., 1998 for an example).

Several authors have analyzed centromeric DNA in an evolutionary context, although this has proved difficult due to the low conservation of these DNAs and their complex sequence evolution. An investigation of chromosome-specific α satellite DNA from the chimpanzee chromosome 4 led Haaf and Willard (1997) to postulate that chromosomally-distinct sets of repeats have undergone lineage specific amplification and/or concerted evolution and that most α satellite sequences do not evolve orthologously. Several chromosome specific subsets of α satellite DNAs have also been identified in humans (Greig et al., 1989; Greig et al., 1993; Verma, 1999), suggesting that while this family of satellite repeats has been conserved within this lineage, there has been extensive chromosome specific diversification. Conversely, studies in plants have shown that some repeated DNA sequences have a conserved centromeric location across a wide range of species (Jiang et al., 1996; Miller et al., 1998).

Recent advancements in the analysis of centromere sequence composition underscores the importance for an evolutionary framework on which to study how centromeric repeats influence chromosome, and subsequently species evolution. While the body of knowledge concerning the evolution of centromere-associated repeated DNA within species groups is growing, studies examining the correlation between these sequences and chromosome evolution within a recently diverged group of organisms are limited. The marsupial genus *Macropus* has been the focus of such an approach in the context of a karyotypically diverse group, the macropodines.

During the late Pleistocene (126–11 ka) the Macropodidae experienced a diversification (Flannery, 1989), ultimately resulting in the three modern subfamilies. Within the fossil record as early as 5 million years ago (MYA) the Macropodine genera *Thylogale*, *Petrogale* and *Macropus* appeared (Archer, 1984). It is during this last major diversification that the evolution of the subfamily Macropodine, and the radiation of the genus *Macropus* (including the "true" kangaroos and wallabies), occurred. The time since the *Macropus* diversification has been relatively short (between 11 and 4 MYA) (Flannery, 1989; Burk and Springer, 2000),

rendering of little use traditional metrics to determine the inter-relatedness within the genus such as classical anatomical (Grigg et al., 1989; Burk and Springer, 2000), DNA–DNA hybridization (Kirsch et al., 1997), serological (Kirsch, 1977; Baverstock et al., 1990) and G-banding chromosome evolution studies (Rofe, 1978, 1979).

As revealed by early cytology, the extensive shuffling of chromosome arms among these species further confounded attempts to derive phylogenetic relationships among the *Macropus*. For instance, a high number of fusions in the swamp wallaby (*Wallabia bicolor*) led scientists to place this species in its own genus. What was clear from these earlier studies, and confirmed in recent studies utilizing chromosome painting techniques (see Rens and Ferguson-Smith, Chapter 2), is that the centromere itself is highly mutable in location, structure and content across marsupials. For example, the size of the centromere is not constant even among such closely related species as the red necked wallaby (*Macropus rufogriseus*) and several of its nearest relatives, the tammar wallaby (*M. eugenii*) and the agile wallaby (*M. agilis*) (Fig. 3.1). It has been estimated that the DNA content of the centromere of the red necked wallaby is roughly 25–35% of the total DNA content within the cell (Dunsmuir, 1976; Bulazel et al., 2006). This DNA is largely heterochromatic, defined by its heavy staining with the C-banding technique (Dunsmuir, 1976; Bulazel et al., 2006) and its high density of H3K9-me2 (Bulazel et al., 2006), a modification to histone 3 that is specific to heterochromatin. While these centromere regions appear large, the majority of this material is pericentric and is thus not involved in the formation of the inner or outer kinetochore during cell division (Bulazel et al., 2006) (Fig. 3.1).

In early efforts to clarify the relationships among kangaroo species, satellite DNAs were a popular target for molecular analyses based on the assumption that reproductive isolation between species would result in lineage-specific evolution of repeat classes. This early work was performed over 35 years ago, before the benefit of modern molecular techniques such as automated sequencing, yet these data laid the foundation for understanding the relationship between centromeric

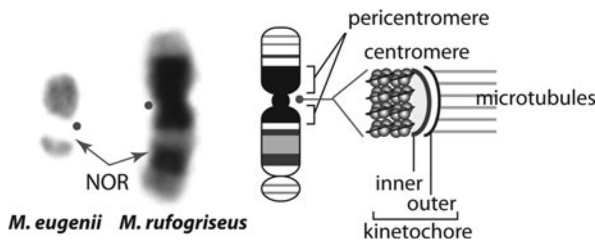


Fig. 3.1 Centromere evolution in macropodines. C-band X chromosomes from *M. eugenii* (*left*) and *M. rufogriseus* (*right*). The NOR is indicated with an *arrow* and the active centromere with a *dot*. To the *right* is an ideogram of the *M. rufogriseus* X chromosome showing the G-bands, active centromere (*dot*) and large pericentric regions (as indicated). To the far right is a schematic close-up of the structure of the centromere with DNA bound to nucleosomes containing either H3 or CENP-A, the inner and outer kinetochore plates and microtubule attachment

DNA and chromosome change. Soon after E.M. Southern published his seminal paper on the sequence and tandem repeat organization and evolution of the guinea pig alpha satellite DNA (Southern, 1970), a satellite sequence from the kangaroo rat, *Dipodomys ordii* was isolated and characterized (Fry et al., 1973), placing this marsupial species among the first mammals for which satellite sequences were analyzed. Using similar centrifugation techniques, satellite DNA was isolated for the red-necked wallaby (*Macropus rufogriseus*), red kangaroo (*M. rufus*) and wallaroo (*M. robustus*) (Dunsmuir, 1976; Dennis et al., 1980; Peacock et al., 1981; Venolia and Peacock, 1981; Elizur et al., 1982).

More recently, Bulazel et al. (2007) used a microdissection and microcloning approach to construct a library of sequences found at the centromere of the X chromosome of *M. rufogriseus*. Within this library, three different satellites were further characterized based on their sequence, function and distribution across the karyotype. Sat23 was found at all centromeres of this species and functions as the primary satellite, with a periodicity of 178 nt and a functional CENP-B DNA binding domain, evidence that it is a functional component of the centromere. Two other satellites, sat1 and B29, were found restricted to the sex chromosomes, although sat1 was also found at the pericentromere of chromosome 2 (Bulazel et al., 2007). This autosomal localization was the first clue that there may be a tighter association between chromosome shuffling and centromere satellites than previously recognized as this location on chromosome 2 is part of a fusion chromosome with the X chromosome and chromosome 7 in another macropodid species, *Wallabia bicolor* (the swamp wallaby) (Toder et al., 1997).

Having a suite of three different satellites, with different distributions within the karyotype of *M. rufogriseus* allowed us to build on, and extend, the studies into the evolution of satellites and chromosomes within macropodidae (Bulazel et al., 2007) initiated decades ago. To this end, an accurate species phylogeny was determined based on both a nuclear (*TRSP*) and mitochondrial gene (*cytb*) for nine *Macropus* species (inclusive of *Wallabia*) (Fig. 3.2). This phylogeny was then compared to that derived from a novel approach using chromosome segments as phylogenetic markers. In this method, each conserved chromosome segment (as defined by Rens et al., 2003) is transformed into a binary code and input into the MGR algorithm (multiple genome rearrangement tool) used primarily to analyze sequence data between different genomes (Bourque and Pevzner, 2002). Using the conserved chromosome segments in this fashion allowed for the derivation of a chromosome phylogeny; in other words, a phylogeny representing the most parsimonious series of rearrangements possible to derive current karyotypes within *Macropus* with no consideration of the true phylogenetic relationships of these species. When the chromosome phylogeny and "true" species phylogeny are compared to one another, it is clear they are not concordant, indicating that similar karyotypes have been derived in distantly related species not by shared ancestry, but by convergent breakpoint reuse. Closer examination of the conserved chromosome segments involved in this reuse showed that only a subset of breakpoints between segments were involved in the derivation of extant karyotypes (C1, C2, C8, C10, C15 and C18) and all involved a centromere. A question emerging from this work was: What is

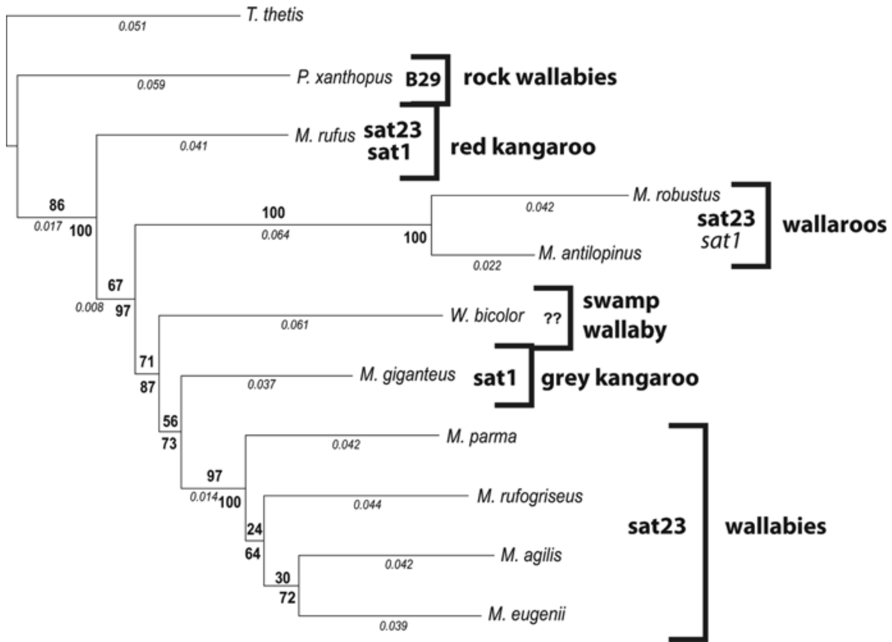


Fig. 3.2 *Macropus* phylogeny and centromere satellites. Phylogeny of *Macropus* (including *Wallabia*) with two outgroup species (*Thylogale thetis* and *Petrogale xanthopus*) as per Bulazel et al. (2007). The predominant centromere satellite is indicated in *bold* for each lineage as are the common names for the lineages. MrBayes credibility values are listed below, ML bootstrap values above the branches, and branch lengths are listed in italics *(italics)*. Note that the sat1 hybridization in wallaroos is limited to only a subset of chromosomes (see text)

the resulting effect on satellite evolution for chromosomes involved in breakpoint reuse?

Tracking the three known satellites across all of these species (Fig. 3.2) showed that sat23, the predominant satellite in *M. rufogriseus*, is not shared as a predominant satellite in all species in this group. Rather, sat1, a satellite limited to sex chromosomes in *M. rufogriseus*, is a predominant satellite in *M. giganteus* (grey kangaroo) and *both* sat23 and sat1 are shared as the predominant satellites in *M. rufus* (red kangaroo). The patterns of satellite sequence mapping for sat23 and sat1 in *M. robustus* and *M. antilopinus* (the wallaroos) were the most informative with respect to the suite of chromosome rearrangements this group has experienced. In both species, sat23 is the predominant satellite, while sat1 is restricted to centromeres of chromosomes 1, 5 and 6, the centromeres that separate syntenic blocks C1, C2, C8, C10, C15 and C18. As these are the only actively rearranging blocks in the karyotype of this lineage, it can be inferred that the retention of this satellite sequence is a consequence of convergent breakpoint reuse.

The “library hypothesis” of satellite evolution posits that related lineages share a collection of repeat sequences that may become preferentially amplified in any of the given species during the normal events of centromere evolution (Salser et al.,

1976). Perhaps in *Macropus* lineages the "library" of satellite sequences include sat1, sat23 and B29 and are involved in a process of accretion into large, satellite arrays. The conservation of satellites after convergent breakpoint reuse indicates that centromeric sequence arrays are not created de novo, but expand and contract from native inhabitants of the centromere regions. Rather than evolution of novel sequences to form new centromere arrays, this study suggests that centromeric satellites remain in the genome, likely at latent centromere locations (Liscinsky et al., 2005) (Fig. 3.3) and experience episodes of expansion and contraction in divergent lineages.

In contrast to the autosomes, the sex chromosomes of the *Macropus* appear to retain tandem arrays of ancestral centromeric material for longer periods of time, although these arrays are no longer homogenized. Some satellites, B29 for example, are shared on the sex chromosomes of all species in this group, indicating an origin for this satellite predating the *Macropus* radiation.

The type of chromosome rearrangement also has some bearing on whether a particular satellite has experienced contraction or expansion in the genome. For example, during a fusion event, chromosome breaks within satellite arrays result in removal of p arms along with a portion of centromeric/pericentric material, ultimately resulting in exposure of the fusion breakpoint and subsequent loss of satellite material. In contrast, duplication events often accompany arm-swapping translocations (Ventura et al., 2003) and may aid in centromere sequence convergence and expansion (Horvath et al., 2003; Bailey et al., 2004; She et al., 2004). For example, segmental duplications preceding a translocation event could increase sequence

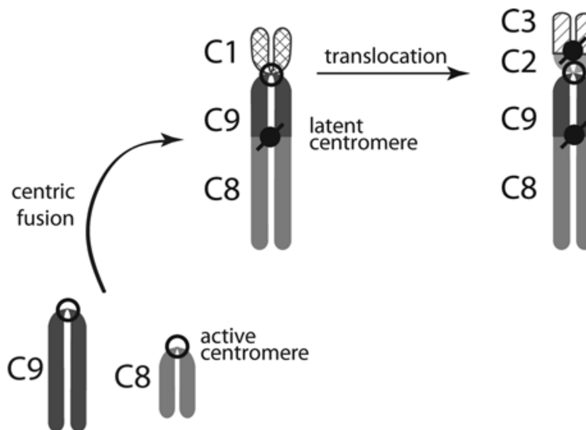


Fig. 3.3 Schematic depicting the derivation of *M. eugenii* chromosome 1 *top left* from a centric fusion of conserved chromosome segments C9 and C8, leading to an inactive, latent centromere at the fusion site. A subsequent arm-swapping translocation is shown to the *right* in the derivation of the *M. giganteus* chromosome 1 *top right*. Active centromeres are depicted as *open circles*; latent centromeres are shown with a *line* through a *filled* in *circle*

identity between sites on non-homologous chromosomes, making such translocation events more likely (Eichler, 1999).

What this rich field of research has shown is that, when examined in the context of species evolution, satellite sequence evolution is found to strictly follow chromosome trajectories. Tracking satellites has uncovered interesting patterns of chromosome change and further supports the thesis that the centromere plays an active, dynamic role in karyotype change.

3.1.2 KERV, Centromeres and Evolutionary Breakpoints

Kangaroo Endogenous Retrovirus (KERV), an endogenous retrovirus (ERV), is resident in centromeres of most marsupial lineages (Ferreri et al., 2004; Liscinsky et al., 2005; Carone and Ferreri, personal communication). While lacking an open reading frame for an envelope protein (*env*, required to form an infectious virion), KERV has a typical ERV structure consisting of open reading frames for *gag, pro, pol,* and *int* and two identical long terminal repeats (LTRs). Initial studies into the genomic distribution of KERV focused on *M. eugenii,* a marsupial model organism for which considerable sequence data has been generated and is the focus of a concerted genome sequencing initiative (Graves et al., 2003).

Within *M. eugenii*, KERV is found concentrated within centromeres, telomeres, and evolutionary breakpoints (EBs) (Ferreri et al., 2004). Further refinement of the localization of KERV copies conducted through a screen of a *M. eugenii* bacterial artificial chromosome (BAC) library with the *gag* region revealed this element to be resident in extremely high copy number in *M. eugenii* (∼10% KERV-positive signal per single BAC filter) (Ferreri et al., 2004). Fluorescence in situ hybridization of 49 KERV-positive BAC clones has further specified the KERV genomic distribution to active centromere regions (pericentric or centric), telomeres or EBs within the *Macropodidae* lineage (Ferreri et al., 2004; Liscinsky et al., 2005; Longo et al., 2009). Taking into account the location of centromeres within all macropod species, 100% of these clones map to active centromere or latent (old) centromere sites (Liscinsky et al., 2005). Whether these sites have the ability to potentiate new centromere sites or re-activate latent sites is a current focus of research. Several lines of evidence indicate that EBs and latent centromere sites within macropods retain many sequence features of active centromere sites, therefore harboring the ability to become active centromeres.

EBs were originally thought to occur at random locations in the genome (Nadeau and Taylor, 1984). However, more recent evidence has suggested that these sites are often conserved even among divergent species (Pevzner and Tesler, 2003; Murphy et al., 2005) and harbor repetitive sequences which may render them prone to breakage (Schibler et al., 2006). The observation of KERV retention at EBs within macropods prompted an investigation of the sequences resident at EBs, centromeres (CENs), and euchromatic regions (EUs). This study has highlighted the sequence signatures at these locations and uncovered several features common to EBs and CENs (Longo et al., 2009). Detailed annotation and comparison of *M. eugenii*

BACs mapped by fluorescence in situ hybridization to CENs, EBs, and EU regions suggests that repetitive elements are enriched at CENs and EBs compared to EU regions. Close examination of the distinct repetitive element classes resident at these locations indicated that long interspersed nucleotide elements (LINEs) and ERVs showed a significantly higher enrichment for CENs and EBs compared to EU regions. In addition, CENs and EBs were relatively devoid of short interspersed nucleotide elements compared to EU regions, suggesting that the sequence composition of CENs and EBs are more similar to each other than to EU regions. Of the ERVs examined, the overwhelmingly most abundant sequence found at CEN and EB regions was KERV, further supporting the non-random distribution of KERV previously seen by BAC mapping (Ferreri et al., 2004; Liscinsky et al., 2005).

3.2 Centromere Function

3.2.1 KERV and Small RNAs

The strict localization of KERV at CENs and EBs suggests that KERV is an element that has been selected for centromeric sites (active or latent) or, at the very least, selectively maintained over significant evolutionary time at these locations. This level of sequence conservation across marsupial lineages and across the karyotype is unprecedented for centromeres, as centromeres are prone to sequence degradation and often harbor divergent sequences even among closely related species, an observation termed the centromere paradox (Henikoff et al., 2001). KERV retention at CENs and EBs is therefore the exception to the rule and has prompted an investigation into whether KERV maintains a functional role in the genome, other than that of a selfish element.

While intact, full-length KERV elements are found at both CEN and EB sites, there is also an enrichment of solo KERV LTRs (kLTRs) at CENs and EBs compared to EU regions (Longo et al., 2009). This observation is intriguing when taking retroviral structure and function into account since the LTRs of retroviruses are known to contain strong promoters that facilitate transcription of these selfish elements upon integration (Coffin et al., 1997). In addition, some retroviral promoters are known to promote transcription bidirectionally (Dunn et al., 2006) and can be responsible for promoting transcription of adjacent surrounding sequences (Danilevskaya et al., 1997; Dunn et al., 2006). Further investigation into the promoter status of the kLTR using a *GFP* transfection assay indicates that the kLTR sequence does, in fact, contain a strong bidirectional promoter (Carone et al., 2009).

Transcription of centromeric sequences has been observed for almost all eukaryotic species examined to date (reviewed in Allshire and Karpen, 2008) and retroelements have been implicated in promoting transcription of these sequences (May et al., 2005; Ugarkovic, 2005), namely satellites (see satellite sequence section) which do not contain known promoter binding sites. High-resolution mapping of kLTR and centromeric satellites (sat23) localized by fiber FISH in *M. eugenii* has

revealed that solo kLTRs are interspersed within sat23 arrays in tammar centromeres (Carone et al., 2009). It is therefore likely that kLTR bidirectional promoters resident at CENs not only promote transcription of KERV elements, but also facilitate the transcription of its nearest resident neighbor, sat23. Interestingly, this structure of interspersed satellite and retroelement sequences has also been observed in plant centromeres (Cheng et al., 2002), indicating this sequence arrangement may be a conserved feature of eukaryotic centromeres (or at least of the core).

Transcripts emanating from centromeric satellites have been linked to centromere function in several species ranging from yeast to human (Reinhart and Bartel, 2002; Volpe et al., 2002; Volpe et al., 2003; Fukagawa et al., 2004; Bouzinba-Segard et al., 2006) through processing by RNA interference (RNAi) machinery into short interfering RNAs (siRNAs). RNAi is best known for its role in regulation of gene expression via small RNA intermediates. Originally discovered in *C. elegans* (Fire et al., 1998), this conserved mechanism has since been reported in species ranging from *Drosophila melanogaster* (Kennerdell and Carthew, 2000) to mouse (Wianny and Zernicka-Goetz, 2000). Within the past 10 years, three main classes of small RNAs have been described with wide ranging targets and effects. The first, microRNAs (miRNA), are generated from hairpin precursors. Processing of hairpin dsRNA into miRNA occurs in two stages: in the nucleus, Drosha processes hairpin RNA to ∼70 nt pre-miRNAs (Lee et al., 2003); following transport to the cytoplasm, the pre-miRNAs are processed by the RNase III-like enzyme Dicer to ∼22 nt mature miRNAs. The second class of small RNA, short interfering RNAs (siRNAs), unlike miRNAs, are generated from long double-stranded RNA (dsRNA) precursors but also processed into ∼21–22 nt by Dicer. Silencing via miRNAs and siRNAs occurs by homologous pairing to target mRNAs of one strand of the Dicer product (the "guide" strand). Targeting of miRNAs/siRNAs to their respective targets is facilitated by the RNA-induced silencing complex (RISC) (Hammond et al., 2000). Alternatively, siRNAs may be recruited by the RNAi induced transcriptional silencing complex (RITS) to facilitate the targeting of epigenetic regulators to centromeric heterochromatin (Verdel et al., 2004). The third class, piwi-interacting RNAs (piRNAs), are a class of Dicer processing-independent small RNAs found in the germline that are involved in transposon silencing and germline development (Klattenhoff and Theurkauf, 2008). These RNAs are larger than their miRNA/siRNA counterparts (24–30 nt) and processed via an unknown mechanism. The specific role of piRNAs is not well understood, but new evidence indicates these small RNAs are maternally inherited and may play a role in hybrid dysgenesis (Brennecke et al., 2008).

In yeast (Verdel et al., 2004) and plants (Topp et al., 2004; May et al., 2005; Neumann et al., 2007) double stranded centromeric satellite transcripts are processed into siRNA and have been shown to recruit epigenetic modifications to pericentric heterochromatin in yeast (Volpe et al., 2002) to propagate epigenetic marks at the centromere. The presence of long satellite transcripts and/or siRNAs processed from mammalian centromeric satellites have been difficult to find, possibly due to the rapid turnover of larger transcripts or an alternative mechanism of processing. However, it is apparent that aberrant regulation of satellite sequences

(Valgardsdottir et al., 2005; Bouzinba-Segard et al., 2006) and abrogation of the RNAi machinery (Fukagawa et al., 2004; Kanellopoulou et al., 2005) in mammals leads to impaired centromere function, indicating that these transcripts are under some mechanism of tight regulation.

Investigations into the role of KERV in *M. eugenii* centromeres has revealed a new mechanism for transcription of centromere satellite and retroviral sequences and a new class of small RNA that may be conserved in other mammalian systems. KERV and sat23 transcripts in *M. eugenii* are processed into a class of small RNA distinct from siRNA, miRNA and piRNAs based on their size (Carone et al., 2009). This new class of small RNAs emanating from satellite and retroviral sequences at the centromere, called centromere repeat associated small interacting RNAs (crasiRNAs), are 34–42 nt and are thought to be processed from KERV and sat23 dsRNA precursors which are transcribed by the kLTR bidirectional promoter (Carone et al., 2009). Further investigation will determine if this is a conserved mammalian mechanism for processing of centromeric transcripts and shed light on the regulatory elements involved in the biogenesis of this new class of small RNA. It is clear that transcription of centromeric sequences is under tight regulation in mammals, including the tammar wallaby, and that the role of retroviral promoters in facilitating this process may be an integral part of mammalian centromere function.

3.3 Hybrid Dysgenesis and Karyotypic Change

Despite the diversification of karyotypes within the *Macropodinae*, reproductive isolation has not formed a complete barrier to interspecific hybridization across all species. Hybrids have been identified in the wild and created in captive colonies (Fig. 3.4a) for many species within both the *Macropus* (including *Wallabia*) and *Petrogale* genera (Sharman et al., 1990; Eldridge and Close, 1992; Lowry et al., 1994; O'Neill et al., 1998; O'Neill et al., 2001; O'Neill et al., 2004; Metcalfe et al., 2007). Such interspecific hybrids have long been an essential tool in the study of genetics and evolutionary biology. As is usually the case in genetic research, abnormal or mutated states offer the best opportunity to gain a better understanding of normal processes. The abnormalities seen most often in interspecific hybrids, known generally as "hybrid dysgenesis" are perturbations of development and reproductive capacity.

The causes of hybrid dysgenesis have been a subject of considerable debate since 1922, when J.B.S. Haldane made his famous observation, known as Haldane's Rule (Haldane, 1922), that in hybrid zones, it is most often the heterogametic sex that is sterile or inviable. The remarkable generality of Haldane's Rule suggests that there is a basic commonality to the genetic defects seen in hybrids. For example, at a gross level, many hybrids suffer from chromosomal aberrations at much higher frequencies than their parent species. One of the primary causes of chromosomal rearrangement, aside from exogenous insults by radiation or chemical mutagens, is the activity of endogenous mobile DNA.

Fig. 3.4 Hybrid dysgenesis in *Macropus*. (**a**) Hybrid animals in a captive colony from *left*: *M. dorsalis* × *M. parma*, *M. rufogriseus bankianus* × *M. rufogriseus rufogriseus, M. rufogriseus banksianus* × *M. agilis*. (**b**) *Left*: SEM image of a chromosome from a *M. rufogriseus* × *M. agilis* hybrid carrying a knob at its centromere (*arrow*). *Right*: Fluorescence in situ hybridization of centromere sequences sat23 and KERV to metaphase chromosomes of the same hybrid (FISH merged image to the *right*, DAPI to the *left*)

It is still unclear whether chromosomal speciation (White, 1978), which can occur very rapidly (Garagna et al., 1997), possibly engendered by inter-species hybridization (O'Neill et al., 1998, 2001, 2004;), has occurred in macropod lineages. However, certain closely-related marsupial clades within this lineage show extreme karyotypic diversity (for example the *Petrogale* (Eldridge and Close, 1993) and *Wallabia* (Rofe, 1978; Hayman et al., 1987; Toder et al., 1997)) and hybrids

among these species exhibit frequent de novo chromosome rearrangements (O'Neill et al., 2001, 2004; Metcalfe et al., 2007).

The molecular mechanism promoting chromosome rearrangements in hybrids is not known, but rearrangements are often observed in conjunction with the mobilization of transposable elements. Classical studies of *Drosophila* (Kidwell and Kidwell, 1976) and *Zea* (collection of papers McClintock, 1987) have established the link between hybrid dysgenesis and transposition. However, the reason why hybrids are specifically prone to these deleterious forces has not been determined. It has been theorized that normal genomes are protected from the de-stabilizing effects of active transposable elements by epigenetic silencing (Yoder et al., 1997). Epigenetic modes of inheritance include alteration of DNA by methylation, patterns of histone acetylation and methylation and small RNA mediated silencing of chromosomal regions.

Using an interspecific hybrid marsupial, a correlation between genomic methylation, transposition and amplification of mobile elements was found associated with chromosome remodeling (O'Neill et al., 1998). In this study, we hypothesized that the failure, in interspecific hybrids, of the host-defense mechanism conferred by DNA methylation could lead to activation of transposable elements and hence, hybrid dysgenesis. Dramatic undermethylation was observed in several hybrids, and one hybrid clearly revealed evidence for activation and amplification of a retroelement that visibly altered centromeres. The cause of this hybrid-specific undermethylation is not yet known, nor is it known whether this phenomenon extends to hybrids of other taxa. This observation provided the first evidence that hybridization-induced mobile element activity occurs in mammals, and suggests DNA methylation is important in this process. Moreover, the rearrangements observed were restricted to centromere locations of the maternally-inherited chromosome complement.

Further studies on macropodid hybrids using a combination of whole chromosome painting techniques and ultrastructural analyses of chromatin indicate that centromeres are a specific target for destabilization that results in retroelement and satellite amplification (namely KERV and B23) accompanied by chromatin remodelling (O'Neill et al., 2001; Metcalfe et al., 2007) (Fig. 3.4b). Given the dynamic role the centromere plays in karyotypic change in macropods, and the link between epigenetic regulation and the KERV element, it is intriguing to postulate that the chromatin remodelling observed in interspecific hybrids may be a direct result of disruption of the processing or biogenesis of the small RNAs emanating from these retroelements.

3.4 Telomeres

Telomeres are another architectural feature of a chromosome that are important to consider when discussing chromosome biology or karyotypic change. A feature that distinguishes natural chromosome ends from damaged DNA resulting from breaks,

telomeres are found at the end of chromosomes, consist of a canonical repeat in most organisms (TTAGGG, often referred to as $(T_2AG_3)_n$) and are maintained by two different mechanisms. The first is a telomerase-mediated mechanism of lengthening, whereby telomeres are lengthened through the reverse transcriptase activity of telomerase after successive rounds of telomere attrition that normally occurs during replication. The second is an alternative lengthening of telomeres (ALT) pathway in which telomere length is maintained through inter-telomere recombination. While telomere studies are valuable for understanding disease progression (in particular tumorigenesis), they have also proven valuable, exciting and even controversial in the chromosome biology of marsupials.

Studies using probes for the telomere repeat in different species clusters of *Petrogale* has shown that telomere sequences are retained at sites of fusions and are an indicator of past chromosome shuffling events. For example, within the *P. penicillata* and *lateralis* complexes, telomere sequences are retained after centric fusion events (Metcalfe et al., 1998; Metcalfe, 2002), indicating a mechanism of chromosome fusion different to that identified in several rodent species (e.g. Garagna et al., 1995).

Several groups have attempted to clarify whether the 2n = 14 complement is indeed ancestral to all marsupials (see previous chapter) or whether it is a complement that has been derived by convergence in many species within six of the seven extant marsupial orders (Rofe and Hayman, 1985; Hayman et al., 1987). Studies of South American marsupials within Didelphidae showed $(T_2AG_3)_n$ sites were restricted to chromosome ends in species carrying the 2n = 22 complement but were found interstitial in species with the 2n = 14 complement (Svartman and Vianna-Morgante, 1998; Carvalho and Mattevi, 2000). These data led the authors to conclude the 2n = 14 complement has been derived through a series of centric fusions, with remnant telomere sequences left at the sites of fusion events. Following this, Pagnozzi et al. (Pagnozzi et al., 2000, 2002) found a correlation between C-band positive material and telomere signal, indicating that interstitial signal may be part of satellite sequences rather than remnants of past telomere locations. A detailed study of this issue was performed by Metcalfe et al. (2004) using representative macropodid species. Detailed analyses of C-banding patterns and $(T_2AG_3)_n$ patterns following known fusion events showed that in cases where telomeric signal is found within C-band positive regions (Fig. 3.5), these signals likely represent part of the satellite suite and not ancestral telomeres, per se. Therefore, interpretation of past telomere locations cannot be conclusively made where signal coincides with increased satellite density (Metcalfe et al., 2004). This has been further supported by the finding that some satellites within *Macropus* bear a striking homology to the telomere repeat sequence (Bulazel et al., 2006).

While few functional studies have been performed on telomeres in marsupials, an interesting observation was reported at a recent conference and summarized in a review by Karlseder and Cooper (2007). Telomere biologist W. Wright discovered that *Vombatus ursinus* (the common, or coarse-haired wombat) lacks

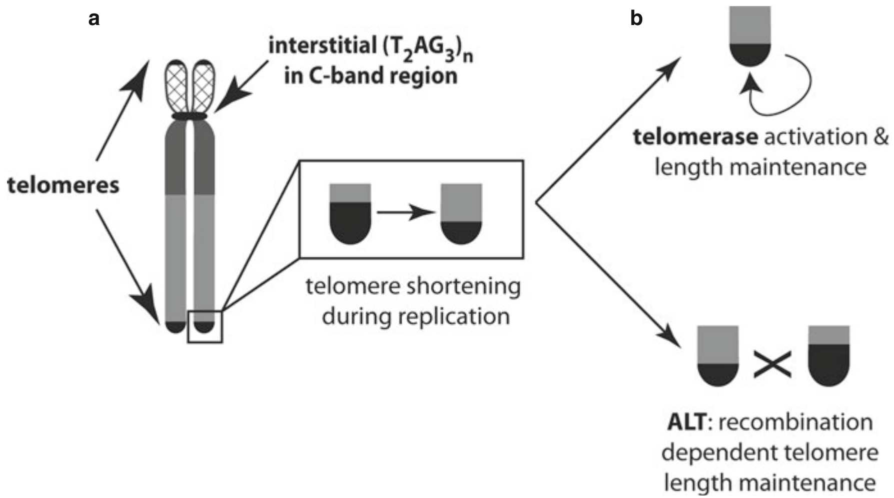


Fig. 3.5 Telomeres in marsupials. Schematic depicting the location of telomeric signal at the ends of chromosome 1 in *M. eugenii* as well as in the C-band positive centromeric regions (*black*). (Blow up to the *right*) Telomeres are shortened during DNA replication but length is maintained by two alternative pathways. The *top* shows the telomerase-dependent pathway whereby length is maintained through the reverse transcriptase activity of telomerase. The *bottom* shows the telomerase-independent pathway of telomere maintenance, alternative lengthening of telomeres (ALT), based on inter-telomere recombination

detectable telomerase activity; instead this species uses the alternative lengthening of telomeres (ALT) pathway for telomere maintenance (Fig. 3.5). This is somewhat surprising as the ALT pathway is utilized by cancer cells during rampant proliferation and cellular immortalization. Studying telomere maintenance in species such as *V. ursinus* may provide invaluable insights into the process of ALT with respect to species evolution, cellular longevity and cancer therapy.

3.5 Looking Forward: Centromere and Chromosome Biology in the Age of Genomics

As the tammar wallaby genome nears completion, this is an exciting time to begin to unravel some of the mysteries of chromosome structure, function and evolution that have fascinated biologists for over 100 years. The small number of chromosomes and high incidence of centric shifts across marsupial genera make marsupials an ideal model for studying the epigenetic regulation of centromere establishment and maintenance as well as the genome shaping capacity for retroelements and non coding RNAs. No doubt genome-scale studies of centromeres, evolutionary breakpoints and telomeres in the wallaby will aide in developing a comprehensive understanding of mammalian evolution.

References

Allshire RC, Karpen GH (2008) Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat Rev Genet 9:923–937.

Archer M (1984) Evolution of arid Australia and its consequences for vertebrates. In: Archer M, Clayton G (eds) Vertebrate Zoogeography and Evolution in Australasia. Southwood Press, Marrickville.

Bailey JA, Baertsch R, Kent WJ, Haussler D, Eichler EE (2004) Hotspots of mammalian chromosomal evolution. Genome Biol 5:R23.

Baverstock PR, Krieg M, Birrell J (1990) Evolutionary Relationships of Australian Marsupials and Assessed by Albumin Immunology. Aust J Zool 37:273–287.

Bourque G, Pevzner PA (2002) Genome-scale evolution: reconstructing gene orders in the ancestral species. Genome Res 12:26–36.

Bouzinba-Segard H, Guais A, Francastel C (2006) Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. Proc Natl Acad Sci USA 103:8709–8714.

Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ (2008) An epigenetic role for maternally inherited piRNAs in transposon silencing. Science 322:1387–1392.

Bulazel K, Metcalfe C, Ferreri G, Yu J, Eldridge M, O'Neill R (2006) Cytogenetic and molecular evaluation of centromere-associated DNA sequences from a marsupial (Macropodidae: *Macropus rufogriseus*) X chromosome. Genetics 172:1129–1137.

Bulazel KV, Ferreri GC, Eldridge MD, O'Neill RJ (2007) Species-specific shifts in centromere sequence composition are coincident with breakpoint reuse in karyotypically divergent lineages. Genome Biol 8:R170.

Burk A, Springer M (2000) Intergeneric relationships among Macropodoidea (Metatheria: Diprotodontia) and the chronicle of kangaroo evolution. J Mammal Evol 7:213–237.

Cambareri EB, Aisner R, Carbon J (1998) Structure of the chromosome VII centromere region in *Neurospora crassa*: degenerate transposons and simple repeats. Mol Cell Biol 18:5465–5477.

Carone DM, Longo MS, Ferreri GC, et al. (2009) A new class of retroviral and satellite encoded small RNAs emanates from mammalian centromeres. Chromosoma 118:113–125.

Carvalho BD, Mattevi MS (2000) $(T2AG3)_n$ telomeric sequence hybridization suggestive of centric fusion in karyotype marsupials evolution. Genetica 108:205–210.

Cheng Z, Dong F, Langdon T, et al. (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell 14:1691–1704.

Clarke L (1990) Centromeres of budding and fission yeasts. Trends Genet 6:150–154.

Coffin JM, Hughes SH, Varmus HE (1997) Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Csink AK, Henikoff S (1998) Something from nothing: the evolution and utility of satellite repeats. Trends Genet 14:200–204.

Danilevskaya ON, Arkhipova IR, Traverse KL, Pardue ML (1997) Promoting in tandem: the promoter for telomere transposon HeT-A and implications for the evolution of retroviral LTRs. Cell 88:647–655.

Dennis ES, Dunsmuir P, Peacock WJ (1980) Segmental amplification in a satellite DNA: restriction enzyme analysis of the major satellite of *Macropus rufogriseus*. Chromosoma 79:179–198.

Dunn CA, Romanish MT, Gutierrez LE, van de Lagemaat LN, Mager DL (2006) Transcription of two human genes from a bidirectional endogenous retrovirus promoter. Gene 366:335–342.

Dunsmuir P (1976) Satellite DNA in the kangaroo *Macropus rufogriseus*. Chromosoma 56: 111–125.

Eichler EE (1999) Repetitive conundrums of centromere structure and function. Hum Mol Genet 8:151–155.

Eldridge MD, Close RL (1992) Taxonomy of rock wallabies, Petrogale (Marsupialia: Macropodidae). I: a revision of the eastern Petrogale with the description of three new species. Aust J Zool 40:605–624.

Eldridge MD, Close RL (1993) Radiation of chromosome shuffles. *Curr Opin Genet Dev* 3: 915–922.
Elizur A, Dennis ES, Peacock WJ (1982) Satellite DNA sequences in the red kangaroo (*Macropus rufus*). Aust J Biol Sci 35:313–325.
Ferreri GC, Marzelli M, Rens W, O'Neill RJ (2004) A centromere-specific retroviral element associated with breaks of synteny in macropodine marsupials. Cytogenet Genome Res 107:115–118.
Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391:806–811.
Flannery TF (1989) Phylogeny of the Macropodoidea; a study in convergence. In: Grigg G, Jarman P, Hume I (eds) Kangaroos, Wallabies and Rat-Kangaroos. Surrey Beatty and Sons, Chipping Norton.
Fry K, Poon R, Whitcome P, et al. (1973) Nucleotide sequence of HS-beta satellite DNA from kangaroo rat *Dipodomys ordii*. Proc Natl Acad Sci USA 70:2642–2646.
Fukagawa T, Nogami M, Yoshikawa M, et al. (2004) Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol 6:784–791.
Garagna S, Broccoli D, Redi CA, Searle JB, Cooke HJ, Capanna E (1995) Robertsonian metacentrics of the house mouse lose telomeric sequences but retain some minor satellite DNA in the pericentromeric area. Chromosoma 103:685–692.
Garagna S, Zuccotti M, Redi CA, Capanna E (1997) Trapping speciation. Nature 390:241–242.
Graves JAM, Wakefield MJ, Renfree MB, Cooper DW, Speed T, Lindblad-Toh K, Lander ES, Wilson RK (2003) Proposal to Sequence the Genome of the Model Marsupial *Macropus eugenii* (Tammar Wallaby). http://wwwgenomegov/Pages/Research/Sequencing/SeqProposals/WallabySEQpdf
Greig GM, Warburton PE, Willard HF (1993) Organization and evolution of an alpha satellite DNA subset shared by human chromosomes 13 and 21. J Mol Evol 37:464–475.
Greig GM, England SB, Bedford HM, Willard HF (1989) Chromosome-specific alpha satellite DNA from the centromere of human chromosome 16. Am J Hum Genet 45:862–872.
Grigg GC, Jarman P, Hume ID (1989) Kangaroos, Wallabies and Rat-Kangaroos. Surrey Beatty & Sons, Chipping Norton.
Haaf T, Willard HF (1997) Chromosome-specific alpha-satellite DNA from the centromere of chimpanzee chromosome 4. Chromosoma 106:226–232.
Haldane JBS (1922) Sex ratio and unisexual sterility in hybrid animals. J Genet 12:101–109.
Hammond SM, Bernstein E, Beach D, Hannon GJ (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature 404:293–296.
Hayman DL, Rofe RH, Sharp PJ (1987) Chromosome evolution in marsupials. Chromosomes Today 9:91–102.
Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293:1098–1102.
Horvath JE, Gulden CL, Bailey JA, et al. (2003) Using a pericentromeric interspersed repeat to recapitulate the phylogeny and expansion of human centromeric segmental duplications. Mol Biol Evol 20:1463–1479.
Jiang J, Nasuda S, Dong F, et al. (1996) A conserved repetitive DNA element located in the centromeres of cereal chromosomes. Proc Natl Acad Sci USA 93:14210–14213.
John B (1988) Heterochromatin Molecular and Structural Aspects. Cambridge University Press, Cambridge, MA.
Kanellopoulou C, Muljo SA, Kung AL, et al. (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 19:489–501.
Karlseder J, Cooper JP (2007) Of wombats and whales: telomere tales in Madrid. Conference on telomeres and telomerase. EMBO Rep 8:542–546.
Kennerdell JR, Carthew RW (2000) Heritable gene silencing in *Drosophila* using double-stranded RNA. Nat Biotechnol 18:896–898.

Kidwell MG, Kidwell JF (1976) Selection for male recombination in *Drosophila melanogaster*. Genetics 84:333–351.

Kirsch JA (1977) The comparative serology of Marsupialia, and a classification of marsupials. Aust J Zool Suppl Ser 52:1–152.

Kirsch JA, Lapointe F, Springer MS (1997) DNA-Hybridization Studies of Marsupials and their Implications for Metatherian Classification. Aust J Zool 45:211–280.

Klattenhoff C, Theurkauf W (2008) Biogenesis and germline functions of piRNAs. Development 135:3–9.

Lee Y, Ahn C, Han J, et al. (2003) The nuclear RNase III Drosha initiates microRNA processing. Nature 425:415–419.

Liscinsky DM, Ferreri GC, Mack JA, Eldridge MD, O'Neill RJ (2005) Retention of latent centromeres in the Mammalian genome. J Hered 96:217–224.

Longo MS, Carone DM, Program NC, Green ED, O'Neill MJ,O'Neill RJ (2009) Distinct retroelement classes define evolutionary breakpoints demarcating sites of evolutionary novelty. BMC Genomics 10:334.

Lowry PS, Eldridge MDB, Johnston PG (1994) Genetic analysis of a female macropodid hybrid (Macropus agilis $\times$ M. rufogriseus) and her backcross offspring. Aust Mamm 18: 79–82.

May BP, Lippman ZB, Fang Y, Spector DL, Martienssen RA (2005) Differential regulation of strand-specific transcripts from Arabidopsis centromeric satellite repeats. PLoS Genet 1:e79.

McClintock B (1987) The Discovery and Characterization of Transposable Elements. Garland Publishing, Inc., New York, NY.

Metcalfe CJ (2002) Telomeres and Chromosome Evolution in Marsupials. Biological Sciences, Macquarie University, Sydney.

Metcalfe CJ, Eldridge MD, Toder R, Johnston PG (1998) Mapping the distribution of the telomeric sequence $(T2AG3)_n$ in the Macropodoidea (Marsupialia), by fluorescence in situ hybridization. I. The swamp wallaby, *Wallabia bicolor*. Chromosome Res 6:603–610.

Metcalfe CJ, Eldridge MD, Johnston PG (2004) Mapping the distribution of the telomeric sequence $(T2AG3)_n$ in the $2n = 14$ ancestral marsupial complement and in the macropodines (Marsupialia: Macropodidae) by fluorescence in situ hybridization. Chromosome Res 12:405–414.

Metcalfe CJ, Bulazel KV, Ferreri GC, et al. (2007) Genomic instability within centromeres of interspecific marsupial hybrids. Genetics 177:2507–2517.

Miller JT, Dong F, Jackson SA, Song J, Jiang J (1998) Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. Genetics 150:1615–1623.

Murphy WJ, Larkin DM, Everts-van der Wind A, et al. (2005) Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. Science 309:613–617.

Nadeau JH, Taylor BA (1984) Lengths of chromosomal segments conserved since divergence of man and mouse. Proc Natl Acad Sci USA 81:814–818.

Neumann P, Yan H, Jiang J (2007) The centromeric retrotransposons of rice are transcribed and differentially processed by RNA interference. Genetics 176:749–761.

O'Neill RJ, O'Neill MJ, Graves JA (1998) Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. Nature 393: 68–72.

O'Neill RJ, Eldridge MD, Graves JA (2001) Chromosome heterozygosity and de novo chromosome rearrangements in mammalian interspecies hybrids. Mamm Genome 12:256–259.

O'Neill RJ, Eldridge MD, Metcalfe CJ (2004) Centromere dynamics and chromosome evolution in marsupials. J Hered 95:375–381.

Pagnozzi JM, De Jesus Silva MJ, Yonenaga-Yassuda Y (2000) Intraspecific variation in the distribution of the interstitial telomeric $(TTAGGG)_n$ sequences in *Micoureus demerarae* (Marsupialia: Didelphidae). Chromosome Res 8:585–591.

Pagnozzi JM, Ditchfield AD, Yonenaga-Yassuda Y (2002) Mapping the distribution of the interstitial telomeric $(TTAGGG)_n$ sequences in eight species of Brazilian marsupials (Didelphidae)

by FISH and the correlation with constitutive heterochromatin. Do ITS represent evidence for fusion events in American marsupials? Cytogenet Genome Res 98:278–284.
Peacock WJ, Dennis ES, Elizur A, Calaby JH (1981) Repeated DNA sequences and kangaroo phylogeny. Aust J Biol Sci 34:325–340.
Pevzner P, Tesler G (2003) Genome rearrangements in mammalian evolution: lessons from human and mouse genomes. Genome Res 13:37–45.
Reinhart BJ, Bartel DP (2002) Small RNAs correspond to centromere heterochromatic repeats. Science 297:1831.
Rens W, O'Brien PC, Fairclough H, Harman L, Graves JAM, Ferguson-Smith MA (2003) Reversal and convergence in marsupial chromosome evolution. Cytogenet Genome Res 102: 282–290.
Rofe R (1978) G-banded Chromosomes and the Evolution of Macropodidae. Aust Mamm 2:53–63.
Rofe RH (1979) G-Banding and Chromosomal Evolution in Australian Marsupials. University of Adelaide, Adelaide.
Rofe R, Hayman D (1985) G-banding evidence for a conserved complement in the Marsupialia. Cytogenet Cell Genet 39:40–50.
Salser W, Bowen S, Browne D, et al. (1976) Investigation of the organization of mammalian chromosomes at the DNA sequence level. Fed Proc 35:23–35.
Schibler L, Roig A, Mahe MF, et al. (2006) High-resolution comparative mapping among man, cattle and mouse suggests a role for repeat sequences in mammalian genome evolution. BMC Genomics 7:194.
Sharman GB, Close RL, Maynes M (1990) Chromosomal evolution, phylogeny and speciation of Rock Wallabies (*Petrogale*: Macropodidae). Aust J Zool 37:351–363.
She X, Horvath JE, Jiang Z, et al. (2004) The structure and evolution of centromeric transition regions within the human genome. Nature 430:857–864.
Singer MF (1982) Highly repeated sequences in mammalian genomes. Int Rev Cytol 76:67–112.
Southern EM (1970) Base sequence and evolution of guinea-pig alpha-satellite DNA. Nature 227:794–798.
Svartman M, Vianna-Morgante AM (1998) Karyotype evolution of marsupials: from higher to lower diploid numbers. Cytogenet Cell Genet 82:263–266.
Toder R, O'Neill RJ, Wienberg J, O'Brien PC, Voullaire L, Marshall-Graves JA (1997) Comparative chromosome painting between two marsupials: origins of an XX/XY1Y2 sex chromosome system. Mamm Genome 8:418–422.
Topp CN, Zhong CX, Dawe RK (2004) Centromere-encoded RNAs are integral components of the maize kinetochore. Proc Natl Acad Sci USA 101:15986–15991.
Ugarkovic D (2005) Functional elements residing within satellite DNAs. EMBO Rep 6: 1035–1039.
Valgardsdottir R, Chiodi I, Giordano M, Cobianchi F, Riva S, Biamonti G (2005) Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. Mol Biol Cell 16:2597–2604.
Venolia L, Peacock WJ (1981) A highly repeated DNA from the genome of the wallaroo (*Macropus robustus robustus*). Aust J Biol Sci 34:97–113.
Ventura M, Mudge JM, Palumbo V, et al. (2003) Neocentromeres in 15q24-26 map to duplicons which flanked an ancestral centromere in 15q25. Genome Res 13:2059–2068.
Verdel A, Jia S, Gerber S (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. Science 303:672–676.
Verma RS (1999) Evolution of the centromeric alpha-satellite DNA sequences of human chromosome 22. Prenat Diagn 19:590–591.
Volpe T, Schramke V, Hamilton GL, et al. (2003) RNA interference is required for normal centromere function in fission yeast. Chromosome Res 11:137–146.
Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Martienssen RA (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297: 1833–1837.

White M (1978) Modes of Speciation. WH Freeman and Co, San Francisco, CA.
Wianny F, Zernicka-Goetz M (2000) Specific interference with gene function by double-stranded RNA in early mouse development. Nat Cell Biol 2:70–75.
Willard HF (1990) Centromeres of mammalian chromosomes. Trends Genet 6:410–416.
Yoder JA, Walsh CP, Bestor TH (1997) Cytosine methylation and the ecology of intragenomic parasites. Trends Genet 13:335–340.

Chapter 4
Marsupial Linkage Maps

Paul B. Samollow

Abstract Linkage maps provide information not only on the linear order of genetic loci on chromosomes; they also reveal the locations and frequencies of crossover events during gametogenesis. As such they have been invaluable resources for genetic mapping for nearly 100 years and continue to be extremely useful for providing insight into the meiotic process and disruptions thereof that physical mapping strategies cannot furnish. Linkage mapping research with marsupial mammals had a late inception due to a long-standing lack of animal resources amenable to the reliable, multigenerational breeding programs necessary for the establishment of genetic lines and tools required for linkage-mapping studies. As this impediment was overcome through the establishment of several self-perpetuating marsupial breeding colonies late in the twentieth century, linkage mapping studies began to uncover evidence of unexpectedly low rates and unusual sex-specific patterns of recombination that run counter to those observed for almost all other vertebrates examined. This chapter describes the history of linkage research in marsupials, summarizes the linkage maps of the tammar wallaby and the gray, short-tailed opossum, considers the value of these linkage maps in connection with the recent opossum and wallaby genome projects, and discusses the extent and implications of low recombination rates and sex-specific recombination differences in these species. The chapter ends with speculation on the potential of the opossum as a model for recombination research.

Keywords Linkage mapping · Marsupial · Opossum · Recombination · Wallaby

---

P.B. Samollow (✉)
Department of Veterinary Integrative Biosciences, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University, College Station, TX 77843, USA; Faculty of Genetics, Texas A&M University, College Station, TX 77843, USA
e-mail: psamollow@cvm.tamu.edu

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_4, © Springer Science+Business Media B.V. 2010

4.1 Introduction

Soon after the rediscovery of Mendel's work at the beginning of the twentieth century, it became clear that the genetic factors controlling variation in different characteristics did not always adhere to Mendel's principle of independent assortment. Rather than assorting as independent entities during gametogenesis, some factors tended to co-segregate with one another more often than expected by chance alone. Viewed from the perspective of the then-novel chromosome theory of inheritance, these observations led to the concept of genetic linkage and the possibility of determining the order and relative distances between factors (now called genes) on chromosomes by assessing the frequency of recombination between them during meiosis (Morgan, 1911a, b, c). The first linkage map was a modest affair proposing the orders and interlocus distances of six sex-linked genes in the fruit fly *Drosophila melanogaster* (Sturtevant, 1913). Thus, linkage mapping became the first analytical approach for studying the structural arrangement of genetic elements in genomes.

The realization that linkage data could provide tools for studying the meiotic behaviour of chromosomes and provide a means for manipulating the genetic content of chromosomes to investigate gene interaction (epistasis) and other genetic phenomena, provided incentive to create extensive linkage maps for as many traits as possible for *Drosophila* and encouraged the building of maps for other model organisms such as corn and mouse, wherein pedigreed stocks, short generation time, and the ability to rear large numbers of progeny made the task relatively straightforward. Modest maps were also constructed for human single-gene traits and disease conditions, and for traits of commercial interest in some agricultural species.

Early linkage work was restricted to co-segregation analysis of morphologic and physiologic variants known to be inherited as simple Mendelian traits; however such variants often display dominance, making discrimination of heterozygotes from one of the homozygous classes impossible. This complication limited crossing schemes to simple pedigrees involving just a few variant traits at a time. Beginning in the 1960s, novel biochemical and molecular genetic methodologies enabled the development of new types of map markers that were abundant in the genomes of virtually all eukaryotic organisms and exhibited codominant phenotypes; e.g., allozymes, restriction-fragment-length polymorphisms, microsatellites, and single-nucleotide polymorphisms, which permitted unambiguous assignment of genotypes regardless of the allelic combination. In most cases these molecular markers are "anonymous", i.e., stretches of DNA with no known connection to any trait or condition; but they do unambiguously flag specific locales in the genome and are highly polymorphic, and so are extremely useful for mapping purposes. This plethora of map markers spurred development of sophisticated computational methods that allowed linkage analyses to be conducted with more complex pedigrees involving large numbers of segregating loci. These advances led to a rapid increase in marker content for the linkage maps of established model species and enabled the construction of maps for species in which no single-gene morphologic or physiologic variants were known.

Marsupials (metatherian mammals) were late arrivals on the linkage-mapping scene. Despite attempts with several species, developing marsupial models amenable to reliable, multigenerational breeding programs was for decades an elusive goal, and the lack of such animal resources precluded the establishment of genetic lines and tools necessary to conduct linkage mapping studies (VandeBerg, 1990; Hope, 1993; Samollow and Graves, 1998). This obstacle was eventually overcome as self-perpetuating colonies of the fat-tailed dunnart (*Sminthopsis crassicaudata*), tammar wallaby (*Macropus eugenii*), and gray, short tailed opossum (*Monodelphis domestica*) were established in the 1980s (Hinds et al., 1990; VandeBerg, 1990; Hope, 1993; VandeBerg and Robinson, 1997). By that time, however, DNA-hybridization-based methods for the physical localization of probes homologous to genes already mapped in human and rodent models had been developed and were being applied with great enthusiasm to a variety of marsupial species, wild and colony-bred alike (Samollow and Graves, 1998; Graves and Westerman, 2002). The ability of these physical methods to locate the positions of highly conserved genes on marsupial chromosomes rapidly, and without the need for genetic variants or breeding procedures, effectively quenched interest in linkage studies as a means to construct marsupial genetic maps for comparative gene mapping research (see Deakin, Chapter 5).

Nevertheless, the availability of marsupial breeding colonies together with the success of human and mouse linkage mapping studies based on the genotyping of highly informative microsatellite polymorphisms proved irresistible, and efforts were undertaken to construct linkage maps for both the wallaby and opossum, the two most intensively used marsupial species for basic biological and biomedically related research. This chapter furnishes an historical overview of linkage mapping in marsupials, describes the current state of their linkage maps, and provides a glimpse of current and possible future applications of these maps in genomic analysis.

4.2 Why Construct Linkage Maps?

A linkage map (also known as a genetic or meiotic recombination map), as defined herein, is a statistical representation of the linear order and distances between specific genetic loci on chromosomes of sexual organisms, based solely on the frequency of co-segregation among these loci during meiosis. Construction of modern linkage maps is accomplished most often through analysis of multi-locus genotypes of offspring produced by parents of known or inferred genetic composition by means of any of several computationally intensive analytical tools that have been devised over the past two decades. Linkage distances are generally given in units known as centiMorgans (cM). One cM corresponds to 1% recombination between two particular loci per meiosis. The genotyping of individual gametes is another linkage mapping approach that enables the multilocus analysis of hundreds or thousands of meiotic products from a single individual, but for mammals it

is essentially limited to male gametes, and its application to large numbers of individuals can be challenging.

The production of families and the genotyping of hundreds of individuals that comprise linkage-mapping panels are both costly and time consuming. One might ask, then, with highly efficient and relatively inexpensive physical mapping methods readily available, why bother with linkage mapping any more? The answer lies in the kind of information generated by linkage data compared with that furnished by physical maps (including full genome sequence assemblies). Physical and linkage maps both reveal the order and relative distances between genetic elements of interest on chromosomes, but only linkage data yield information on the behaviours of chromosomes during meiosis.

As is the case for physical maps, the order of loci on a linkage map reflects the order of loci on the chromosome. However, the distances between loci on a linkage map (map distances) do not necessarily correspond to the physical distances (number of DNA bases) between the loci, but rather reflect the frequency with which crossing-over occurs between the loci during meiosis. There is of course a correlation between physical distance and linkage distances, but this correlation is a loose one and its strength may vary between regions on the same chromosome, between chromosomes, between individuals, and between species. Thus, whereas physical maps yield information about the arrangement and the genomic coordinates (in base pairs or other physical metrics) between features on chromosomes, linkage maps provide information concerning the likelihood that a crossover event will occur in a particular chromosomal region defined by flanking marker loci. Used together, physical and linkage maps enable the estimation of recombination events per physical distance, usually expressed as map distance per megabase (Mb) pairs, or simply cM/Mb. This is the physical recombination rate.

Linkage maps can be used as independent datasets to corroborate marker order developed from physical mapping procedures, support the construction of genome sequence assemblies, and refine the order and orientation of large-scale structures in existing genome assemblies. In addition, they uniquely provide information on the distribution and frequency of recombination events across the genome, thereby providing a means to examine the molecular mechanisms that promote or inhibit the formation of chiasmata (the physical manifestation of meiotic crossing over) in different regions of chromosomes, which in turn can influence the success or failure of homologous chromosomes to disjoin normally at anaphase I. Finally, linkage maps enable the localization of quantitative trait loci (QTLs) that influence normal and abnormal physiologic and developmental variation, disease susceptibilities, and traits of economic importance.

4.3 Pioneering Linkage Studies in Marsupials

The earliest evidence regarding genetic linkage in marsupials arose from investigations of expression patterns of marsupial homologues of genes known to be X-linked in eutherian ("placental") mammals. In keeping with Ohno's (1967) conjecture

that the gene content of X chromosomes should be highly conserved across all mammals, simultaneous publications reported that electrophoretic mobility variants (allozymes) of glucose-6-phosphate dehydrogenase (G6PD) (Richardson et al., 1971) and phosphoglycerate kinase A (PGK-A) (Cooper et al., 1971) were inherited as X-linked traits in macropodid (kangaroo family) intra- and inter-species crosses, and that the genes encoding these enzymes (*G6pd* and *Pgk-A*, respectively) were subject to X-chromosome inactivation, as is true in eutherian mammals. Surprisingly, data from female offspring heterozygous at these loci failed to reflect the random pattern of X-linked gene inactivation seen in eutherian females, but instead indicated that it was exclusively the allele inherited from the father that was inactivated, regardless of which species or subspecies served as the sire in the cross. This unexpected, non-random X-inactivation pattern, which became known as paternal X-chromosome inactivation to distinguish it from the eutherian pattern of random X-chromosome inactivation (see Hore et al., Chapter 12 and Al Nadaf et al., Chapter 13), arguably represents the earliest observation of genomic imprinting in animals.

These findings spurred a great deal of activity to assess the expression of X-linked genes in a variety of Australian marsupial species and eventually an American one, and concurrently yielded the first chromosomal gene assignments accomplished by genetic analysis in marsupials (reviewed by VandeBerg et al., 1987; Cooper et al., 1990, 1993). However, limited offspring numbers and lack of appropriate multilocus variation in these species precluded the construction of linkage maps. Defining the locations of these genes on marsupial X chromosomes was ultimately accomplished through physical mapping methods.

The emphasis of early marsupial gene mapping was decidedly comparative. The general objective was to better understand genome evolution by exploiting eutherian/marsupial differences as probes to explore the phylogenetic nuances of sex determination, gene expression, and other genetic processes related to the evolution of gene structure and function (Samollow and Graves, 1998; Graves and Westerman, 2002). The fundamental strategy was to produce genetic maps that could be used to investigate the degree to which syntenic relationships were conserved or altered among species and between major mammalian clades. The emphasis was on highly conserved loci, especially protein-coding genes, the homologies of which could be inferred across vast evolutionary distances. For reasons already mentioned, physical mapping approaches overshadowed linkage analysis as the favoured means to determine syntenic relationships on marsupial chromosomes. This led to rapid progress in the construction of marsupial gene maps, but furnished no information on the behaviour of marsupial chromosomes during gametogenesis.

The first linkage estimates for a marsupial were derived from a combined cytologic and family study of the fat-tailed dunnart, *Sminthopsis crassicaudata* (family Dasyuridae). Linkage analysis defined two linkage groups (LG I and LG II), each containing three genes (Table 4.1), and indicated that recombination between these genes was much less frequent in female dunnarts than in males. Four pairwise gene combinations exhibited no recombination in females but showed modest to high recombination frequencies (rf: 0.16–0.37) in males, while two other pairwise

Table 4.1 Results from early linkage studies in marsupials: linkage groups and sex-specific recombination fractions

Species (linkage group)	Marker interval	Maternal recombinants[a]	Female recombination fraction	Paternal recombinants[b]	Male recombination fraction	References
S. crassicaudata (LG I)	*ADA – GPI*	0/211	0.00	88/236	0.37	Bennett et al. (1986)
	ADA – PI	0/61	0.00	26/78	0.33	Bennett et al. (1986)
	GPI – PI	0/11	0.00	29/116	0.25	Bennett et al. (1986)
S. crassicaudata (LG II)	*SOD – TRF*	0/169	0.00	55/335	0.16	Bennett et al. (1986)
	SOD – PGD	6/97	0.06	99/197	0.50	Bennett et al. (1986)
	TRF – PGD	52/419	0.12	172/334	0.51	Bennett et al. (1986)
	PGD – HBE	[c]	0.00	[c]	"unlinked"	Hope (1993)
	SOD – HBE	[c]	0.07	–[d]	–	Hope (1993)
M. eugenii (LG I)	*GPI – PI*	–	–	[c]	0.27	van Oorschot and Cooper (1989)
M. eugenii (LG II)	*pB72 – LALBA*	–	–	[c]	0.06	McKenzie et al. (1993)
M. eugenii (LG III)	*LLP – LPL*	–	–	[c]	0.28	McKenzie et al. (1993)
	LLP – pB65	–	–	[c]	0.07	McKenzie et al. (1993)
	LPL – pB65	–	–	[c]	0.30	McKenzie et al. (1993)
M. eugenii (X chr)	*RNR – G6PD*	0/69	0.00	–	–	McKenzie et al. (1996)
M. domestica (LG 1)	*AK1 – PI*	[c]	0.09	[c]	0.48	van Oorschot et al. (1992)
	AK1 – PI	–	–	10/36	0.28	Perelygin et al. (1996)
	AK1 – U15557	0/29	0.00	1/36	0.03	Perelygin et al. (1996)
	PI – U15557	4/51	0.08	9/36	0.25	Perelygin et al. (1996)
M. domestica (LG 2)	*C6 – C7*	[c]	0.00	[c]	0.00	van Oorschot et al. (1993)
	C6/C7 – GPT	[c]	0.17	[c]	0.41	van Oorschot et al. (1993)

[a](Observed offspring haplotypes attributable to maternal recombination events)/(total offspring observed)
[b](Observed offspring haplotypes attributable to paternal recombination events)/(total offspring observed)
[c]Raw data not listed in referenced article
[d]No data available for this interval/sex combination

combinations that had minimal recombination in females (rf: 0.06, 0.12) exhibited independent assortment (were unlinked) in males (Bennett et al., 1986). This situation was unexpected because in humans and other eutherians females generally exhibited higher recombination rates than males. The accompanying cytologic examination of meiotic cells revealed lower chiasma frequency in females than in males and a strong clustering of chiasmata in terminal (sub-telomeric) regions of female, but not male, autosomes, suggesting a causal relationship between chiasma characteristics and recombination rates in the dunnart (Bennett et al., 1986). The discovery several years later (Hope, 1993) of linkage between the genes *HBE* and *PI* enlarged LG II and corroborated the initial findings of strong heterochiasmy (sex difference in recombination) by showing that *HBE* and *PI* did not recombine with one another in female dunnarts (rf = 0.00), but assorted independently in males.

Following upon studies in the dunnart, another cytologic examination found very similar meiotic characteristics (lower frequency and restricted distribution of female chiasmata) in the distantly related grey, short tailed opossum, *Monodelphis domestica* (family Didelphidae), leading to the prediction that females of this South American marsupial would also exhibit reduced recombination in linkage studies (Hayman et al., 1988). The prediction was borne out by the first linkage data produced for opossums (van Oorschot et al., 1992, 1993; Perelygin et al., 1996) (Table 4.1). Female opossums exhibited low to moderate recombination rates in the intervals *PI* – *U5557* (rf = 0.08), *AK1* – *PI* (rf = 0.09), and C6/C7-GPT (rf = 0.17), whereas the corresponding rates for males were much higher (0.25, 0.48, and 0.41, respectively).

The only other marsupial for which linkage data have been published is the tammar wallaby, *Macropus eugenii* (family Macropodidae). Early linkage data for this species showed that the *GPI* – *PI* linkage, previously noted in the dunnart, was also present in the wallaby, providing some of the earliest data on synteny conservation for deeply divergent marsupial clades (van Oorschot and Cooper, 1989). The *GPI* – *PI* linkage defined wallaby LG I (Table 4.1). Subsequent studies (McKenzie et al., 1993) led to the discovery of two additional linkage groups, LG II (anonymous marker *pB72* and *LALBA*), LG III (*pB65*, *LLP*, and *LPL*), and very tight linkage between *RNR* and *G6PD* on the X chromosome (McKenzie et al., 1996). Sex-specific data were not reported in these studies.

None of these pioneering efforts established the order of genetic markers within their respective linkage groups, so did not create linkage maps. Nonetheless, they formed the foundation for linkage work upon which linkage maps would ultimately be built. First, they proved the practicality of conducting large family studies with marsupials, which encouraged the characterization of genetic diversity within pedigreed marsupial research colonies. These characterizations fostered greater interest in the genetic features of marsupials as tools for examining the molecular underpinnings of the distinctive reproductive and developmental features that distinguish marsupial and eutherian mammals. Equally important, these investigations were the first to call attention to a very unusual pattern of heterochiasmy in which females show much lower levels of recombination than males. As discussed below, this is an exceedingly rare pattern in vertebrates, and certain features of the pattern may be unique to marsupials.

4.4 Linkage Maps of Marsupials

4.4.1 The Tammar Wallaby Linkage Map

The first marsupial linkage map was produced by Kyall Zenger and colleagues for the tammar wallaby (Zenger et al., 2002). By that time, physical mapping investigations had already demonstrated the power of non-genetic approaches for determining synteny relationships and linear orders of highly conserved (type-1) coding genes on marsupial chromosomes. The impetus for constructing a wallaby linkage map was interest in its use as a tool for localizing unidentified genes (QTLs) that influenced variation in physiological, developmental, and health-related traits in the wallaby, which was well established as one of the most important and intensively used marsupial research models in the world.

Lacking inbred strains of this species, the mapping strategy was based on creating a highly informative mapping panel through crosses involving wallabies descended from different natural populations that were geographically distant from one another and, thus, expected to be strongly genetically differentiated. The F_1 hybrids so produced would be expected to be heterozygous at many genetic loci. The F_1 could then be backcrossed to one or the other parental stock to produce highly informative backcross progeny. It had been shown previously (McKenzie et al., 1993) that crosses between distantly related tammar wallaby subspecies from Garden Island (GI), Western Australia and Kangaroo Island (KI), South Australia produced fertile F_1 offspring and could be used for creating such a panel. Using this strategy, Zenger and colleagues analyzed 242 backcross offspring produced by matings between 25 hybrid parents crossed to six GI and KI mates.

The wallaby mapping panel was genotyped for variation at 64 genetic loci including type-1 coding genes, anonymous microsatellite markers, and other anonymous DNA sequences. Of these, 60 markers showed convincing pairwise linkage with at least one other marker and could be clustered into nine distinct linkage groups (Table 4.2). Markers within each linkage group were ordered and interlocus distances were estimated to yield nine multipoint linkage-group maps ranging from 176.5 to 15.7 cM sex-average length. Summed over all linkage groups, the total map was 828.8 cM long. Previous physical mapping studies with some of the same type-1 loci enabled anchoring of three of the linkage groups to specific autosomes and the identification of an X-chromosome linkage group. Two of these linkage groups anchored to the same autosome, bringing the number of linkage groups into agreement with the $n = 8$ haploid complement of the species. Inheritance patterns indicated that the five unanchored linkage groups were autosomally located. The autosomal linkage groups accounted for 771.7 cM of the total map length and the X-chromosome linkage group for 57.1 cM. Markers placed uniquely (i.e., with non-overlapping confidence ranges) defined 37 interlocus intervals.

Sex-specific recombination rates were inferred by comparing the lengths of the maps produced by recombination occurring in female parents vs. male parents. The overall length of the female autosomal map was only 78% that of the male map (F/M ratio = 0.78), and this difference was highly significant ($P < 0.001$ that the true

Table 4.2 Characteristics of the tammar wallaby first-generation linkage map

Linkage group	Chromosomal assignment[a]	No. of markers mapped	No. of type-I markers	Sex-average length (cM)	No. of interlocus intervals	No. of sex-divergent intervals	F/M length ratio
1	A (Chr 2)	12	–	176.5	10	3	0.64**
2	Chr 1 (Chr 4)	13	1	151.5	9	2	0.84*
3[b]	Chr 1	8	3	107.0	4	1	0.66*
4[c]	Chr 3	7	2	103.5	4	0	[d]
5	A (Chr 7)	6	2	100.4	3	0	[d]
6	A (Chr5)	4	–	76.9	2	1	0.49***
7	Chr X	5	4	57.1	2	0	[d]
8	A (Chr 6)	3	–	40.2	2	0	[d]
9	A (Chr 1)	2	–	15.7	1	1	0.51*
Total; autosomes	–	55	8	771.7	35	8	–
Total; all	–	60	12	828.8	37	8	0.78***

[a]Non-parenthetical assignments were inferred by Zenger et al. (2002) from previous in situ hybridization studies of type-I genes present the linkage group. A = autosomal, chromosome not identified. Assignments in parentheses are from the CompLDB website (http://cmap.sg.angis.org.au/cmap/)
[b]LG III of McKenzie et al. (1993)
[c]LG II of McKenzie et al. (1993)
[d]Not significantly different from 1.00. Ratio not listed by Zenger et al. (2002)
*probability (F/M ratio = 1.0) < 0.05
**probability (F/M ratio = 1.0) < 0.01
***probability (F/M ratio = 1.0) < 0.001

F/M = 1.0). Eight of the 37 interlocus intervals, accounting for 26.2% of the autosomal map, exhibited significant heterochiasmy, with the female having the shorter interval in each case (Table 4.2); and each of the linkage groups that contained these significant intervals were themselves significantly shorter in females than in males. Remarkably, when the eight significant intervals were removed, none of the linkage groups exhibited heterochiasmy and the overall autosomal map lengths of the two sexes converged (F/M =1.01). This suggested that reduced recombination does occur in wallaby females, but is probably localized to specific chromosomal regions rather than a whole-chromosome or genome-wide phenomenon.

The proportion of the wallaby genome covered by this map was not estimated from the mapping data, but an estimate based on chiasma counts from spermatogenic cells suggested that the wallaby recombinational map was ~1,172 cM, implying that the map covered at least 71% of the genome. The 828.8 cM sex-average map and the 1,172 cM male cytologic estimate were among the shortest known for vertebrates at that time, and shorter than any known for mammals with linkage maps of comparable or greater extent. The short map was not the result of a diminutive genome; the wallaby genome, at ~3.6 gigabases (Gb), is larger than that of many species that have much larger linkage maps. For example, it is ~5%

larger than the human genome, which has a sex-average recombinational length of ∼3,790 cM (Matise et al., 2007). The short map instead reflects an overall low rate of recombination in the wallaby genome relative to other vertebrate genomes. As we will see later, low recombination rate is also characteristic of the opossum genome.

Building on the initial map, a second-generation wallaby linkage map has been compiled and will soon be published (Chenwei Wang, personal communication). The second-generation map comprises 150 loci including the 60 markers placed on the original map, four original markers that could not be placed on the original map, and 86 new microsatellite loci. Judged from fluorescence in situ hybridization (FISH) analysis using hybridization probes containing linkage map markers, the second-generation map covers ∼85% of the wallaby genome. Definitive chromosomal assignments have been made for all linkage groups. The second-generation map is larger, exceeding 1,000 cM, but remains comparatively short by mammalian standards (Chenwei Wang, personal communication). It still exhibits sexual dimorphism in overall map length, and male map length exceeds female map length for several, but not all of the linkage groups, reinforcing the original impression that reduced female recombination is localized rather than a genome-wide characteristic of the wallaby genome.

4.4.2 Gray, Short-Tailed Opossum Linkage Maps

As for the wallaby, the initial motivation for linkage mapping in the opossum was to generate a resource that could serve as a framework for QTL mapping. An additional impetus was to determine the magnitude and describe the genomic distribution of the striking intersexual difference in recombination rates observed in early opossum linkage data as a means to better understand the meiotic characteristics of the species.

Lacking inbred strains, the mapping strategy for the opossum was similar to that described for the wallaby. Two independent linkage maps have been constructed for the opossum using different mapping panels: the GMBX and BBBX panels. Three random-bred laboratory stocks designated, Pop1, Pop3, and Pop5 were used to establish the panels. Pop1 and Pop3 founders originated from two distantly separated locations in Brazil; Pop5 founders were captured in central Bolivia. Each mapping panel was constructed by first crossing parents from two different stocks (GMBX: Pop1 × Pop3; BBBX: Pop1 × Pop5) to create an F_1 generation. For each panel, F_1 individuals of both sexes were backcrossed to Pop1 mates to produce the backcross generation (Pop3 and Pop5 mates were unavailable for the reciprocal backcrosses). The BBBX-based map is the more highly developed of the two and serves as the basis for ongoing mapping studies, but the general features of the GMBX map bear discussion to provide a basis for appreciating the similarity of results obtained from the two separate mapping studies.

The GMBX panel comprises 356 individuals from 30 three-generation families. The mapping dataset includes genotypes for 83 polymorphic loci including

a mixture of type-1 coding genes, anonymous microsatellite loci, and anonymous random-amplified polymorphic DNA markers (RAPDs). These markers clustered at high confidence into nine distinct linkage groups, in agreement with the $n = 9$ haploid number of the species (Samollow et al., 2004). Three X-linked loci were excluded from analysis leaving an autosomal dataset of 80 marker loci in eight linkage groups (Table 4.3). Several marker pairs did not exhibit recombination in either sex, and some loci could not be ordered at high confidence in their linkage group maps. The successfully ordered, recombining markers defined 56 unambiguous interlocus intervals. These intervals covered a total sex-average length of only 633.0 cM. Adding estimates for the unmapped regions beyond the ends of the autosomal linkage groups and a "best guess" for the length of the X-chromosome map, the full length sex-average map was estimated to be a mere 890.7 cM, the shortest known for any vertebrate for which there is a substantially complete map. At ~3.6 Gb, the opossum genome is at least as large as that of the wallaby and larger than those of humans and most other eutherian mammals; and evidence presented by Samollow and colleagues (2004) argued that a large fraction of the genome was covered by the linkage map. Thus, the short map length was not attributable to small genome size or poor genome coverage, but rather was indicative of a very low recombination rate in opossums.

Turning to sex-specific maps, the GMBX female-specific map was only half the length of the male map (443.1 vs. 884.6 cM) (F/M = 0.501) and the length of each male linkage-group map exceeded that of the corresponding female map (Table 4.3). Of the 56 recombining intervals, 22, representing 45.9% of the total map length, differed significantly in length between the sexes, with 20 of them longer in the male map than in the female map. Contrasting with the situation in the wallaby, removal of the 22 significant intervals did not equalize the lengths of the female and male maps, which implies that the recombinational differences underlying the sexual dimorphism in map lengths in the opossum are not localized to a few small dimorphic regions or just a few chromosomes, but are widespread across the genome.

Overall then, results from the GMBX mapping analysis indicated that the opossum has a very low meiotic recombination rate and that females have a substantially lower rate of recombination than males, a pattern that runs counter that observed in eutherian mammals and non-mammalian vertebrates, in which recombination rates are either similar in both sexes or biased toward higher rates in females.

Based on inferential evidence it was reasoned that the GMBX map covered all eight opossum autosomes and encompassed a large proportion of the autosomal genome (Samollow et al., 2004). Even so, the map was based on only 80 loci, and the formal possibility remained that substantial regions of some chromosomes might have been overlooked. Thus, it seemed prudent to increase the number of markers on the map to see if its very unusual characteristics would persist with greater map density. Specifically, we wondered if adding more markers would lead to a substantial increase in map length, and if so, what effect this increase would have on the disparity in sex-specific map lengths (Samollow et al., 2007). This issue arises from observations that female, but not male, chiasmata are concentrated at chromosome

Table 4.3 Characteristics of the opossum GMBX linkage map

Linkage group (autosomal)	No. of markers mapped	No. of markers ordered	No. of interlocus intervals[a]	Map length (cM) Sex average	Female	Male	F/M ratio	F/M probability[b]	No. of intervals exhibiting significant heterochiasmy[c]: female map/male map
LG1	16	14	11	113.4	51.2	172.9	0.296	$2e^{-7}$	0/7
LG2	8	7	4	56.4	18.9	100.4	0.188	$1e^{-12}$	0/4
LG3	15	13	11	139.6	111.2	194.9	0.571	$1e^{-11}$	1/3
LG4	13	12	10	80.7	48.4	105.5	0.458	$2e^{-5}$	1/3
LG5	6	6	4	41.6	40.5	46.8	0.865	0.29	0/0
LG6	9	8	7	70.6	56.5	82.9	0.682	0.003	0/0
LG7	7	6	3	64.9	61.4	73.4	0.837	0.19	0/1
LG8	7	7	6	65.8	55.0	107.8	0.510	0.41	0/2
TOTAL	80	73	56	633.0	443.1	884.6	0.501	–	2/20

[a] Intervals with non-zero recombination in at least one sex
[b] Probability that F/M = 1.00 across entire linkage group; see Samollow et al. (2004) for test details
[c] Defined as a statistically significant departure from F/M = 1.00; see Samollow et al. (2004) for test details

ends in the opossum (Hayman et al., 1988) and dunnart (Bennett et al., 1986). That being the case, newly added markers falling in subtelomeric regions beyond the ends of the existing linkage groups could disproportionately increase the length of the female map relative to the male map. Another and perhaps more important question was whether the heterochiasmy observed in the GMBX map mirrored a fundamental opossum characteristic or rather might have been peculiar to the specific cross used to create the mapping panel. This possibility argued that building a new and more densely populated linkage map should be done using a different mapping panel.

The BBBX panel includes 29 Pop1 and 8 Pop5 grandparents, 33 F_1 and 33 Pop1 mates, and 468 third-generation offspring for a total of 571 individuals. Map construction with this panel is ongoing, but only about half of the data have been published. Herein I summarize the general characteristics of the initial BBBX linkage map based on 150 mapped markers (Samollow et al., 2007) and update the general trends based on inspection of additional unpublished data. Because the BBBX map is based almost entirely on microsatellite markers, only 34 of the GMBX markers overlap with the BBBX study and no attempt has been made to integrate GMBX and BBBX maps. However, the positions of the 34 overlapping markers are fully consistent between the two maps.

The salient features of the GMBX map, i.e., short sex-average length and the very strong, male-biased heterochiasmy, are fully evident in the BBBX mapping data, indicating that the unusual meiotic features inferred from the GMBX-based map are not idiosyncratic to a specific mapping panel, but are fundamental features of the opossum genome (Table 4.4). The sum of the eight BBBX autosomal linkage-group maps is only 714.8 cM. Adding estimates for unmapped regions yields a full autosomal map-length estimate of 866.2 cM, which is remarkably similar to the 890.7 cM sex-average estimate from the GMBX map. The length of the male-specific map exceeds that of the female map for every linkage group, and in all cases the difference is highly statistically significant (Table 4.4). Summing all linkage groups, the female- and male-specific maps cover 515.5 and 948.2 cM (579.4 and 1,064.9 cM after correcting for unmapped regions), respectively, yielding an F/M ratio of 0.544, which is in close agreement with the 0.501 F/M ratio of the GMBX sex-specific maps. Markers exhibiting X-linked inheritance were not included in the mapping analysis.

The chromosomal affiliation of each linkage group, its orientation on the corresponding chromosome, and the extent of chromosomal coverage by each linkage group were determined by FISH mapping using BAC clones that encompassed particular marker regions of interest, including the terminal markers of each linkage group. Individual linkage groups were estimated to cover between ~72.9 and ~97.5% of their respective chromosomes, and total coverage was estimated at ~86.4% of the autosomal genome. This physical estimate agreed well with a statistical approach that yielded an estimate of ~89.4% coverage. Averaging the FISH-based and statistical estimates suggests that the published BBBX linkage map encompasses ~87.8% of the opossum autosomal genome.

The virtually identical results from the GMBX and BBBX studies, which used different genetic crosses and largely non-overlapping genetic marker sets, clearly

Table 4.4 Characteristics of the opossum BBBX linkage map

Chromosome	Linkage group	No. of markers	No. of markers ordered	No. of intervals[a]	Map length (cM)			F/M ratio	F/M proba-bility	No. of intervals exhibiting significant hetero-chiasmy[b]: female map/male map	Estimated chromo-some size (Mb)[c]	Percentage of chromo-some covered by LG[d]	Recomb. rate (cM/Mb)[e]
					Sex average	Female	Male						
1	LG1	38	36	35	181.9	154.5	221.7	0.697	$2e^{-64}$	3/22	748	97.5	0.26
2	LG3	27	23	22	120.6	85.1	159.5	0.534	$6e^{-39}$	0/10	542	93.8	0.24
3	LG2	20	16	15	78.7	44.3	120.2	0.369	$8e^{-33}$	0/9	528	84.2	0.17
4	LG4	17	17	16	71.1	42.8	101.2	0.423	$5e^{-25}$	0/6	435	74.2	0.19
5	LG8	13	13	12	67.2	44.8	91.8	0.488	$2e^{-17}$	0/6	305	77.8	0.26
6	LG7	8	8	7	61.3	36.3	91.5	0.397	$2e^{-22}$	0/4	292	89.5	0.25
7	LG5	16	16	15	68.4	53.1	84.2	0.631	$1e^{-7}$	0/5	261	86.6	0.30
8	LG6	11	11	10	65.6	54.6	78.1	0.699	$2e^{-18}$	1/4	313	72.9	0.24
(X)	–	–	–	–	–	–	–	–	–	–	79	–	(0.44)[f]
TOTAL	–	150	140	132	714.8	515.5	948.2	0.544	–	4/66	3503	86.4	0.23[g]

[a]Intervals with non-zero recombination in at least one sex
[b]Defined as a statistically significant departure from F/M = 1.00; see Samollow et al. (2004) for test details
[c]Sequenced bases plus spanned gaps; Mikkelsen et al. (2007); Supplementary Information
[d]Estimated % of the chromosome lying between terminal FISH signals for the linkage group; see Samollow et al. (2007) for details
[e]Calculated from data of Mikkelsen et al. (2007); Supplementary Information, and Samollow et al. (2007), sex-average map
[f]Estimated assuming a minimum of one recombinant per meiotic bivalent (females only) – see text or Mikkelsen et al. (2007)
[g]Autosomes only

show that the short map length and strong, genome-wide and male-biased heterochiasmy are not artifacts of the particular population crosses employed. Moreover, inclusion of ∼88% of the autosomal genome in the BBBX map greatly diminishes the likelihood that these results can be attributable to gross under representation of large portions of the genome.

Since publication of the BBBX map nearly 150 more markers have been added, many chosen by inspection of the opossum genome assembly (Mikkelsen et al., 2007) specifically to fill gaps in the map. Preliminary assessment of the data suggest there has been some expansion of female linkage-group maps due to the targeted addition of markers in subtelomeric regions, which would be expected to affect the female map length more than the male map length. This sex-specific differential growth of linkage-group ends is expected to increase the F/M ratio slightly, but the difference in sex-specific map lengths remains strikingly large. As a result of sex-specific expansions the overall sex-average map has also grown, but it is likely to remain under 1,000 cM.

We are also developing a linkage map for the opossum X chromosome. On the presumption that all chromosomes must undergo at least one recombinational event to assure proper disjunction at meiosis I, the diminutive opossum X (Table 4.4) was estimated to require a minimum recombination rate of 0.44 cM/Mb, a rate much higher than that for the autosomes (Mikkelsen et al., 2007). Preliminary data from ∼30 X-linked loci bear out this prediction and suggest that the X-chromosome recombination rate is likely to be considerably higher than the theoretical minimum (K. C. Douglas and P. B. Samollow, unpublished data).

4.5 Implications of Low Recombination Rates

The wallaby and opossum genomes have some of the lowest recombination rates among vertebrates. The opossum estimated sex-average map length of 866 cM, for example, falls far below the ∼1,630 cM sex-average mouse linkage map (Shifman et al., 2006), which is the smallest known among eutherian species. As previously noted, the human genome is slightly smaller than that of the opossum, but its linkage map spans ∼3,790 cM (Matise et al., 2007). Among other vertebrates with substantially complete linkage maps none have been reported to have shorter maps than the opossum. The combination of short linkage map and large genome translates to a very low average physical recombination rate of only ∼0.23 cM/Mb for the opossum autosome set, a rate far lower than that for mouse which, at ∼0.63 cM/Mb, has the lowest rate known among eutherian mammals. Rates for the individual opossum autosomes vary between ∼0.17 and ∼0.30 cM/Mb (Table 4.4). The corresponding genome-wide rate for the wallaby, based on the published map (Zenger et al., 2002), is ∼0.33 cM/Mb. Although higher than that of the opossum, this is still very low by mammalian standards.

This low recombination rate may account for the unusual base composition of the opossum genome relative to that of other fully sequenced and annotated

amniote genomes. The opossum autosomes average only 37.7% G and C nucleotides (G+C), contrasting with levels of 40.9-41.8% for human, mouse, dog, and chicken (Mikkelsen et al., 2007), and 45.5% in the platypus (Warren et al., 2008). Mirroring the low G+C content, the opossum autosomes also possess $\leq$ 50% the average CpG dinucleotide content of other sequenced amniote genomes. A growing body of evidence suggests that G+C and CpG contents of a chromosomal region are strongly influenced by the interplay of two opposing mutational processes; a general GC $\rightarrow$ AT mutation bias that reduces G+C content, and biased gene conversion which increases G+C content through biased AT $\rightarrow$ GC mismatch repair during recombination-mediated gene conversion events (Hogstrand and Bohme, 1999; Galtier et al., 2001; Duret et al., 2006). This model predicts that G+C and CpG contents will be positively correlated with recombination rate on a regional basis within chromosomes and averaged across entire chromosomes. Averaged across whole genomes, differences in species-specific recombination rate would be expected to drive interspecific diversity in overall G+C and CpG levels. The low G+C and CpG content of opossum autosomes are consistent with this model. Also consistent with the model, the G+C and CpG fractions of the opossum X chromosome are much higher than those of the autosomes (40.9 and 1.4%, respectively), which is particularly interesting in light of the fact that X chromosomes of eutherians tend to have substantially lower G+C and CpG levels than their autosomes (Mikkelsen et al., 2007).

4.6 Sex-Role Reversal in Recombination Rates

In addition to the species already discussed, it appears that recombination occurs more frequently in male than female meiosis in at least one other marsupial, the brush-tailed possum, *Trichosurus vulpecula* (family Phalangeridae). Cytologic data from this Australian species indicate that female meiocytes have reduced frequency and distal clustering of chiasmata relative to those of males (Hayman and Rodger, 1990). The only contrary finding so far has been from the brush-tailed bettong, *Bettongia penicillata* (family Potoroidae), another Australian marsupial, which shows no obvious difference in sex-specific chiasma number or distribution (Hayman et al., 1990). Unfortunately there are no linkage data for the possum or bettong. Of the two species with linkage maps, the situation in the opossum is the most extreme with the female linkage map barely half the size of the male map (54.5%). This degree of heterochiasmy is equaled in humans, house cats, and domestic pigs (Matise et al., 2007; Menotti-Raymond et al., 2009; Vingborg et al., 2009) and considerably surpassed in some fish species (Singer et al., 2002; Kai et al., 2005; Moen et al., 2008; Kucuktas et al., 2009); however, in all such cases it is the female that has the longer map. Indeed, whenever heterochiasmy is encountered, it is almost invariably the female that has the higher recombination rate.

Among non-marsupial vertebrates only two convincing cases of male-biased heterochiasmy have been reported (other reported cases, based on pseudotestcross strategies and other single-sibship F_1 mapping designs, are not considered here

because such approaches are unreliable for discriminating genuine sex-specific recombination rates from normal, interindividual variation in recombination rates that is not sex related. In addition, maps based on such designs often exhibit poor overlap in genomic coverage between the parents and may not be representative across the genome generally). Female collared flycatchers (*Ficedula albicollis*) and domestic sheep (*Ovis aries*) have shorter linkage maps than their male counterparts. Overall F/M for the flycatcher (a bird) is ~0.82 but, as was the case for the wallaby, not all linkage groups exhibit heterochiasmy (Backstrom et al., 2008). All other bird species examined either display no heterochiasmy or have higher female than male recombination rates (Calderon and Pigozzi, 2006; Akesson et al., 2007; Stapley et al., 2008; Groenen et al., 2009; Jaari et al., 2009).

The situation in sheep is particularly interesting. Sheep exhibit an overall F/M map ratio of ~0.85 (Maddox et al., 2001 and Jill Maddox, personal communication). Across most of the map, the distances between markers are very similar for both sexes (little or no heterochiasmy), but the male recombination rate rises dramatically in portions of the map corresponding to subtelomeric and pericentric regions of the chromosomes, resulting in longer male linkage-group maps and overall map length (Jill Maddox, personal communication). The low overall F/M ratio of sheep thus reflects a largely sex-equal recombination pattern punctuated by much higher male rates in subtelomeric and pericentric regions only. Increasing male recombination (reduced F/M ratio) in subtelomeric regions is characteristic of eutherians, and possibly vertebrates generally, occurring in humans, mice, and pigs (Broman et al., 1998; Kong et al., 2002; Shifman et al., 2006; Vingborg et al., 2009) and in several fish species (Singer et al., 2002; Reid et al., 2007; Kochakpour and Moens, 2008; Moen et al., 2008), so sheep are typical in this regard. The rise in male recombination in pericentric regions of sheep is unusual however, being opposite of what is known based on humans and mice, wherein the F/M ratio in these regions generally increases sharply due to a drop in male recombination and concomitant rise in female recombination.

In any case, the recombination pattern underlying the low F/M ratio in sheep is quite distinct from that in the opossum, wherein males have higher recombination across almost the entire genome. Equally important, the opossum map does not exhibit the eutherian pattern of sharply increased male relative to female recombination rate (reduction in F/M ratio) in subtelomeric regions. Indeed, the pattern in the opossum is reversed, with the F/M ratio rising in subtelomeric regions. Compared with the male map, inter-locus distances in regions corresponding to the midsections of chromosomes appear highly compressed in the female map while those corresponding to the ends of chromosomes are expanded (Fig. 4.1). This pattern, observed on several chromosomes (Samollow et al., 2007), is expected if female opossums regularly exhibit strong clustering of recombination events at the ends of chromosomes but males do not, as suggested by the cytologic observations of Hayman and colleagues (1988). Finally, it does not appear that opossum males have reduced pericentromeric recombination rates relative to females, as is seen in mouse and human maps, although rigorous conclusions must await more precise localization of centromeres on the linkage maps. Overall then, the sex-roles for the meiotic

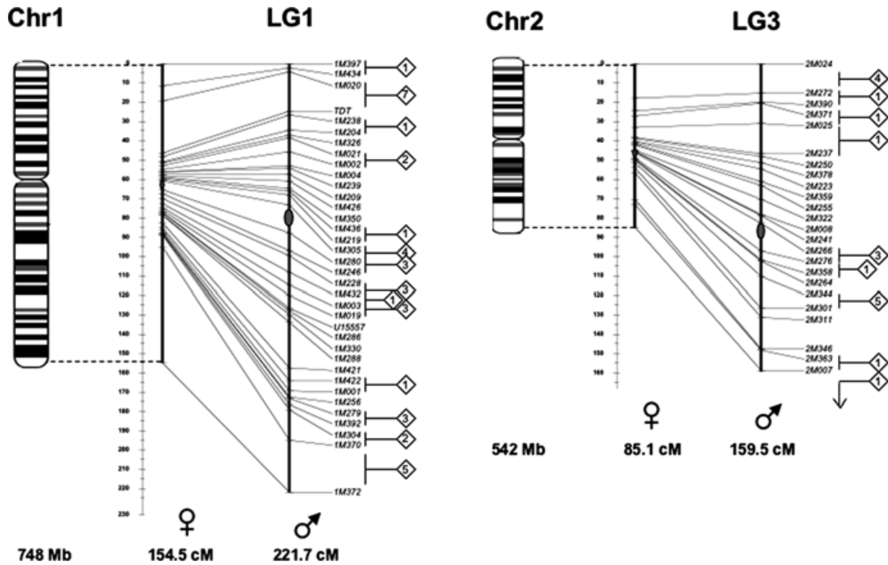


Fig. 4.1 Chromosome ideograms and sex-specific linkage maps for Chromosomes 1 and 2 (Linkage Groups 1 and 3) of the 150-marker BBBX *Monodelphis domestica* linkage map. For each pair of maps, the ideogram is to the *left*, female map in the *center*, and male map on the *right*. *Gray ovals* indicate approximate positions of the centromeres. *Diamonds* indicate the number and locations of unpublished markers added to the map since publication of the 150 marker map (Samollow et al., 2007). *Scale bar* is in centiMorgans (cM) with zero corresponding to the p terminus of the chromosome. Horizontal broken lines connecting the *top* and *bottom* of the female maps to the ideograms indicate the physical positions of FISH hybridization signals for BACs containing the terminal linkage group markers indicated. Lengths of chromosomes (in Mb) and linkage groups (in cM) are indicated below the corresponding diagrams. Information for individual map markers may be found in Samollow et al. (2007)

recombination behaviour of opossum chromosomes appear reversed relative to those of eutherians, inasmuch as the female opossum has lower overall recombination than the male and has strongly elevated recombination in subtelomeric regions, both of which are characteristic of eutherian males.

4.7 Applications of the Linkage Maps

4.7.1 Map Integration

The wallaby linkage map has been integrated with extensive location data from several kinds of physical mapping studies and sequence data from the analysis of wallaby BACs to produce a comprehensive "integrated map" of the wallaby genome. Further, by merging the integrated map with data from the annotated opossum genome assembly, a "virtual map" of gene locations throughout the

wallaby genome has also been produced (Liao et al., 2007). The integrated and virtual maps are accessible at the Comparative Location Database website (http://compldb.angis.org.au/). These resources will provide the essential template for guiding assembly and annotation of the recently sequenced wallaby genome (http://www.genome.gov/12512299), which at 2X genomic coverage cannot be assembled adequately from the raw data alone.

The opossum genome, at 6.7X depth coverage, was assembled through de novo assembly algorithms, with large-scale assembly of scaffolds assisted by the opossum linkage map and the mapping of BAC clones corresponding to regions contained within specific scaffolds (Duke et al., 2007; Mikkelsen et al., 2007). The accuracy of the opossum sequence assembly (MonDom 5) is very high in the range of kilobases (Kb) to a few Mb (Mikkelsen et al., 2007), which has enabled in-depth analyses of gene structure, gene content, non-coding element content, and other characteristics. However, the large-scale structure contains some uncertainties. Specifically, we have detected several examples wherein the linkage map and genome assembly disagree with regard the order of larger-scale genomic regions (1 to ∼175 Mb). Figure 4.2 compares marker order on the BBBX Chromosome 1 linkage map (150-marker map) to that on the Chromosome 1 build of the MonDom5 assembly. Overall agreement between map and assembly is very good; only two minor disputes are evident at the level of resolution afforded by the 150-marker map. Comparisons between the updated map (∼280 confidently placed markers) and MonDom5 assembly reveals substantially more disagreement, because increased map density is able to discriminate order disputes on a much finer scale. Together with additional mapping resources under development, the linkage map will aid resolution of such disputes and contribute to substantially improved large-scale accuracy of the opossum genome assembly.

4.7.2 QTL Mapping

The original impetus for constructing linkage maps for both wallaby and opossum was to enable QTL mapping. A recent linkage analysis of low-density-lipoprotein cholesterol (LDL-C) phenotypes in the opossum has furnished proof-of-principle for this objective by defining a chromosomal region containing a QTL that influences variation in serum LDL-C response to a high-fat, high-cholesterol diet, and a second region containing a different QTL that influences LDL-C variation on a low-fat, low-cholesterol diet (C. M. Kammerer, D. L. Rainwater, N. Gouin, A. S. Dressen, M. Jasti, K. C. Douglas, P. Ganta, J. L. VandeBerg, and P. B. Samollow, unpublished manuscript). The $\log_{10}$odds ratios (LOD scores) for these QTLs are highly significant, showing convincingly that genes controlling these important physiologic traits are located within the respective QTL regions. Coupled with the ability to inspect the annotated genome assembly, quantitative-trait linkage analysis based on the opossum linkage map will enable improved focus in searches for candidate genes for heritable variation in a variety of physiologic and developmental traits in this model species.

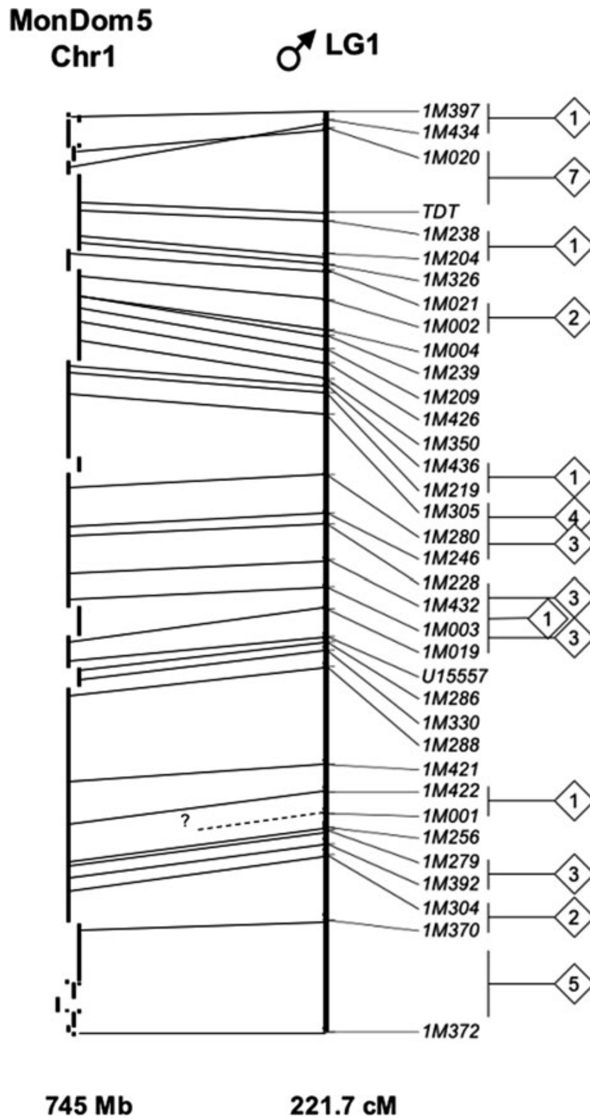


Fig. 4.2 Comparison of Linkage Group 1 map-marker order to the locations of the corresponding marker sequences on the opossum genome assembly (MonDom5). Chromosome 1 assembly is on the *left*, LG1 map on the *right*. *Vertical bars* comprising the chromosome represent individual genome assembly scaffolds (supercontigs). Marker symbols on the linkage-group map and locations indicated by diamonds are as described for Fig. 4.1. Positions at which lines cross indicate disputes between the genome assembly and linkage map. The topmost discrepancy (involving markers *1M434* and *1M020*) most likely represents scaffold misplacement. The second dispute (markers *1M004* and *1M239*) represents an order error for very closely spaced markers (<1 cM) on the 150-marker map. The order of these markers "flipped" after addition of new markers to the map, resolving the dispute (not shown). The sequence for *1M001* cannot be located by *in silico* searches of the MonDom5 assembly suggesting that it is located in an assembly gap

4.7.3 Determinants of Sex-Specific Recombination Rates

A substantial body of theory suggests that the frequency of meiotic recombination is under genetic control and can be adjusted by natural selection to optimize the balance between maintenance and destruction of favourable and unfavourable haplotypes from one generation to the next (reviewed by Otto and Lenormand, 2002; Rice, 2002; Barton and Otto, 2005; Otto and Gerstein, 2006; Coop and Przeworski, 2007; see also Martin et al., 2006; Gandon and Otto, 2007). The increasing availability of quality linkage maps and genome sequence data have more recently enabled evaluation of neutral models of recombination rate evolution as well as the examination of phylogenetic history and the effects of genomic substructure and redundancy on recombination rates (reviewed by Coop and Przeworski, 2007; de Visser and Elena, 2007; Dumont and Payseur, 2008). Finally, physical constraints on meiotic recombination rates and patterns in assuring proper chromosomal disjunction, as distinct from their influence in shaping haplotype structure and diversity, are also receiving increasing scrutiny (see Coop and Przeworski, 2007).

Corollary to these general theories of recombination rate, several hypotheses and numerous variants of them have been elaborated to explain why males and females of the same species should have different recombination rates (reviewed by Hedrick, 2007). None of these concepts is adequate to explain all or even most such cases, but an important outcome of these modeling efforts is recognition that the forces that shape recombination rates can act differentially on the sexes, resulting in very distinct male and female recombination rates and patterns (Lenormand, 2003; Lenormand and Dutheil, 2005).

Despite this extensive theoretical development, we do not know why any particular species has the recombination rate it does, or why marsupials in particular should have reduced recombination. In contrast, good progress is being made toward understanding the proximate molecular mechanisms that determine where and how often meiotic recombination events take place (Morelli and Cohen, 2005; Hunt, 2006; Moens et al., 2007; Petkov et al., 2007; and general reviews by Petronczki et al., 2003; Kauppi et al., 2004; Ivanovska and Orr-Weaver, 2006). Illuminating the proximal factors that influence sex-specific differences in chiasma number and distribution can enable the testing of more general hypotheses regarding the physical basis of crossover interference and the relationships between chiasma number and location vis-à-vis sex-specific differences in the risk of nondisjunction. For example, there is mounting evidence that large-scale genomic structure, local DNA sequence characteristics, and particularly variation in chromatin configuration (epigenetic states) have influences on recombination rates and locations (Lercher and Hurst, 2003; Sandovici et al., 2006; Buard and de Massy, 2007; Ng et al., 2009; Sigurdsson et al., 2009). In this context, the extreme heterochiasmy seen in the opossum can provide a powerful system with which to explore the interaction of sex and meiotic chromosome behavior. For example, if sex-specific differences in chromatin structure (histone modifications, DNA methylation state) differentially influence vulnerability to double-strand DNA breaks or their ultimate resolution (reciprocal exchange vs. non-exchange) in mammalian meiocytes, then the genomic

distributions and intensities of sex-specific epigenetic states in the early meiotic cells of male and female opossums can be expected to be predictably different (reversed) from those of a eutherian model species, such as mouse. The unusual recombinational features of the opossum genome can thus furnish a naturally occurring, comparative model for testing hypotheses about epigenetic influences on molecular processes that promote or inhibit chiasma formation and distributions, particularly between the sexes.

Acknowledgments The author's work is supported in part by grant RR014214 from the National Center for Research Resources of the National institutes of Health (USA).

References

Akesson M, Hansson B, Hasselquist D, Bensch S (2007) Linkage mapping of AFLP markers in a wild population of great reed warblers: importance of heterozygosity and number of genotyped individuals. Mol Ecol 16:2189–2202.

Backstrom N, Karaiskou N, Leder EH, et al. (2008) A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive synteny and gene-order conservation during 100 million years of avian evolution. Genetics 179:1479–1495.

Barton NH, Otto SP (2005) Evolution of recombination due to random drift. Genetics 169: 2353–2370.

Bennett JH, Hayman DL, Hope RM (1986) Novel sex differences in linkage values and meiotic chromosome behaviour in a marsupial. Nature 323:59–60.

Broman KW, Murray JC, Sheffield VC, White RL, Weber JL (1998) Comprehensive human genetic maps: individual and sex-specific variation in recombination. Am J Hum Genet 63:861–869.

Buard J, de Massy B (2007) Playing hide and seek with mammalian meiotic crossover hotspots. Trends Genet 23:301–309.

Calderon PL, Pigozzi MI (2006) MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. Chromosome Res 14:605–612.

Coop G, Przeworski M (2007) An evolutionary view of human recombination. Nat Rev Genet 8:23–34.

Cooper DW, Johnston PG, VandeBerg JL, Robinson ES (1990) X-chromosome inactivation in marsupials. In: Graves JAM, Hope R, Cooper DW (eds) Mammals from Pouches and Eggs: Genetics, Breeding and Evolution of Marsupials and Monotremes. CSIRO, Melbourne.

Cooper DW, Johnston PG, Watson JM, Graves JAM (1993) X-inactivation in marsupials and monotremes. Sem Dev Biol 4:117–128.

Cooper DW, VandeBerg JL, Sharman GB, Poole WE (1971) Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. Nat New Biol 230:155–157.

de Visser JA, Elena SF (2007) The evolution of sex: empirical insights into the roles of epistasis and drift. Nat Rev Genet 8:139–149.

Duke SE, Samollow PB, Mauceli E, Lindblad-Toh K, Breen M (2007) Integrated cytogenetic BAC map of the genome of the gray, short-tailed opossum, *Monodelphis domestica*. Chromosome Res 15:361–370.

Dumont BL, Payseur BA (2008) Evolution of the genomic rate of recombination in mammals. Evolution 62:276–294.

Duret L, Eyre-Walker A, Galtier N (2006) A new perspective on isochore evolution. Gene 385: 71–74.

Galtier N, Piganeau G, Mouchiroud D, Duret L (2001) GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. Genetics 159:907–911.

Gandon S, Otto SP (2007) The evolution of sex and recombination in response to abiotic or coevolutionary fluctuations in epistasis. Genetics 175:1835–1853.

Graves JA, Westerman M (2002) Marsupial genetics and genomics. Trends Genet 18:517–521.

Groenen MA, Wahlberg P, Foglio M, et al. (2009) A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. Genome Res 19:510–519.

Hayman D, Moore H, Evans E (1988) Further evidence of novel sex differences in chiasma distribution in marsupials. Heredity 61:455–458.

Hayman DL, Rodger JC (1990) Meiosis in male and female *Trichosurus vulpecula* (Marsupialia). Heredity 64:251–254.

Hayman DL, Smith MJ, Rodger JC (1990) A comparative study of chiasmata in male and female *Bettongia penicillata* (Marsupialia). Genetica 83:45–49.

Hedrick PW (2007) Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. Evolution 61:2750–2771.

Hinds LA, Poole WE, Tyndale-Biscoe CH, van Oorschot RAH, Cooper DW (1990) Reproductive biology and the potential for genetic studies in the tammar wallaby, *Macropus eugenii*. Aust J Zool 37:223–234.

Hogstrand K, Bohme J (1999) Gene conversion of major histocompatibility complex genes is associated with CpG-rich regions. Immunogenetics 49:446–455.

Hope RM (1993) Selected features of marsupial genetics. Genetica 90:165–180.

Hunt PA (2006) Meiosis in mammals: recombination, non-disjunction and the environment. Biochem Soc Trans 34:574–577.

Ivanovska I, Orr-Weaver TL (2006) Histone modifications and the chromatin scaffold for meiotic chromosome architecture. Cell Cycle 5:2064–2071.

Jaari S, Li MH, Merila J (2009) A first-generation microsatellite-based genetic linkage map of the Siberian jay *(Perisoreus infaustus*): insights into avian genome evolution. BMC Genomics 10:1.

Kai W, Kikuchi K, Fujita M, et al. (2005) A genetic linkage map for the tiger pufferfish, *Takifugu rubripes*. Genetics 171:227–238.

Kauppi L, Jeffreys AJ, Keeney S (2004) Where the crossovers are: recombination distributions in mammals. Nat Rev Genet 5:413–424.

Kochakpour N, Moens PB (2008) Sex-specific crossover patterns in Zebrafish (*Danio rerio*). Heredity 100:489–495.

Kong A, Gudbjartsson DF, Sainz J, et al. (2002) A high-resolution recombination map of the human genome. Nat Genet 31:241–247.

Kucuktas H, Wang S, Li P, et al. (2009) Construction of genetic linkage maps and comparative genome analysis of catfish using gene-associated markers. Genetics 181:1649–1660.

Lenormand T (2003) The evolution of sex dimorphism in recombination. Genetics 163:811–822.

Lenormand T, Dutheil J (2005) Recombination difference between sexes: a role for haploid selection. PLoS Biol 3:e63.

Lercher MJ, Hurst LD (2003) Imprinted chromosomal regions of the human genome have unusually high recombination rates. Genetics 165:1629–1632.

Liao W, Collins A, Hobbs M, Khatkar MS, Luo J, Nicholas FW (2007) A comparative location database (CompLDB): map integration within and between species. Mamm Genome 18: 287–299.

Maddox JF, Davies KP, Crawford AM, et al. (2001) An enhanced linkage map of the sheep genome comprising more than 1000 loci. Genome Res 11:1275–1289.

Martin G, Otto SP, Lenormand T (2006) Selection for recombination in structured populations. Genetics 172:593–609.

Matise TC, Chen F, Chen W, et al. (2007) A second-generation combined linkage physical map of the human genome. Genome Res 17:1783–1786.

McKenzie LM, Collet C, Cooper DW (1993) Use of a subspecies cross for efficient development of a linkage map for a marsupial mammal, the tammar wallaby (*Macropus eugenii*). Cytogenet Cell Genet 64:264–267.
McKenzie LM, Johnston PG, Eldridge MD, Cooper DW (1996) Close linkage between *RNR* and *GPD* genes on the tammar wallaby (*Macropus eugenii*) X chromosome. Cytogenet Cell Genet 72:197–199.
Menotti-Raymond M, David VA, Schaffer AA, et al. (2009) An autosomal genetic linkage map of the domestic cat, *Felis silvestris catus*. Genomics 93:305–313.
Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.
Moen T, Hayes B, Baranski M, et al. (2008) A linkage map of the Atlantic salmon (*Salmo salar*) based on EST-derived SNP markers. BMC Genomics 9:223.
Moens PB, Marcon E, Shore JS, Kochakpour N, Spyropoulos B (2007) Initiation and resolution of interhomolog connections: crossover and non-crossover sites along mouse synaptonemal complexes. J Cell Sci 120:1017–1027.
Morelli MA, Cohen PE (2005) Not all germ cells are created equal: aspects of sexual dimorphism in mammalian meiosis. Reproduction 130:761–781.
Morgan TH (1911a). Random segregation versus coupling in Mendelian inheritance. Science 34:384.
Morgan TH (1911b). Chromosomes and associative inheritance. Science 34:636–638.
Morgan TH (1911c). An attempt to analyze the constitution of the chromosomes on the basis of sex-limited inheritance in Drosophila. J Exp Zool 11:365–413.
Ng SH, Madeira R, Parvanov ED, Petros LM, Petkov PM, Paigen K (2009) Parental origin of chromosomes influences crossover activity within the *Kcnq1* transcriptionally imprinted domain of *Mus musculus*. BMC Mol Biol 10:43.
Ohno S (1967) Sex-Chromosomes and Sex-Linked Genes. Springer, Berlin.
Otto SP, Gerstein AC (2006) Why have sex? The population genetics of sex and recombination. Biochem Soc Trans 34:519–522.
Otto SP, Lenormand T (2002) Resolving the paradox of sex and recombination. Nat Rev Genet 3:252–261.
Perelygin AA, Samollow PB, Perelygina LM, Cherry LM, Mahaney SM, VandeBerg JL (1996) A new DNA marker, *U15557*, is linked to protease inhibitor and adenylate kinase-1 in the laboratory opossum, *Monodelphis domestica*. Anim Genet 27:113–116.
Petkov PM, Broman KW, Szatkiewicz JP, Paigen K (2007) Crossover interference underlies sex differences in recombination rates. Trend Genet 23:539–542.
Petronczki M, Siomos MF, Nasmyth K (2003) Un menage a quatre: the molecular biology of chromosome segregation in meiosis. Cell 112:423–440.
Reid DP, Smith CA, Rommens M, Blanchard B, Martin-Robichaud D, Reith M (2007) A genetic linkage map of Atlantic halibut (*Hippoglossus hippoglossus* L.). Genetics 177:1193–1205.
Rice WR (2002) Experimental tests of the adaptive significance of sexual recombination. Nat Rev Genet 3:241–251.
Richardson BJ, Czuppon AB, Sharman GB (1971) Inheritance of glucose-6-phosphate dehydrogenase variation in kangaroos. Nat New Biol 230:154–155.
Samollow PB, Gouin N, Miethke P, et al. (2007) A microsatellite-based, physically anchored linkage map for the gray, short-tailed opossum (*Monodelphis domestica*). Chromosome Res 15:269–281.
Samollow PB, Graves JAM (1998) Gene maps of marsupials. ILAR J 39:203–224.
Samollow PB, Kammerer CM, Mahaney SM, et al. (2004) First generation linkage map of the gray, short-tailed opossum, *Monodelphis domestica*, reveals genome-wide reduction in female recombination rates. Genetics 166:307–329.
Sandovici I, Kassovska-Bratinova S, Vaughan JE, Stewart R, Leppert M, Sapienza C (2006) Human imprinted chromosomal regions are historical hot-spots of recombination. PLoS Genet 2:e101.

Shifman S, Bell JT, Copley RR, et al. (2006) A high-resolution single nucleotide polymorphism genetic map of the mouse genome. PLoS Biol 4:e395.

Sigurdsson MI, Smith AV, Bjornsson HT, Jonsson JJ (2009) HapMap methylation-associated SNPs, markers of germline DNA methylation, positively correlate with regional levels of human meiotic recombination. Genome Res 19:581–589.

Singer A, Perlman H, Yan Y, et al. (2002) Sex-specific recombination rates in zebrafish (*Danio rerio*). Genetics 160:649–657.

Stapley J, Birkhead TR, Burke T, Slate J (2008) A linkage map of the zebra finch *Taeniopygia guttata* provides new insights into avian genome evolution. Genetics 179:651–667.

Sturtevant AH (1913) The linear arrangement of six sex-linked factors in Drosophila, as shown by their mode of inheritance. J Exp Zool 14:43–59.

VandeBerg JL (1990) The gray, short-tailed opossum (*Monodelphis domestica*) as a model didelphid for genetic research. Aust J Zool 37:235–247.

VandeBerg JL, Robinson ES (1997) The laboratory opossum (*Monodelphis domestica*) in laboratory research. ILAR Journal 38:4–12.

VandeBerg JL, Robinson ES, Samollow PB, Johnston PG (1987) X-linked gene expression and X-chromosome inactivation: marsupials, mouse, and man compared. Isozymes Curr Top Biol Med Res 15:225–253.

van Oorschot RA, Birmingham V, Porter PA, Kammerer CM, VandeBerg JL (1993) Linkage between complement components 6 and 7 and glutamic pyruvate transaminase in the marsupial *Monodelphis domestica*. Biochem Genet 31:215–222.

van Oorschot RAH, Cooper DW (1989) A conserved linkage between PI and GPI in a marsupial mammal, the tammar wallaby (*Macropus eugenii*). Cytogenet Cell Genet 51:1097.

van Oorschot RA, Porter PA, Kammerer CM, VandeBerg JL (1992) Severely reduced recombination in females of the South American marsupial *Monodelphis domestica*. Cytogenet Cell Genet 60:64–67.

Vingborg RK, Gregersen VR, Zhan B, et al. (2009) A robust linkage map of the porcine autosomes based on gene-associated SNPs. BMC Genomics 10:134.

Warren WC, Hillier LW, Marshall Graves JA, et al. (2008) Genome analysis of the platypus reveals unique signatures of evolution. Nature 453:175–183.

Zenger KR, McKenzie LM, Cooper DW (2002) The first comprehensive genetic linkage map of a marsupial: the tammar wallaby (*Macropus eugenii*). Genetics 162:321–330.

Chapter 5
Physical and Comparative Gene Maps in Marsupials

Janine E. Deakin

Abstract Comparative gene mapping in marsupials is responsible for many advances in our understanding of the events occurring during mammalian genome evolution. Over the past few years, the ease and speed at which genes can be physically mapped in marsupials has resulted in moderately dense physical maps for the South American opossum and the tammar wallaby. These maps have enabled genome sequence assemblies to be anchored to chromosomes and facilitated detailed comparative studies into genome evolution. The physical assignment of genes to marsupial chromosomes has resulted in many interesting and unexpected findings, including the discovery of novel genes and the absence of others, as well as providing insight into the evolution of epigenetic phenomena of genomic imprinting and X chromosome inactivation. Expanding comparative maps to include other distantly related marsupials is now possible and will be important for an accurate reconstruction of the ancestral marsupial karyotype.

Keywords Chromosome · Cytogenetic · Genome evolution · Physical map · Comparative genomics

5.1 Introduction

For over 30 years, a major goal in marsupial genetics and genomics studies has been to construct detailed gene maps for several distantly related species, in order to compare gene arrangement between different marsupials, as well as between marsupials and other vertebrates. Gene mapping data complements the comparative

---

J.E. Deakin (✉)
ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia
e-mail: janine.deakin@anu.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_5, © Springer Science+Business Media B.V. 2010

chromosome painting work described by Rens and Ferguson-Smith (Chapter 2), which has provided an excellent overview of global chromosome homologies amongst marsupials and shown the highly conserved nature of the marsupial karyotype. However, gene maps can provide information on gene order and internal rearrangements, which is beyond the sensitivity of chromosome painting. Marsupial gene maps assist in determining the events which have shaped the mammalian genome since the divergence of mammals and birds. They provide information on gene order and genome rearrangements among marsupials, enabling the evolutionary history of their genomes to be traced and the relationship between gene arrangement and function to be determined.

There have been several different methods used to construct gene maps in marsupials, including somatic cell hybrids, linkage (genetic) mapping and physical (cytogenetic) mapping. Marsupial-rodent somatic cell hybrids, the original method used for gene mapping, were problematic to work with and very unstable (reviewed in Graves et al., 1989). However, they were useful for assigning genes to the X chromosome in marsupials (Dawson and Graves, 1984; Donald and Hope, 1981; Graves et al., 1979). Linkage mapping (discussed by Samollow in the previous chapter) is able to provide information on gene order within a linkage group but requires physical mapping to assign these linkage groups to chromosomes. In addition, linkage mapping, for the most part, consists of anonymous sequence markers rather than actual genes, limiting the value of these maps in comparative genomic studies. Physical mapping by fluorescence in situ hybridisation (FISH) overcomes the limitations of somatic cell hybrids and linkage mapping. It is the most appropriate method for localising a gene of interest to a chromosome and provides a fairly precise location of the gene, which is important for comparing genome organisation across taxa.

5.2 Species Used for Physical Mapping

In the past, the construction of gene maps in marsupials was a challenging and slow process. Early work on gene maps in marsupials saw a handful of genes assigned to chromosomes, by various techniques, in 17 different species (Samollow and Graves, 1998). With different genes mapped in these species and so few genes in total, construction of comparative maps was difficult. To overcome part of this problem, three model marsupials were proposed for genetic studies and the construction of detailed gene maps; the tammar wallaby (*Macropus eugenii*), the South American opossum (*Monodelphis domestica*) and the fat-tailed dunnart (*Sminthopsis crassicaudata*) (Bennett et al., 1989; Hinds et al., 1989; VandeBerg, 1989). These species were chosen due to the availability of captive breeding colonies with some genetic linkage data available, and they represented three of the five major marsupial orders (Hope and Cooper, 1989). However, most cytogenetic mapping in a dasyurid species has been carried out in the striped-face dunnart (*S. macroura*) rather than *S. crassicaudata.*

5.3 Physical Mapping in the Pre-genomics Era

The first genes physically localised in a marsupial were the globin genes *HBB* and *HBA* of the Eastern quoll (*Dasyurus viverrinus*) (Wainwright and Hope, 1985). Radioactive in situ hybridisation (RISH), a technique which detected the location of radioactively labelled probes on chromosomes by autoradiography, was used to map the quoll cDNA probes for the *HBA* and *HBB* genes (Wainwright and Hope, 1985). Due to the difficulty in isolating marsupial sequences, heterologous probes to highly conserved genes, often human or mouse cDNA clones, were used to localise most genes by this technique. Unfortunately, RISH was inherently floored by the use of these heterologous probes, which also detected processed pseudogenes and paralogues and relied on statistical analysis to provide an approximate position on the chromosome. A number of the localisations made using this approach have now been shown to be incorrect by fluorescence in situ hybridisation (FISH) (Deakin et al., 2008).

FISH revolutionised physical mapping in marsupials in the late 1990s. This approach relies on the labelling of large homologous probes, either directly or indirectly, with a fluorochrome. Following hybridisation, fluorescent signals are visualised on chromosomes using a fluorescence microscope, overcoming many of the problems associated with RISH. It has the added advantage of allowing two or more probes, labelled with different fluorochromes, to be mapped in one experiment, permitting the relative order of genes on a chromosome to be determined. A limiting factor for this technique prior to the marsupial genome projects was obtaining the large homologous probes for hybridisation. This involved screening lambda genomic libraries, with inserts of approximately 15 kb, using human or mouse cDNA clones. Screening with these heterologous probes was often frustrating and resulted in the isolation of false positives. Alternatively, species-specific probes for library screening could be generated by PCR with degenerate primers, but unfortunately these primers often did not result in the amplification of the gene of interest. More recently, the construction of BAC (Bacterial Artificial Chromosome) libraries for several marsupials with inserts on average ∼150 kb have proven to be a tremendous resource for physical mapping studies. Libraries used for physical mapping purposes include two male tammar wallaby libraries [Me_VIA library (Sankovic et al., 2005) and ME_KBa (Arizona Genomics Institute)], male (VMRC6) and female (VMRC18) *M. domestica* libraries (BACPAC Resources) and a *Didelphis virginiana* (LBNL-3) library (BACPAC Resources). The larger insert size of the BAC clones compared to lambda clones makes the hybridisation more efficient and FISH signals easier to detect under the fluorescence microscope.

During the pre-genomics era, genes were chosen for mapping based on interest, such as those on the sex chromosomes in human. With different genes being chosen for mapping in different species, the construction of large-scale comparative maps between marsupials and human and other vertebrates was difficult. Despite this drawback, gene mapping still proved to be extremely valuable for determining the relationship between gene arrangement and gene function. One renowned example involves the discovery of the mammalian sex-determining gene. The region

responsible for testis determination on the human Y chromosome was narrowed down to a 35 kb region after many years of deletion mapping analysis on sex-reversed humans (Vergnaud et al., 1986). One zinc finger protein gene cloned from this region, *ZFY*, was put forward as a candidate testis-determining gene. Mapping of the marsupial *ZFY* orthologue to an autosome provided the first evidence that it was not the gene responsible for testis determination (Sinclair et al., 1988). Later, the *SRY* gene was isolated from this same region of the human Y (Sinclair et al., 1990), and evidence for this gene playing a role in testis development was demonstrated by the development of testes in XX mice with a *Sry* transgene (Koopman et al., 1991). *SRY* was subsequently mapped to the marsupial Y chromosome, confirming it as the sex-determining gene in therian (marsupial and eutherian) mammals (Foster et al., 1992). The hypothesis to explain the evolution of *SRY* from an X-borne copy was proposed after a related gene (*SOX3*) was mapped to the marsupial X chromosome (Foster and Graves, 1994).

5.4 Physical Mapping in the Genomics Era

Following the sequencing of the human genome, a number of other vertebrates were chosen for genome sequencing to aid in the annotation of the human genome and to facilitate comparative studies into genome evolution, structure and function (Green, 2007; http://www.genome.gov/25521739). Two marsupials were among the list of mammals chosen for genome sequencing; the South American grey short-tailed opossum and the tammar wallaby. Although, genome sequence is unarguably an extremely valuable resource, the assignment of sequence to chromosomes is critical if this resource is to be used to its full potential (Lewin et al., 2009). In both marsupial genome projects, physical mapping has played an important role in anchoring sequence to chromosomes. With the emergence of Devil Facial Tumour Disease (DFTD) in the Tasmanian devil population (see Bender, Chapter 23), a third marsupial genome project has commenced. This genome is being sequenced entirely by next generation sequencing technology and, once again, physical mapping will need to be an integral part of the assembly process if this sequence is to be anchored to chromosomes. Fortunately, the devil belongs to the dasyuridae family and will now replace the dunnart as the third model marsupial for gene mapping and genomics studies.

With the sequencing of these marsupial genomes, the construction of physical maps has moved away from the mapping of genes haphazardly as a result of a particular interest in certain genes, to being able to construct maps encompassing the entire genome. For the first time, detailed maps comparing the genome organisation across multiple vertebrate taxa, including marsupials, have been constructed (Deakin et al., 2008; Deakin et al., in preparation; Mikkelsen et al., 2007; Mohammadi et al., 2009) providing new insight into genome evolution.

5.4.1 The Important Role of Gene Mapping in the Opossum Genome Project

In addition to the important role it plays in assigning sequence to chromosomes, gene mapping assists in determining the quality of the genome assembly and detects regions with potential assembly errors. Gene mapping can help in overcoming inherent problems with genome sequence assembly in highly repetitive regions, in regions where haplotype differences in gene copy number pose a problem, or areas enriched in genes which have evolved via segmental duplication.

The opossum was the first marsupial to have its genome deeply sequenced (Mikkelsen et al., 2007). In the absence of a radiation hybrid map or a reference genome sequence from a closely related species, physical mapping was seen as the most efficient means of anchoring sequence to chromosomes (Duke et al., 2007). By mapping 381 BACs, identified from BAC end sequences located at the ends of sequence scaffolds, 97% of the genome was assigned to opossum chromosomes (Duke et al., 2007). A further 103 Mb of genome sequence remained unanchored (Duke et al., 2007). Unfortunately, genes of interest falling into this unanchored category, or those simply missing from the assembly altogether, have had to be localised as the need arises.

There are a number of examples for the opossum genome where gene mapping has played a vital role. In fact many of the opossum genome studies mentioned elsewhere in this book have relied to some extent on gene mapping. Examples of the importance of gene mapping include aiding in genome assembly, anchoring genes from large gene families to chromosomes, isolating genes absent from the assembly and confirming that novel sequences in the assembly are true and not the result of assembly errors.

The first large region of the opossum genome (∼4 Mb) to be annotated was the Major Histocompatability Complex (MHC) (Belov et al., 2006 and reviewed in Cheng et al., Chapter 16). This analysis was carried out on preliminary assemblies MonDom1 and MonDom2. Genes within the MHC can be divided into three classes with molecules encoded by genes belonging to two of these classes (MHC Class I and II genes) involved in presenting antigens to T cells. MHC Class I and II genes evolve via segmental duplication. Genome assembly programs often struggle to accurately assemble such regions. Haplotype differences further confound the assembly process when it is not possible to use an inbred individual for genome sequencing, as is the case for the opossum genome project.

In the MonDom1 assembly, genes from the MHC region were distributed across five different sequence scaffolds. With MHC Class I genes *UB/UC* and *UG* mapping to different regions of opossum chromosome 2 (Belov et al., 2006), it was imperative that the MHC-containing sequence scaffolds be localised in order to determine whether the region had been split due to chromosomal rearrangement or whether genes had been moved by a transposition event. BACs isolated from

the ends of these scaffolds mapped to the same region on the long arm of chromosome 2 as the MHC Class I gene *UG*. Aided by mapping information, this region was assembled onto a single scaffold in the subsequent MonDom2 assembly (Belov et al., 2006). The movement of *UB* and *UC* away from the main MHC cluster may be linked to marsupial-specific repeats found flanking these genes (Belov et al., 2006).

It is a difficult task to prove that a gene not found in a genome assembly is truly absent from the genome, and not simply missing from the assembly due to a difficulty in sequencing or assembly of the region containing the gene. One such example of a gene missing from the assembly is *XIST* (X-Inactive Specific Transcript), a non-coding RNA critical to the process of X chromosome inactivation in humans and other eutherian mammals (see Al Nadaf et al., Chapter 13). Searches of the assembly failed to find any sequence similar to *XIST* in the opossum genome (Davidow et al., 2007; Hore et al., 2007; Shevchenko et al., 2007). Genes flanking *XIST* in humans are adjacent in chicken and frog, yet these same genes map to different ends of the opossum X chromosome, suggesting that the region was disrupted in the marsupial lineage and acquired *XIST* after marsupial/eutherian divergence (Davidow et al., 2007; Hore et al., 2007; Shevchenko et al., 2007).

Another gene absent from the assembly was Insulin-like growth factor II (*IGF2*), a gene subject to genomic imprinting in both marsupials and eutherians (for more details on genomic imprinting refer to Hore et al., Chapter 12). Since cDNA sequence for this gene had previously been obtained (O'Neill et al., 2000), its absence from the assembly was surprising. A BAC clone containing *IGF2* was subsequently isolated and mapped to opossum chromosome 5, proving that it is only missing from the assembly but very much a part of the opossum genome (Lawton et al., 2007).

The discovery of a novel T-cell receptor (TCR) gene sequence in the genome assembly, with sequence similarity to *TCRD* (one of four TCR genes found in all jawed vertebrates), could have easily been disregarded as an assembly error or simply a divergent *TCRD* (Baker et al., 2005; Parra et al., 2007). However, physical mapping of all opossum TCR loci showed that this novel sequence was found on chromosome 3 (Parra et al., 2007) and did not co-localise with other TCR genes (Deakin et al., 2006). This novel TCR gene (*TCRM*) appears to be common to all marsupials, having been isolated from three distantly related marsupial species (refer to Baker et al., Chapter 17 for more details).

Anchoring the opossum genome sequence to chromosomes enabled the identification of 372 large blocks of genes with conserved synteny between opossum and human. Extending this comparison to other eutherians provided insight into eutherian genome evolution and has been used to reconstruct the karyotype of the boroeutherian ancestor (ancestor of most eutherian species) (Mikkelsen et al., 2007). It has also facilitated further "e-painting" studies (Kemkemer et al., 2009) such as the one mentioned in Rens and Ferguson-Smith (Chapter 2).

5.4.2 The Wallaby Genome and a New Strategy for Physical Map Construction

The tammar wallaby genome has now been sequenced to a depth of two-fold coverage in a combined effort by the Australian Genome Research Facility and the Baylor College of Medicine Human Genome Sequencing Center. Scaffold sequences for this light coverage genome are a great deal smaller and the assembly contains many more gaps (http://www.ensembl.org/Macropus_eugenii/Info/Index) than the average found in the opossum assembly. Obviously using the same approach as that used to anchor sequence to opossum chromosomes would be a laborious and costly task for the more than 400,000 scaffolds of the wallaby genome. However, the assignment of sequence to chromosomes is imperative if this sequence is to be used to answer questions regarding genome evolution (Lewin et al., 2009).

An alternative approach to mapping every sequence scaffold is to identify large conserved blocks of genes between opossum and human genomes and to map genes at either end of these blocks (Fig. 5.1). Information from the opossum genome can then be used to infer the location of genes within these blocks, circumventing the need to map every wallaby sequence scaffold. This approach has been made feasible with the availability of a high density wallaby BAC library and the development of a library screening technique using short, homologous probes called overgos. Overgo probes consist of two 24 bp oligonucleotides overlapping by eight base pairs, which are radioactively labelled by an end-filling reaction. Overgos are designed by extracting sequence from the wallaby genome project for the gene of interest, thereby resulting in a species-specific probe. These probes can be pooled to screen the library for multiple genes at once. BACs isolated in this way are confirmed positive for the gene of interest by a second round of screening on dot blots with overgos to individual genes. Dot blots have proven to be a very accurate method for confirming the presence of the gene of interest in a BAC clone but, as an additional level of confirmation, individual overgos have also been used to obtain sequence from positive BAC clones (Deakin et al., 2008).

The first two regions of the wallaby genome to be densely mapped were those containing genes found on the human X chromosome. Earlier gene mapping and chromosome painting studies had shown that the human X chromosome could be divided into an ancient conserved region, with genes from this region on the X in therian mammals, and a region added to the eutherian X after the divergence of marsupials and eutherians (Glas et al., 1999; Graves, 1995; Wilcox et al., 1996). Previous mapping experiments had shown that genes from the recently added region were found on wallaby chromosome 5 (Delbridge and Graves, 2004; Fitzgerald et al., 1993; Pask et al., 1997; Sinclair et al., 1988; Toder and Graves, 1998; Waters et al., 2003; Watson et al., 1992; Wilcox et al., 1996). In the opossum assembly, these genes are found in two separate blocks, one on chromosome 4 and another on chromosome 7. Overgos designed for human X genes found in the block on chromosome 4 were used to screen the wallaby BAC library. A dense map of this “neo-X” region was generated, with 52 out of a total of 58 genes from opossum chromosome 4 with a homologue on the human X being localised to the short arm of wallaby

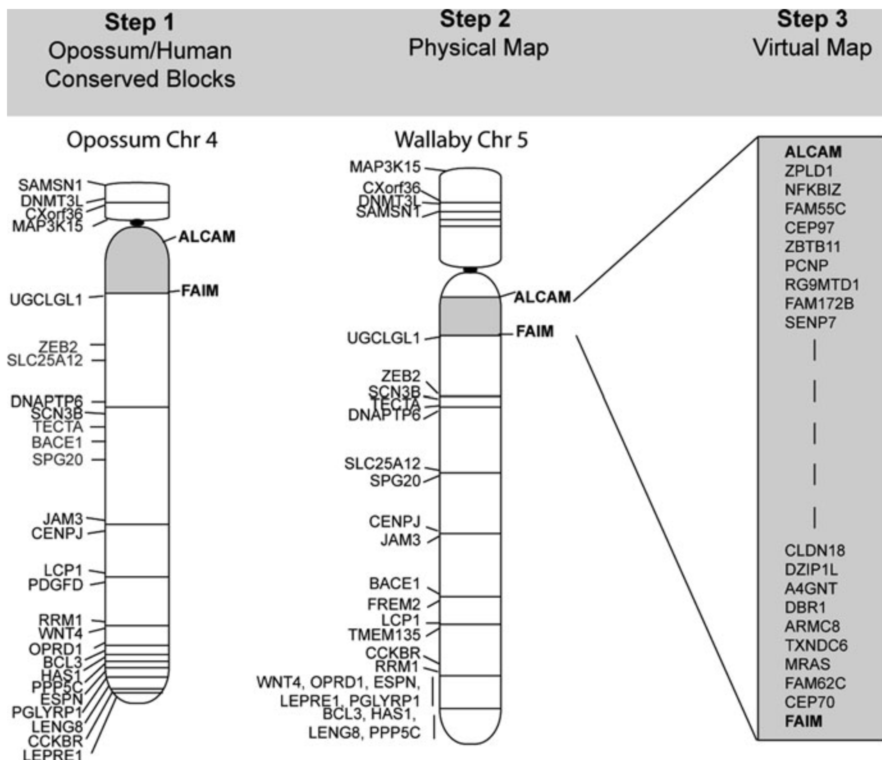


Fig. 5.1 Strategy used to generate physical and virtual maps of tammar wallaby genome. Step 1 is to identify conserved blocks of genes between human and opossum. Step 2 is to map a gene from either end of the block onto wallaby chromosomes (e.g. *ALCAM* and *FAIM* from the human/opossum conserved block containing human chromosome 3 genes). A virtual map is constructed in Step 3 by extrapolating from the opossum genome assembly that genes within this conserved block will be together as a block in wallaby

chromosome 5 (Deakin et al., 2008). The remaining six genes were not isolated during the BAC library screening process, perhaps due to being absent from the BAC library or a problem with the overgos used to screen for these genes. Nonetheless, this dense map provided support for the idea that, by mapping the ends of evolutionary conserved blocks, the location of the genes occurring within the block can be inferred.

The strategy for mapping the ends of evolutionary conserved blocks was then used to construct a complete map of wallaby chromosome 5. Chromosome painting had demonstrated that wallaby chromosome 5 shared homology with opossum chromosomes 4 and 7, corresponding to conserved segments C11 and C12 (Rens et al., 2003). By comparing the gene content of the opossum chromosomes to the arrangement of these genes in humans, 15 large evolutionary conserved blocks were

identified. Genes from the ends of these conserved blocks were mapped to chromosome 5. Four large-scale rearrangements have taken place between wallaby and opossum including two inversions which have occurred on the long arm of chromosome 5. Breakpoints for these inversions were narrowed down to within a three Megabase region on opossum chromosome 4 between genes *SPG20* and *FREM2* and a four Megabase region between *TECTA* and *BACE1* on opossum chromosome 4. The region corresponding to opossum chromosome 7 encompasses the pericentric region of wallaby chromosome 5, spanning from *RPL31* to *CPOX*. The centromere was determined to lie between two human X genes, *EGFL6* on the short arm and *CTPS2* on the long arm (Deakin et al., 2008).

By completing the map of chromosome 5, it was discovered that the "neo X" region is split into two regions by genes found on human chromosomes 2 and 15 (Deakin et al., 2008). The arrangement of these conserved gene blocks has particular relevance to the evolution of Prader-Willi/Angelman syndrome imprinted domain on human chromosome 15. In humans, this region consists of the maternally silenced (imprinted) *SNRPN* gene associated with Prader-Willi syndrome and the paternally imprinted *UBE3A* gene connected with Angelman syndrome. In the tammar wallaby, these two genes are found on different chromosomes, with *SNRPN* located on chromosome 1 and *UBE3A* part of the conserved block of human chromosome 15 genes found on chromosome 5, and neither gene is imprinted (Rapkins et al., 2006). Genes from human chromosomes X, 2 and 15 are found together in the chicken as well (Edwards et al., 2007), indicating that the organisation of the conserved gene blocks in wallaby represents the ancestral arrangement. The Prader-Willi/Angelman region assembled by fusing the regions containing the *SNPRN* and *UBE3A* genes, acquiring retrotransposed copies of genes and clusters of small nucleolar RNAs from other parts of the genome after the divergence of marsupials and eutherians (Rapkins et al., 2006).

About 20% of the protein coding genes making up the conserved region of the X were mapped on the wallaby X chromosome. A comparison of gene order between opossum, wallaby and human demonstrated a high level of rearrangement between the X chromosomes of these species, a feature undetectable by chromosome painting. The high degree of rearrangement between the marsupial X chromosomes was surprising given the high level of conservation of gene order on the X amongst eutherians, a feature attributed to the mode of X chromosome inactivation in eutherians, which would be disrupted by rearrangement (Deakin et al., 2008). This suggests that X inactivation in marsupials is achieved independent of gene position. Although the factor(s) controlling X inactivation in marsupial is yet to be elucidated, the extensive rearrangement of X-borne genes between species and absence of *XIST* suggests that control of X inactivation in marsupials will be quite different to that of eutherian mammals.

The physical map of chromosome 5, with the ends of evolutionary conserved blocks and genes near breakpoints mapped, made it possible to construct a virtual map of the chromosome. By extrapolating from the opossum genome, an additional 2,320 protein-coding genes were able to assigned to wallaby chromosome 5 (Deakin et al., 2008). This efficient physical mapping strategy was further applied to the long

of chromosome 6, a region which shares homology with the long arm of opossum chromosome 7 (Mohammadi et al., 2009). This chromosome arm is represented by 18 conserved blocks (Mohammadi et al., 2009; Deakin et al., 2008). Gene order for these blocks is well conserved between the wallaby and opossum, with only two small interstitial inversions (Mohammadi et al., 2009) and the movement of *MINA* to a different end of the chromosome (Deakin et al., 2008) being the only detectable differences in gene order between these species.

With the proven success of this strategy, this work was extended to the remaining wallaby chromosomes and a physical map of the entire genome has been constructed (Deakin et al., in preparation). Good physical maps are proving invaluable for whole genome sequencing projects by providing the framework for genome annotation and comparative genomic analysis. Physical maps provide an independent means to check the accuracy of the assembly, providing information on the order of scaffolds on a chromosome and are the only way to anchor sequence scaffolds to chromosomes (Lewin et al., 2009).

The complete physical map of the tammar wallaby genome has permitted comparative maps to be constructed (Fig. 5.2), where the genome organisation between different species can be compared. This allows the evolution history of a chromosome or specific regions of interest to be traced and the reconstruction of ancestral karyotypes.

5.4.3 Genes that Break the Rules

Although the conserved block mapping strategy works for most genes, there are always some that break the rules. MHC Class I genes in the wallaby are one such example. While constructing a BAC contig spanning the entire wallaby MHC, BACs containing MHC Class I genes were isolated and physically mapped using FISH. It was expected that these genes would be located within the MHC region on the long arm of wallaby chromosome 2 but these genes astoundingly localised to every autosome, mapping eight out of ten cases at the telomeres (Deakin et al., 2007). An endogenous retroviral element (KERV, see Carone and O'Neill, Chapter 3) has been implicated in the dispersal of these class I genes, perhaps by facilitating illegitimate recombination events (Deakin et al., 2007; Siddle et al., 2009; see also Cheng et al., Chapter 16). This extensive dispersal of MHC Class I genes has not been reported for any other vertebrate and would have remained undiscovered without the physical mapping of these BACs.

5.4.4 Integrating the Linkage and Physical Maps

With two different types of maps available for both the opossum and wallaby, it would obviously be advantageous to integrate the two data sets and ultimately incorporate this integrated map with the genome sequence data to obtain the best genome

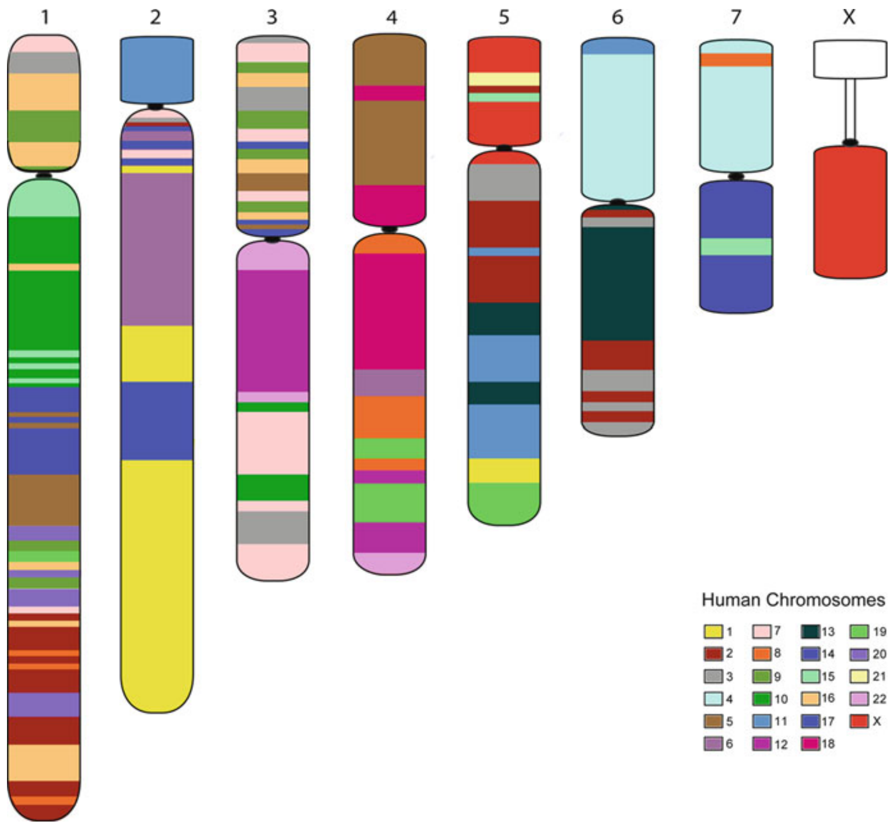


Fig. 5.2 Wallaby-human comparative map. The seven autosomes and the X chromosome of the wallaby have been *colour-coded* to show the opossum-human conserved blocks determined by FISH mapping. Each *colour* represents a different human chromosome

assembly possibly. A basic level of map integration has been carried out for the opossum by isolating BACs for markers at the ends of linkage groups and mapping these by FISH to opossum chromosomes (Samollow et al., 2007). These 34 BACs have been integrated into the opossum cytogenetic map (Duke et al., 2007). Unfortunately, due to a lack of sequence data from these BACs, it has not been possible to determine their position within the genome assembly (Duke et al., 2007). Likewise, integration of internal anonymous sequence markers has been difficult without their position being known in relation to genes.

Integration of the wallaby linkage and physical maps has been facilitated by physically mapping many of the linkage markers to wallaby chromosomes. Some markers were identified from physically mapped BACs or from genome sequence surrounding physically mapped genes (Wang et al., in preparation), meaning that there are many more internal markers which have been physically mapped than for the opossum and thereby simplifying the integration process. Mapping genes by

both techniques also serves as a form of quality control. Overgos used to isolate BACs for physical mapping were designed from sequence generated by the wallaby genome project, making it possible to subsequently incorporate genome assembly data into the integrated map. This will greatly enhance the genome assembly.

5.5 Mapping in Other Marsupials

The development of this efficient mapping strategy makes it possible to rapidly construct gene maps for other marsupials for which sequence information and a BAC library are available. At present there are two additional species for which physical maps could be readily generated: the Tasmanian devil (*Sacrophilus harrisii*) and the Northern brown bandicoot (*Isoodon macrourus*).

The plight of the Tasmanian devil has recently received a vast amount of attention due to Devil Facial Tumour Disease (DFTD) disease, which is decimating the devil population (for more details refer to Bender, Chapter 23). As a result, the devil is the third marsupial to have its genome and transcriptomes sequenced (see Papenfuss et al., Chapter 6) and two BAC libraries have been constructed providing the basic resources required for cytogenetic mapping. With data from the opossum and wallaby genomes, we are now able to determine blocks of genes conserved between these two species and map the ends of these blocks rather than the opossum/human conserved blocks, reducing the total number of genes required to be mapped. The added advantage of constructing such a map in the devil is that these same genes can be mapped in the different strains of the DFTD tumour and assist in the characterisation of the disease at the genetic level. Work on the devil physical map is progressing quickly and should be complete in early 2010 (Deakin, unpublished).

The Northern brown bandicoot (*Isoodon macroura*), belonging to the family Peramelidae, is another interesting species for which a BAC library is available. The position of bandicoots in marsupial phylogeny has always been a little puzzling. They possess features in common with dasyurids and macropodids but molecular studies suggest they fall between the American and Australian marsupial radiations (Springer et al., 1998). There is only very limited sequence data for this species in the form of ESTs (Baker et al., 2007), which would make designing species-specific overgo probes rather challenging. However, it may be possible to design universal overgos using conserved sequences (Sullivan et al., 2008).

Bandicoots possess a couple of distinct features, which make them very interesting animals to study. Their complex allantoic placenta means that they have the most invasive placenta among marsupials. This almost eutherian-like placenta means that they would be a very valuable species in which to study genomic imprinting (for more information in genomic imprinting in marsupials refer to Hore et al., Chapter 12). In a bizarre twist on X chromosome inactivation, a mechanism employed to equalise the expression of genes on the X chromosome between females with two copies of the X chromosome and males with just one, bandicoots eliminate one sex chromosome, one X in females and the Y in males, from somatic

tissues. Very little is known about this fascinating phenomenon, but a detailed map of the X chromosome would be a good starting point.

Like many other marsupials, bandicoots have a diploid chromosome number of 2n = 14. G-banding of marsupials with a 2n = 14 number suggests that the slight variations in banding patterns seen among different species are due to internal chromosome rearrangements, such as inversions or transpositions, and not inter-chromosomal arrangements (Hayman, 1990). Hence it would be expected that the composition of the 2n = 14 karyotype would be very similar to that of the dasyurids. A good physical map of the bandicoot genome would test this hypothesis regarding the 2n = 14 karyotype and provide insight into marsupial genome evolution.

5.6 Conclusion

Physical mapping of marsupial genomes has progressed in leaps and bounds over the last few years, particularly since the release of marsupial genome sequence. Physical maps have been used to anchor assemblies of the opossum and wallaby genomes to chromosomes, enabling comparative studies to be carried out. The efficient strategy developed for mapping the wallaby genome can now be used to generate maps for other divergent species, enabling the ancestral marsupial karyotype to be reconstructed. Information on the location of genes within the genome is key to answering questions in genome evolution, making these maps an essential component of any marsupial sequencing project.

References

Baker ML, Indiviglio S, Nyberg AM, et al. (2007) Analysis of a set of Australian northern brown bandicoot expressed sequence tags with comparison to the genome sequence of the South American grey short tailed opossum. BMC Genomics 8:50.

Baker ML, Osterman AK, Brumburgh S (2005) Divergent T-cell receptor delta chains from marsupials. Immunogenetics 57:665–673.

Belov K, Deakin JE, Papenfuss AT, et al. (2006) Reconstructing an ancestral mammalian immune supercomplex from a marsupial major histocompatibility complex. PLoS Biol 4:e46.

Bennett JH, Breed WG, Hayman DL, Hope RM (1989) Reproductive and genetic-studies with a laboratory colony of the dasyurid marsupial *Sminthopsis crassicaudata*. Aust J Zool 37: 207–222.

Davidow LS, Breen M, Duke SE, et al. (2007) The search for a marsupial XIC reveals a break with vertebrate synteny. Chromosome Res 15:137–146.

Dawson GW, Graves JAM (1984) Gene mapping in marsupials and monotremes. I. The chromosomes of rodent-marsupial (*Macropus*) cell hybrids, and gene assignments to the X chromosome of the grey kangaroo. Chromosoma 91:20–27.

Deakin JE, Koina E, Waters PD, et al. (2008) Physical map of two tammar wallaby chromosomes: a strategy for mapping in non-model mammals. Chromosome Res 16:1159–1175.

Deakin JE, Parra ZE, Graves JAM, Miller RD (2006) Physical mapping of T cell receptor loci (TRA@, TRB@, TRD@ and TRG@) in the opossum (*Monodelphis domestica*). Cytogenet Genome Res 112:342 K.

Deakin JE, Siddle HV, Belov K, Graves JAM (2007) Class I genes have split from the MHC in the tammar wallaby. Cytogenet Genome Res 116:205–211.

Delbridge ML, Graves JAM (2004) Assignment of the eukaryotic translation initiation factor (EIF2S3) to tammar wallaby chromosome 5p by in situ hybridization. Cytogenet Genome Res 107:139.

Donald JA, Hope RM (1981) Mapping a marsupial X chromosome using kangaroo-mouse somatic cell hybrids. Cytogenet Cell Genet 29:127–137.

Duke SE, Samollow PB, Mauceli E, Lindblad-Toh K, Breen M (2007) Integrated cytogenetic BAC map of the genome of the gray, short-tailed opossum, *Monodelphis domestica*. Chromosome Res 15:361–370.

Edwards CA, Rens W, Clarke O, et al. (2007) The evolution of imprinting: chromosomal mapping of orthologues of mammalian imprinted domains in monotreme and marsupial mammals. BMC Evol Biol 7:157.

Fitzgerald J, Wilcox SA, Graves JA, Dahl HH (1993) A eutherian X-linked gene, PDHA1, is autosomal in marsupials: a model for the evolution of a second, testis-specific variant in eutherian mammals. Genomics 18:636–642.

Foster JW, Brennan FE, Hampikian GK, et al. (1992) Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. Nature 359:531–533.

Foster JW, Graves JAM (1994) An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. Proc Natl Acad Sci USA 91:1927–1931.

Glas R, Marshall Graves JAM, Toder R, Ferguson-Smith M, O'Brien PC (1999) Cross-species chromosome painting between human and marsupial directly demonstrates the ancient region of the mammalian X. Mamm Genome 10:1115–1116.

Graves JAM (1995) The evolution of mammalian sex chromosomes and the origin of sex determining genes. Philos Trans R Soc Lond B Biol Sci 350:305–311.

Graves JA, Chew GK, Cooper DW, Johnston PG (1979) Marsupial – mouse cell hybrids containing fragments of the marsupial X chromosome. Somatic Cell Genet 5:481–489.

Graves JAM, Sinclair AH, Spencer JA (1989) Marsupial gene mapping and the evolution of mammalian sex chromosome form and function. Aust J Zool 37:365–383.

Green P (2007) 2x genomes – does depth matter? Genome Res 17:1547–1549.

Hayman DL (1990) Marsupial cytogenetics. Aust J Zool 37:331–349.

Hinds LA, Poole WE, Tyndale-Biscoe CH, van Oorschot RAH, Cooper DW (1989) Reproductive biology and the potential for genetic studies in the tammar wallaby, *Macropus eugenii*. Aust J Zool 37:223–234.

Hope RM, Cooper DW (1989) Marsupial and monotreme breeding in wild and captive populations: towards a laboratory marsupial. Aust J Zool 37:157–160.

Hore TA, Koina E, Wakefield MJ, Graves JAM (2007) The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. Chromosome Res 15:147–161.

Kemkemer C, Kohn M, Cooper DN, et al. (2009) Gene synteny comparisons between different vertebrates provide new insights into breakage and fusion events during mammalian karyotype evolution. BMC Evol Biol 9:84.

Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R (1991) Male development of chromosomally female mice transgenic for Sry. Nature 351:117–121.

Lawton BR, Obergfell C, O'Neill RJ, O'Neill MJ (2007) Physical mapping of the IGF2 locus in the South American opossum *Monodelphis domestica*. Cytogenet Genome Res 116:130–131.

Lewin HA, Larkin DM, Pontius J, O'Brien SJ (2009) Every genome sequence needs a good map. Genome Res 10.1101/gr.094557.109

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Mohammadi A, Delbridge ML, Waters PD, Graves JAM (2009) Conservation of a chromosome arm in two distantly related marsupial species. Cytogenet Genome Res 124:147–150.

O'Neill MJ, Ingram RS, Vrana PB, Tilghman SM (2000) Allelic expression of IGF2 in marsupials and birds. Dev Genes Evol 210:18–20.

Parra ZE, Baker ML, Schwarz RS, et al. (2007) A unique T cell receptor discovered in marsupials. Proc Natl Acad Sci USA 104:9776–9781.

Pask A, Toder R, Wilcox SA, Camerino G, Graves JAM (1997) The candidate sex-reversing DAX1 gene is autosomal in marsupials: implications for the evolution of sex determination in mammals. Genomics 41:422–426.

Rapkins RW, Hore T, Smithwick M, et al. (2006) Recent assembly of an imprinted domain from non-imprinted components. PLoS Genet 2:e182.

Rens W, O'Brien PC, Fairclough H, et al. (2003) Reversal and convergence in marsupial chromosome evolution. Cytogenet Genome Res 102:282–290.

Samollow PB, Gouin N, Miethke P, et al. (2007) A microsatellite-based, physically anchored linkage map for the gray, short-tailed opossum (*Monodelphis domestica*). Chromosome Res 15:269–281.

Samollow PB, Graves JAM (1998) Gene Maps of Marsupials. ILAR J 39:203–224.

Sankovic N, Bawden W, Martyn J, Graves JAM, Zuelke K (2005) Construction of a marsupial bacterial artificial chromosome library from the model Australian marsupial, the tammar wallaby (*Macropus eugenii*). Aust J Zool 53:389–393.

Shevchenko AI, Zakharova IS, Elisaphenko EA, et al. (2007) Genes flanking Xist in mouse and human are separated on the X chromosome in American marsupials. Chromosome Res 15: 127–136.

Siddle HV, Deakin JE, Coggill P, et al. (2009) Genomic structure and function of MHC-linked and un-linked Class I genes in the wallaby. BMC Genomics 10:310.

Sinclair AH, Berta P, Palmer MS, et al. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature 346:240–244.

Sinclair AH, Foster JW, Spencer JA, et al. (1988) Sequences homologous to ZFY, a candidate human sex-determining gene, are autosomal in marsupials. Nature 336:780–783.

Springer MS, Westerman M, Kavanagh JR (1998) The origin of the Australasian marsupial fauna and the phylogenetic affinities of the enigmatic monito del monte and marsupial mole. P Roy Soc Lond B Bio 265:2381–2386.

Sullivan RT, Morehouse CB, Comparative Sequencing NISC, Program TJW (2008) Uprobe 2008: an online resource for universal overgo hybridisation-based probe retrieval and design. Nucleic Acids Res 36:W149–W153.

Toder R, Graves JAM (1998) CSF2RA, ANT3, and STS are autosomal in marsupials: implications for the origin of the pseudoautosomal region of mammalian sex chromosomes. Mamm Genome 9:373–376.

VandeBerg JL (1989) The gray short-tailed opossum (*Monodelphis domestica*) as a model didelphid species for genetic research. Aust J Zool 37:235–247.

Vergnaud G, Page DC, Simmler MC, et al. (1986) A deletion map of the human Y chromosome based on DNA hybridization. Am J Hum Genet 38:109–124.

Wainwright B, Hope R (1985) Cloning and chromosomal location of the alpha- and beta-globin genes from a marsupial. Proc Natl Acad Sci USA 82:8105–8108.

Waters PD, Sankovic N, Kirby PJ, Delbridge ML, Graves JAM (2003) Assignment of the thymosin beta 4 X/Y chromosome (TMSB4X/Y) gene to tammar wallaby chromosome 5p by fluorescence in situ hybridisation. Cytogenet Genome Res 103:203F.

Watson JM, Spencer JA, Graves JAM, Snead ML, Lau EC (1992) Autosomal localization of the amelogenin gene in monotremes and marsupials: implications for mammalian sex chromosome evolution. Genomics 14:785–789.

Wilcox SA, Watson JM, Spencer JA, Graves JAM (1996) Comparative mapping identifies the fusion point of an ancient mammalian X-autosomal rearrangement. Genomics 35:66–70.

Part III
Marsupial Genome Sequencing

Jennifer A. Marshall Graves

The most spectacular advances in marsupial genetics and genomics since the publication of "Mammals from Pouches and Eggs" (Graves et al., 1990) have been engendered by the ability to obtain DNA sequence from study species, and, more recently, the availability of whole genome sequence for two marsupial species.

Every aspect of marsupial genetics and genomics has been hugely advanced by this new source of data. Initially, investigations of marsupial genes depended on cloning and Southern or Northern blot analysis of small fragments of DNA. Gene mapping in the 1980s and even the 1990s depended largely on identification and hybridisation of homologous sequence (usually human cDNA), which often identified the location of pseudogenes or paralogues. Large scale and global analyses of genome structure, repeat frequency and gene arrangements especially within gene families, was impossible. Even with the advent of PCR, obtaining marsupial sequence was guesswork based on human and chicken gene sequence, and the hit rate was usually below 20%.

A big advance in the early 2000s was the construction of a BAC library from the tammar wallaby (Sankovic et al., 2005), and its use both for mapping and for investigating genes and gene families with limited BAC sequencing (e.g. de Leo et al., 2005, globin genes). Several marsupial BAC libraries are now available for tammar wallaby, devil and opossum.

With the sequencing of the human genome in 1999, mammalian genetics and genomics entered a new era in which not only genes and gene families could be investigated, but also global analyses could be performed to survey the large scale genome landscape, and to investigate the frequency and arrangement of repetitive sequences and transcribed non-coding sequences that provided new data that altered our concept of the mammalian genome. The development of short read sequencing methods greatly increased the ease of resequencing, introducing the possibility of examining different individuals (normal or diseased), but with the caveat that a good shotgun sequence was required for the purposes of alignment.

A full genome sequence of a marsupial seemed, in the early 2000s, to be an impossible dream. However, I was approached in 2003 to write a "White paper" proposing the sequencing of a marsupial genome by the National Institutes of Health of the USA. A proposal was generated by a consortium of marsupial

(largely Australian) geneticists, featuring the tammar wallaby (Graves et al., 2003). However, the decision was made to sequence "an American marsupial", and the Brazilian short-tailed grey opossum *Monodelphis domestica* was chosen, and subsequently sequenced by Sanger shotgun sequencing (Mikkelsen et al., 2007). The tammar genome has now been sequenced lightly by the shotgun method, but several other sources of data need to be integrated.

In this part, Chapter 6 by Papenfuss et al. describe the history of marsupial genome sequencing, concentrating on the bioinformatics challenges they have posed. Analysis of the opossum sequence taught us several lessons about annotating sequence from an unknown genome for which there is no real reference. Papenfuss et al. discuss some of the bioinformatics challenges that have had to be surmounted in annotating the opossum, and more recently the tammar genome. As described by these authors, many of the commonly used bioinformatics packages are not appropriate or are inefficient, and it has proved necessary to develop new tools to get the most out of these genomes. Novel solutions will be particularly necessary for the so-called hybrid genomes which have been sequence lightly by the Sanger method, but for which transcriptome and short-read sequence data are available, such as the genome of the Tasmanian devil. The ultimate objective would be to develop methods for putting a de novo genome together solely from short-read sequence, as claimed for the giant panda (Li et al., 2009).

One of the paramount reasons for whole genome sequencing of a range of animals is to make sense of the "dark matter" that comprises more than 95% of the mammalian genome (Wakefield and Papenfuss, Chapter 7). Repeat analysis using sequence from the opossum genome established that the large size of the opossum genome was accounted for by the high proportion of repetitive sequence. In particular it was important to examine the distribution of LINE elements on the autosomes and the X chromosomes, these elements have been implicated in X chromosome inactivation. It was found that LINEs are not accumulated on the opossum X, which suggested that these elements were exapted to a new function relatively recently in eutherian mammals. Indeed, one of the most important contributions of the opossum genome was a new understanding of how ancient and recent exaptation of transposable sequences have contributed to building up regulatory pathways.

In addition, marsupial genome sequence made it possible to analyse the classes of miRNAs that might have regulatory activity, and the long non-coding transcribed sequences, such as the *XIST* gene that controls X chromosome inactivation. Given that *XIST* is absent from the marsupial X (Hore et al., 2007), it is interesting to investigate when this class of regulator evolved and how such sequences took over roles in regulation.

With the next generation of single molecule DNA sequencing, it is likely that all marsupials (and monotremes and eutherian mammals) will be sequenced (Genome 10K, 2009), at least to two-fold, providing comparative genomics with the power to resolve these longstanding questions.

References

De Leo AA, Wheeler D, Lefevre C, et al. (2005) Sequencing and mapping hemoglobin gene clusters in the Australian model dasyurid marsupial *Sminthopsis macroura*. Cytogenet Genome Res 108:333–341.

G10K Community of Scientists: Coordinators, Haussler D, O'Brien SJ, Ryder OA (2009) A proposal to obtain whole genome sequence for 10,000 vertebrate species. J Hered 100:675–680.

Graves JAM, Hope RM, Cooper DW (eds) (1990) Mammals from Pouches and Eggs: Genetics, Breeding and Evolution of Marsupials and Monotremes. CSIRO Press, Melbourne.

Graves JAM, Wakefield MJ, Renfree MB, et al. (2003) Proposal to sequence the genome of the model marsupial *Macropus eugenii* (Tammar Wallaby). http://wwwgenomegov/Pages/Research/Sequencing/SeqProposals/WallabySEQpdf

Hore TA, Koina E, Wakefield MJ, Graves JAM (2007) The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. Chromosome Res 15:147–162.

Li R, Fan W, Tian G, Zhu H, He L, Cai J, et al. (2009) The sequence and de novo assembly of the giant panda genome. Nature 463:311–317.

Mikkelsen TS, Wakefield MJ, Aken B, et al (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Sankovic N, Bawden W, Martyn J, Graves JAM, Zuelke K (2005) Construction of a marsupial BAC library from the model Australian marsupial Tammar wallaby (*Macropus eugenii*). Aust J Zool 53:389–393.

Chapter 6
Marsupial Sequencing Projects and Bioinformatics Challenges

Anthony T. Papenfuss, Arthur Hsu, and Matthew Wakefield

Abstract The arrival of next or second generation sequencing has ushered in a new era of marsupial genomics, where large-scale sequencing of marsupial transcriptomes, and soon perhaps genomes, is within the scope of many independent laboratories. This promises to reveal much about the biology of marsupial genomes and provides opportunities for comparison with eutherian genomes. These comparisons will highlight both the conserved features that are critical, as well as important differences where marsupials and eutherians have chosen different evolutionary paths. Here we describe the current state of marsupial genomic sequencing projects and resources, including available genome and transcriptome sequences. We also survey a number of useful bioinformatics tools, particularly those that we have utilized on marsupial, or sometimes monotreme, genomic data and found useful. Finally, some of the challenges met in dealing with, largely next generation, marsupial sequence are described – experience that we think is also relevant to other non-model organisms.

Keywords Marsupial genomes · Transcriptomics · Next generation sequencing · Bioinformatics

6.1 Marsupial Sequencing Projects – The State of Play

The first marsupial genome to be sequenced was that of the South American short-tailed grey opossum, *Monodelphis domestica*. This was performed at the Broad Institute and published in 2007 (Mikkelsen et al., 2007). The opossum genome was

A.T. Papenfuss (✉)
Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
e-mail: papenfuss@wehi.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_6, © Springer Science+Business Media B.V. 2010

Sanger-sequenced[1] to ~7x coverage or draft quality. This has produced a generally very good and highly usable assembly, reflected by the large contig N50 of 108 kb and scaffold N50 of 60 Mb. The N50 is a measure of how well the assembly has come together. It is the length of the contig or scaffold such that 50% of nucleotides in the assembly are from shorter sequences.

The Ensembl consortium (http://www.ensembl.org) annotated the opossum genome. Unfortunately at the time, and still in late 2009, no large-scale opossum transcriptomic data was available. As a consequence, the annotation was largely based on transcript and protein sequences from distantly-related species (e.g. human) (Hubbard et al., 2007). This has resulted in an annotation that is of good quality for well-conserved genes, but one that misannotates or has failed to identify many highly divergent genes. Immune genes are particularly affected by this problem and this has motivated several more careful analyses of these genes (e.g. Belov et al., 2006; Wong et al., 2006; Belov et al., 2007). Bioinformatics approaches to address the challenge of finding highly divergent genes are described in the last section of this chapter.

The opossum genome was selected by the National Human Genome Research Institute over the bid of a friendly rival, the Tammar wallaby, *Macropus eugenii*, a small kangaroo species which is an Australian model marsupial. The wallaby has been Sanger-sequenced to 2x coverage and is now assembled and annotated, and due for publication in 2010. The low-coverage sequencing of the wallaby genome has produced a highly fragmented assembly. Currently its scaffold N50 is in the range 10–20 kb, though some efforts are underway to improve this. Additional bioinformatics challenges are introduced by such a fragmented genome. For example, genes may be split across multiple contigs or scaffolds. To deal with this, Ensembl uses a gene-scaffold construct for low-coverage genomes (Hubbard et al., 2007), which is a concatenation of contigs and scaffolds with order and orientation inferred by alignment to a gene sequence in a high quality reference genome such as human.

The opossum and wallaby represent key models for marsupial research. However, the recent emergence of Devil Facial Tumour Disease (DFTD), a contagious cancer affecting the Tasmanian devil, *Sarcophilus harrisii*, which is spread as an allograft by biting (Pearse and Swift, 2006 and reviewed in Bender, Chapter 23), has motivated the sequencing of this marsupial's genome. Two individuals are being sequenced using the Roche 454 platform.[2] This cannot come too soon, as the devil population is crashing as a result of the disease, which has spread to more than 65% of Tasmania (McCallum et al., 2007) and kills infected devils in around 6 months (Lachish et al., 2007). It is estimated that the devil will be extinct in the wild within 25–35 years (McCallum et al., 2007). In the meantime, the transcriptome[3] of the

[1] Sanger-sequencing is capillary-based DNA sequencing that producing long (600–1,000 nt), high quality reads.

[2] The Roche 454 is a next generation sequencing platform that produces hundreds of thousands of 200–600 nt long reads.

[3] A transcriptome is the set of all transcripts or expressed genes in a tissue.

DFTD tumour has been sequenced (Murchison et al., 2010), along with a representative normal tissue (testis) using the Roche 454 platform. This has provided important insights into the tumour, including the identification of the tissue of origin.

A number of next generation sequencing projects are now planned for the devil. Sequencing of cDNAs from spleen and other immunological tissues using the Roche 454 platform will enhance our understanding of immune genes in the devil and provide an important resource for understanding immune evasion of the DFTD tumour. Multiple tumour strains have now been observed in the environment (Pyecroft et al., 2007), suggesting that the tumour is evolving, perhaps even to evade immune recognition in resistant hosts. Illumina-based[4] RNA-seq[5] on multiple tumour strains is also planned. This will shed light on key differences, via their expression profiles, between tumour strains.

Short read sequencing is also currently underway for several projects, in particular a number of marsupial Y chromosomes, which are comparatively small (P. Waters, personal communication). This approach to sequencing and de novo assembly has been spearheaded by the panda genome. It is too soon to say what the resulting assembly quality will be like using this approach on micro-dissected or flow-sorted chromosomes, or indeed on whole mammalian genomes.

These and other significant marsupial genomic projects and resources are detailed in Table 6.1.

Table 6.1 Marsupial genomic projects and resources

Resource	Description	References
Opossum (*Monodelphis domestica*) genome	Draft quality genome, ~7x Sanger sequencing	Mikkelsen et al. (2007)
Tammar wallaby (*Macropus eugenii*) genome	~2x coverage genome, Sanger sequencing	Baylor College of Medicine Australian Genome Resource Facility
Tasmanian devil (*Sarcophilus harrisii*) genome	Reportedly 4–8x coverage of two individuals using 454 sequencing	In progress, V. Hayes (Children's Cancer Research Institute Australia)
Tasmanian devil transcriptome	454 sequencing of DFTD tumour & testis cDNA; spleen cDNA (underway) RNA-seq on tumour strains	http://pylons.bioinf.wehi.edu.au/diablome
Tammar wallaby transcriptome	cDNAs from multiple tissues, including thymus and milk, sequenced using Sanger & 454	
Bandicoot (*Isoodon macrourus*) ESTs	~1,319 ESTs	Baker et al. (2007)

[4]The Illumina GA2 is a next generation sequencing platform that produces millions of short reads (32–100 nt).

[5]RNA-seq is expression analysis based on sequencing transcripts and counting reads mapping to each gene.

Table 6.1 (continued)

Resource	Description	References
Possum (*Trichosurus vulpecular*) ESTs	~150,000 ESTs	http://www.possumbase.org.nz
Dunnart (Sminthopsis *crassicaudata*) ESTs	~2,000 ESTs	Kullberg et al. (2007)
Thylacine mitochondrial genome	Two complete sequences	GenBank accession: NC_011944 & FJ515781
Marsupial Y chromosomes		In progress, Dr Paul Waters (ANU)

6.2 Bioinformatics Tools

Here we briefly describe a few bioinformatics tools that have been indispensable to us in analysing marsupial genomic and transcriptomic sequence data. The survey is necessarily brief. We do not discuss, for example, tools for phylogenetic analysis. Nonetheless, we think it useful, particularly in setting the scene for discussion of bioinformatics challenges.

6.2.1 Sequence Alignment

The Basic Local Alignment Search Tool or BLAST (Altschul et al., 1990) is one of the most widely used sequence alignment programs. The original BLAST paper has been cited 26,400 times and the gapped BLAST paper a similar number again. These publications are probably dramatically under-cited, due to the common but inappropriate trend to not cite software. BLAST implements local alignment, making use of heuristics to increase search speed over the exhaustive Smith-Waterman algorithm, and returning a list of High-scoring Segment Pairs (HSPs), which may, depending on the output format, be chained into "hits". BLAST comes in five basic "flavours":

BLASTN (search a *nucleotide* database with a *nucleotide* sequence),
BLASTP (search a *protein* database with a *protein* sequence),
BLASTX (search a *protein* database using a *translated nucleotide* query),
TBLASTN (search a *translated nucleotide* database using a *protein* query) and
TBLASTX (search a *translated nucleotide* database using a *translated nucleotide* query).

A number of related BLAST tools such as PSI-BLAST and RPS-BLAST perform more sophisticated types of searches. In terms of individual queries, we have made more use of BLAST than any other software tool. Applications of BLAST have

included work annotating the opossum MHC and immunome, the whole opossum and tammar genomes, and the DFTD transcriptome. The capacity to perform large-scale BLAST searches in parallel on a computer-cluster is important for the analysis of large transcriptomic and genomic datasets.

The BLAST-like Alignment Tool or BLAT (Kent, 2002) is similar to BLAST, but uses an index to accelerate alignments. It is much faster and more conservative than BLAST, and therefore suited for the purpose of genome annotation. The Ensembl genome browser supports both BLAST and BLAT against the two currently available marsupial genome assemblies (http://www.ensembl.org/Multi/blastview).

Exonerate (Slater and Birney, 2005) is a generic alignment tool that implements many alignment models, such as ungapped, local, global, or spliced (cDNA to genome). We have found its flexibility as a single tool capable of many styles of alignment to be highly useful.

It should also be noted that these aligners are developed for alignment of longer sequences and are not efficient for the alignment of millions of short reads (25–100 nt) such as those produced by Illumina GA2 sequencer, although they are still applicable for the longer reads produced by the Roche 454 sequencer. Recently, a number of aligners have become available that are specifically designed for short reads. These include Bowtie (Langmead et al., 2009), SHRIMP (Rumble et al., 2009) and SOAP (Li et al., 2008).

6.2.2 Gene Prediction

Gene predictors integrate signals such as start, stop and splice sites and information about sequence content to predict the locations of protein-coding genes. Sequence content sensors need to be trained for each species they are applied to, but there are no gene predictors trained on marsupial genomes, so gene prediction tools developed for the human genome have frequently been used on the opossum and wallaby genomes.

To improve their accuracy, some gene predictors use extrinsic data, such as homologous sequences, to guide their annotations. These have been called informant-based gene predictors and can be helpful in compensating for an inaccurate sequence content sensor based upon another species. GenomeScan (Yeh et al., 2001) is one such gene predictor (http://genes.mit.edu/genomescan.html). It makes use of a homologous protein sequence alignment to the genome to improve the likelihood of correctly identifying exons. We have made extensive use of GenomeScan in marsupials (e.g. Belov et al., 2006; Wong et al., 2006; Belov et al., 2007).

The advent of next generation sequencing means that species-specific transcript data is cheaper and far more accessible, so that the need for gene prediction is decreasing, though it may remain important for lowly expressed or tissue-specific transcripts. Additionally, there is still some need for informant-based predictors that can utilize partial transcript sequences, such as those obtained by sequencing a transcriptome using the Roche 454 platform, to improve exon identification.

6.2.3 Sequence Assembly

Current sequence assembly methods fall into two camps: *overlap-layout-consensus* or *de Bruijn graph* approaches. In general, the overlap-layout-consensus methods are most suited for long read sequence assembly, like traditional Sanger and Roche 454 sequences. On the other hand, de Bruijn graphs are suitable for assembling millions of short reads such as those from the Illumina GA2.

Some long read sequence assembly programs, which we have found useful, are CAP3 (Huang and Madan, 1999), which can be used mostly for Sanger sequences but is also applicable to Roche 454 reads, and Newbler, which is optimised for assembling Roche 454 reads using flowgram data,[6] sequence quality and reads. We have also made use of de Bruijn graph-based assembly programs like Velvet (Zerbino and Birney, 2008) and ABySS (Simpson et al., 2009) for short-read assemblies.

6.3 Bioinformatics Challenges

6.3.1 Dealing with Next Generation Data

Dealing with the deluge of raw data is the primary challenge of working with next generation sequences. The shear volume of data produced in a sequencing run calls for new features in analysis tools. Specialized data handling tools are the key to utilizing next generation data. Typical operations are sequence extraction, alignment (often in parallel on computer clusters), and flexible handling of alignment output and annotation data. Once this investment in a framework for analysing next generation data has been made, analysis becomes more or less straightforward given the appropriate computing hardware.

Within a decade, sequencing the genome of any species of interest to high quality will be cheap and affordable by a single laboratory. This will mean that any transcriptomic analysis can be done within the context of the genome of that particular species. Today, however, we are at an odd transitional stage in the evolution of sequencing, where it is relatively easy to construct large transcriptomic datasets in species for which there is no genome sequence from a closely related species. In the marsupial world, the Tasmanian devil is currently a splendid example of this, with the nearest sequenced genomes being the low coverage and fragmented tammar wallaby and the more distantly-related opossum. Distinguishing these cases from true model organisms like the mouse, we half-jokingly refer to wallaby and opossum as "model non-models," while the devil is a "non-model non-model."

The analysis of non-model transcriptomic data, such as that from the Tasmanian devil, presents additional challenges: should we annotate by aligning cross-species

[6] A flowgram is the signal intensity data from a Roche 454 sequencer. It is analogous to the chromatogram in Sanger sequencing. The signal intensity is proportional to the number of the base of the same type added in each sequencing step.

to a marsupial genome, or align to a protein dataset (say human)? Which species should we use? Should we assemble first, then annotate? There may be many reasonable answers to these questions and judging which is optimal can be difficult or impossible. Our best advice is to pick a reasonable solution and try it. For the devil, we aligned reads against the opossum genome in order to determine transcript abundance and the assembled contigs against human proteins to make use of the high quality human annotation.

6.3.2 Next Generation Transcriptome Assembly

The problem of assembling reads obtained from mixed sequencing platforms with their corresponding different lengths and types of sequencing errors, is largely unsolved. Most labs working with next generation data develop some type of ad hoc assembly/annotation pipeline. In this section, we describe in detail methods that we have used to clean, assemble and annotate the transcriptome of the tammar wallaby.

We obtained two Roche 454 FLX sequencing runs from cervical and thoracic thymus and a run on a normalized mixture of wallaby tissues. The average read lengths from these runs were 169 and 191 nt respectively. In addition, 63,893 Sanger reads have been sequenced from a variety of tissues, with average read length of 728 nt, and 14,837 milk ESTs were generated by CRC for Innovative Dairy Products (Lefevre et al., 2007).

Sequences were cleaned, trimmed and repeat masked using SeqClean (Pertea, 2009), Lucy (Chou and Holmes, 2001) and RepeatMasker (http://www.repeatmasker.org) respectively. One measure of a good quality assembly is larger contig N50 length, more assembled contigs and fewer unplaced reads. In our trials, Lucy was found to have minimal effect on assembly statistics. This is not entirely surprising as it is designed to trim low quality sequences based on the Sanger phred quality scores and may not be entirely appropriate for cleaning based on next generation phred-like quality scores. The trimmed and cleaned reads then were subjected to a read length filter where reads shorter than 40 bp were discarded. Following sequence cleaning, reads were pooled and assembled using CAP3 (Huang and Madan, 1999). The choice of using CAP3 was also evaluated against another assembly created with Newbler, which yielded a shorter contig N50. Again this is unsurprising as Newbler is known to be conservative. In our final assembly, the CAP3 assembler produced about 48,000 contigs with over 200,000 unplaced reads.

The contigs, which probably represent partial transcripts, were then aligned to the human RefSeq[7] protein. An E-value[8] cut-off of 10^{-5} was used to ensure high similarity. We detected about 23,000 distinct genes in the alignment.

[7]NCBI Reference Sequence collection (http://www.ncbi.nlm.nih.gov/refseq).

[8]The E-value is the number of hits expected by chance when searching a database of a particular size. Note that you should use proper scientific notation to write E-values in publications, not the common, but dreadful computer shorthand (e.g. 1e-5).

The assembled transcriptome can then be queried via the annotation to rapidly identify genes of interest and provide a minimally redundant set of annotated transcripts suitable for designing high quality custom long oligo microarrays.

6.3.3 Finding Divergent Gene Sequences in Marsupial Genomes

The problem of finding divergent genes arises in many areas, for example: when searching for genes that evolve by duplication and divergence or those under strong positive selection. Many genes involved in immunity are rapidly evolving and it can be difficult to identify orthologues between distantly related species.

There are several methods to performing more sensitive genome searches. These start with simply using peptide or translated BLAST searches, rather than BLASTN. Next, one can use a more sensitive scoring matrix. The default scoring matrix used by NCBI BLAST, when performing peptide or translated searches, is BLOSUM62. This was constructed from ungapped segments taken from multiple alignments of proteins with at least 62% identity, while in the BLOSUM45 matrix the similarity drops to 45%. This increases the scores of divergent genes, enhancing the sensitivity of the BLAST search. For details see Korf et al. (2003), and for an application see Wong et al. (2006). Another approach is to narrow the search to a targeted region by identifying conserved flanking genes. Opossum interleukin 2 and 4, and interferon-gamma were identified in this way (Wong et al., 2006).

We have developed and applied more sophisticated approaches to searching for divergent genes in the opossum and tammar wallaby genomes, and in other organisms, based around profile hidden Markov models (e.g. Eddy, 1998). Profile hidden Markov models (HMMs) are statistical models of a multiple sequence alignment. They are frequently used to represent protein domains. The databases Pfam (Finn et al., 2008) and Superfamily (Gough et al., 2001) contain models of thousands of protein domains and are useful resources. The main tool used for searching with profile HMMs is called HMMer (http://hmmer.janelia.org). It can be used to search proteins, translated open reading frames and six frame translations of genomes. An important point with hmmer searches is that two types of models exist; alignment is always local with respect to the sequence, that is to say that if any part of the model matches the sequence it will be regarded as a hit, but may be global ("ls model") or local ("fs model") with respect to the model. A local model will fragment and identify what part of the model is aligning to the sequence. This is useful when a domain consists of multiple sub-domains or when searching with a profile HMM representing a protein family.

Our approach to sensitive genome searching is summarized in Fig. 6.1. The domain organization of the protein family of interest is first characterized. We then search the genome using profile HMMs representing each domain. The results are then searched for domains occurring in genomic regions in the correct orientation and order, possibly also incorporating score. The final step can be achieved in an excel spreadsheet, or using a specialized alignment model. If the protein family

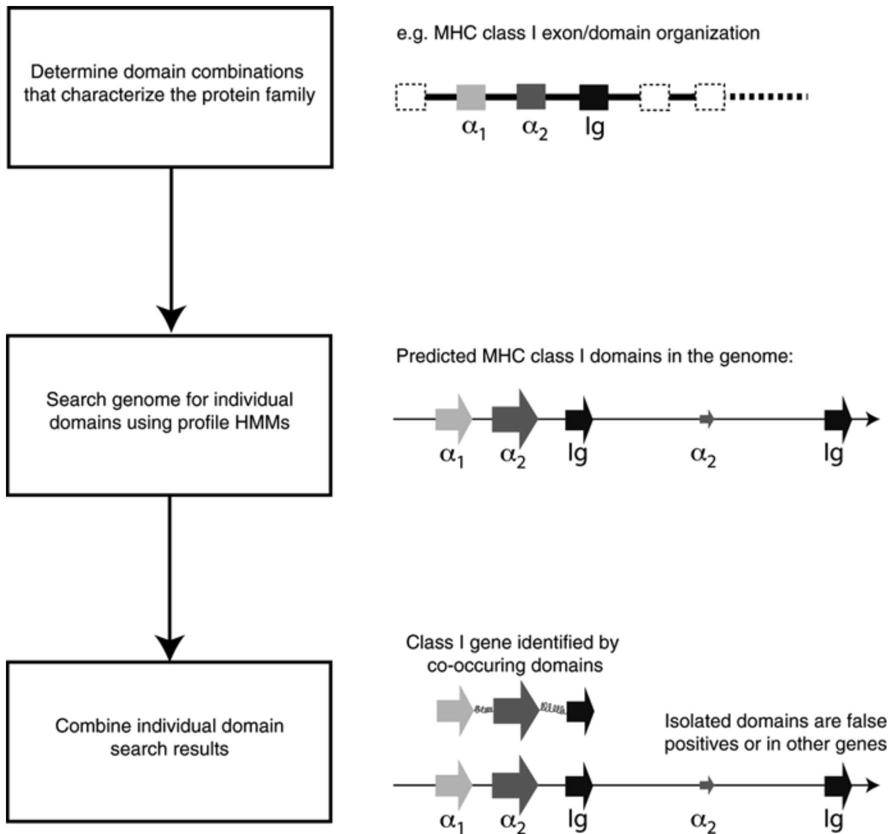


Fig. 6.1 Illustration of profile hidden Markov model (HMM) based approach to sensitive gene searching (*left*). Example of searching for MHC class I genes (*right*)

is characterized by just one domain, then gene prediction around each hit may be necessary. We illustrate the idea of this approach with two different examples.

MHC class I proteins are expressed on the surface of all nucleated cells and present antigens to T-cells. They are involved in recognition of self and non-self. Members of the MHC class I gene family can also play other roles including non-immune function. Class I genes have a signal peptide encoded in exon 1, α_1 and α_2 domains (together making up the antigen-presenting domain) are encoded by exons 2 and 3 respectively, and exon 3 encodes an immunoglobulin domain. Subsequent exons may contain a transmembrane domain and a conserved C-terminal domain. To identify all MHC class I genes in the opossum, we searched the six frame translation of the genome with profile HMMs representing the antigen-presenting domain (Supfam: 0045513), the immunoglobulin domain (Pfam: PF07654) and the C-terminal domain (Pfam: PF06623) (see Fig. 6.1, right hand side). A local alignment model (fs) is again used for the antigen-presenting domain, so that it will

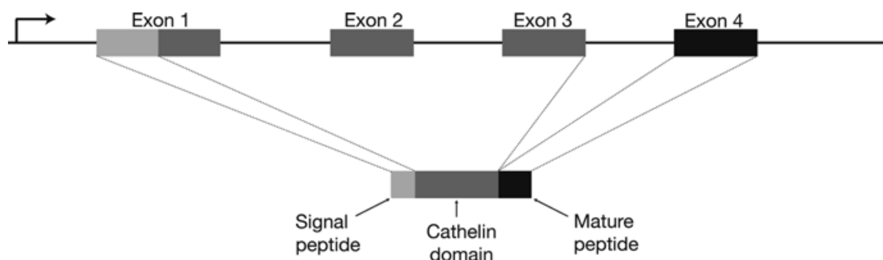


Fig. 6.2 Cathelicidin gene (*top*) and protein (*bottom*) domain organization

split up to identify exons 2 and 3, however the immunoglobulin and C-terminal domains are global alignment models (ls model). Using the default E-value cutoff (E-value$\leq$10), hmmer yields more than 10,000 matches to the antigen-presenting domain (divided roughly evenly between α_1 and α_2 domain hits), more than 2,700 matches to the immunoglobulin domain and around 34,000 matches to the C-terminal domain. Most of these will be false positives. A more conservative E-value threshold would reduce these numbers. However, we retained all hits and developed a heuristic alignment algorithm to identify genomic regions containing an α_1, α_2 and Ig domain in the correct orientation and order, and separated by reasonable intron-like distances. Additionally, C-terminal domains were also included when present and within a reasonable "gene-like" distance. This identified an additional 13 class I genes or pseudo-genes compared to the Ensembl annotation. Expression of a limited number of these has been tested and confirmed (unpublished data).

Cathelicidins are antimicrobials encoded by 4 exon genes. The mature peptide is encoded by the fourth exon and is highly divergent, while exon one encodes a signal peptide and exons one to three encode the cathelin domain (Fig. 6.2). Pfam has a profile HMM representing part of the cathelin domain (Pfam: PF00666). We searched the opossum genome using TBLASTN with the human cathelicidin protein sequence and found only one good hit (E-value $<10^{-4}$). We then searched the six frame translation of the opossum genome with the cathelin profile HMM using hmmer (fs alignment model). The domain model splits up and aligns locally to exons one and two, plus many false positives. We then looked for genomic regions containing matches to exons one and two. Manually curated gene prediction was used to identify the gene structure. In this way we found 12 putative cathelicidin genes, suggesting a large expansion of these genes in the marsupials. See Belov et al. (2007) for further details.

6.4 Conclusion

It is a fertile time for marsupial genomics and sequencing projects. Next generation sequencing is rapidly expanding the amount of marsupial genomic sequence data available – driven by the final stages of the wallaby genome project and the tragic

story of the devil tumour. Third generation or "single molecule" sequencers will bring costs down a further order of magnitude and the deluge of data from this interesting group of mammals may really begin.

This chapter has discussed some of the bioinformatics challenges that arise when dealing with marsupial genomic sequence resources. It has also discussed a number of tools that are useful in dealing with data of this type. It is important to realize that not all bioinformatics problems have good packaged solutions or can be solved with pipelines of existing tools and there is a strong need for bioinformaticians with skills in mathematics, statistics and computer science, to develop new tools where required.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410.

Baker ML, Indiviglio S, Nyberg AM, et al. (2007) Analysis of a set of Australian northern brown bandicoot expressed sequence tags with comparison to the genome sequence of the South American grey short tailed opossum. BMC Genomics 8:50.

Belov K, Deakin JE, Papenfuss AT, et al. (2006) Reconstructing an ancestral mammalian immune supercomplex from a marsupial major histocompatibility complex. PLoS Biol 4:e46.

Belov K, Sanderson CE, Deakin JE, et al. (2007) Characterization of the opossum immune genome provides insights into the evolution of the mammalian immune system. Genome Res 17: 982–991.

Chou H-H, Holmes MH (2001) DNA sequence quality trimming and vector removal. Bioinformatics 17:1093–1104.

Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14:755–763.

Finn RD, Tate J, Mistry J, et al. (2008) The Pfam protein families database. Nucleic Acids Re 36:D281–D288.

Gough J, Karplus K, Hughey R, Chothia C (2001) Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure. J Mol Biol 313:903–919.

Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res 9:868–877.

Hubbard TJ, Aken BL, Beal K, et al. (2007) Ensembl 2007. Nucleic Acids Res 35:D610–D617.

Kent WJ (2002) BLAT–the BLAST-like alignment tool. Genome Res 12:656–664.

Korf I, Yandell M, Bedell J (2003) BLAST. O'Reilly and Associates, Sebastapol.

Kullberg M, Hallström B, Arnason U, Janke A (2007) Expressed sequence tags as a tool for phylogenetic analysis of placental mammal evolution. PLoS One 2:e775.

Lachish S, Jones M, McCallum H (2007) The impact of disease on the survival and population growth rate of the Tasmanian devil. J Anim Ecol 76:926–936.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25.

Lefevre CM, Digby MR, Whitley JC, Strahm Y, Nicholas KR (2007) Lactation transcriptomics in the Australian marsupial, Macropus eugenii: transcript sequencing and quantification BMC Genomics 8:417.

Li R, Li Y, Kristiansen K, Wang J (2008) SOAP: short oligonucleotide alignment program. Bioinformatics 24:713–714.

McCallum H, Tompkins D, Jones M, et al. (2007) Distribution and impacts of Tasmanian Devil Facial Tumour Disease. EcoHealth 4:318–325.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Murchison EP, Tovar C, Hsu A, et al. (2010). The Tasmanian devil transcriptome reveals Schwann cell origins of a clonally transmissible cancer. Science 327(5961):84–87.

Pearse AM, Swift K (2006) Allograft theory: transmission of devil facial-tumour disease. Nature 439:549.

Pertea G (2009) http://compbio.dfci.harvard.edu/tgi/software, Retrieved 20 August 2009, from Dana Farber-Cancer Institute Software Tools.

Pyecroft S, Pearse A, Loh R, et al. (2007) Towards a case definition for devil facial tumour disease: what is it? EcoHealth 4:346–351.

Rumble SM, Lacroute P, Dalca AV, Fiume M, Sidow A, Brudno M (2009) SHRiMP: accurate mapping of short color-space reads. PLoS Comput Biol 5:e1000386.

Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I (2009) ABySS: a parallel assembler for short read sequence. Genome Res 19(6):1117–1123.

Slater GS, Birney E (2005) Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6:31.

Wong ES, Young LJ, Papenfuss AT, Belov K (2006) In silico identification of opossum cytokine genes suggests the complexity of the marsupial immune system rivals that of eutherian mammals. Immunome Res 2:4.

Yeh RF, Lim LP, Burge CB (2001) In silico identification of opossum cytokine genes suggests the complexity of the marsupial immune system rivals that of eutherian mammals. Genome Res 11:803–816.

Zerbino DR, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829.

Chapter 7
Insight into Evolution of Gene Regulation Networks from the Opossum Genome

Matthew J. Wakefield and Anthony T. Papenfuss

Abstract The Brazilian short-tailed grey opossum *Monodelphis domestica* was the first marsupial genome to be sequenced. The high quality draft genome sequence has provided significant new understanding of mammal genome evolution, suggesting that innovation in protein coding genes occurs primarily by diversification of existing gene families and that truly novel protein coding genes are rare. The opossum genome also highlights the magnitude of the role transposable elements have had in shaping gene regulatory networks, including X chromosome inactivation. The rate of innovation of new conserved non-coding elements is 20 fold higher than for protein coding genes, and a substantial portion of the novel eutherian specific conserved non-coding elements can be attributed to arising from transposable elements. Combined with insights into the role of recombination on genome composition and structure, the opossum genome has provided unique insight into the forces that shaped the genomes of all mammals.

Keywords Brazilian short-tailed grey opossum · Gene evolution · Gene regulation · Transposable elements · Exaptation

7.1 Introduction

The first marsupial to be sequenced to high quality draft status was the Brazilian short grey-tailed opossum *Monodelphis domestica*. The major motivation for sequencing a marsupial genome was to use comparisons with other mammalian species to explore the evolutionary history of the mammalian genome. Evolutionary comparisons provide insight into conserved features that are present in all mammals, as well as the unique features that have evolved independently in the eutherian "placental" mammals and the metatherian "marsupial" mammals. Marsupial mammals

M.J. Wakefield (✉)
Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
e-mail: wakefield@wehi.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_7, © Springer Science+Business Media B.V. 2010

provide a powerful addition to the comparative genomics arsenal due to their great evolutionary distance from eutherian mammals, while sharing the majority of common mammalian traits (Wakefield and Graves, 2003; Wakefield and Graves, 2005).

One of the biggest surprises of the Human and subsequent genome projects was the small number of protein coding genes compared to the initial estimates. Prior to the sequencing of the human genome, estimates of gene numbers ranged as high as 100,000, while the final number of genes in the finished human genome was 19,599 with an upper bound of 25,000 (IHGSC, 2004). Comparisons with single celled eukaryotes, insects, and vertebrates led to the conclusion that gene number does not scale with organism complexity (Lander et al., 2001). The lack of correlation between complexity and gene number has lead to the hypothesis that organism complexity is governed by changes in regulatory programs, altering the combination of genes expressed and their response to the environment. Proposed mechanisms for this complex regulation include evolution of transcription factors and their binding sites, and non-coding RNA regulation of transcripts and non-coding RNA regulation of translation into protein (Mattick, 2003; Davidson, 2006; Chen et al., 2007). The importance of regulation in accounting for the discrepancy in complexity is supported by the complex transcription of single gene loci that occurs in mammals (Carninci et al., 2005).

Comparative genomics has been recognized as a powerful tool for identifying functional regions of the human genome. The first mammalian comparative genome analysis was that of the mouse genome, chosen both to provide sequence comparison with the human and due to its importance as an experimental model (Mouse Genome Sequencing Consortium et al., 2002). Comparative power was greatly enhanced by the addition of the dog genome as a third divergent comparator (Lindblad-Toh et al., 2005). Sequencing of the same small regions from the genomes of multiple species, to provide a representation of the power added in comparative analysis, showed that for non protein coding regions marsupials and monotremes were the most powerful single species comparisons (Chapman et al., 2003; Margulies et al., 2005; Wakefield and Graves, 2005). This analysis of power resulted in the eventual selection of platypus as the monotreme to be sequenced to draft quality (Warren et al., 2008), *Monodelphis domestica* to draft quality as the representative marsupial (Mikkelsen et al., 2007), and light two-fold coverage of the tammar wallaby (*Macropus eugenii*) as the representative Australian marsupial and within-marsupial comparator.

The analysis of the opossum genome was therefore focused on the questions of how non-coding regions of the genome evolve and function, and how features unique to eutherian mammals evolved.

7.2 Opossum Genome Features

The opossum genome was sequenced using the shotgun cloning methodology in which small random fragments of DNA are cloned into bacterial vectors, purified individually, and sequenced using Sanger dideoxy fluorescent sequencing on

automated capillary sequencers. The majority of clones were 1–2 kb in size supplemented with 5–10 kb clones and BAC ends to aid scaffolding across contig breaks and anchor scaffolds to chromosomes. In total 38.8 million Sanger sequencing reads were performed to produce 3,475 Mb of sequence with a redundancy of 6.8 fold, N50 contig length of 108 kb and N50 scaffold length of 59.8 Mb (Mikkelsen et al., 2007). Ninety-seven percent of the genome sequence was anchored to chromosomes using fluorescent in situ hybridisation and linkage mapping (Duke et al., 2007; Samollow et al., 2007).

Unusual features of the opossum genome include a large euchromatic genome size of 3.47 Gbp (20% larger than the previous largest sequenced mammal, human with 2.88 Gbp), a low G+C content, low rate of segmental duplications, and a high 52% of the genome derived from transposable elements.

In common with other marsupials, the *Monodelphis domestica* genome has large chromosomes and overall low recombination rates. All eight of the opossum autosomes are larger than any human chromosomes; the smallest opossum autosome, at 258 Mb, is 11 Mb longer than human chromosome 1. Recombination averages 0.2–0.3 centimorgans per megabase (cM/Mb), between half and 1/15th the rate seen in other amniotes, which range from 0.5–3 cM/Mb (see Samollow, Chapter 4). The large size of the *Monodelphis domestica* chromosomes is proposed to drive the overall low recombination rates.

The base composition of the opossum genome is significantly different from that of other sequenced mammal genomes; a G+C density of 37.7% being significantly lower than the 40.9–41.8% observed in other sequenced amniotes (Mikkelsen et al., 2007). This compositional shift in nucleotides away from G+C towards A+T may also be explained by the lower recombination rate, which is proposed to alter the balance between repair and mutation. Mutation is biased towards elevating A+T and thought to remain constant at an absolute level, while the reduced recombination rate would decrease the rate of gene conversion which is biased towards increasing G+C levels (Goodstadt et al., 2007). This conclusion is supported by the low frequency of CpG dinucleotides in the genome (0.9% vs. 1.7–2.2% in other sequenced amniotes), reflecting the high rate of mutation by deamination of methyl cytosine to thymine being less effectively opposed due to lower recombination mediated repair in *Monodelphis domestica* (Goodstadt et al., 2007).

The lower recombination rate may also account for other features of the opossum genome. The large size of the opossum genome relative to human reflects the accumulation of transposable elements, which is consistent with a lower rate of recombination removing fewer new insertions. There is also a greater diversity of transposable elements which may be explained by a reduced level of non-homologous recombination which is thought to be a major homogenizing factor for repeated sequences (Gentles et al., 2007). A lower level of non-homologous recombination would also explain the low level of segmental duplications in the opossum genome (Mikkelsen et al., 2007).

Other differences in the genome metabolism of the opossum, such as changes in the rate of mutation repair pathways that are independent of recombination, and alterations in the biosynthetic pools of bases, may contribute to the differences in

genome composition. Some known changes, such as the duplication of the DNA methyl transferase *DNMT1*, would be predicted to have effects on DNA metabolism (Mikkelsen et al., 2007).

As further recombination mapping in opossum is undertaken the higher resolution measurement of recombination rates local to a particular genome region will provide new opportunities for studying the interplay between composition, epigenetic structure, mutation, repair and recombination in greater detail.

7.3 Protein Coding Genes

The initial annotation of the opossum genome predicted 18,648 protein coding genes. Unambiguous one-to-one orthologues between human and opossum genes account for ∼70% of these genes. The opossum protein coding gene catalogue supports the hypothesis that evolution of protein coding genes occurs mainly via expansion, neofunctionalization (Fig. 7.1) and loss in gene families. Of opossum genes ∼5% have been amplified in human, ∼2.5% amplified in opossum, and ∼2.5% variable multicopy in both. 14% have human homologues with unclear orthology.

The naïve expectation for marsupial specific traits is that novel marsupial specific protein-coding genes would be found to explain these traits. There are only 18 well

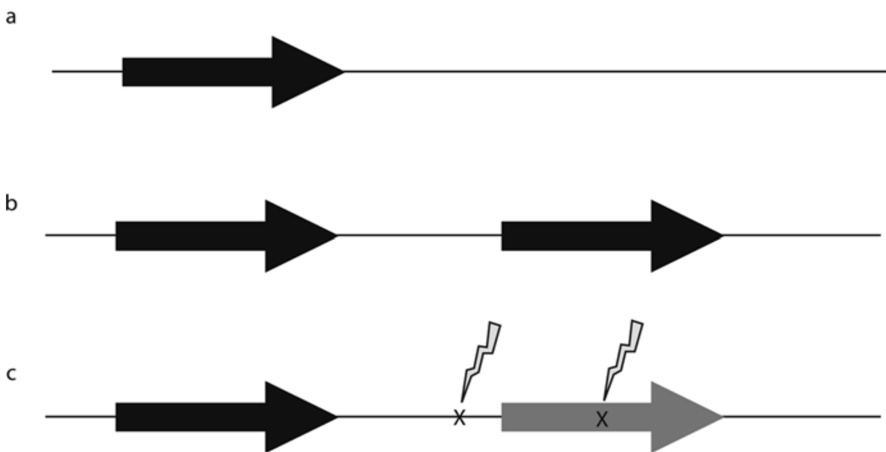


Fig. 7.1 Evolution of new genes by duplication and neofunctionalization. Evolution of new genes in opossum as in other mammals has occurred predominantly by expansion of existing gene families. Initially a gene (**a**) is duplicated by errors in homologous recombination (often in existing tandem arrays of similar genes) or retrotransposition. Initially the duplication creates functional redundancy (**b**). This functional redundancy allows mutations to occur in one copy with no effect on phenotype. In the majority of cases mutations will cause loss of function and creation of a non-coding pseudogene. In a minority of cases mutations in one copy will alter the function of the gene to create a novel gene (**c**). Additional protein coding region and promoter mutations will subsequently refine the function of this new gene

supported opossum specific genes when compared to eutherian mammals, which are ancient genes retained in opossum and lost in the eutherian lineage (Mikkelsen et al., 2007). The method of annotation of the opossum genome did not use opossum transcriptome data but relied on homology to known proteins from other species, primarily human. Due to the restriction to only finding genes that have some similarity to known genes, the gene annotation is therefore potentially biased away from detecting truly novel genes; however, novel genes in other species where transcript based annotation has been undertaken have been observed to be mainly detectable by homology and this bias is likely to be minor. Taking this bias into consideration the observed low rate of novel genes strongly suggests that true innovation in protein coding genes is rare relative to regulatory changes and diversification by duplication. The estimate that 15% of genes have been involved in one or more gene duplication or gene conversion event is strongly reduced when contrasted with the 20% duplication rate in human. This reduction in duplication and gene conversion is consistent with the reduced rate of segmental duplication implicating the lower opossum recombination rate as a driver of reduced innovation in protein coding genes (Goodstadt et al., 2007).

The diversification that has occurred in the opossum lineage is mainly in gene families involved in environmental interactions. These include eye lens proteins, taste, odorant and pheromone receptors. Diet related genes such as those involved in toxin degradation and gastric enzymes are another rapidly evolving class, as well as those involved in the adaptive immune system (Belov et al., 2007), which is examined in detail elsewhere in this volume (Cheng et al., Chapter 16; Baker et al., Chapter 17; Morris et al., Chapter 18).

Interestingly one rapidly evolving family in opossum and other mammals is that of the KRAB zinc fingers, transcription factors that have rapid positive selection for adaptive changes in the zinc finger DNA binding site that mediates the specificity of these transcription factors. KRAB zinc fingers therefore provide a link between the evolution of protein coding genes and regulatory elements (Shannon et al., 2003; Mikkelsen et al., 2007).

The rapid birth and death of genes in existing multi gene families, and the functional categories involved, mirror those observed previously in eutherian mammals (Waterston et al., 2002; Emes et al., 2003; Gibbs et al., 2004; Chimpanzee Sequencing and Analysis Consortium et al., 2005; Goodstadt et al., 2007). Two notable expansions that have not previously been observed in vertebrates may have significant impacts on marsupial biology. Multiple duplications have occurred in the *SMG5* and *SMG6* genes which are involved in nonsense mediated decay (Mikkelsen et al., 2007). The *SMG* proteins are components of P-bodies which in addition to nonsense mediated decay play roles in miRNA and RNAi processes (Rehwinkel et al., 2005), raising the intriguing possibility of marsupial specific innovation in miRNA processing. Another marsupial specific duplication is the presence of two tandem paralogues of the *DNMT1* cytosine methyl transferase gene. In mouse *DNMT1* has been shown to be essential for the maintenance of random X inactivation in the embryo but not imprinted extraembryonic tissue inactivation (Sado et al.,

2000), and duplication of the *DNMT1* gene may be involved marsupial specific difference in X inactivation and differences in nucleotide composition of the opossum genome.

7.4 Evolution of Non Protein Coding Regulatory Sequences

The significant paradigm shift that occurred as the consequence of the genomics revolution at the turn of the twenty-first century was an understanding of the transcriptional complexity of the genome. A corollary of the transcriptional complexity is the requirement for complex regulatory mechanisms to orchestrate this complexity.

Conservation of sequence is a powerful tool for identifying regions of the genome that are important for function. Over time, sequences will be subject to random mutation at a similar rate across the genome, and a highly similar rate in local regions of the genome. If a mutation occurs in a non-functional region of the genome, there will be no phenotypic consequence for this mutation and natural selection will be unable to act on this mutation. These neutral mutations will therefore accumulate randomly in non-functional DNA. In locations that are functional, either due to the coding for a protein or serving a regulatory function, mutations will have a phenotypic consequence. As the phenotype is changed by the mutation natural selection is able to act on these mutations. In a very tiny minority of cases the mutation will cause a change that is beneficial and eventually come to dominate in the population; however, in most cases the mutation will be deleterious and will be quickly eliminated from the population. This removal of deleterious mutations is known as purifying selection.

As a consequence of the different fates of mutation in functional and non-functional regions of the genome, functional DNA sequence can be identified by the higher sequence conservation between species. This conservation effect is strongest in protein coding genes, as mutations that truncate a protein have a high phenotypic consequence and are strongly selected against, but is also apparent in genome features that are permissive of sequence variation such as transcription factor binding sites.

In comparisons between the human genome and those of mouse and dog it was discovered that the majority of evolutionarily conserved sequence occurs outside protein coding genes (Lindblad-Toh et al., 2005). Only 1–2% of the genome in mammals is identified as being protein coding. By comparing to ancient transposable elements and degenerate codon positions to estimate the neutral mutation rate, 5–6% of the genome is observed to be under purifying selection. Given that only 1–2% of this 5–6% conserved sequence is protein coding the observed conservation suggests that 3–5% of the genome is conserved non-coding sequence and putative regulatory regions of the genome.

The proportion of the genome under purifying selection may be underestimated due to the difficulty in defining truly neutral alignable sequence. There are multiple

evolutionary forces acting on the sequences that are usually used to calibrate the neutral rate of no selection for function. For example synonymous sites in proteins are known to be under selection for codon usage and context dependent effects such as CpG mutation rates (Huttley et al., 2000; Lindsay et al., 2008), while the use of ancient repeats is confounded by the homogenising effect of non-homologous recombination and gene conversion, and possible selection (Elder and Turner, 1995; Brookfield, 2005a; Kamal et al., 2006). If these evolutionary forces acting on the genome constrain the sequences used for calibration, causing the underestimation of the rate of change in neutral sequence, the real portion of the genome under selection for unknown functions may be significantly higher than the current estimate of 5–6%.

One potential origin for the excess conservation observed is transposable elements. Transposable elements are usually viewed as genomic parasites, hitchhiking in the genome with no significant function other than the selfish replication and propagation of their own sequence. The maintenance and flourishing diversity of repeat element flora in the genome requires only that they have no strong negative selective effect (i.e. do not kill the host prior to reproduction), and replicate faster than the host (Brookfield, 2005b). Although this neutral model is sufficient to explain the presence of repeat elements in the genome, it fails to capture the potential consequences and evolutionary possibilities that the presence of repeated sequences enables.

The potential role of transposable elements as a fundamental component in gene regulation and regulatory evolution has been apparent since the pioneering work of Barbara McClintock (McClintock, 1950). This early work demonstrated the ability of mobile element transposition to bring repressors or enhancers of transcription into proximity with new genes, altering the pattern of gene expression in the organism. This form of regulatory evolution mediated by transposons has been demonstrated to occur in many important evolutionary systems, including the highly economically important evolution of insecticide resistance (Chung et al., 2007). Although the mobility of transposable elements provides an important evolutionary mechanism, the transposition of retroelement like repeats that replicate by producing and inserting new copies of the element in the genome has had a significantly more profound impact on mammalian evolution. By replicating highly similar sequences and distributing copies of this sequence in lightly biased but essentially random locations throughout the genome transposable elements provide a “battery” of similar target sequences that can be exapted to co-regulate multiple genes (Fig. 7.2) (Britten and Davidson, 1969).

The precise mechanism of exaptation can take several forms, but are unified by the key feature of taking neutrally evolved patterns in the genome created by the transposable elements and using them as cis recognition sites for protein or small RNA based transcriptional regulation (Feschotte, 2008).

One of the earliest genomics studies to provide evidence that some transposable elements are under functional selection was a comparison between human and mouse genomes (Silva et al., 2003). This study provided support for conservation of the MIR and L2 families of transposons in excess of the background for human

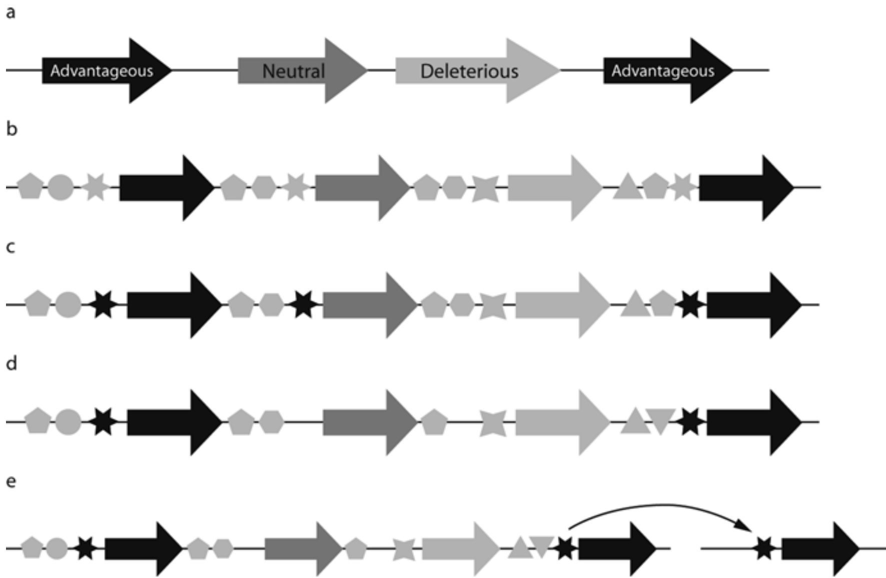


Fig. 7.2 Evolution of regulatory networks by exaptation of repeats. (**a**) As an example of exaptation of repeat sequences as a multigene regulatory network consider the case where expression of two genes simultaneously (*black*) is advantageous, expression of another gene is deleterious (*light grey*), and expression of a fourth gene has no effect on fitness (*dark grey*). (**b**) Transposable elements inserted in the genome in a lightly biased but essentially random pattern provide a battery of identical sequences with different distributions around the genes under selection. The distribution of the repeat elements occurs in the absence of any selection for function in a regulatory network. (**c**) If a mutation of a *trans* acting factor occurs such that it recognizes one of the repeat elements a new multi gene regulatory network is formed. In this example the *star element* becomes an enhancer activating the two advantageous genes and the neutral gene. Equally the *trans* factor could be a repressor, a micro RNA or other regulatory factor. Note that unlike *cis* element mutation, *trans* element mutation and exaptation of existing repeat patterns only requires one mutation to target multiple genes. (**d**) If a new regulatory network contains advantageous and neutral genes it will be selected for. This selection will maintain the portion of the repeat elements recognised by the *trans* factor around genes under selection. Where no *trans* factor binds, or where the expression of the gene is not under selection (neutral gene – *dark grey*) random mutation and drift will eventually result in the loss of the repeat element. (**e**) In cases where the transposable element is still active transposition can result in the insertion of new copies of the *cis* recognition sites recruiting additional genes to the regulatory network

and mouse. Further studies on specific example enhancers provided functional evidence that transposon derived sequences that are conserved in excess of background are distal enhancers capable of regulating tissue specific expression, and identified ancient repeat elements in other species that support a transposable element origin of these sequences (Bejerano et al., 2006; Nishihara et al., 2006). For additional examples see the review by Feschotte (2008).

The completion of the *Monodelphis* genome sequence allowed comparisons with human, mouse, dog, rat and chicken as an outgroup, greatly improving power to identify elements that are conserved across mammals. This, for the first time,

provided a powerful genome wide examination of the degree of conservation and evolutionary history of conserved non-coding elements. By using the opossum genome as an outgroup to identify new, eutherian specific, conserved non-coding elements the degree this form of regulatory innovation and the origin of these elements could be determined. A surprisingly high 20% of conserved non-coding elements in eutherian mammals were found to be recent eutherian innovations, post-dating the marsupial – eutherian divergence. When the origin of these elements was investigated 16% of these elements were found to be derived from clearly recognizable transposons (Kamal et al., 2006; Gentles et al., 2007; Mikkelsen et al., 2007). This rate of innovation in regulatory sequences was found to be 20 fold higher than the 1% observed for protein coding genes.

In addition to the elements which have become evolutionarily constrained in the eutherian lineage, a substantial number of ancient transposable elements are present in all mammals and are more conserved than neutral sequence. The sequencing of the *Monodelphis* genome aided in identifying 83 new multicopy elements. These elements have 18,290 copies in *Monodelphis* and 11,488 copies in human. Of the sites conserved in human, 3,512 copies are located in previously identified evolutionary conserved regions (Gentles et al., 2007). As 4.75% of the human genome is identified as evolutionarily conserved regions, the occurrence of 30% of the insertion sites of these 83 elements in conserved non coding regions strongly supports a transposable element origin of many of these regulatory regions (Gentles et al., 2007). Some of these elements have been shown to be more highly conserved between opossum and human than within opossum (Gentles et al., 2007) suggesting there is more selection on a subset of these elements in their current genomic location than there was selection on the element during its transposition throughout the genome. This strongly supports Britten and Davidson's thesis that exaptation of transposable elements is driven by the selective advantage provided by multiple copies, and that transposition precedes exaptation in the majority of cases. The selection after transposition hypothesis is further supported by conservation of flanking sequence of the elements that is stronger between species than between copies of the element within a species in one t-RNA SINE element studied in detail (Gentles et al., 2007). The greater conservation between species demonstrates that the selection is acting on the elements in their inserted locations is greater than the selection on the element at the time of transposition, consistent with neutral dispersal and initial mutational drift, and subsequent selection of a subset of elements for function dampening decay.

7.5 Exaptation of Line Elements in XCI

X chromosome inactivation is a large scale epigenetic silencing mechanism that equalizes the amount of gene product between male and female mammals. In eutherian mammals X inactivation is mediated by the *XIST* gene and in somatic tissues the X chromosome to be inactivated is randomly chosen to be either the maternally or

paternally derived chromosome. In marsupials the paternal X chromosome is always inactivated and inactivation occurs without the involvement of the *XIST* gene (Hore et al., 2007; see also Al Nadaf et al., Chapter 13).

One particular family of transposable elements, L1 LINEs, are concentrated on the eutherian X (enrichment from 17 to 29%) and have been proposed to act as booster elements for the spread of X chromosome inactivation (Lyon, 1998). This hypothesis is supported by the correlation between the L1 density in the different evolutionary blocks of the X (Graves, 1995), this density being higher in the ancient conserved region of the human X chromosome represented by the marsupial X (in which nearly all genes are inactivated), than in the region of the X chromosome that has been added since the divergence from marsupials (which is autosomal in marsupials and contains many genes that escape inactivation). Additionally X inactivation does not spread into autosomal translocations that are deficient in L1 (Bailey et al., 2000). However, at a local scale there is no simple association between genes that escape X chromosome inactivation and percentage of L1 element coverage (Carrel and Willard, 2005), although classifiers using L1 LINEs combined with other sequence can predict inactivation status with 80% accuracy (Wang et al., 2006). The opossum genome provided an opportunity to examine the accumulation of LINE elements and determine whether this enrichment represented an exaptation event prior to the entire mammalian radiation or at the base of the eutherian radiation after divergence from marsupials.

Comparing L1 LINE density between the opossum autosomes and X chromosome indicates a small enrichment on the marsupial X from 19.2 to 21.5%, a 1.15 fold enrichment (Mikkelsen et al., 2007). The magnitude of this difference is far less than the 1.7 fold enrichment in human, and reveals major differences in the processes that are acting to enrich L1 elements on the marsupial and human X chromosomes.

Whereas the significantly high concentration of L1 LINEs on the ancient region of the eutherian X suggests selection for a role in X inactivation, neutral factors not related to X inactivation can readily account for the lower L1 concentration on the marsupial X as many autosomal regions of similar size have higher L1 concentration. This leads to the conclusion that L1 LINE accumulation does not play a vital role in marsupial X inactivation. While the difference in L1 accumulation on the eutherian and marsupial X might be due to unknown lineage-specific differences in neutral factors, it is more likely that they relate to fundamental differences in the mechanism of X chromosome inactivation in the two lineages (Mikkelsen et al., 2007).

Marsupial X chromosome inactivation is less stable than inactivation in eutherians and is associated with reduced or absent promotor methylation of inactivated genes, and a reduction in histone acetylation on the inactive X chromosome (Wakefield et al., 1997). The repressive polycomb PRC2 histone H3 lysine 27 trimethylation mark has been proposed to be absent from the inactive tammar wallaby X chromosome (Koina et al., 2009); however, contrary observations have recently been reported for opossum (Mahadevaiah et al., 2009). Further work will

therefore be required before the any role for L1 accumulation in the development of the multilayered X inactivation mechanism can be determined.

X chromosome inactivation therefore appears to be a striking case of exaptation at the base of the eutherian lineage. L1 elements accumulated neutrally on the ancient X chromosome in a common therian ancestor, and this continued on the marsupial X, without selection for a role in X inactivation. Paternal X inactivation in marsupials requires no choice or counting mechanism. Inactivation could be accurately targeted to the evolving heteromorphic region of the X chromosome simply by identifying and silencing regions that are unpaired in male meiosis. Such an unpaired DNA silencing mechanism would have had fitness benefits prior to sex chromosome evolution in silencing chromosomal abnormalities and transposable elements and is seen in meiotic sex chromosome inactivation (Turner et al., 2002), so need not have co-evolved with the heteromorphic state.

In the eutherian lineage, however, these enriched L1 LINEs have apparently been co-opted into a chromosome-wide X inactivation mechanism after the evolution of the *XIST* gene. This exaptation may have been enhanced (or even precipitated) by the addition of the L1-rich XAR. The inherently unstable *XIST* homologous region XICHR is hypothesized to have undergone pseudogenisation (Duret et al., 2006) and invasion of repetitive elements (Hore et al., 2007) to create a proto-*XIST* RNA that interacted with nearby L1 sequences on the X chromosome, promoting stabilization of inactivation by inducing repeat silencing molecular mechanisms. This interaction with L1 LINEs may account for the constrained re-arrangement of the X in eutherian mammals relative to Australian and American marsupials, and actively selected for further accumulation of the critical L1 sequences (Mikkelsen et al., 2007; Warren et al., 2008). Further evolution of the proto-*XIST* and the surrounding locus then gave rise to the female specific expression, counting and choice mechanisms present in the modern eutherian XCI system from this basic exaptation of repeats origin.

7.6 Conclusion

The large chromosomes and reduced recombination rate of the opossum genome has provided an illuminating contrast to other mammals that highlights important processes of genome evolution. The reduced rate of segmental duplication has suppressed the contribution of duplication and diversification of protein coding genes, suppressed clearance of transposable elements resulting in an increased genome size, and altered the balance of mutation and repair resulting in reduced G+C content.

Although the phenomenon of exaptation of repeats was known and demonstrated prior to the sequencing of the opossum genome, this evolutionary process was often considered a quirky minor player in how vertebrate complexity evolved. The lasting legacy of the opossum genome analysis will be the identification of multiple examples of the exaptation of transposable elements, and an indication of the scale of contribution exaptation has made throughout evolution.

This has provided an appreciation of the important role the critical twin features of diverse sequence, and its distribution in variable repetitive copy density throughout the genome, has played in the evolution of regulatory complexity. The examples of ancient mammal wide repeats and recent exaptation of transposable elements in the eutherian lineage provide evidence of both the ancient and evolutionarily recent contributions of this mechanism. Excitingly, this new appreciation of how networks of gene regulation can form via exaptation provides testable hypotheses on the nature of regulatory network evolution. Exapted networks must initially start with the random distribution of elements that exists prior to selection. Therefore, recently evolved networks will tend to regulate more genes than the phenotype requires. Over time mutation will remove non-essential genes from the network as there is no selection for their retention. This predicts that more ancient networks will be smaller and contain a higher proportion of genes essential for the phenotype.

Many questions remain unanswered about how complex regulatory networks evolve, and the ancient and recent networks identified by the opossum genome will continue to enhance our understanding of our symbiotic relationship with our genomes parasites.

Glossary

N50 The size of scaffold or contig for which half of all bases are in scaffolds or contigs that are smaller.

Neofunctionalization The mutation of a duplicated copy of a gene to undertake a completely novel function (distinct from subfunctionalization where the role of the gene in specific tissues, or one function of a multidomain protein is split between the duplicate copies of the gene).

Exaptation The co-option by selection for a new function of a structure that has evolved under selective pressure for an unrelated function or under neutral conditions with no functional fitness advantage (Gould and Vrba, 1982).

LINE Long Interspersed Nuclear Element, a family of retrotransposons common in mammals which is transcribed as RNA and intergrates as additional reverse transcribed copies in genome.

References

Bailey JA, Carrel L, Chakravarti A, Eichler EE (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci USA 97:6634–6639.

Bejerano G, Lowe CB, Ahituv N, et al. (2006) A distal enhancer and an ultraconserved exon are derived from a novel retroposon. Nature 441:87–90.

Belov K, Sanderson CE, Deakin JE, et al. (2007) Characterization of the opossum immune genome provides insights into the evolution of the mammalian immune system. Genome Res 17: 982–991.

Britten RJ, Davidson EH (1969) Gene regulation for higher cells: a theory. Science 165:349–357.

Brookfield JF (2005a). Evolutionary forces generating sequence homogeneity and heterogeneity within retrotransposon families. Cytogenet Genome Res 110:383–391.

Brookfield JF (2005b). The ecology of the genome – mobile DNA elements and their hosts. Nat Rev Genet 6:128–136.

Carninci P, Kasukawa T, Katayama S, et al. (2005) The transcriptional landscape of the mammalian genome. Science 309:1559–1563.

Carrel L, Willard HF (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434:400–404.

Chapman MA, Charchar FJ, Kinston S, et al. (2003) Comparative and functional analyses of LYL1 loci establish marsupial sequences as a model for phylogenetic footprinting. Genomics 81: 249–259.

Chen JM, Cooper DN, Chuzhanova N, Férec C, Patrinos GP (2007) Gene conversion: mechanisms, evolution and human disease. Nat Rev Genet 8:762–775.

Chung H, Bogwitz MR, McCart C, et al. (2007) Cis-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the Drosophila melanogaster insecticide resistance gene Cyp6g1. Genetics 17:1071–1077.

Chimpanzee Sequencing and Analysis Consortium et al. (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437:69–87.

Davidson EH (2006) The Regulatory Genome: Gene Regulatory Networks in Development and Evolution. Academic, Burlington and San Diego.

Duke SE, Samollow PB, Mauceli E, Lindblad-Toh K, Breen M (2007) Integrated cytogenetic BAC map of the genome of the gray, short-tailed opossum, *Monodelphis domestica*. Chromosome Res 15:361–370.

Duret L, Chureau C, Samain S, Weissenbach J, Avner P (2006) The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. Science 312:1653–1655.

Elder JF Jr, Turner BJ (1995) Concerted evolution of repetitive DNA sequences in eukaryotes. Q Rev Biol 70:297–320.

Emes RD, Goodstadt L, Winter EE, Ponting CP (2003) Comparison of the genomes of human and mouse lays the foundation of genome zoology. Hum Mol Genet 12:701–709.

Feschotte C (2008) Transposable elements and the evolution of regulatory networks. Nat Rev Genet 9:397–405.

Gentles AJ, Wakefield MJ, Kohany O, et al., (2007) Evolutionary dynamics of transposable elements in the short-tailed opossum *Monodelphis domestica*. Genome Res 17:992–1004.

Gibbs RA, Weinstock GM, Metzker ML, et al. (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428:493–521.

Goodstadt L, Heger A, Webber C, Ponting CP (2007) An analysis of the gene complement of a marsupial, Monodelphis domestica: evolution of lineage-specific genes and giant chromosomes. Genome Res 17:969–981.

Gould SJ, Vrba ES (1982) Exaptation; a missing term in the science of form. Paleobiology 8:4–15.

Graves JAM (1995) The evolution of mammalian sex chromosomes and the origin of sex determining genes. Philos Trans R Soc Lond B Biol Sci 350:305–311.

Hore TA, Koina E, Wakefield MJ, Graves JAM (2007) The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. Chromosome Res 15:147–161.

Huttley GA, Jakobsen IB, Wilson SR, Easteal S (2000) How important is DNA replication for mutagenesis? Mol Biol Evol 17:929–937.

IHGSC (2004) Finishing the euchromatic sequence of the human genome. Nature 431:931–945.

Kamal M, Xie X, Lander ES (2006) A large family of ancient repeat elements in the human genome is under strong selection. Proc Natl Acad Sci USA 103:2740–2745.

Koina E, Chaumeil J, Greaves IK, Tremethick DJ, Graves JA (2009) Specific patterns of histone marks accompany X chromosome inactivation in a marsupial. Chromosome Res 17:115–126.

Lander ES, Linton LM, Birren B, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921.

Lindblad-Toh K, Wade CM, Mikkelsen TS, et al. (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438:803–819.

Lindsay H, Yap VB, Ying H, Huttley GA (2008) Pitfalls of the most commonly used models of context dependent substitution. Biol Direct 3:52.

Lyon MF (1998) X-chromosome inactivation: a repeat hypothesis. Cytogenet Cell Genet 80: 133–137.

Mahadevaiah SK, Royo H, VandeBerg JL, McCarrey JR, Mackay S, Turner JM (2009) Key features of the X inactivation process are conserved between marsupials and eutherians. Curr Biol 19:1478–1484.

Margulies EH, NISC Comparative Sequencing Program, Maduro VV, et al. (2005) Comparative sequencing provides insights about the structure and conservation of marsupial and monotreme genomes. Proc Natl Acad Sci USA 102:3354–3359.

Mattick JS (2003) Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. Bioessays 25:930–939.

McClintock B (1950) The origin and behavior of mutable loci in maize. Proc Natl Acad Sci USA 36:344–355.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Mouse Genome Sequencing Consortium Waterston RH, Lindblad-Toh K, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562.

Nishihara H, Smit AF, Okada N (2006) Functional noncoding sequences derived from SINEs in the mammalian genome. Genome Res 16:864–874.

Rehwinkel J, Behm-Ansmant I, Gatfield D, Izaurralde E (2005) A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. RNA 11:1640–1647.

Sado T, Fenner MH, Tan SS, Tam P, Shioda T, Li E (2000) X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. Dev Biol 225:294–303.

Samollow PB, Gouin N, Miethke P, et al. (2007) A microsatellite-based, physically anchored linkage map for the gray, short-tailed opossum (*Monodelphis domestica*). Chromosome Res 15:269–281.

Shannon M, Hamilton AT, Gordon L, Branscomb E, Stubbs L (2003) Differential expansion of zinc-finger transcription factor loci in homologous human and mouse gene clusters. Genome Res 13:1097–1110.

Silva JC, Shabalina SA, Harris DG, Spouge JL, Kondrashovi AS (2003) Conserved fragments of transposable elements in intergenic regions: evidence for widespread recruitment of MIR- and L2-derived sequences within the mouse and human genomes. Genet Res 82:1–18.

Turner JM, Mahadevaiah SK, Elliott DJ, et al. (2002) Meiotic sex chromosome inactivation in male mice with targeted disruptions of Xist. J Cell Sci 115:4097–4105.

Wakefield MJ, Graves JAM (2003) The kangaroo genome. Leaps and bounds in comparative genomics. EMBO Rep 4:143–147.

Wakefield MJ, Graves JAM (2005) Marsupials and monotremes sort genome treasures from junk. Genome Biol 6:218.

Wakefield MJ, Keohane AM, Turner BM, Graves JA (1997) Histone underacetylation is an ancient component of mammalian X chromosome inactivation. Proc Natl Acad Sci USA 94: 9665–9668.

Wang Z, Willard HF, Mukherjee S, Furey TS (2006) Evidence of Influence of Genomic DNA Sequence on Human X Chromosome Inactivation. PLoS Comput Biol 2:e113.

Warren WC, Hillier LW, Graves JAM, et al. (2008) Genome analysis of the platypus reveals unique signatures of evolution. Nature 453:175–183.

Part IV
Marsupial Sex Chromosomes

Paul D. Waters

Comparing sex chromosome systems between distantly related mammalian species has been a powerful tool in helping to solve mysteries surrounding the origin and evolution of the mammal X and Y. Comparisons of model eutherian sex chromosomes (generally human and mouse) with the sex chromosomes of distantly related marsupial representatives have revealed many similarities, but also some striking differences. Although much of the eutherian X chromosome is homologous to the marsupial X, there is a region of the eutherian X (generally most of the short arm) that is autosomal in marsupials, indicating that the eutherian sex chromosomes were augmented by an addition prior to eutherian radiation (Graves, 1995), and accounts for the size difference observed between the eutherian (~5% of the haploid genome) and marsupial (~3% of the haploid genome) Xs.

In this part, Patel and colleagues (Chapter 8) review the structure of the marsupial X chromosome, which has become possible at the sequence level thanks to the release of the opossum (*Monodelphis domestica*) genome (Mikkelsen et al., 2007) and upcoming release of the tammar wallaby genome. In contrast to the eutherian X chromosome, which is well conserved across almost all eutherian lineages, the X chromosomes of American and Australian marsupials have suffered several independent rearrangements (Deakin et al., 2008). Hemizygosity (one copy) of the X in males, and meiotic sex chromosome inactivation (MSCI), have been major forces in shaping the gene content of the X chromosome, which happens to be enriched for genes involved in brain development and reproduction. In Chapter 9, Delbridge examines the gene content of the X chromosome and discusses how these forces have resulted in gene trafficking to and from the X, and how the X might possibly be involved in speciation.

It is now apparent that the therian mammal (eutherian and marsupial) sex chromosomes arose after monotremes diverged from therians, but before therian radiation (Veyrunes et al., 2008). As different as the X and Y chromosomes are in therian mammals, they evolved from an homologous pair of autosomes after the Y obtained the dominant male determining gene *SRY*. Genes beneficial to males accumulated near this new testis-determining factor (TDF), and recombination was suppressed with the X so that this cassette of male beneficial genes was only ever inherited with the TDF. In the absence of recombination, mutated Y chromosomes

in the population could not be repaired by recombination with the X, and the Y began to lose gene function and degrade.

In eutherian mammals all that remains of once extensive homology between the X and Y is small pseudoautosomal region (PAR). Within this PAR there is an obligatory recombination event at male meiosis, which is responsible for the proper pairing and segregation of the X and Y. However, marsupial sex chromosomes do not have a PAR, instead pairing and segregation of the X and Y during male meiosis is controlled by a completely different mechanism. Fernández-Donoso and colleagues (Chapter 10) review the intriguing process whereby proteins involved in synaptanemal complex formation (Page et al., 2006) come together to form a structure called the dense plate, which associates with the X and Y, and controls their proper segregation into daughter cells.

The final chapter of this part, by Murtagh et al., reviews the marsupial Y chromosome and Y chromosome evolution. This chromosome is generally small in eutherian mammals, and often minute (10 Mb) in marsupials. As such the marsupial Y has been suggested to represent the minimal mammalian Y (Toder et al., 2000). Although genome sequence is available for two marsupial species, there is little information about the Y because females were chosen for sequencing to avoid the problems associated with Y chromosome assembly. Murtagh and colleagues describe current knowledge of the novel gene content of the tammar wallaby Y chromosome. All genes so far discovered on the wallaby Y have partners on the X that, interestingly, almost all have a role in the vertebrate brain or central nervous system. Most genes on the marsupial Y are missing from the eutherian Y, and probably gained male-specific functions only in marsupials. There are only a few genes conserved on the Y chromosome of all therian mammals, an obvious example being the testis-determining factor *SRY*.

The four chapters in this part highlight the importance of data gathered from a range of distantly related mammals, and other vertebrates, in understanding the structure and organization of sex chromosomes. There is a shortfall in Y chromosome sequence data from a host of mammalian species, and completed X chromosome assemblies are missing from many species. Filling in these gaps is critical to a deeper understanding of sex chromosome evolution.

References

Deakin JE, Koina E, Waters PD, et al. (2008) Physical map of two tammar wallaby chromosomes: a strategy for mapping in non-model mammals. Chromosome Res 16:1159–1175.

Graves JAM (1995). The origin and function of the mammalian Y chromosome and Y-borne genes – an evolving understanding. Bioessays 17:311–320.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial Monodelphis domestica reveals innovation in non-coding sequences. Nature 447:167–177.

Page J, Viera A, Parra MT, et al. (2006) Involvement of synaptonemal complex proteins in sex chromosome segregation during marsupial male meiosis. PLoS Genet 2:e136.

Toder R, Wakefield MJ, Graves JAM (2000) The minimal mammalian Y chromosome – the marsupial Y as a model system. Cytogenet Cell Genet 91:285–292.

Veyrunes F, Waters PD, Miethke P, et al. (2008) Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res 18:965–973.

Chapter 8
Organization and Evolution of the Marsupial X Chromosome

Hardip R. Patel, Margaret L. Delbridge, and Jennifer A.M. Graves

Abstract Marsupial genomics and genetics provide invaluable data for comparative analysis because these mammals are distantly related cousins of humans and other eutherian mammals (~148 million years), but they are more closely related to eutherians compared to monotremes (~166 million years) and birds (~310 million years). This divergence time places them uniquely on the phylogenetic tree whereby comparative gene mapping and genome sequences of marsupials enable us to trace the evolution of the therian X chromosome. It is shown by comparative gene mapping that the marsupial X chromosome is homologous to the long arm and pericentric region of the short arm of the human X chromosome (X conserved region). However, it is homologous to autosomes in the platypus and the chicken, suggesting that a novel XY system evolved in therian mammals after they diverged from monotremes 166 million years ago but before marsupials diverged from eutherians 148 million years ago. Apart from the X conserved region, eutherian mammals have an autosomal addition on the X chromosome (X added region) since their divergence from marsupials. Although the marsupial X chromosome is similar in composition to that of the eutherian X chromosome, it has accumulated numerous rearrangements, has become GC rich and has acquired a higher synonymous substitution rate since divergence of marsupial and eutherian mammals. Similarly, the marsupial X chromosome undergoes X inactivation like eutherian mammals, but the mechanisms, which regulates X inactivation, is very different from that observed in eutherian mammal, and may be the ancestral state. The *XIST* gene responsible for X inactivation in eutherian mammals is absent from the marsupial lineage suggesting that the evolution of *XIST* in early eutherians set up selection for accumulation of LINE sequences that aided the spread of X inactivation. Unlike the LINE accumulation on the eutherian X chromosome, the marsupial X has accumulated and expanded a microRNA family that could be involved in the marsupial X inactivation. This chapter provides a comprehensive

H.R. Patel (✉)
ARC Centre for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia
e-mail: hardip.patel@anu.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_8, © Springer Science+Business Media B.V. 2010

review of the organization and evolution of the marsupial X chromosome and the implications this may have for X inactivation.

Keywords Chromosome painting · Comparative genomics · Genome evolution · Karyotype · Sex chromosomes

8.1 Introduction

Sex chromosomes have evolved multiple times during the vertebrate evolution. In other words, different vertebrate groups have different sex chromosome pairs harbouring genes that trigger genetic sex determination (GSD) of the embryo. Mammals subscribe to a conserved XY system of male heterogamety in which the *SRY* gene on the Y chromosome is the male-determining gene at least in therian mammals. In contrast, birds and snakes show a ZW system of female heterogamety, although their ZW pairs are non-homologous. Similarly, many different species of fish, amphibians and reptiles exhibit either XY or ZW systems. Although sex chromosomes are strongly dimorphic (morphologically distinguishable) in mammals and birds, they are homomorphic (morphologically and genetically similar to each other) in many fish and amphibians with GSD. It should be noted that some fish and reptilian species determine sex via environmental cues such as temperature (ESD – environmental sex determination or TSD – temperature sex determination).

These quite different sex chromosome systems are the result of independent but parallel evolutionary processes. First a novel sex-determining locus is acquired on one member of a pair of autosomes. This is followed by accumulation of sex-specific genes that confer an advantage to the heterogametic sex, and then homologous recombination is decreased in this region by selection for the preservation of a sex-specific block of genes (Rice, 1987b). The result is progressive degeneration of the sex-specific element (Y or W) (reviewed in Graves, 2006).

For instance, in mammals, acquisition of the testis-determining *SRY* gene on a proto-Y chromosome was followed by accumulation of male-advantage (e.g. spermatogenesis) genes to favour male fitness. This non-recombining region accumulated deletions as well as beneficial mutations because variation on the male-specific Y is high and selection is inefficient (Muller, 1918; Charlesworth, 1978). The degradation of the Y chromosome is exacerbated by a one-way genetic drift ("Muller's ratchet"; Felsenstein, 1974) and inefficient selection on a non-recombining chromosome ("genetic hitchhiking"; Rice, 1987a; Graves, 2006). On an evolutionary time scale, the proto-Y chromosome becomes fixed for male beneficial recessive alleles at several loci and the rest of the Y degenerates by deletion and accumulation of repetitive sequences, becoming the small heterochromatic Y chromosome that we observe in all mammals.

This evolutionary history explains why the eutherian X and Y chromosomes are so different in size, gene content and function. In humans and other eutherians, the X chromosome is a medium-sized chromosome, bearing more than 1,000 genes with

a wide range of functions. It is highly conserved among eutherians, amounting to about 5% of the haploid genome and bearing essentially the same suite of genes, as predicted by Ohno (1967). The Y chromosome on the other hand is much smaller, heterochromatic, and bears few active protein-coding genes (45, of which 27 are in the male-specific region), which are largely devoted to male-specific functions. Although the X and Y are so different, relics of ancient homology are apparent in the short pseudoautosomal region, in which the X and Y pair and recombine at meiosis (Rappold, 1993; Graves et al., 1998), and presence of the X copies of Y genes (Skaletsky et al., 2003).

Comparative gene mapping in fish, birds, reptiles, monotremes, marsupials and eutherians have shown that sex chromosomes have evolved independently from different autosomes in different vertebrate lineages. Thus sex-determining loci were acquired by different autosomes at different time points during evolution (Ohno, 1967; Veyrunes et al., 2008; Delbridge et al., 2009). The analysis of the marsupial sex chromosomes has been crucial to our understanding of the evolution of mammalian X and Y chromosomes.

8.2 Marsupial Karyotypes and Sex Chromosomes

The marsupial Y chromosome harbours an *SRY* gene (Foster et al., 1992), which is thought to be the male-determining gene, although direct evidence from sex reversal or transgenesis is lacking. Marsupial sex chromosomes are generally smaller than eutherian sex chromosomes because eutherian sex chromosomes are fused with an autosomal region since their divergence from marsupials 148 million years ago (MYA) (Spencer et al., 1991a; Graves, 1995).

The karyotype of marsupials is extraordinarily well conserved, even between American and Australian marsupials that diverged 83 MYA (see Rens and Ferguson-Smith, Chapter 2). Chromosome painting shows that marsupial karyotypes can be divided into 19 blocks of conserved synteny even between the most divergent karyotypes (2n = 10/11 to 2n = 32) (Rens et al., 2003). A common 2n = 14 karyotype was probably ancestral to Australian marsupials and through a series of centric fusion rearrangements, these 19 blocks of conserved synteny make up the diverse karyotype observed in different marsupial species (Rofe and Hayman, 1985; Svartman and Vianna-Morgante, 1998).

Chromosome painting revealed that the euchromatic region of the marsupial X chromosome is completely conserved between marsupial species (Glas et al., 1999a; Rens et al., 2003). Exceptions are species with large heterochromatic additions to the X (e.g. *Macropus parryi* and *Antechinus apicalis*) (Young et al., 1982), or the many species in which there have been translocations or fusions between sex chromosomes and autosomes (Toder et al., 1997a). X-autosome fusions and translocation are observed also in eutherians such as short-tailed bats and several mouse species (Jacobs, 2004; Noronha et al., 2009), but they seem to be particularly widespread in marsupials. For instance the swamp wallaby *Macropus bicolor* has a karyotype of 2n = 10 in females and 2n = 11 in males due to an XY_1Y_2 system generated

by fusion of the ancestral X to an autosome, leaving an unpaired autosome that is homologous to tammar wallaby chromosomes 2 and 7 (Toder et al., 1997a). At diplotene of male meiosis in an XY_1Y_2 system, the fused autosomal region on the X forms chiasmata with the unpaired autosome (Y_2). Y_1 and Y_2 chromosomes segregate towards one pole and the X chromosome towards the opposite pole, ensuring the production of balanced gametes.

8.3 Marsupial X Chromosome Structure and Organization

Early studies of marsupial sex chromosomes focused on their size, morphology, and visible landmarks such as the pairing segment and the nucleolar organizing region. More recently, gene mapping and chromosome painting have been used to investigate organization and homology. The ultimate information on the X chromosome has been gathered from the genome sequence of representative marsupial species, the opossum (*Monodelphis domestica*) representing American marsupials, and the model kangaroo, tammar wallaby (*Macropus eugenii*), representing Australian marsupials. These two species represent the most distantly related clades of the marsupials that last shared a common ancestor ~83 MYA (Bininda-Emonds et al., 2007). Therefore, phylogenetically they are approximately as closely related to each other, as are human and mice. The opossum genome was sequenced to approximately 6× coverage and most large contigs have been assigned to chromosomes (Mikkelsen et al., 2007). The tammar wallaby genome has been sequenced to approximately 2× coverage and has been recently assembled into contigs of ~2,600 bp N50 size, most of which are yet to be assigned to chromosomes.

8.3.1 X Chromosome Size and Gene Content

The marsupial X chromosome is smaller than the X of humans and other eutherians (Hayman et al., 1982). The human X is about 155 Mb (~5% of the haploid genome) and harbours about 1,000 protein-coding genes. In contrast, the marsupial X is about 80 Mb (~3%) and harbours 442 genes. The marsupial Y can be as small as 10 Mb (Toder et al., 2000), and to date has seven genes assigned to it (reviewed Murtagh et al., Chapter 11).

The conserved region of the human X chromosome (XCR), which is equivalent to the marsupial X chromosome, contains 678 protein-coding genes compared to 442 protein-coding genes on the opossum X chromosome (Hubbard et al., 2009). Part of the discrepancy is accounted for by the expansion in primates of the cancer/testis antigen gene family, which contains approximately 100 members in the human XCR region (Ross et al., 2005; Delbridge and Graves, 2007). The human XCR and the opossum X chromosome share at least 322 orthologous genes (Vilella et al., 2009) supporting the hypothesis that most genes on the therian X chromosome are well conserved.

8.3.2 The Marsupial X Chromosome Centromere

Differences in centromere position make the marsupial X quite variable morphologically between species. For instance, the X chromosome is metacentric in some species of opossums (genus *Gracilinanus*), submetacentric in the slender mouse opossum (*Marmosops parvidens*) and the tammar wallaby (*M. eugenii*), and acrocentric in the water opossum (*Chironectes minimus*) (Yunis et al., 1972; Glas et al., 1999b; Carvalho et al., 2002).

The centromere plays a critical role during cell division by participating in kinetechore formation and attachment to the spindle, ensuring regular segregation. Disturbance in centromere function results in improper segregation of the chromosomes during division, which can result in aneuploidy. In eukaryotes, centromeres are made up of arrays of repetitive DNA sequences called satellite DNA (Craig et al., 1999; see also Carone and O'Neill, Chapter 3). In most marsupial species, the X chromosome contains a heterochromatic centromere that stains positive by C-banding (Bulazel et al., 2006). The amounts of heterochromatin differ greatly between species (e.g. the almost identical $2n = 14$) karyotypes of dasyurids are greatly different in heterochromatin content (Young et al., 1982).

The DNA component of marsupial centromeres is now well understood in kangaroos (Bulazel et al., 2006). For instance, the centromeric sequences of all chromosomes in the red-necked wallaby (*Macropus rufogriseus*) are made up of an α-like satellite, Mrb-sat23 (178 bp AT-rich repeat unit). The sex chromosomes harbour two additional repeats Mrb-sat1 (342 bp AT-rich repeat unit) and Mrb-B29 (very short AT-rich tandem repeats). The Mrb-sat1 sequence is specific to the X chromosome, but is not confined to its centromere; it also extends into the short and long arm of the chromosome. Mrb-B29 localizes to the whole Y chromosome, the centromere of the X chromosome and the pericentric region of chromosome 2. The Mrb-sat23 sequence is found on all chromosomes, including the centromere and pericentric region of the X chromosome, and is extensively amplified on the Y chromosome. It contains a CENP-B DNA binding domain (CENP-B box), which may bind the CENP-B protein associated with kinetechore formation. The role of the centromeric X and the Y shared repeats in marsupials needs further investigation.

8.3.3 Absence of Pseudoautosomal Region

During meiosis, homologous chromosomes pair synapse and undergo recombination. In eutherian mammals a small terminal region of homology is shared by the X and Y chromosomes, which pair to form obligatory chiasma during meiosis (Burgoyne, 1982). Segments of the X chromosome that pair with the Y chromosome escape X inactivation in females. Since they behave much like autosomes, they are called pseudoautosomal regions (PARs). There is an obligate crossover within this small region (2.6 Mb in humans), so the recombination rate is extremely high in males, but not in females.

Surprisingly, marsupial sex chromosomes have no homologous pseudoautosomal region, although the X and the Y chromosome may share heterochromatic regions (Toder et al., 1997b; Graves et al., 1998) and there are at least five genes shared between the X and Y chromosomes in tammar wallaby (Delbridge et al., 1997; Pask et al., 2000; Waters et al., 2001). Despite the lack of a PAR, marsupial sex chromosomes are seen to pair end-to-end before pachytene of meiosis, and they segregate correctly (Sharp, 1982). However, no synaptonemal complex is present, and the observation that either end of the chromosome can participate in this pairing (Page et al., 2003) indicates that pairing does not depend on regions of homology between the X and Y chromosome. Instead, the segregation of marsupial X and Y chromosomes during meiosis is achieved by structural and biochemical modifications of axial elements of the sex chromosomes (discussed in detail by Fernández-Donoso et al., Chapter 10).

8.3.4 Nucleolus Organizer Region (NOR)

Some marsupial species have a heterochromatic nucleolar organizer region on the X chromosome, which gives the appearance of a secondary constriction and satellite (Johnston et al., 1984; Dhaliwal et al., 1988; Alsop et al., 2005). NORs contain tandem repeats of ribosomal RNA (rRNA) genes, and they participate in nucleolus formation and ribosome synthesis during cell growth.

The NOR on the X chromosome is very evident in unbanded chromosome preparations of all the kangaroos (Family Macropodidae), and several South American species including the opossum *M. domestica* (Graves, 1967; Johnston et al., 1984; Hayman, 1990; Schmid et al., 2002; Alsop et al., 2005). In other Australian marsupials, such as the dasyurids, the NOR is on a small satellited short arm of chromosome 5. It has been proposed that the X and chromosome 5 of an ancestral kangaroo about 20 MYA underwent Robertsonian translocation, to place the NOR on the X chromosome in the entire family (Toder et al., 1997b).

There appears to be no nucleolus organizer on the kangaroo Y chromosome, except perhaps in the pretty face wallaby, so that females have twice as many rDNA genes as males (Robins et al., 1984). Since ribosome number greatly affects growth, it might be expected that some dosage compensation system operate to equalize the dose of ribosomal genes in males and females. Indeed, the morphology of the NORs, which is indicative of transcriptionally active or inactive NORs, on the two X chromosomes in females is strikingly different (Graves, 1967; Alsop et al., 2005). However, silver staining, which visualises the protein component on metaphase chromosomes shows that the NOR is actively transcribed from both X chromosomes in female kangaroos (Dhaliwal et al., 1988) and in female opossums (Merry et al., 1983; Svartman and Vianna-Morgante, 2003), suggesting that this region escapes X-inactivation and the dose of rRNA is not equal between males and females. In many other species (e.g. frogs, platypus), the NOR-bearing autosome is regularly dimorphic, but both appear to be active, so it is likely that the NOR on the marsupial X is regulated by some separate and very ancient mechanism.

8.3.5 Repeated Sequences and Isochore Structure of the Marsupial X Chromosome

Isochores are defined as stretches of DNA with a homogenous base composition (such as high GC), and examining changes in the conservation of isochores can be used to trace evolutionary changes in chromosomes. Vertebrate genomes are mosaics of isochores, and the base composition of isochores reflects the base composition of the coding sequences present in these isochores. Isochore structures are observed in mammalian, avian and even crocodilian genomes, suggesting the genomes of these groups can be compartmentalized into stretches of homogenous base compositions (Bernardi, 2000; Chojnowski et al., 2007). Isochore structures can be determined by analysing genomic DNA sequences (Costantini et al., 2006).

The human X chromosome contains relatively low GC (39%) compared to the genome average of 41% (Ross et al., 2005). A low GC content is correlated with a low gene density and low substitution rate, which is consistent with the human X being a comparatively gene-poor chromosome. In contrast, the opossum X chromosome contains an unusually high GC content. The GC content of the X chromosome (40.9%) is significantly higher than the average of autosomal DNA (37.7%) (Mikkelsen et al., 2007).

The human X chromosome can be divided into 175 isochores (122 in XCR and 53 in XAR) that belong to five major families *viz.* L1 = ≤36.5% GC, L2 = 36.5–40.5% GC, H1 = 40.5–46.5% GC, H2 = 46.5–53.5% GC, and H3 = >53.5% GC (Fig. 8.1) (Costantini et al., 2006; Costantini et al., 2009). The human XCR contains 47 isochores belonging to H1, H2 and H3 families, and covering 24.9 Mb of the human X chromosome. On the other hand, the opossum X chromosome can be divided into 104 isochores (Costantini et al., 2009). Given that the opossum X chromosome is entirely homologous to the human X conserved region (XCR), it would be expected that the number of isochores on the opossum X chromosome would be the same as for the conserved region of the human X. The difference between the numbers of isochores on the opossum X (104) and the human XCR (122) are inconsistent with the expectation that isochore structures are preserved during evolution. The opossum X chromosome contains 54 isochores belonging to H1, H2 and H3 family covering 33.5 Mb of the opossum X chromosome. This is more than on the human X (47 covering 24.9 Mb), suggesting that the opossum X chromosome contains more high GC content regions than the human XCR.

What does this difference in isochore structure imply? Does this higher GC content imply more CpG islands in the promoter regions of the opossum X chromosome genes? It will be important to study the correlation between higher GC content of the opossum X chromosome with the CpG island density, and subsequently the methylation patterns to examine if the marsupial X inactivation process is controlled by epigenetic factors like DNA methylation of the CpG islands in the promoter region (Hellman and Chess, 2007) (see Al Nadaf et al., Chapter 13 for details on the role of DNA methylation as a mechanism of X inactivation and dosage compensation).

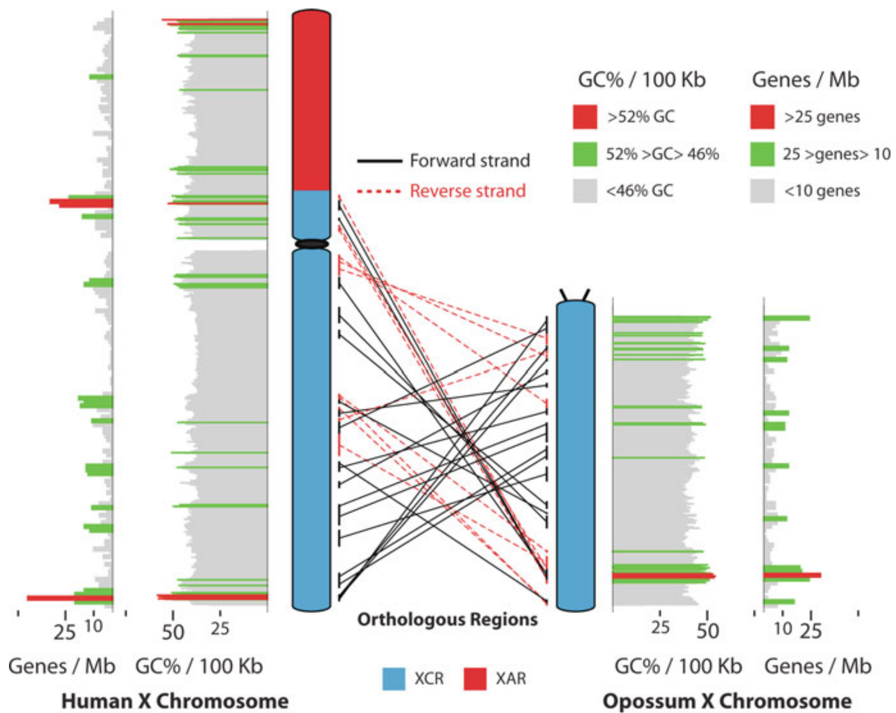


Fig. 8.1 The comparison of the human X chromosome and the opossum X chromosome showing rearrangements of homologous syntenic blocks. The GC content and gene density are displayed along the length of chromosome. A positive correlation can be observed between GC content and gene density

8.3.6 Non-coding Transcripts from the Marsupial X Chromosome

An interesting feature of the marsupial X chromosome is the transcription of large numbers of microRNA molecules. MicroRNAs are small (21–25 bp) non-coding RNAs found in a wide variety of organisms including plants, worms and mammals. In the eukaryotic genome they are involved in post-transcriptional regulation. Some microRNAs are evolutionarily conserved and others have evolved independently in different lineages.

The opossum X chromosome harbours one family of microRNAs that expanded only in marsupials as a result of a series of duplication events (Devor and Samollow, 2008; Devor et al., 2009). Mdo-miR-1544 and Mdo-miR-1545 are two marsupial specific microRNAs that map to a 100 kb cluster of 39 microRNAs on the opossum X chromosome. Some of these X chromosome specific microRNAs do not form the usual hairpin structures, but have evolved to form highly divergent mature sequences. The opossum cluster of microRNAs is the largest microRNA cluster known on any mammalian X chromosome, and could possibly be involved in a unique form of dosage compensation via post-transcriptional regulation of X chromosome transcripts.

8.3.7 Gene Maps and Evolution of Marsupial Sex Chromosomes

Chromosome painting and gene mapping have been used to investigate homology between X chromosomes of different species. Physical localization of DNA fragments containing known genes has been accomplished by fluorescent in situ hybridisation (FISH), in which large insert bacterial artificial chromosomes (BAC) clones are tagged with a fluorochrome and hybridized to homologous sequences in chromosomes immobilized on a microscope slide. This technique has been used to map 73 genes onto the tammar wallaby X chromosome and compare gene content and order with the chromosomes of opossum, other mammals and other vertebrates (Deakin et al., 2008).

8.3.7.1 Identification of Evolutionary Layers on the Eutherian X Chromosome

Comparisons between the gene content of the human and marsupial X provided the first clues that the human X is composed of two distinct evolutionary layers. Marsupials (metatherian mammals) diverged from eutherian mammals 148 MYA (Bininda-Emonds et al., 2007). Therian mammals (marsupials and eutherians) shared a common ancestor with monotreme mammals 168 MYA. This deep divergence allowed an examination of the origin of the mammalian X chromosome.

The eutherian X chromosome is exceptionally conserved in gene content, as originally proposed by "Ohno's Law". Comparative gene mapping confirmed that the eutherian X chromosome has a virtually identical gene content in species as diverse as human, mouse, horse, cat, cow and elephant (Murphy et al., 1999; Band et al., 2000; Swinburne et al., 2006; Rodriguez Delgado et al., 2009).

Early comparative gene mapping showed that ten genes from the long arm of the human X chromosome (Xq) were also located on the X chromosome in different species of marsupials, consistent with Ohno's Law (Spencer et al., 1991b). The conservation of the therian X chromosome was also illustrated by chromosome painting using a fluorochrome labelled tammar wallaby X chromosome probe that hybridized to the long arm and the pericentric region of the short arm of the human X chromosome (Glas et al., 1999a). These gene mapping and chromosome painting studies have shown that the long arm of the human X chromosome is conserved on the X chromosome of all therian mammals (marsupials and eutherians). The region homologous to the marsupial X chromosome in humans is thus referred to as the X conserved region (XCR) (Graves, 1995).

However, genes from the short arm of the human X chromosome (Xp) were shown to be autosomal in marsupials (Spencer et al., 1991a) suggesting that either the X chromosome of eutherian mammals gained an autosomal region, or marsupials lost part of the X chromosome after the marsupials and eutherian mammals split 148 MYA (Graves, 1995).

Further insight into the origins of the therian X were gained by comparative gene mapping in platypus, one of the three extant species of monotreme mammal, whose genome was recently sequenced (Warren et al., 2008). These studies confirmed that the XCR region of the eutherian X chromosome is entirely conserved on the

platypus chromosome 6; similar to the X chromosome of marsupials (Waters et al., 2005; Veyrunes et al., 2008). However, the short arm of the eutherian X chromosome is conserved on two autosomes in platypus, chromosomes 15 and 18 (Veyrunes et al., 2008), and autosomes in marsupials (Spencer et al., 1991a; Mikkelsen et al., 2007; Deakin et al., 2008). The autosomal location of the genes from the short arm of the eutherian X chromosome revealed that eutherian X chromosome gained an autosomal region since their divergence from marsupials 148 MYA. This region was termed as X added region (XAR) (Graves, 1995).

The presence of human X chromosome genes on two chicken chromosomes (XCR on chicken 4p and XAR on chicken 1q) favoured the hypothesis that the region of the human X chromosome homologous to an autosome in marsupials was added in the eutherian lineage. The boundary of the autosomal addition to the eutherian X chromosome was determined to be within the cytogenetic band Xp11.23 on the human X chromosome by gene mapping and chromosome painting (Wilcox et al., 1996; Glas et al., 1999a). More recently, the complete sequence of the entire genomes of chicken, opossum and human were made available and the precise boundary between XAR and XCR was refined and shown to lie between the *RGN* and *NDUFB11* genes at 46.83–46.88 Mb from the telomere of the short arm of the human X chromosome by cross species comparisons (Ross et al., 2005; Mikkelsen et al., 2007).

The homology of the therian XCR with an autosome in platypus, as well as birds, implies that therian mammal sex chromosomes evolved quite recently from an autosome, represented by platypus chromosome 6. Platypus sex chromosomes proved to have homology with the bird Z chromosome, suggesting that their multiple XY complex system is a novel version of an ancient reptilian ZW system (Graves, 2008; Veyrunes et al., 2008).

Detailed comparisons of sequence on the human X chromosome with the chicken genome suggested that the human X chromosome was made up of not two (the XAR and the XCR), but three distinct evolutionary blocks (Kohn et al., 2004; Ross et al., 2005). This third block of the human X chromosome was proposed to consist of genes from the pericentric region in the cytogenetic band Xp11 and a gene-dense region in the cytogenetic band Xq28. Genes from this putative third block on the eutherian X chromosome all lay on the marsupial X (so were part of the XCR), but shared homology with the chicken chromosome 12, rather than with chicken 4p (Kohn et al., 2004). Orthologues of human Xp11 and Xq28 genes also mapped to platypus chromosome 6q along with other XCR genes (Fig. 8.2) (Delbridge et al., 2009), suggesting that the putative third block was added before monotremes diverged from therians. However, comparisons of the homologues of the putative third evolutionary block between the human, opossum, rat and chicken genomes revealed that the chicken genes were paralogues that were wrongly assigned as a third evolutionary layer. The true orthologues were missing from the chicken genome assembly. This refutes the hypothesis that the human X chromosome is composed of three evolutionary blocks, and confirms the original hypothesis that two ancient genome regions, now represented by XAR and XCR, became fused early in the eutherian lineage (Graves, 1995).

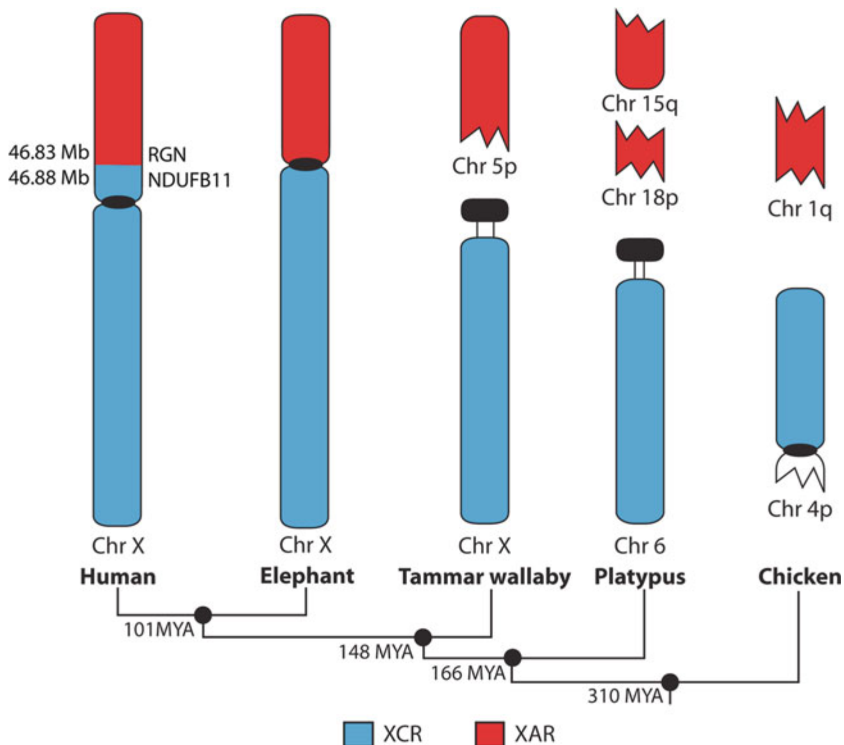


Fig. 8.2 Comparative analysis of therian mammal X chromosome shows the two distinctive blocks of shared homology in human, tammar wallaby, platypus and chicken; the X-conserved region (XCR = *light grey* for *black* and *white* and *blue* for color image) and the X-added region (XAR = *dark grey* for *black* and *white* and *red* for color image). XCR is conserved on the X chromosome in all eutherians and marsupials. XAR is autosomal in marsupials and was added to the eutherian X chromosome after the marsupial eutherian divergence, but before the eutherian radiation. The therian mammal X chromosome shares no homology with platypus and chicken sex chromosomes, indicating independent origin of the therian X chromosome 148–166 million years ago. Comparative analysis of the human X chromosome with marsupial and chicken genomes places the boundary between XCR in the cytogenetic band Xp11 between RGN and NDUFB11 (46.83–46.88 Mb)

Recent gene mapping experiments (Rodriguez Delgado et al., 2009) shows that the elephant X chromosome is collinear with the human X, except that the centromere position is different, implying that a centric shift occurred in one or other lineage after afrotherians diverged from other eutherians. The position of the elephant centromere corresponds to human Xp11.23, the fusion point of XCR and XAR, suggesting that the original fusion was a Robertsonian fusion between an ancient, marsupial-like X and an autosome.

8.3.7.2 Gene Arrangement on the Marsupial X Chromosome

Eutherian mammals not only share similar gene complement on the X chromosome, but even the gene order appears to be conserved in the X chromosome of eutherians;

for instance ten marker genes lie in the same order along the X chromosome in elephant and human, representing the most divergent eutherians (Rodriguez Delgado et al., 2009). This conservation may be the result of selection against rearrangements that could disrupt the chromosome-wide X chromosome inactivation system in eutherians,

However, comparisons between human and marsupial X chromosomes show that, although the gene content is conserved, there have been at least 15 internal rearrangements in the opossum lineage (Mikkelsen et al., 2007). With the construction of a dense physical map of the tammar X chromosme (Deakin et al., 2008), it now becomes apparent that there are many rearrangements between marsupial and eutherian X chromosomes, as well as between the X chromosomes of different marsupial species.

One major rearrangements between the tammar wallaby X chromosome and the opossum X chromosome involves the genes *PSMD10*, *TBC1D8B* and *IL1RAPL2*, which are close to each other on the tammar wallaby X chromosome but localize to three distinct locations on the opossum X chromosome (Deakin et al., 2008). The arrangement of these three genes in the tammar wallaby is similar to the arrangement on the human X chromosome, indicating that the tammar wallaby retains the ancestral arrangement and the opossum has a derived arrangement.

In other instances, the tammar wallaby X chromosome, as well as the opossum X, is rearranged compared to the human X chromosome. Famously, a search for the *XIST* gene that controls X chromosome inactivation in eutherians, revealed that flanking markers were some distance apart in both marsupial species (Hore et al., 2007). A group of genes that includes *STARD8*, *TEX11*, *OGT*, *XIST*, *SLC16A2*, *RLIM* and *PGK1*, are present in that order on the human X chromosome from 67.8 to 77.3 Mb in the cytogenetic band Xq13 (Fig. 8.3). Their marsupial orthologues are

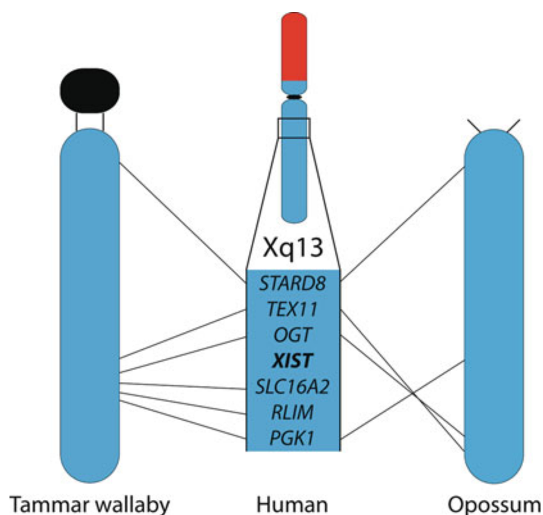


Fig. 8.3 Comparative genetic maps of six genes flanking *XIST* from human Xq13. The *XIST* gene is not present in marsupials but evolved in the eutherian lineage. The disruption of relative gene order between marsupial species and between marsupials and human implies that there is an alternative X-inactivation mechanism in marsupials

discontiguous on both the tammar wallaby and the opossum X chromosome (Deakin et al., 2008). *STARD8* maps close to the centromere on the tammar wallaby and opossum X chromosomes (7.3 Mb). *TEX11*, *OGT*, *SLC16A2*, *RLIM* (*RNF12*) and *PGK1* genes are located more distally on the Xq in both marsupial species; *OGT* and *TEX11* are located at ~70 Mb and *PGK1* is located at 56.3 Mb on the opossum X chromosome, whereas all three genes (*OGT*, *TEX11* and *PGK1*) are located together, distal to the *STARD8* on the tammar wallaby X chromosome. It is evident from these comparative mapping studies that marsupial X chromosomes have undergone several rearrangements since their divergence from eutherian mammals, and that many rearrangements are specific for different lineages. In comparison to the conservatism of the gene order on the X of eutherian mammals, this propensity for large-scale lineage-specific rearrangements on marsupial X chromosomes might relate to the different X inactivation systems discussed later in this chapter.

8.3.8 Mutation Rates on the Marsupial X Chromosome

The rate of evolution is different for the X chromosome, the Y chromosome and the autosomes (Charlesworth et al., 1987; Rice, 1987a; Rice, 1987b) because of the different proportions of their cycles that are spent in the hostile environment of the testis, where chromosomes go through many more divisions, and DNA repair is less effective. The Y is in the testis at every generation, by definition, whereas the X are in a testis only two-thirds as frequently as the autosomes. This produces a bias in mutation rate of Y>A>X; thus the X would be expected to change more slowly.

However, selection acts differently for genes on the X chromosome, autosomes and the Y chromosome. The effects of new rare recessive autosomal mutations in heterozygotes are masked by the presence of the dominant allele in randomly mating populations. In contrast, recessive mutations arising on the X chromosome are hemizygous in males and therefore natural selection would rapidly fix such recessive mutations if they are beneficial to males, or remove them from the population if deleterious to males. As long as they are rare, there will be no homozygous females for such rare mutations, so selection against a female phenotype is irrelevant. Thus, male advantage mutations are more rapidly fixed on the X chromosome than on the autosomes.

This is called the fast-X theory whereby genes on the X chromosome are biased towards rare male beneficial mutations. This fast-X theory can be assessed by comparing the rates of nucleotide substitutions at the codon level. The ratio ($\omega = dn/ds$) of non-synonymous substitutions (nucleotide substitutions in the codon triplet that change the amino acid, *dn*) to synonymous substitutions (nucleotide substitutions in the codon triplet that does not change the amino acid, *ds*) is used to calculate the rate of mutation at a locus, across a gene set or an entire chromosome. The genes on the X chromosome should exhibit higher ω than the autosomal genes in the population when averaged over the entire genome for the fast-X theory to be correct. The fast-X theory has been tested in many species with contradictory evidence (Torgerson

and Singh, 2003; Thornton et al., 2006; Mank et al., 2007). Specifically, mutation rates in *Drosophila spp.* do not support the hypothesis that beneficial mutations go to fixation faster on the X chromosome than on other chromosomes, whereas there is evidence that beneficial mutations are fixed faster on the X chromosome of mammals and the Z chromosome of birds (Torgerson and Singh, 2003; Thornton et al., 2006; Mank et al., 2007).

The fast-X theory has not been tested in the marsupial lineage. However, it is noted that genes on the marsupial X chromosome exhibit decreased ω compared to autosomal genes, owing to increased synonymous substitution rate (Goodstadt et al., 2007). This decreased ω is also observed in subtelomeric regions of the marsupial autosomes suggesting that fast-X theory does not apply to the marsupial X chromosome or substitution rate ratio (ω) measure is not a reliable predictor of the fast-X theory.

8.3.9 Marsupial X Chromosome Sequence and X Inactivation

In eutherian mammals with XX females and XY males, one of the X chromosomes is transcriptionally silenced to maintain an equal dose of genes in males and females in somatic cells. This silencing phenomenon is called X chromosome inactivation and balancing of dose is called dosage compensation (see Al Nadaf et al., Chapter 13 for detailed discussion about the X inactivation). In human and mouse X inactivation is random; either the paternally derived or maternally derived X chromosome is silenced early in embryogenesis, and the clonal replicates of the cells are propagated throughout development to produce a mosaic phenotype (Wutz and Gribnau, 2007). However, in marsupials, X inactivation occurs, but is paternally imprinted in the embryo such that the paternally derived X is always the inactive one (VandeBerg et al., 1983). Additionally, marsupial X inactivation is incomplete, tissue-specific, and seems to occur by a stochastic mechanism. The genetic and molecular basis of the X inactivation in marsupials will be discussed in following sections.

8.3.10 Somatic X Inactivation in Marsupials

X inactivation in human and mouse is a whole X phenomenon regulated by the *XIST* gene at the X inactivation centre (XIC), and suggested to be propagated by LINE elements that act as “booster stations”. The concentration of LINE1 elements on the opossum X was found to be 22%, which is similar to the XAR of the human X (Mikkelsen et al., 2007). In contrast, the concentration of LINE1 elements in the human XCR, the region homologous to the opossum X chromosome, is significantly higher (33%). Similarly, dog X (34%), rat X (36%), and mouse X (32%) also have higher concentration of the LINE1 elements. This suggest that there has been an accumulation of LINE1 elements on the eutherian X chromosome to propagate and

maintain X inactivation. Lack of this accumulation of LINE1 elements, absence of XIST gene (Hore et al., 2007) and rearrangement of XIST locus (Deakin et al., 2008) in marsupials suggests alternative mechanisms of the X chromosome inactivation.

The absence of the XIST and consequent spreading of the X inactivation locus in marsupials might also explain why the marsupial X chromosome is much more rearranged in different species than is the eutherian X chromosome. It might also explain why compound X chromosomes occur in many marsupial families. Evidence from replication studies shows that the condensation and late replication of a compound X is confined to the original X (Hayman and Sharp, 1981) and thus only the original X might be subject to X inactivation.

The X inactivation in mouse and human involves DNA methylation and several histone modifications at a molecular level. In marsupials, the X inactivation is accompanied by loss of active histone modifications, but apparently not the acquisition of inactive histone marks (Koina et al., 2009), and does not involve DNA methylation. This molecular evidence also suggests alternative mechanisms of the X inactivation in marsupials.

Marsupial X inactivation provides some clues to how the complex inactivation system evolved in mammals, because in mouse and cattle (but curiously, not human) X inactivation is paternal and incomplete in the extraembryonic membranes, does not involve DNA methylation and does not require *XIST* (reviewed in Huynh and Lee, 2001). Paternal X inactivation may therefore represent the original mechanism by which one X was silenced in females. Sequence analysis of the opossum X chromosome produced no evidence that LINEs have accumulated (Mikkelsen et al., 2007), suggesting that LINE elements are not involved in X inactivation in marsupials, but accumulated once *XIST* had evolved a function to coordinate and control spreading of inactivation. However, the expansion of a microRNA family on the marsupial X chromosome suggests that epigenetic factors or post-transcriptional silencing might be involved in X inactivation in marsupials.

Curiously, given the rather leaky inactivation in marsupials, some species have gone to the extreme of eliminating the inactive X from the somatic cells (Hayman and Martin, 1965). Germline cells of the Southern brown bandicoot (*Isoodon obesulus*), the northern brown bandicoot (*I. macrourus*) and the long-nosed bandicoot (*Perameles nasuta*) carry both a full complement of sex chromosomes (XX or XY). However, in somatic cells, males lose their Y chromosome and females one of their X chromosomes so that both are XO. The Y chromosome is eliminated from a proportion of cells in the male greater gliders as well (*Petauroides volans*) (Murray and McKay, 1979). Though not unique to marsupials, such elimination systems are rare in eutherians, and the creeping vole (*Microtus oregoni*), is the only eutherian mammal in which tissue-specific chromosomal loss is observed (Ohno et al., 1963). The loss of the Y in males and the second X in females of these species implies that Y-specific genes are unnecessary for somatic development, and that gene dosage is either compensated in females, or is not critical for development.

8.3.11 Meiotic Sex Chromosome Inactivation

The eutherian X chromosome also undergoes inactivation during male meiosis in a process called meiotic sex chromosome inactivation (MSCI). Both the X and Y chromosomes become transcriptionally inactivated in primary spermatocytes during MSCI (Richler et al., 1992), and this occurs even in the absence of *XIST* (Turner et al., 2002). MSCI occurs in marsupials, as well as eutherians.

An important consequence of MSCI is that inactivation of X and Y chromosome genes that are essential during spermatogenesis are copied or retrotransposed to autosomes. Two recent reports on the relatively high retrotransposition rate of genes from the X chromosome to autosomes (retrogenes), compared to autosome to autosome retrotransposition, suggests that metatherians and eutherians may compensate for the transcriptional silencing of essential genes on the X and Y chromosomes during meiosis by having the "daughter genes" as retrogenes expressed from the autosomes (Emerson et al., 2004; Potrzebowski et al., 2008). Copying of essential genes onto autosomes has occurred in the marsupial lineage; for instance, marsupials, as well as eutherians have a testis-specific autosomal copy of *PGK1* (Vandeberg et al., 1980; Mccarrey et al., 1992). It is thus observed that the X chromosome inactivation has shaped the entire genome whereby testis-specific essential genes have moved to autosomes and absence of *XIST* locus controlling eutherian mammal X inactivation has allowed for more rearrangements on the marsupial X chromosome compared to the eutherian X chromosome.

The similarities with paternal X inactivation in somatic cells has led to the proposal that MSCI was the system from which X inactivation evolved. It has been suggested to have evolved concurrently with the degradation of the Y chromosomes, as it would be a disadvantage for sperm to differ in the presence of genes important in meiosis (for details see Delbridge, Chapter 9). This system then was expected to produce X inactivation in the embryo and extraembryonic tissues.

We suggest, therefore, that X inactivation in a therian ancestor was originally paternal and incomplete, and was accomplished on a gene-by-gene basis by a mechanism that included loss of active histone modifications, but not accumulation of inactive marks, or DNA methylation. With the evolution of *XIST* early in the eutherian lineage, coordinate control by spreading from the *XIST* locus was evolved in the embryo, but not the embryonic membranes, which retained the original mechanism. Comparison with the region orthologous to the *XIST* region in chicken and frog suggested that the *XIST* gene evolved by pseudogenization of a protein coding gene after eutherians diverged from the common therian ancestor 148 MYA (Duret et al., 2006), although the minimal sequence overlap suggests it may have contributed little more than a transcription start site (Hore et al., 2007). Since X inactivation in marsupials seems not to be a whole X phenomenon, LINE booster elements are suggested to have been added after the evolution of *XIST* to enable X inactivation to spread (Mikkelsen et al., 2007). Molecular mechanisms such as DNA methylation and accumulation of inactive histone marks were also added to render inactivation more complete and stable in the eutherian lineage.

8.4 Conclusion

Studies of the marsupial X chromosome have been particularly revealing of the process by which an autosomal pair took on the functions of an XY pair of sex chromosomes; in particular sex determination and dosage compensation. Homology of the marsupial X to part of the eutherian X establishes a conserved therian X chromosome, which is present as an autosome in the genomes of birds and even monotreme mammals. Acquisition of the testis-determining gene *SRY* by the proto-Y chromosome occurred in a therian ancestor between 166 and 148 MYA, and was followed by degradation of the Y and the evolution of X chromosome inactivation in females. Marsupials, having diverged from eutherians 148 MYA, retain some characteristics of the original therian X; its smaller size and gene content, its propensity to rearrange internally and fuse with autosomes, its absence of LINE elements, and an ancestral form of paternal and partial X inactivation. Features of the X that evolved specifically in marsupials include high GC and microRNA contents. Features that evolved specifically on the eutherian X include Robertsonian fusion with an autosome, centromeric shift, acquisition of *XIST* and accumulation of mechanisms that ensured spreading of inactivation in *cis*.

References

Alsop A, Miethke P, Rofe R, et al. (2005) Characterizing the chromosomes of the Australian model marsupial *Macropus eugenii* (tammar wallaby). Chromosome Res 13:627–636.

Band MR, Larson JH, Rebeiz M, et al. (2000) An ordered comparative map of the cattle and human genomes. Genome Res 10:1359–1368.

Bernardi G (2000) Isochores and the evolutionary genomics of vertebrates. Gene 241:3–17.

Bininda-Emonds ORP, Cardillo M, Jones KE, et al. (2007) The delayed rise of present-day mammals. Nature 446:507–512.

Bulazel K, Metcalfe C, Ferreri GC, Yu J, Eldridge MD, O'Neill RJ (2006) Cytogenetic and molecular evaluation of centromere-associated DNA sequences from a marsupial (Macropodidae: *Macropus rufogriseus*) X chromosome. Genetics 172:1129–1137.

Burgoyne PS (1982) Genetic homology and crossing over in the X and Y chromosomes of mammals. Hum Genet 61:85.

Carvalho BD, Oliveira LFB, Nunes AP, Mattevi MS (2002) Karyotypes of nineteen marsupial species from Brazil. J Mammal 83:58–70.

Charlesworth B (1978) Model for evolution of Y chromosomes and dosage compensation. Proc Natl Acad Sci USA 75:5618–5622.

Charlesworth B, Coyne JA, Barton NH (1987) The relative rates of evolution of sex-chromosomes and autosomes. Am Nat 130:113–146.

Chojnowski JL, Franklin J, Katsu Y, et al. (2007) Patterns of vertebrate isochore evolution revealed by comparison of expressed mammalian, avian, and crocodilian genes. J Mol Evol 65:259–266.

Costantini M, Clay O, Auletta F, Bernardi G (2006) An isochore map of human chromosomes. Genome Res 16:536–541.

Costantini M, Cammarano R, Bernardi G (2009) The evolution of isochore patterns in vertebrate genomes. BMC Genomics 10:538.

Craig JM, Earnshaw WC, Vagnarelli P (1999) Mammalian centromeres: DNA sequence, protein composition, and role in cell cycle progression. Exp Cell Res 246:249–262.

Deakin JE, Koina E, Waters PD, et al. (2008) Physical map of two tammar wallaby chromosomes: a strategy for mapping in non-model mammals. Chromosome Res 16:1159–1175.

Delbridge ML, Graves JA (2007) Origin and evolution of spermatogenesis genes on the human sex chromosomes. Soc Reprod Fertil Suppl 65:1–17.
Delbridge ML, Harry JL, Toder R, et al. (1997) A human candidate spermatogenesis gene, RBM1, is conserved and amplified on the marsupial Y chromosome. Nat Genet 15:131–136.
Delbridge ML, Patel HR, Waters PD, McMillan DA, Marshall Graves JA (2009) Does the human X contain a third evolutionary block? Origin of genes on human Xp11 and Xq28. Genome Res 19:1350–1360.
Devor EJ, Samollow PB (2008) In vitro and in silico annotation of conserved and nonconserved microRNAs in the genome of the marsupial *Monodelphis domestica*. J Hered 99:66–72.
Devor EJ, Peek AS, Lanier W, Samollow PB (2009) Marsupial-specific microRNAs evolved from marsupial-specific transposable elements. Gene 448:187–191.
Dhaliwal MK, Pathak S, Shirley LR, Flanagan JP (1988) Ag-nor Staining in the Bennett Wallaby, *Macropus rufogriseus* – Evidence for Dosage Compensation. Cytobios 56:29–38.
Duret L, Chureau C, Samain S, Weissenbach J, Avner P (2006) The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. Science 312:1653–1655.
Emerson JJ, Kaessmann H, Betran E, Long M (2004) Extensive gene traffic on the mammalian X chromosome. Science 303:537–540.
Felsenstein J (1974) The evolutionary advantage of recombination. Genetics 78:737–756.
Foster JW, Brennan FE, Hampikian GK, et al. (1992) Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. Nature 359:531–533.
Glas R, Graves JAM, Toder R, Ferguson-Smith M, O'Brien PC (1999a). Cross-species chromosome painting between human and marsupial directly demonstrates the ancient region of the mammalian X. Mamm Genome 10:1115–1116.
Glas R, Leo A, Delbridge M, et al. (1999b). Chromosome painting in marsupials: genome conservation in the Kangaroo family. Chromosome Res 7:167–176.
Goodstadt L, Heger A, Webber C, Ponting CP (2007) An analysis of the gene complement of a marsupial, *Monodelphis domestica*: evolution of lineage-specific genes and giant chromosomes. Genome Res 17:969–981.
Graves JAM (1967) DNA synthesis in chromosomes of cultured leucocytes from two marsupial species. Exp Cell Res 46:37–57.
Graves JAM (1995) The origin and function of the mammalian Y chromosome and Y-borne genes – an evolving understanding. Bioessays 17:311–320.
Graves JAM (2006) Sex chromosome specialization and degeneration in mammals. Cell 124:901.
Graves JAM (2008) Weird animal genomes and the evolution of vertebrate sex and sex chromosomes. Annu Rev Genet 42:565–586.
Graves JAM, Wakefield MJ, Toder R (1998) The origin and evolution of the pseudoautosomal regions of human sex chromosomes. Hum Mol Genet 7:1991–1996.
Hayman DL (1990) Marsupial cytogenetics. Aust J Zool 37:331–349.
Hayman DL, Martin PG (1965) Sex chromosome mosaicism in the marsupial genera *Isoodon* and *Perameles*. Genetics 52:1201–1206.
Hayman D, Sharp P (1981) Verification of the structure of the complex sex chromosome system in *Lagorchestes conspicillatus* Gould (Marsupialia: Mammalia). Chromosoma 83:263–274.
Hayman DL, Ashworth LK, Carrano AV (1982) The relative DNA contents of the eutherian and marsupial X-chromosomes. Cytogenet Cell Genet 34:265–270.
Hellman A, Chess A (2007) Gene body-specific methylation on the active X chromosome. Science 315:1141–1143.
Hore TA, Koina E, Wakefield MJ, Graves JAM (2007) The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. Chromosome Res 15:147–161.
Hubbard TJ, Aken BL, Ayling S, et al. (2009) Ensembl 2009. Nucleic Acids Res 37:D690–D697.
Huynh KD, Lee JT (2001) Imprinted X inactivation in eutherians: a model of gametic execution and zygotic relaxation. Curr Opin Cell Biol 13:690–697.

Jacobs D (2004) The evolution of a neo-XY1Y2 sex chromosome system by autosome–sex chromosome fusion in *Dundocoris nodulicarinus* Jacobs (Heteroptera: Aradidae: Carventinae). Chromosome Res 12:175–191.

Johnston PG, Davey RJ, Seebeck JH (1984) Chromosome homologies in *Potorous-Tridactylus* and *Potorous-Longipes* (Marsupialia, Macropodidae) based on G-Banding patterns. Aust J Zool 32:319–324.

Kohn M, Kehrer-Sawatzki H, Vogel W, Graves JAM, Hameister H (2004) Wide genome comparisons reveal the origins of the human X chromosome. Trends Genet 20:598.

Koina E, Chaumeil J, Greaves IK, Tremethick DJ, Graves JAM (2009) Specific patterns of histone marks accompany X chromosome inactivation in a marsupial. Chromosome Res 17:115–126.

Mank JE, Axelsson E, Ellegren H (2007) Fast-X on the Z: rapid evolution of sex-linked genes in birds. Genome Res 17:618–624.

Mccarrey JR, Berg WM, Paragioudakis SJ, et al. (1992) Differential transcription of Pgk genes during spermatogenesis in the mouse. Dev Biol 154:160–168.

Merry DE, Pathak S, VandeBerg JL (1983) Differential NOR activities in somatic and germ cells of *Monodelphis domestica* (Marsupialia, Mammalia). Cytogenetic Genome Res 35: 244–251.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Muller HJ (1918) Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethal factors. Genetics 3:422–499.

Murphy WJ, Sun S, Chen Z-Q, Pecon-Slattery J, O'Brien SJ (1999) Extensive conservation of sex chromosome organization between cat and human revealed by parallel radiation hybrid mapping. Genome Res 9:1223–1230.

Murray JD, McKay GM (1979) Y Chromosome mosaicism in pouch young of the marsupial, greater glider (Marsupialia: Petauridae). Chromosoma 72:329–334.

Noronha RCR, Nagamachi CY, O'Brien PCM, Ferguson-Smith MA, Pieczarka JC (2009) Neo-XY body: an analysis of XY1Y 2 meiotic behavior in Carollia (Chiroptera, Phyllostomidae) by chromosome painting. Cytogenetic Genome Res 124:37–43.

Ohno S (1967) Sex Chromosome and Sex-Linked Genes. Springer-Verlag, Berlin.

Ohno S, Jainchill J, Stenius C (1963) The Creeping Vole (*Microtus Oregoni*) as a gonosomic mosaic. I. the Oy/Xy constitution of the male. Cytogenetics 14:232–239.

Page J, Berrios S, Rufas JS, et al. (2003) The pairing of X and Y chromosomes during meiotic prophase in the marsupial species Thylamys elegans is maintained by a dense plate developed from their axial elements. J Cell Sci 116:551–560.

Pask A, Renfree MB, Graves JAM (2000) The human sex-reversing ATRX gene has a homologue on the marsupial Y chromosome, ATRY: implications for the evolution of mammalian sex determination. Proc Natl Acad Sci USA 97:13198–13202.

Potrzebowski L, Vinckenbosch N, Marques AC, Chalmel F, Jegou B, Kaessmann H (2008) Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. PLoS Biol 6:e80.

Rappold GA (1993) The pseudoautosomal regions of the human sex chromosomes. Hum Genet 92:315–324.

Rens W, O'Brien PCM, Fairclough H, Harman L, Graves JAM, Ferguson-Smith MA (2003) Reversal and convergence in marsupial chromosome evolution. Cytogenet Genome Res 102:282–290.

Rice WR (1987a). Genetic hitchhiking and the evolution of reduced activity of the Y sex chromosome. Genetics 116:161–167.

Rice WR (1987b). The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. Evolution 41:911–914.

Richler C, Soreq H, Wahrman J (1992) X-inactivation in mammalian testis is correlated with inactive X-specific transcription. Nat Genet 2:192–195.

Robins AJ, Hayman DL, Wells JRE (1984) Ribosomal gene reiteration in a marsupial species with an X-linked nucleolar organizer. Aust J Biol Sci 37:211–215.

Rodriguez Delgado CL, Waters PD, Gilbert C, Robinson TJ, Graves JAM (2009) Physical mapping of the elephant X chromosome: conservation of gene order over 105 million years. Chromosome Res 17:917–926.

Rofe R, Hayman D (1985) G-banding evidence for a conserved complement in the Marsupialia. Cytogenet Cell Genet 39:40–50.

Ross MT, Grafham DV, Coffey AJ, et al. (2005) The DNA sequence of the human X chromosome. Nature 434:325.

Schmid M, Feichtinger W, Steinlein C, Haaf T, Visbal Garcia R, Fernandez Badillo A (2002) X chromosomes of American marsupials contain minimal amounts of euchromatin. Cytogenet Genome Res 99:315–322.

Sharp P (1982) Sex chromosome pairing during male meiosis in marsupials. Chromosoma 86: 27–47.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837.

Spencer JA, Sinclair AH, Watson JM, Graves JAM (1991a). Genes on the short arm of the human X chromosome are not shared with the marsupial X. Genomics 11:339–345.

Spencer JA, Watson JM, Graves JAM (1991b). The X chromosome of marsupials shares a highly conserved region with eutherians. Genomics 9:598–604.

Svartman M, Vianna-Morgante AM (1998) Karyotype evolution of marsupials: from higher to lower diploid numbers. Cytogenet Cell Genet 82:263–266.

Svartman M, Vianna-Morgante AM (2003) Conservation of chromosomal location of nucleolus organizer in American marsupials (Didelphidae). Genetica 118:11–16.

Swinburne JE, Boursnell M, Hill G, et al. (2006) Single linkage group per chromosome genetic linkage map for the horse, based on two three-generation, full-sibling, crossbred horse reference families. Genomics 87:1–29.

Thornton K, Bachtrog D, Andolfatto P (2006) X chromosomes and autosomes evolve at similar rates in Drosophila: no evidence for faster-X protein evolution. Genome Res 16:498–504.

Toder R, O'Neill R, Wienberg J, O'Brien P, Voullaire L, Graves JAM (1997a). Comparative chromosome painting between two marsupials: origins of an XX/XY1Y2 sex chromosome system. Mamm Genome 8:418–422.

Toder R, Wienberg J, Voullaire L, O'Brien PCM, Maccarone P, Graves JAM (1997b). Shared DNA sequences between the X and Y chromosomes in the tammar wallaby – evidence for independent additions to eutherian and marsupial sex chromosomes. Chromosoma 106: 94–98.

Toder R, Wakefield MJ, Graves JAM (2000) The minimal mammalian Y chromosome – the marsupial Y as a model system. Cytogenet Cell Genet 91:285–292.

Torgerson DG, Singh RS (2003) Sex-linked mammalian sperm proteins evolve faster than autosomal ones. Mol Biol Evol 20:1705–1709.

Turner JM, Mahadevaiah SK, Elliott DJ, et al. (2002) Meiotic sex chromosome inactivation in male mice with targeted disruptions of Xist. J Cell Sci 115:4097–4105.

Vandeberg JL, Cooper DW, Sharman GB, Poole WE (1980) Somatic expression and autosomal inheritance of phosphoglycerate kinase-B in kangaroos. Genetics 95:413–424.

VandeBerg JL, Johnston PG, Cooper DW, Robinson ES (1983) X-chromosome inactivation and evolution in marsupials and other mammals. Isozymes Curr Top Biol Med Res 9:201–218.

Veyrunes F, Waters PD, Miethke P, et al. (2008) Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res 18:965–973.

Vilella AJ, Severin J, Ureta-Vidal A, Heng L, Durbin R, Birney E (2009) EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. Genome Res 19:327–335.

Warren WC, Hillier LW, Graves JAM, et al. (2008) Genome analysis of the platypus reveals unique signatures of evolution. Nature 453:175–183.

Waters PD, Duffy B, Frost CJ, Delbridge ML, Graves JAM (2001) The human Y chromosome derives largely from a single autosomal region added to the sex chromosomes 80–130 million years ago. Cytogenet Cell Genet 92:74–79.

Waters PD, Delbridge ML, Deakin JE, et al. (2005) Autosomal location of genes from the conserved mammalian X in the platypus (Ornithorhynchus anatinus): implications for mammalian sex chromosome evolution. Chromosome Res 13:401–410.

Wilcox SA, Watson JM, Spencer JA, Graves JAM (1996) Comparative mapping identifies the fusion point of an ancient mammalian X-autosomal rearrangement. Genomics 35:66–70.

Wutz A, Gribnau J (2007) X inactivation Xplained. Curr Opin Genet Dev 17:387–393.

Young GJ, Graves JAM, Barbieri I, Woolley PA, Cooper DW, Westerman M (1982) The chromosomes of dasyurids (Marsupialia). In: Archer M (ed) Carnivorous Marsupials. Royal Zoological Society of New South Wales, Sydney.

Yunis E, Ramirez E, Cayon J, Hernandez J (1972) The chromosomes of the Didelphids Caluromys Lanatus Illiger and Chironectes minimus Zimmerman (Marsupialia: Didelphidae). Aust J Zool 20:265–269.

Chapter 9
Gene Content of the Mammalian X Chromosome

Margaret L. Delbridge

Abstract The human and mouse X chromosomes have an atypical gene content. These X chromosomes contain more genes involved in reproduction and brain function than would be expected by chance alone. The accumulation of these types of genes may be facilitated by the structure of the X chromosome, which may be conducive to the amplification of genes in particular regions. In addition, as the X chromosome is hemizygous in males and is silenced during male meiosis, this affects both the fixation of male-beneficial and female-beneficial genes, and the expression of X-linked genes during meiosis. A requirement for the expression of X-linked genes during spermatogenesis may have driven retrotransposition of genes both on and off the X, and this has had a role in shaping the gene content of the mammalian X chromosome. Characterisation of the structure and gene content of the marsupial X chromosome is essential to understand whether the features of structure and expression observed on the human and mouse X chromosomes are fundamental to all therian X chromosomes. This will indicate whether the therian X chromosome plays a special role in reproduction and brain functions, and possibly speciation.

Keywords Mammalian X chromosome · Meiotic sex chromosome inactivation · Sex and reproduction related genes · Mental retardation genes

9.1 Introduction

The mammals are comprised of three major lineages; the eutherians ("placental" mammals) which radiated up to 105 million years ago; the metatherians, or marsupials, which diverged from the eutherians approximately 148 million years ago;

M.L. Delbridge (✉)
ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia
e-mail: margaret.delbridge@anu.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_9, © Springer Science+Business Media B.V. 2010

and the prototheria, or monotremes, which diverged from the theria (the eutheria and metatheria together) approximately 166 million years ago. The mammals diverged from the next closest group of vertebrates, the birds and reptiles, approximately 310 million years ago (Bininda-Emonds et al., 2007), and so the marsupials and monotremes represent two groups that fill the evolutionary gap between the birds and reptiles and the eutherian mammals, and are useful for assessing the conservation of differences observed amongst the eutherian mammals.

Almost all mammals have an XX female: XY male sex chromosome system, in which a gene on the Y chromosome initiates male development. The X chromosome gene content of different eutherian mammals is highly conserved. Mapping human X-borne genes in distantly related mammals showed that only part of the X chromosome is conserved between the eutherians and marsupials. The long arm and pericentric region of the human X chromosome is equivalent to the marsupial X chromosome and is therefore at least 148 million years old. This region of the human X chromosome is known as the X conserved region (XCR) (Graves, 1995). However, genes on the short arm of the human X chromosome are autosomal in two distantly related marsupials, the South American grey short-tailed opossum (*Monodelphis domestica*) (Mikkelsen et al., 2007) and the tammar wallaby (*Macropus eugenii*) (Deakin et al., 2008), implying the addition of a region to the sex chromosomes between 148 and 105 million years ago. This region of the human X chromosome is known as the X added region (XAR) and it is exposed to the same forces that shaped the conserved region (Graves, 2006).

The therian X and Y chromosomes have a different origin to the monotreme sex chromosome system. The platypus (our model monotreme) sex chromosome system consists of five X chromosomes ($X_1X_2X_3X_4X_5$) and five Y chromosomes ($Y_1Y_2Y_3Y_4Y_5$). Females have two copies of each X chromosome and males have one copy of each of the X chromosomes and each of the Y chromosomes (Rens et al., 2004). The different X and Y chromosomes form a translocation chain during male meiosis in the order $X_1Y_1X_2Y_2$ etc. The X's and Y's then segregate alternately to form balanced gametes (Grützner et al., 2004). The platypus sex chromosomes share no homology with the therian sex chromosomes, instead they share extensive homology with the chicken Z chromosome, indicating that the platypus X and Y chromosomes have a different origin from the therian X and Y chromosomes (Veyrunes et al., 2008). The building blocks of the human X chromosome are represented by autosomal blocks on platypus chromosomes 6 (XCR), and 15 and 18 (XAR) (Waters et al., 2005; Hore et al., 2007; Veyrunes et al., 2008), and are present as two major autosomal regions on chromosomes 4 (XCR) and 1 (XAR) in chickens (Kohn et al., 2004).

The marsupial X chromosome, therefore, represents a minimal therian X chromosome. Homology between the X and Y chromosomes supports the hypothesis that the therian X and Y chromosomes differentiated from a homologous autosomal pair (the proto-X and proto-Y) when one member acquired a male-determining allele. *SRY* is the male-determining gene in therian mammals. *SRY* is found on the Y chromosome of all therian mammals except for a few rodent species that have recently lost their Y chromosome (Soullier et al., 1998; Vogel et al., 1998;

Sutou et al., 2001). Following the acquisition of *SRY*, suppression of recombination between the X and Y preserved *SRY* and the nearby accumulation of genes beneficial to males on the Y chromosome. This has led to the loss of active genes and accumulation of mutations on the Y chromosome in the region where recombination has been suppressed. Progressive degradation of genes and sequences on the Y chromosome has resulted in the heterologous X and Y chromosomes observed in most extant species, such that most genes on the X chromosome have no homologue on the Y chromosome (Graves, 2006).

9.2 Y Chromosome Gene Content

The gene content of the X and Y chromosomes has been shaped by this process of restriction of recombination and the associated loss of Y genes. The therian Y chromosome contains only a small number of genes, most of which have a role in male-specific functions such as sex determination and spermatogenesis. These Y genes have differentiated from their more widely expressed X homologues. For example, in humans, the putative gonadoblastoma gene *TSPY*, has a more restricted expression pattern than its X homologue *TSPX*, which is involved in the cell cycle (Delbridge et al., 2004). Similarly, the putative spermatogenesis gene *RBMY*, and the testis-determining factor *SRY* have differentiated from their widely expressed homologues on the X chromosome *RBMX* and *SOX3* respectively (Foster and Graves, 1994; Delbridge et al., 1999). Gene mapping has shown that there are different suites of Y genes that have been retained in different lineages on the Y (reviewed in Waters et al., 2007).

9.3 X Chromosome Gene Content

In contrast to the specialisation of genes on the Y chromosome, the majority of genes on the human X chromosome are well conserved, with genes from the XCR found on the X chromosome in all therians. There also is a subset of more specialised genes on the human and mouse X chromosomes that appear to have undergone amplification, which may have occurred out of the necessity for their expression during the repression of the X during male meiosis.

The human X chromosome is approximately 155 Mb long and contains about 1,300 genes. The majority of genes from the XCR and XAR regions of the human X chromosome have a variety of housekeeping functions, and these genes are almost entirely conserved between the X chromosomes of eutherian species. In the marsupials, these genes are also conserved on the autosomal region that represents the XAR and the X chromosome, representing the XCR (Graves, 2006). Genome sequencing projects have now confirmed the gene content of the human X chromosome is not representative of all types of genes, but in fact contains an atypical bias in the gene content that is different to that found on the autosomes. This bias

is not confined to the human X chromosome as it is also observed on the mouse X chromosome; there is enrichment for genes that are involved in sex and reproduction (Wang et al., 2001; Ross et al., 2005; Koslowski et al., 2006), and brain development (Wilda et al., 2000; Zechner et al., 2001).

9.4 Sex and Reproduction-Related Genes Within Large Inverted Repeats

The X chromosomes of humans and mice are enriched for genes that are involved in sex and reproduction. Early spermatogonial genes, sex and reproduction related (SRR) and prostate-specific genes are over-represented on the X in humans and mice (Saifi and Chandra, 1999; Wang et al., 2001). In contrast, a paucity of genes involved in late spermatogenesis on the human and mouse X chromosomes has also been observed (Khil et al., 2004; Ross et al., 2005). This description of the gene content of the human and mouse X chromosomes relies largely on the analysis of conserved single copy genes. However, a major feature of the human X chromosome that was revealed by sequencing (Ross et al., 2005) is the presence of a number of large inverted repeat sequences (arm length between 9 and 142 kb) concentrated in the Xp11, Xq22 and Xq28 regions of the human X chromosome (Warburton et al., 2004). A more recent report has detailed the existence of similar large repeat sequences that make up approximately 12% of the mouse X chromosome (Mueller et al., 2008). Analysis of marsupial X chromosome sequences (opossum and tammar wallaby) is currently underway, to determine if these large inverted repeats are a feature of all therian X chromosomes. These large inverted repeats found on the human and mouse X may be analogous to similar large palindromic sequences found on the human Y chromosome in Yq11. On the Y chromosome, these palindromes have been proposed to allow gene conversion between the different members of multicopy testis-specific gene families within them, such as *TSPY* and *RBMY*, to preserve the integrity and function of genes with a crucial male specific role in the absence of meiotic recombination between homologues (Rozen et al., 2003). Arm-to arm gene conversion is unlikely to be a critical mechanism to maintain gene integrity on the X chromosome, as X chromosomes do undergo crossing over in females, but there is evidence that it does occur (Small et al., 1997).

Similarly to the Y chromosome, the inverted repeat structures on the human and mouse X chromosomes harbour multicopy testis-specific gene families (Mueller et al., 2008; Warburton et al., 2004). The human X chromosome repeats contain large numbers of CT (cancer-testis) antigen genes. These genes are multicopy gene families that are normally expressed in the testis, but are reactivated in some cancers (Simpson et al., 2005). The expression of these genes in early spermatogenesis is consistent with the prevailing view that the X chromosome is enriched for genes involved in spermatogenesis prior to meiosis (reviewed in Delbridge and Graves, 2007). However, the X was thought to be depleted in genes expressed after meiosis (Khil et al., 2004), and this was attributed to the effects of meiotic sex chromosome inactivation (MSCI), which silences the X and Y chromosomes during meiosis.

9.5 Meiotic Sex Chromosome Inactivation

In eutherians, the X and Y chromosomes undergo meiotic sex chromosome inactivation (MSCI) during spermatogenesis in males. This is a temporary silencing of the sex chromosomes during the first meiotic prophase. Eutherian MSCI is initiated when the BRCA1 protein recruits the ATR kinase to phosphorylate the nucleosomal core histone H2AX. Early marks of MSCI in the mouse are the association of H2AX and the heterochromatin associated proteins (HP1β and HP1γ)with the synapsed X and Y chromosomes, a structure known as the XY body (Mahadevaiah et al., 2001; Namekawa et al., 2006). The repression of transcription is largely maintained following meiosis in round spermatids, due to the persistence of post-meiotic sex chromatin (PMSC), and the associated H2AX, HP1β and HP1γ. Transcription of both the sex chromosomes and autosomes alike ceases during spermiogenesis when the genome undergoes extensive chromatin condensation and packaging with protamines, rather than histones, in sperm (Namekawa et al., 2006).

MSCI and PMSC also occur during spermatogenesis in marsupials (Hornecker et al., 2007; Namekawa et al., 2007). Levels of expression of X-linked housekeeping genes decline in meiotic spermatocytes in the opossum (*M. domestica*) similar to the reduction in levels of expression observed in mouse (Namekawa et al., 2006, 2007). This reduction in expression of housekeeping genes coincides with an increase in the level of expression of *BRCA1* and *ATR*, and the association of H2AX, HP1β and HP1γ proteins with the X and Y chromosomes in early pachytene. In the opossum, the association of these MSCI marks occurs prior to the association of the XY body at the dense plate. This key difference between eutherian and marsupial MSCI implies that MSCI in all therians is triggered by the presence of unpaired DNA rather than paired elements. The silencing of the XY body continues through meiosis II (spermatocytes and round spermatids) and into the post meiotic period. Overall then, MSCI and PMSC follow a similar pattern in all therian mammals (Fig. 9.1) (Franco et al., 2007; Hornecker et al., 2007; Namekawa et al., 2007).

9.6 Postmeiotic Expression of Testis-Specific X Chromosome genes

In eutherians, however, there is a specific derepression of some X-linked genes following meiosis. In mouse, there is a specific reactivation of X-linked spermatogenesis genes during spermatogenesis in the round spermatids, following the meiotic division (Mueller et al., 2008). The X-linked genes that are re-activated are mostly multicopy testis-specific genes, whose levels of expression are increased to levels comparable to the expression levels of autosomal genes. A number of single copy X-linked genes are also derepressed, but their post-meiotic expression levels are not comparable to their pre-meiotic expression levels (Mueller et al., 2008). The gene products from the X chromosome are also accessible to the Y bearing round spermatids, due to the continued close association of the tetrad of spermatids. The

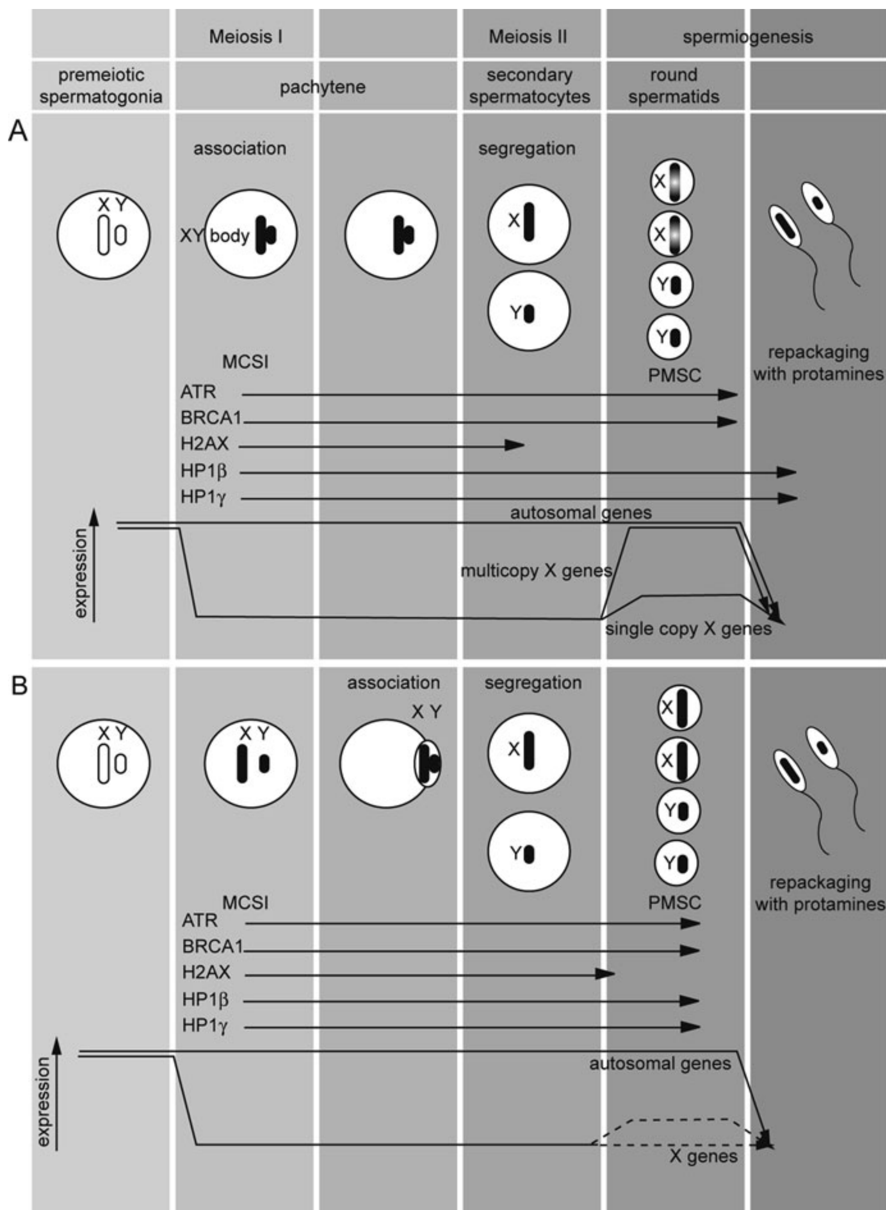


Fig. 9.1 Comparison between X-linked gene expression during meiotic sex chromosome inactivation in (**a**) mice and (**b**) marsupials. During meiosis I the X and Y chromosomes associate in the XY body in mice at an earlier stage than in marsupials. In mice, silencing of X-linked genes during MSCI occurs at the same time as the appearance of the MSCI proteins ATR, BRCA1, H2AX, HP1β and HP1γ. In marsupials, association of the X and Y chromosomes at the dense plate occurs after the appearance of the MSCI proteins. Silencing of X-linked genes is maintained in secondary spermatocytes and round spermatids. In mice, multicopy X-linked genes and a small number of

mouse X chromosome, therefore, also contains a much higher than expected proportion of genes that are expressed during spermatogenesis following meiosis (Mueller et al., 2008). There is evidence that this also occurs in humans, as the *SPANX* gene family from human Xq27 is expressed in late spermatogenesis (Kouprina et al., 2004, 2005).

It has been reported that this reactivation is not seen in the round spermatids of the opossum, where germ cell specific X-linked gene expression is not detected in the presence of high expression levels of the *ATR* gene (Hornecker et al., 2007) and the silencing marks HP1β and HP1γ, indicate the silencing of the opossum X and Y chromosomes in round spermatids (Fig. 9.1) (Namekawa et al., 2007). In contrast, reactivation of expression has been observed for a number of single copy X-linked genes in opossum round spermatids (Mahadevaiah et al., 2009). This reactivation occurs in only a proportion of X bearing spermatids (ranging from 28 to 71%) (Mahadevaiah et al., 2009), and it is not clear whether the levels of gene derepression are comparable to pre-meiotic levels. The reactivation of these single copy genes in the opossum could be the equivalent of the low level reactivation of single copy X-linked genes observed in the mouse. The continued repression of germ-cell specific X-linked genes may be a feature of marsupial MSCI (Hornecker et al., 2007), but this conclusion is based on data for a small number of genes, so surveying a larger number of X-linked genes could help to resolve this question. Characterisation of the X chromosome in both opossum and tammar wallaby is currently underway to identify single copy and multi-copy testis specific genes on the marsupial X chromosome. Investigation of the expression patterns of these genes will be an essential contribution to determine how similar the patterns of MSCI and PMSC are in marsupial and eutherian mammals.

It is now known that the ZW chromosome pair is also transiently silenced during meiosis in female chickens. The presence of MSCI in a species with female heterogamety is evidence that MSCI is not specific to spermatogenesis. The method of pairing between the Z and W in chicken oogenesis is different to the method of pairing of the X and Y during therian spermatogenesis, and it may be that the form of MSCI observed in birds evolved independently from MSCI in mammals. Silencing of the sex chromosomes during meiosis could be a mechanism by which recombination is suppressed between the heterologous chromosome pair (Schoenmakers et al., 2009).

In the mature sperm of eutherians, most of the genome is silenced, sex chromosomes and autosomes alike, due to the packaging of the DNA into a different kind of chromatin, where protamines rather than histones are used. It has been suggested

Fig. 9.1 (Continued) single copy X-linked genes escape the silencing of PMSC in round spermatids, despite the continued presence of MSCI proteins. This reactivation of X-linked gene expression has been observed in marsupials for some single copy X-linked genes, but has not been observed for germ cell-specific genes. All genes are silenced in the mature sperm when the DNA is repackaged with protamines rather than histones

that X chromosome silencing during spermatogenesis might persist after fertilisation, to facilitate dosage compensation by transmission of a partially inactivated paternal X chromosome to the eutherian embryo. The transmission of a pre-inactivated X might form the basis of imprinted X inactivation in some mammals (Huynh and Lee, 2003). Imprinted X inactivation, such as is found in marsupials where it is always the paternal X chromosome that is inactivated independently of *XIST/TSIX* expression, has been proposed to be an ancestral form of X inactivation (reviewed in Al Nadaf et al., Chapter 13). Imprinted X inactivation has also been suggested to have evolved independently in eutherians, possibly following the evolution of the more stable, random X inactivation that depends on *XIST/TSIX* expression (reviewed in Al Nadaf et al., Chapter 13; Heard and Disteche, 2006).

9.7 Retrotransposition of Genes On and Off the X Chromosome

Specific silencing of X-linked genes during spermatogenesis has been proposed to drive the retrotransposition events that generate testis-expressed copies of X-linked genes (e.g. *PGK*, *PDHA1* and *RBMX*) on autosomes (Dahl et al., 1990; McCarrey et al., 1996; Lingenfelter et al., 2001). This data has largely been collected for genes that are single copy on the X chromosome. This mechanism has been proposed to compensate for the loss of activity of essential X-linked genes during spermatogenesis (Emerson et al., 2004). Further studies have substantiated that retrotransposition off the X chromosome occurs in marsupials and eutherians, and this supports the idea that, as the therian X and Y chromosomes evolved from their autosomal pair, the evolution of MSCI followed the spread of non-recombination and was the driving force for retrotransposition off the X chromosome to compensate for silencing of X chromosome genes during spermatogenesis (Potrzebowski et al., 2008).

However, retrotransposition seems to occur in both directions on the human X chromosome. Many X-linked single copy testis-specific genes do not contain introns, and so probably arose on the X by retrotransposition of a processed copy of an autosomal source gene. For example, the *ATXN3L* and *YY2* genes are likely to be derived from intron-containing paralogues on autosomes (Loriot et al., 2003; Emerson et al., 2004). More than 75% of the genes generated by retrotransposition on the human X chromosome are also found on the mouse X chromosome, which implies that these retrotransposition events must have occurred prior to the divergence of primates and rodents approximately 105 million years ago (Emerson et al., 2004).

The formation of cruciforms or other secondary structures by the large inverted repeat sequences on the X that harbour many of the testis-specific gene families, may have contributed to the duplication of the genes within them, but also might permit escape from MSCI of those gene families that are crucial in late spermatogenesis (Skaletsky et al., 2003; Warburton et al., 2004). Within the repeat structures on the X chromosome many of the gene families have been amplified by retrotransposition as well as gene duplication, such as has occurred in the expansion of the

MAGE gene family on the human X chromosome (Chomez et al., 2001; Bertrand et al., 2004). In addition, it is suggested that gene amplification of the testis-specific gene families facilitates higher gene expression, by the additive effects of a small amount of expression from a large number of gene copies from a generally repressed X chromosome following meiosis (Mueller et al., 2008). This may be another mechanism to compensate for the loss of expression of essential single copy genes during spermatogenesis. There must also be additional mechanisms that accomplish reactivation of X-linked genes post-meiotically, as a number of single copy X genes are also reactivated.

It is interesting that the testis-specific genes within the inverted repeat sequences on the human and mouse X comprise two different subsets of genes (Warburton et al., 2004; Mueller et al., 2008), suggesting that amplification has occurred independently in different lineages.

9.8 X-Linked Genes with a Function in the Brain

The eutherian X chromosome, therefore, is enriched in spermatogenesis genes that have a role both before and after meiosis. In addition to this, there is evidence that the human X chromosome also contains a disproportionately large number of genes involved in brain development. The special role of the X chromosome in brain function is supported by the fact that there are three times as many X-linked mental retardation (XLMR) disorders than other abnormal X-linked phenotypes in humans. Although X-linked inheritance of mental retardation disorders is easily detected in families, it has been shown that the ease of detection has not created an ascertainment bias on the X chromosome (Ropers and Hamel, 2005). XLMR genes are spread over the entire human X chromosome, with a higher density in Xp11 and Xq28 (Chiurazzi et al., 2008). Interestingly, these two regions are also where the testis-specific genes on the X and the large inverted repeats are concentrated. It is striking that many genes that are expressed in the brain are also expressed in the testis (Graves et al., 2002; Lubs et al., 1996). Perhaps this is how MSCI and PMSC can affect the accumulation of X-linked genes involved in brain function, through their effects on gene expression in the testis. The accumulation of brain and testis genes on the mammalian X chromosome might mean that the mammalian X has a role in speciation as does the *Drosophila* X chromosome, which is known to be involved in speciation and has an important effect on hybrid sterility (Orr and Coyne, 1989).

The mammalian X chromosome has an accumulation of genes with a role in brain development or function, which may affect traits such as intelligence (Zechner et al., 2001). The accumulation of "intelligence" genes was speculated to be because intelligence is a male-advantage trait, or perhaps because intelligence is selected for by females, as an indicator of beneficial genes (Zechner et al., 2001). The X chromosome is transmitted two-thirds of the time through females and only one-third of the time through males. As a consequence, there would be more time for selection

to act on mutations that are of benefit to females. As long as the benefit of these genes to females was greater than the disadvantage of these genes to males, positive selection would result in an accumulation of female-beneficial mutations on the X chromosome. In contrast, the hemizygosity of the X chromosome in males allows male beneficial recessive or partially recessive alleles to be fixed more rapidly on the X chromosome relative to the autosomes, which can lead to accumulation of male-advantage genes on the X chromosome. Therefore the X chromosome can contain both female-advantage and male-advantage genes, depending on the underlying genetic variation. Human and mouse X chromosomes show an enrichment for both female and male-advantage genes, but this is not universal. For example, in *Drosophila* and *Caenorhabditis elegans*, the X chromosomes are depleted in male-specific genes and enriched for female biased genes, whereas in the chicken, in which the female is the heterogametic sex (ZW), the Z chromosome shows enrichment for male biased genes and depletion of female-biased genes (Gurbich and Bachtrog, 2008).

Genes that have a role in either female or male reproduction could clearly be classified as female or male-beneficial alleles that could be accumulated on the human X chromosome (Koslowski et al., 2006). It is more speculative whether this mechanism might result in the accumulation of genes that are involved in brain function on the X chromosome, where beneficial recessive alleles of genes with a role in the brain give a reproductive advantage to males, assuming females preferentially select more intelligent or resourceful males (Zechner et al., 2001). It has been hypothesised that these genes may have evolved on the X chromosome because X-linked genes evolve faster than autosomal genes under positive selection, and/or they may be subject to the effects of sexual antagonism (Vicoso and Charlesworth, 2006). In addition to there being a higher number of genes involved in brain function located on the X, there is also higher expression of these genes in the brains of both sexes (Nguyen and Disteche, 2006a, b). It is suggested that mechanisms such as chromatin modifications that alter gene transcription levels in the brain cells would favour the recruitment of X-linked genes into brain-specific functions because of their increased expression (Nguyen and Disteche, 2006b).

Investigation of the location and expression patterns of homologues of human XLMR genes in marsupials showed that the XLMR genes are located on the equivalent regions in marsupials to the XCR and XAR of the human X chromosome, suggesting that they have been located in the same genomic regions for more than 148 million years. As many of these marsupial XLMR homologues are widely expressed, this suggests that although some X-linked genes have had a role in the brain throughout therian evolution, others have undergone a narrowing of their expression pattern, and a change of function, to play a larger role in the brain in humans (Delbridge et al., 2008). This is consistent with observations that expression patterns of some XLMR genes are similar in older regions of the brain (such as the cerebellum) in mice and chickens, but are markedly different in younger regions of the brain such as the telencephalon. The testis and telencephalon of the mouse have specific expression of XLMR genes, which is absent from these tissues in the chicken. This suggests that a location on the X chromosome in mammals

may have accelerated selection for the more specialised functions in cognition and reproduction that have been acquired by XLMR genes in eutherians (Kohn et al., 2007).

9.9 Conclusion

The mammalian X chromosome has a large number of genes involved in sex and reproduction, and also brain function, suggesting that it may play a special role in these functions. Further analysis of the genes on the X chromosome in different species, and particularly marsupials, may indicate whether it is possible that their concentration on the X, and the lineage-specific amplification of testis and cognition genes within the large inverted repeats on the X chromosome, is related to a special role of the therian X chromosome in speciation.

References

Bertrand M, Huijbers I, Chomez P, De Backer O (2004) Comparative expression analysis of the *MAGED* genes during embryogenesis and brain development. Dev Dyn 230:325–334.

Bininda-Emonds ORP, Cardillo M, Jones KE, et al. (2007) The delayed rise of present-day mammals. Nature 446:507–512.

Chiurazzi P, Schwartz CE, Gecz J, Neri G (2008) XLMR genes: update 2007. Eur J Hum Genet 16:422–434.

Chomez P, De Backer O, Bertrand M, et al. (2001) An overview of the *MAGE* gene family with the identification of all human members of the family. Cancer Res 61:5544–5551.

Dahl HH, Brown RM, Hutchison WM, Maragos C, Brown GK (1990) A testis-specific form of the human pyruvate dehydrogenase E1 alpha subunit is coded for by an intronless gene on chromosome 4. Genomics 8:225–232.

Deakin JE, Koina E, Waters PD, et al. (2008) Physical map of two tammar wallaby chromosomes: a strategy for mapping in non-model mammals. Chromosome Res 16:1159–1175.

Delbridge ML, Graves JAM (2007) Origin and evolution of spermatogenesis genes on the human sex chromosomes. Soc Reprod Fertil Suppl 65:1–17.

Delbridge ML, Lingenfelter PA, Disteche CM, Graves JAM (1999) The candidate spermatogenesis gene *RBMY* has a homologue on the human X chromosome. Nat Genet 22:223–224.

Delbridge ML, Longepied G, Depetris D, et al. (2004) *TSPY*, the candidate gonadoblastoma gene on the human Y chromosome, has a widely expressed homologue on the X – implications for Y chromosome evolution. Chromosome Res 12:345–356.

Delbridge ML, McMillan DA, Doherty RJ, Deakin JE, Graves JAM (2008) Origin and evolution of candidate mental retardation genes on the human X chromosome (MRX). BMC Genomics 9:65.

Emerson JJ, Kaessmann H, Betran E, Long M (2004) Extensive gene traffic on the mammalian X chromosome. Science 303:537–540.

Foster JW, Graves JAM (1994) An *SRY*-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. Proc Natl Acad Sci USA 91:1927–1931.

Franco MJ, Sciurano RB, Solari AJ (2007) Protein immunolocalization supports the presence of identical mechanisms of XY body formation in eutherians and marsupials. Chromosome Res 15:815–824.

Graves JAM (1995) The origin and function of the mammalian Y chromosome and Y-borne genes–an evolving understanding. Bioessays 17:311–320.
Graves JAM (2006) Sex chromosome specialization and degeneration in mammals. Cell 124: 901–914.
Graves JAM, Gecz J, Hameister H (2002) Evolution of the human X–a smart and sexy chromosome that controls speciation and development. Cytogenet Genome Res 99:141–145.
Grützner F, Rens W, Tsend-Ayush E, et al. (2004) In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. Nature 432:913–917.
Gurbich TA, Bachtrog D (2008) Gene content evolution on the X chromosome. Curr Opin Genet Dev 18:493–498.
Heard E, Disteche CM (2006) Dosage compensation in mammals: fine-tuning the expression of the X chromosome. Genes Dev 20:1848–1867.
Hore TA, Koina E, Wakefield MJ, Graves JAM (2007) The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. Chromosome Res 15:147–161.
Hornecker JL, Samollow PB, Robinson ES, Vandeberg JL, McCarrey JR (2007) Meiotic sex chromosome inactivation in the marsupial *Monodelphis domestica*. Genesis 45:696–708.
Huynh KD, Lee JT (2003) Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. Nature 426:857–862.
Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD (2004) The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. Nat Genet 36:642–646.
Kohn M, Kehrer-Sawatzki H, Vogel W, Graves JAM, Hameister H (2004) Wide genome comparisons reveal the origins of the human X chromosome. Trends Genet 20:598–603.
Kohn M, Kehrer-Sawatzki H, Steinbach P, Graves JAM, Hameister H (2007) Recruitment of old genes to new functions: evidences obtained by comparing the orthologues of human XLMR genes in mouse and chicken. Cytogenet Genome Res 116:173–180.
Koslowski M, Sahin U, Huber C, Tureci O (2006) The human X chromosome is enriched for germline genes expressed in premeiotic germ cells of both sexes. Hum Mol Genet 15: 2392–2399.
Kouprina N, Mullokandov M, Rogozin IB, et al. (2004) The *SPANX* gene family of cancer/testis-specific antigens: rapid evolution and amplification in African great apes and hominids. Proc Natl Acad Sci USA 101:3077–3082.
Kouprina N, Pavlicek A, Noskov VN, et al. (2005) Dynamic structure of the *SPANX* gene cluster mapped to the prostate cancer susceptibility locus HPCX at Xq27. Genome Res 15:1477–1486.
Lingenfelter PA, Delbridge ML, Thomas S, et al. (2001) Expression and conservation of processed copies of the *RBMX* gene. Mamm Genome 12:538–545.
Loriot A, Boon T, De Smet C (2003) Five new human cancer-germline genes identified among 12 genes expressed in spermatogonia. Int J Cancer 105:371–376.
Lubs HA, Chiurazzi P, Arena JF, et al. (1996) XLMR genes: update 1996. Am J Med Genet 64:147–157.
Mahadevaiah SK, Turner JM, Baudat F, et al. (2001) Recombinational DNA double-strand breaks in mice precede synapsis. Nat Genet 27:271–276.
Mahadevaiah SK, Royo H, Vandeberg JL, et al. (2009) Key features of the X inactivation process are conserved between marsupials and eutherians. Curr Biol 19:1478–1484.
McCarrey JR, Kumari M, Aivaliotis MJ, et al. (1996) Analysis of the cDNA and encoded protein of the human testis-specific PGK-2 gene. Dev Genet 19:321–332.
Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.
Mueller JL, Mahadevaiah SK, Park PJ, et al. (2008) The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. Nat Genet 40:794–799.
Namekawa SH, Park PJ, Zhang LF, et al. (2006) Postmeiotic sex chromatin in the male germline of mice. Curr Biol 16:660–667.

Namekawa SH, VandeBerg JL, McCarrey JR, Lee JT (2007) Sex chromosome silencing in the marsupial male germ line. Proc Natl Acad Sci USA 104:9730–9735.

Nguyen DK, Disteche CM (2006a) Dosage compensation of the active X chromosome in mammals. Nat Genet 38:47–53.

Nguyen DK, Disteche CM (2006b) High expression of the mammalian X chromosome in brain. Brain Res 1126:46–49.

Orr HA, Coyne JA (1989) The genetics of postzygotic isolation in the *Drosophila virilis* group. Genetics 121:527–537.

Potrzebowski L, Vinckenbosch N, Marques AC, et al. (2008) Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. PLoS Biol 6:e80.

Rens W, Grützner F, O'Brien PC, et al. (2004) Resolution and evolution of the duck-billed platypus karyotype with an $X_1Y_1X_2Y_2X_3Y_3X_4Y_4X_5Y_5$ male sex chromosome constitution. Proc Natl Acad Sci USA 101:16257–16261.

Ropers HH, Hamel BC (2005) X-linked mental retardation. Nat Rev Genet 6:46–57.

Ross MT, Grafham DV, Coffey AJ, et al. (2005) The DNA sequence of the human X chromosome. Nature 434:325–337.

Rozen S, Skaletsky H, Marszalek JD, et al. (2003) Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature 423:873–876.

Saifi GM, Chandra HS (1999) An apparent excess of sex- and reproduction-related genes on the human X chromosome. Proc R Soc Lond B Biol Sci 266:203–209.

Schoenmakers S, Wassenaar E, Hoogerbrugge JW, et al. (2009) Female meiotic sex chromosome inactivation in chicken. PLoS Genet 5:e1000466.

Simpson AJ, Caballero OL, Jungbluth A, Chen YT, Old LJ (2005) Cancer/testis antigens, gametogenesis and cancer. Nat Rev Cancer 5:615–625.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837.

Small K, Iber J, Warren ST (1997) Emerin deletion reveals a common X-chromosome inversion mediated by inverted repeats. Nat Genet 16:96–99.

Soullier S, Hanni C, Catzeflis F, Berta P, Laudet V (1998) Male sex determination in the spiny rat *Tokudaia osimensis* (Rodentia: Muridae) is not *Sry* dependent. Mamm Genome 9:590–592.

Sutou S, Mitsui Y, Tsuchiya K (2001) Sex determination without the Y chromosome in two Japanese rodents *Tokudaia osimensis osimensis* and *Tokudaia osimensis* spp. Mamm Genome 12:17–21.

Veyrunes F, Waters PD, Miethke P, et al. (2008) Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res 18:965–973.

Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet 7:645–653.

Vogel W, Jainta S, Rau W, et al. (1998) Sex determination in *Ellobius lutescens*: the story of an enigma. Cytogenet Cell Genet 80:214–221.

Wang PJ, McCarrey JR, Yang F, Page DC (2001) An abundance of X-linked genes expressed in spermatogonia. Nat Genet 27:422–426.

Warburton PE, Giordano J, Cheung F, Gelfand Y, Benson G (2004) Inverted repeat structure of the human genome: the X-chromosome contains a preponderance of large, highly homologous inverted repeats that contain testes genes. Genome Res 14:1861–1869.

Waters PD, Delbridge ML, Deakin JE, et al. (2005) Autosomal location of genes from the conserved mammalian X in the platypus (*Ornithorhynchus anatinus*): implications for mammalian sex chromosome evolution. Chromosome Res 13:401–410.

Waters PD, Wallis MC, Graves JAM (2007) Mammalian sex–Origin and evolution of the Y chromosome and *SRY*. Semin Cell Dev Biol 18:389–400.

Wilda M, Bachner D, Zechner U, et al. (2000) Do the constraints of human speciation cause expression of the same set of genes in brain, testis, and placenta? Cytogenet Cell Genet 91:300–302.

Zechner U, Wilda M, Kehrer-Sawatzki H, et al. (2001) A high density of X-linked genes for general cognitive ability: a run-away process shaping human evolution? Trends Genet 17:697–701.

Chapter 10
Marsupial Sex Chromosome Behaviour During Male Meiosis

Raúl Fernández-Donoso, Soledad Berríos, Julio S. Rufas, and Jesús Page

Abstract The absence of homology between sex chromosomes in marsupials strongly influences their behaviour during male meiosis. The highly differentiated X and Y chromosomes perform a precise and specific meiotic program that includes pairing and segregation, but lacks the usual mechanisms of synapsis, recombination and chiasma formation that occur in the autosomes and also in the sex chromosomes of eutherian mammals. The most relevant feature of marsupial male meiosis is the development of a synaptonemal complex-derived structure called the dense plate (DP). This structure maintains the association between the asynaptic and achiasmatic sex chromosomes during the first meiotic division, contributing to the proper segregation of sex chromosomes into daughter cells. Comparison of the meiotic mechanism present in marsupials with those present in some eutherian mammals opens new perspectives concerning the origin of sex chromosomes and sex chromosome segregation in the ancestor of marsupials and placental mammals. Similarly, recent characterisation of the mechanisms involved in the inactivation of sex chromosomes during marsupial meiosis has led to the idea that somatic inactivation of sex chromosomes in mammals may have originated from the more ancient and conserved mechanism of meiotic sex chromosome inactivation. This clearly places marsupial meiosis at the heart of the discussion concerning sex chromosome evolution and the origin of gene dosage compensation in mammals.

Keywords Dense plate · Inactivation · Marsupial · Meiosis · Sex chromosomes · Synapsis

---

J. Page (✉)
Laboratorio C-001, Departamento de Biología, Universidad Autónoma de Madrid, Edificio de Ciencias Biológicas, Cantoblanco, Madrid 28049, Spain
e-mail: jesus.page@uam.es

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_10, © Springer Science+Business Media B.V. 2010

10.1 Marsupial Sex Chromosomes

Marsupials are characterised by a striking conservation of their karyotypes. Chromosome numbers range from 2n = 10 to 2n = 32 in the approximate 270 recognised species, though there is a predominance of 2n = 14 and 2n = 22 (Sharman, 1973; Hayman, 1990). Chromosome morphology is also very similar, even between distantly related species (Fig. 10.1). The evolution of marsupial karyotypes has been attributed to the action of two main processes: chromosome fusion, which reduces chromosome number, and chromosome fission, which contributes to an increase in chromosome number (Svartman and Vianna-Morgante, 1998).

A second feature of marsupial karyotypes is the reduced size of sex chromosomes, particularly the Y chromosome. Except for a few species in which one or both sex chromosomes have been translocated to an autosomal pair, most species bear sex chromosomes that are much smaller than their eutherian (placental) counterparts (Hayman, 1990; Graves, 1996; Toder et al., 2000). The proportion of the X chromosome is estimated to represent about 3% of the haploid genome, whereas the Y chromosome represents less than 1% (Graves, 1995). The reduction in size is accompanied by poor gene content. This is especially notable for the Y chromosome, which bears the male sex determinant *SRY* and very few other genes (Toder et al., 2000). Indeed, different reports have indicated that the Y chromosome is dispensable for the somatic line and can be lost in some cell types (Gallardo and Paterson, 1987; Hayman, 1990; Palma, 1995; Watson et al., 1998).

A third feature of marsupial karyotypes is the absence of a region of homology between the X and Y chromosomes (Graves and Watson, 1991). This is distinct from that observed in eutherian mammals, in which the X and Y chromosomes share a homologous segment called the pseudoautosomal region (PAR), and is the consequence of the complex evolutionary history of mammalian sex chromosomes. It has been proposed that X and Y chromosomes originated from an ordinary pair of autosomes in the reptilian ancestor of mammals through the acquisition by the Y chromosome of a male determining factor (Ohno, 1967). The suppression of recombination around this locus, putatively the *SRY* locus, could have promoted the genetic isolation and subsequent attrition of the Y chromosome, resulting in the highly differentiated sex chromosomes found today (Graves and Watson, 1991; Graves et al., 1998; Charlesworth et al., 2005; Wallis et al., 2008). In eutherian mammals, recurrent translocation events of autosomal segments to both sex chromosomes have occurred. These autosomal segments make up the current PAR, which have contributed to the partial restoration of homology between the X and Y chromosomes (Graves et al., 1998; Skaletsky et al., 2003). In marsupials, such translocation events are rare, and except for a few species that possess large PARs of autosomal origin, which usually comprise the nucleolar organising regions (NORs), most species show a complete lack of homology between sex chromosomes (Hayman, 1990; Graves, 1996; Graves et al., 1998). Therefore, while eutherian sex chromosomes may have engaged in recurrent cycles of addition/attrition, marsupial sex chromosomes would have undergone very few or no addition events and extreme attrition of the Y chromosome. Nevertheless, it is

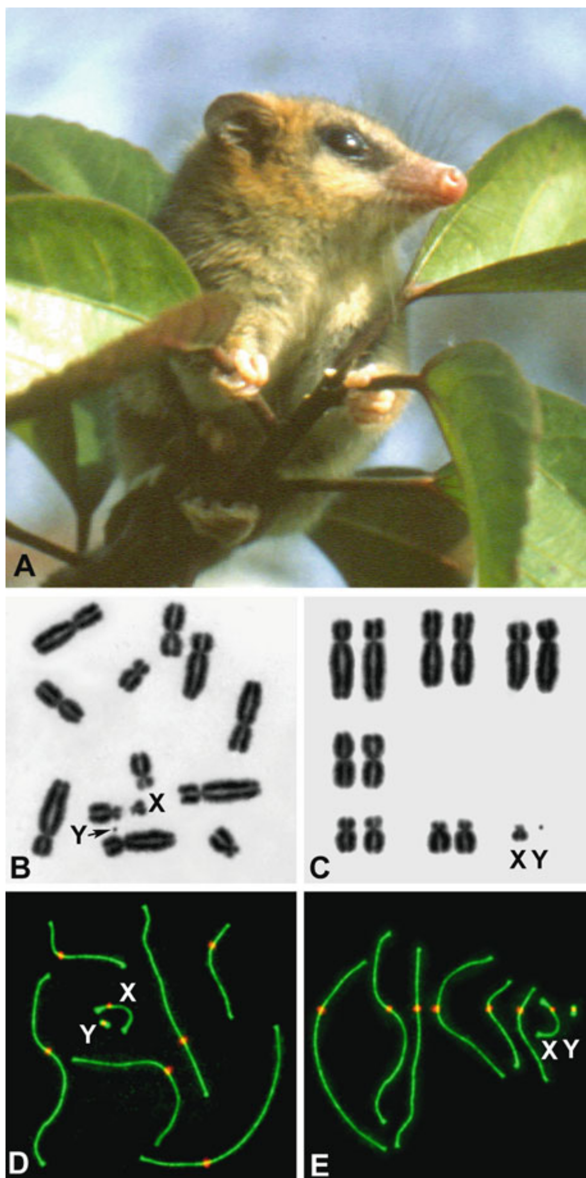


Fig. 10.1 The typical 2n = 14 marsupial karyotype. (**a**) The species *Dromiciops gliroides*, the only extant member of the Order Microbiotheria. (**b**) A somatic metaphase plate of *D. gliroides*. The minute Y chromosome is usually located close to one of the largest chromosomes and to one of the small nucleolar chromosomes. (**c**) Arranged karyotype from the previous metaphase plate showing the standard 2n = 14 chromosome number found in many marsupial species. The two small pairs in the same line as the sex chromosomes are the nucleolar chromosomes. (**d**) A spread pachytene spermatocyte stained for SYCP3 protein (*green*) and centromeres (*red*). The signal in

important to highlight that all marsupial species studied to date retain the Y chromosome, whereas some cases of complete loss of the Y chromosome have been described among eutherian mammals (Fredga and Bulmer, 1988).

Recent advances in the characterisation of monotreme sex chromosomes indicate that the evolutionary scenario described above is only valid for species included in the therian clade. Sex chromosomes in the monotreme clade, comprised of platypus and echidnas, have a different evolutionary origin. Monotremes present a complex system of sex chromosomes that is composed of five X and five Y chromosomes in the platypus (Grützner et al., 2004; Rens et al., 2004), and five X and four Y chromosomes in the echidna (Rens et al., 2007). Interestingly, monotremes do not have an *SRY* gene (Wallis et al., 2007), and their sex chromosomes share homology with avian sex chromosomes (Rens et al., 2007; Veyrunes et al., 2008), a puzzling situation that opens numerous questions about the origin of sex chromosomes and sex determination in mammals and their reptilian relatives (Wallis et al., 2008).

10.2 Sex Chromosome Behaviour During Meiosis in Mammals

Meiosis is a specialised kind of cell division that comprises two consecutive cell divisions without a duplication of the DNA between them, resulting in haploid daughter cells. The fusion of gametes during fertilization restores diploidy. A specific series of events takes place during meiosis, namely: pairing, synapsis, recombination, and segregation of homologous chromosomes. During prophase of the first meiotic division, homologous chromosomes associate in pairs to form the bivalents, and their association is maintained and stabilised by the synaptonemal complex (SC), a proteinaceous structure assembled between the two homologues. The SC consists of two lateral elements (LEs), one per homologue, called axial elements (AEs) before synapsis is completed, and a series of transverse filaments (TFs) that emanate from each LE and imbricate in a zipper-like manner, organising the central element (CE) (von Wettstein et al., 1984; Heyting, 1996; Page and Hawley, 2004). The molecular components of the SC are well characterised in mammals. SYCP2 and SYCP3 are the primary protein components of the LE (Lammers et al., 1994; Offenberg et al., 1998). SYCP1 is the main component of the TFs and the CE (Meuwissen et al., 1992), and SYCE1, SYCE2 and Tex12 are components located only on the CE (Costa et al., 2005; Hamer et al., 2006).

Pairing, i.e., the close association between homologous chromosomes, and synapsis, the formation of the SC between them, are dependent on the previous

Fig. 10.1 (continued) *green* outlines the trajectory of the synaptonemal complex. Sex chromosomes (X, Y) appear in the same area, but there is not contact between them. (**e**) Arranged meiotic karyotype from the same spermatocyte shown in (**d**). The tiny Y chromosome is visualised as submetacentric. To see original colors consult the PDF version

initiation of recombination/repair events during the early stages of the first meiotic prophase (Zickler and Kleckner, 1999). At the beginning of meiosis, during leptotene, massive DNA double strand breaks (DSBs) are induced by the SPO11 protein (Keeney et al., 1997; Romanienko and Camerini-Otero, 2000). The repair of these DSBs through a homologous recombination repair mechanism triggers a search for homologous DNA segments, leading to the close association of homologous chromosomes that culminates during zygotene with the assembly of the SC CE between their LEs. Once homologous chromosomes are completely synapsed, some of these DSBs are repaired during pachytene as reciprocal interchange events, leading to the formation of crossovers between homologues, whose cytological manifestation are chiasmata. Homologous chromosomes closely tied by chiasmata orientate each one to an opposite spindle pole during metaphase-I. Each homologue then segregates to a different pole at the beginning of anaphase-I, resulting in a reductional division event that halves the number of chromosomes in the daughter cells.

Under this general model, the processes of synapsis, recombination and chiasma formation, along with the presence of sister chromatid cohesion mechanisms, are required to ensure proper chromosome orientation and segregation during the first meiotic division (Roeder, 1997). In eutherians, all of these events initially described for autosomes also occur for the highly differentiated X and Y chromosomes during male meiosis. Studies developed in the late sixties showed that sex chromosomes in male mice, and other mammals, have a short pairing region at the distal ends, in which the SC is assembled and recombination and chiasma formation occur (Solari, 1969; Solari, 1970; Solari, 1974). The presence of a region of synapsis, thereafter called the PAR, has been subsequently identified in most eutherians. Nevertheless, it was also found that some species strikingly lack a synapsis region. Examples are mainly found among rodents, including gerbils (Solari and Ashley, 1977; Ashley and Moses, 1980; Ratomponirina et al., 1986) and voles (Wolf et al., 1988; Ashley et al., 1989; Carnero et al., 1991; Jimenez et al., 1991; Borodin et al., 1995). However, the most widespread case of PAR absence is found in marsupials. Several studies have shown that this genetic condition results in the unique behaviour of sex chromosomes during meiosis.

10.3 Sex Chromosome Pairing in Marsupials

The pioneering study of Solari and Bianchi on the South American marsupial *Monodelphis dimidiata* (Solari and Bianchi, 1975) uncovered surprising features of the meiotic behaviour of sex chromosomes. The X and Y chromosomes showed a delayed pairing compared to autosomes, and a complete SC between them was absent. The distance between their AEs was always greater than the normal distance between LEs in synapsed chromosomes, and no CE was observed. The asynaptic relationship between sex chromosomes in this marsupial contrasted with the behaviour reported in eutherian sex chromosomes (Solari, 1974). Surprisingly, an

electron dense structure, called the dense plate (DP), was organised in the region of association of the sex chromosomes to the nuclear envelope during prophase of the first meiotic division (Solari and Bianchi, 1975) (Fig. 10.2). A role for this structure in maintaining the asynaptic association of sex chromosomes during prophase-I was then proposed. Subsequent studies on both American and Australian marsupials showed that the delayed pairing and asynapsis were features conserved among very distant marsupial species, indicating that they are likely of ancient origin (Pathak

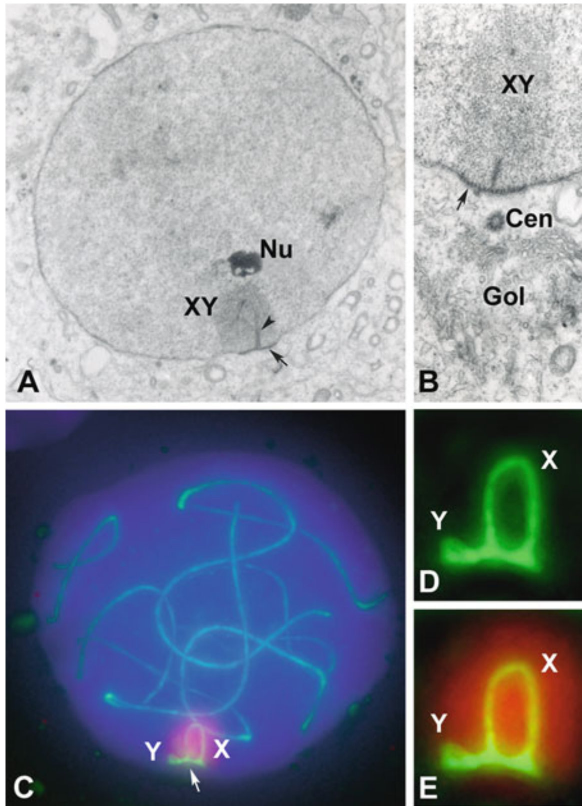


Fig. 10.2 Sex chromosome location and organisation in pachytene spermatocytes. (**a**) Electron microscopy image of a *D. gliroides* spermatocyte section at pachytene. The sex body (XY) is visible at the nuclear periphery. The trajectory of one axial element, presumably that of the X chromosome, is seen inside the sex body (*arrowhead*). An electron dense structure corresponding to the DP (*arrow*) is present in the region of sex chromosome association to the nuclear envelope. The nucleolus (Nu) lies in the vicinity of the sex body. (**b**) The DP at higher magnification. The centrioles (Cen) and the Golgi apparatus (Gol) are close to the region where the sex chromosomes appear inside the nucleus. (**c**–**e**) A spermatocyte of *Thylamys elegans* stained for SYCP3 (*green*), γH2AX (*red*), and DNA (*blue*). The sex chromosomes appear at the nuclear periphery, and their axial elements extend over the nuclear envelope to form the DP (*arrow*). Sex chromosomes are the only chromatin region labelled with γH2AX, one of the main markers of meiotic sex chromosome inactivation. To see original colors consult the PDF version

et al., 1980; Sharp, 1982; Roche et al., 1986; Seluja et al., 1987). In some of these reports, another electron dense structure called "balloon" was described associated to the chromosomal ends of X and Y chromosomes. The coincidence in their location and electron dense behaviour indicated that the DP and the balloons could be the same structure (Sharp, 1982; Roche et al., 1986; Seluja et al., 1987; Solari, 1993).

The use of electron microscopy has revealed other relevant features of sex chromosome behaviour. The most remarkable is the association of nuclear and cytoplasmic structures to the sex chromosomes. For example, the nucleolus has been repeatedly observed in the vicinity of the sex chromosomes, even in those species in which sex chromosomes do not bear nucleolar organising regions (NORs) (Fernandez-Donoso et al., 1979) (Fig. 10.2). However, the relationship between the nucleolus and the sex body is not exclusive to marsupials, since it is also common in many eutherian species (Tres, 2005), including man (Berrios and Fernandez-Donoso, 1990). Unfortunately, the causes and consequences of this association remain elusive.

As for cytoplasmic structures, the centrioles and the Golgi apparatus are frequently located on the cytoplasmic side of the nucleus region where sex chromosomes are found (Solari and Bianchi, 1975; Roche et al., 1986; Page et al., 2003). The positioning of centrioles has been correlated in many biological models with the polarisation of the chromosome telomeres that occurs at the beginning of meiosis, in the so-called "bouquet" configuration (Zickler and Kleckner, 1998). As such it is likely that this association may play a role in the polarisation and pairing of sex chromosomes in marsupials (Page et al., 2003). This mechanism could replace the usual way of chromosome pairing that, in most eutherian mammals, relies on the recognition of homologous sequences. Furthermore, it could act at different stages of meiosis, thus explaining the late pairing of sex chromosomes.

Further data on the behaviour and dynamics of sex chromosomes in South American marsupials have been obtained by means of SC protein immunolocalisation. The use of antibodies against SC proteins corroborated the absence of SC assembly between sex chromosomes, and importantly showed a relationship between the proteins that participate in the organisation of the AEs and the DP (Page et al., 2003). SYCP3, the primary protein component of the AEs/LEs, also localised on the DP (Figs. 10.2 and 10.3). The precise characterisation of sex chromosome pairing indicates that the DP is formed during mid-pachytene as an extension of the ends of the sex chromosome AEs over the nuclear envelope (Fig. 10.3). The location of SYCP3 in both the DP and the balloons establishes that they are indeed the same structure (Page et al., 2003). The formation of the DP also involves an uncharacterised phosphoprotein, but does not rely on other known chromosomal proteins like the cohesins, which organise a cohesin axis along meiotic chromosomes (Page et al., 2003; Page et al., 2005; Page et al., 2006a).

Although these studies have been exclusively performed in South American marsupials, it is likely that their conclusions could be valid for Australian marsupials as well. One of the species studied, *Dromiciops gliroides*, is of special relevance in this sense. *D. gliroides* (Fig. 10.1) is the only extant member of the order Microbiotheria,

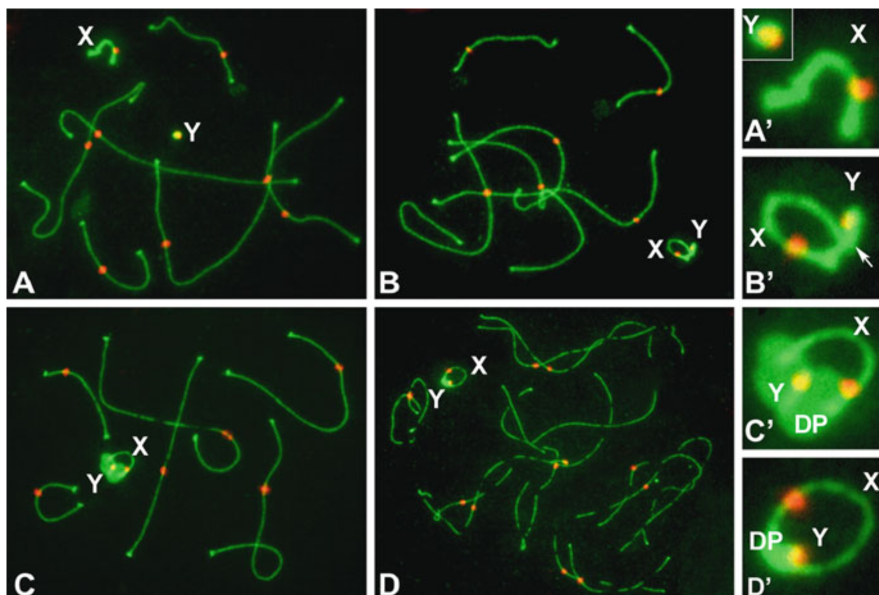


Fig. 10.3 Sex chromosome pairing and formation of the dense plate in *D. gliroides*. SYCP3 protein (*green*) and centromeres (*red*). (**a**–**a'**) Early Pachytene. Sex chromosomes lie apart from each other with thickened axial elements. All the autosomal chromosomes are synapsed. (**b**–**b'**) Mid pachytene. Sex chromosomes approach each other and contact at their ends (*arrow*). The axial elements start to become thinner. (**c**–**c'**) Late pachytene. Sex chromosomes are associated. Their axial elements are thinner and a dish-like structure, the dense plate (DP), is observed. The ends of both sex chromosomes remain associated to this dense plate. (**d**–**d'**) Diplotene. Autosomes have started to desynapse. Sex chromosomes remain associated through the dense plate (DP), which at this stage has become smaller. To see original colors consult the PDF version

and although its biogeographical distribution is restricted to the temperate rain forest in southern Chile and Argentina, phylogenetic studies have demonstrated that it is more closely related to Australian marsupials than it is to American marsupials (Kirsch et al., 1991; Szalay, 1994; Spotorno et al., 1997; Palma and Spotorno, 1999; Amrine-Madsen et al., 2003; Nilsson et al., 2003). Thus, the studies on this living fossil are especially valuable and extend the above results to Australian marsupials (Page et al., 2005).

Therefore, under the present view, the process of sex chromosome pairing in marsupial males appears as a conserved and regulated process, which although is always delayed in relation to autosomes, occurs in a precise way and at a precise time during pachytene. Furthermore, it is associated with the polarisation of other cellular structures, and involves the development of a specific structure, the DP, derived from the components of the SC AEs/LEs. At this point, an interesting question is whether the development of the DP could be a marsupial singularity. As mentioned above, some eutherian species have asynaptic sex chromosomes. In some of them, gerbils for instance, specific modifications of the sex chromosome AEs have been

also described (Ratomponirina et al., 1986; de la Fuente et al., 2007). However, these modifications appear at late stages of first meiotic prophase, usually during diplotene, and are not as conspicuous and well developed as the DP. Therefore, development of the DP, as it has been described, seems to be exclusive to marsupials.

10.4 Sex Chromosome Segregation in Marsupials

The asynaptic nature of marsupial sex chromosomes has been accepted as evidence that sex chromosomes are also achiasmatic (Hayman, 1990; Graves and Watson, 1991). If this is the case, then it would be expected that sex chromosomes appear as univalents during the first meiotic metaphase. However, data concerning this issue are contradictory. The seminal work of Koller showed that in five marsupial species studied, sex chromosomes consistently appeared associated at metaphase-I and performed a normal reductional segregation (Koller, 1936). However, later studies observed that a high incidence of sex chromosome univalents could be found in some of the same species studied by Koller (Hayman, 1990). These discrepancies have been difficult to solve because the studies on sex chromosome segregation in marsupials have been very limited. Furthermore, the work of Koller and others clearly predates the idea that marsupial sex chromosomes could be asynaptic, and therefore it is not surprising that sex chromosomes were thought to be tied by a chiasma (Koller, 1936; Solari and Bianchi, 1975). A recent report has proposed a plausible model that can reconcile these apparently dissimilar interpretations. The study of sex chromosome segregation in the South American marsupial *Thylamys elegans*, revealed that the DP that organises during first meiotic prophase is maintained until first meiotic division, providing a means for sex chromosomes to remain associated (Page et al., 2006b) (Fig. 10.4). This report also showed that the association of sex chromosomes in this species was achiasmatic, since sex chromosome segregation preceded the release of sister chromatid cohesion, and that the DP remained joined to the end of sex chromosomes throughout the process of orientation and stabilisation of these chromosomes at the metaphase-I plate. Once the sex chromosomes were properly oriented, each one facing an opposite cell pole, the DP detached from them and they initiated a precocious segregation. Therefore, this model accounts for the regular segregation of sex chromosomes during first meiotic division (Koller, 1936), and also explains why in some instances, especially when harsh chromosome spreading techniques have been used, sex chromosomes may appear as univalents (Hayman, 1990).

Intriguingly, the DP adopted a SC-like structure at the end of prophase-I, incorporating the SYCP1 protein of the SC central element. This organisation has also been observed by electron microscopy in *M. dimidiata* (Solari and Bianchi, 1975), indicating that this late pattern of transformation of the DP at the end of prophase-I could be a general process, at least in South American marsupials. The presence of the SC or SC-derived structures that maintain the association of achiasmatic chromosomes was already known for some organisms, such as butterflies and beetles

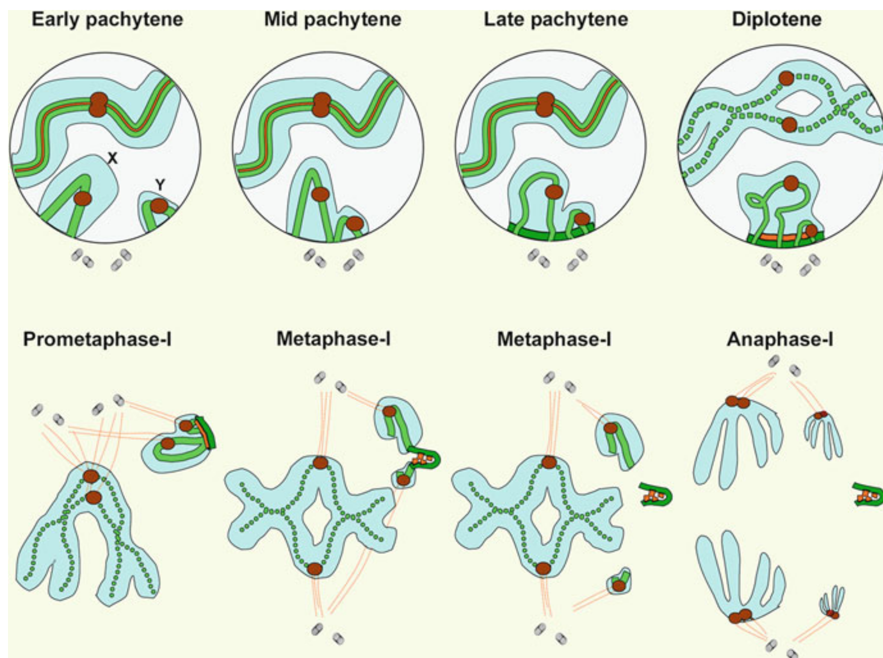


Fig. 10.4 Schematic representation of sex chromosome pairing and segregation during marsupial meiosis. The AEs/LEs of the synaptonemal complex (*green*), CE (*red*), centromeres (*brown*), centrioles (*grey*), and microtubules (*orange*) are indicated. Sex chromosomes are separated during early pachytene, and they pair during mid pachytene. The dense plate develops during mid and late pachytene and plays a role in maintaining the association between sex chromosomes in the absence of a complete synaptonemal complex, and the occurrence of recombination between sex chromosomes. During diplotene, the dense plate incorporates components of the synaptonemal complex central element and bends to form a synaptonemal complex-like structure. The dense plate remains associated to the sex chromosomes during the process of orientation and stabilisation in the metaphase-I spindle. Thereafter, it detaches from the sex chromosomes and they initiate their migration to opposite cell poles. Thus, the dense plate also serves the purpose of ensuring the proper segregation of sex chromosomes during first meiotic division. To see original colors consult the PDF version

(Wahrman et al., 1973; Rasmussen, 1977; Wolf, 1994), but no examples have been described for vertebrates. The association between achiasmatic sex chromosomes in eutherian mammals has been primarily attributed to telomere and/or heterochromatin association (Solari and Ashley, 1977; Ashley and Moses, 1980; John and King, 1985; Wolf et al., 1988; Ashley, 1994). However, a recent report has shown that association of the achiasmatic sex chromosomes in gerbils is also mediated by a SC-derived aggregate organised between the sex chromosomes at the end of prophase-I (de la Fuente et al., 2007), and similar structures are probably present also in voles (J. Page., unpublished). Although the structure and organisational pattern seem to differ between the DP and the structures found in eutherians, which are devoid of SC CE components, these studies suggest that the SC is a structure with

surprising plasticity that, besides its role in synapsis, may also play a crucial role in chromosome segregation when recombination and chiasma are absent between meiotic chromosomes.

10.5 Sex Chromosome Inactivation

Transcriptional silencing is a characteristic feature of sex chromosomes during meiosis. It was recognised early that sex chromosomes appear as a compact chromatin mass during prophase of the first meiotic division, and that this chromatin does not transcribe (Henderson, 1963; Das et al., 1965; Monesi, 1965; Solari, 1969). This process is referred to as meiotic sex chromosome inactivation (MSCI). Although it was first believed that mammalian sex chromosomes form a separate vesicle in the meiotic nucleus, as in some insects, it was eventually realised that this is not the case, and the term sex body was coined to describe the special configuration of sex chromosomes in mammals (Solari, 1974). The reason why sex chromosomes remain inactive has been a matter of debate for years (Handel and Hunt, 1992; McKee and Handel, 1993), but it is currently accepted that it is due to the presence of unsynapsed AEs at non-homologous regions of both the X and the Y chromosomes (Turner et al., 2004; Turner, 2007). It is now established that MSCI is a special case of a more general process called meiotic silencing of unsynapsed chromatin (MSUC) that affects both autosomes and sex chromosomes in a wide range of organisms, from worms to vertebrates (Schimenti, 2005; Turner et al., 2005; Kelly and Aramayo, 2007). In eutherian mammals, MSUC is initiated by the persistence of recombination/repair components like BRCA1 and ATR on unsynapsed sex chromosomal AEs during prophase-I. This induces a deep chromatin remodelling in these regions, which involves: (1) modification of histones, the most relevant being the phosphorylation of histone variant H2AX at serine 139 (γH2AX) (Mahadevaiah et al., 2001; Fernandez-Capetillo et al., 2003), ubiquitination of H2A (Baarends et al., 2005), and methylation and deacetylation of H3 and H4 (Khalil et al., 2004; Khalil and Driscoll, 2006); (2) the incorporation of specific histones like macroH2A (Hoyer-Fender et al., 2000) and H3.3 (van der Heijden et al., 2007); (3) the specific incorporation of some proteins like XY77 (Kralewski and Benavente, 1997), HP1 (Motzkus et al., 1999), XMR (Escalier and Garchon, 2000), SUMO-1 (Rogers et al., 2004), Maelstrom (Costa et al., 2006) and others (Hoyer-Fender, 2003; Handel, 2004); and (4) the association of non coding RNAs like *XIST* (Ayoub et al., 1997) and micro RNAs (Marcon et al., 2008). Some of these chromatin modifications are lost throughout meiosis, whereas others are maintained until postmeiotic stages of spermatid differentiation (Namekawa et al., 2006; Turner et al., 2006).

Although MSCI has been intensively studied in eutherian models, early studies have shown that marsupial sex chromosomes similarly formed a sex body during meiosis (Solari and Bianchi, 1975; Fernandez-Donoso et al., 1979) (Fig. 10.2). More recently, it has been reported that marsupial sex chromosomes are subjected to MSCI, and that some of the chromatin modifications found in eutherians, like

γH2AX, H3 methylation, and HP1 incorporation, also occur in South American marsupials (Franco et al., 2007; Hornecker et al., 2007; Namekawa et al., 2007) (Fig. 10.2). Likewise, postmeiotic sex chromosome inactivation has been reported in marsupials (Hornecker et al., 2007; Namekawa et al., 2007), suggesting that some of the mechanisms that govern MSCI are conserved in therian mammals. Nevertheless, the pattern of loading and release of some MSCI markers differs between eutherians and marsupials, since some MSCI markers that disappear from eutherian sex chromosomes at the end of first meiotic prophase are maintained in marsupials until spermatid differentiation (Namekawa et al., 2007). Initial reports indicated that postmeiotic sex chromosome inactivation could be more stringent in marsupials than in eutherians (Hornecker et al., 2007). However, other authors have reported contrasting results on the analysis of postmeiotic sex chromosome inactivation in marsupials, indicating that the X chromosome is reactivated in the spermatids (Mahadevaiah et al., 2009). Obviously, more studies are necessary to elucidate this issue.

The existence of specific mechanisms of sex chromosome inactivation during spermatogenesis may have interesting consequences for the understanding of somatic X chromosome inactivation (XCI). It is well established that one of the X chromosomes in mammalian females is inactivated in the somatic cell line (Lyon, 1961). XCI in eutherian mammals occurs mostly at random in the embryo tissues, so that the X chromosome from either paternal or maternal origin can be inactivated, but it exclusively affects the paternally derived X chromosome in the extra-embryonic tissues of mice and cattle. The initiation of XCI relies on the activation of the *XIST* locus and the coating of one of the X chromosomes with this non-transcribed RNA, which subsequently leads to the incorporation of a plethora of chromatin modifications that stabilises and maintains the silencing of the X chromosome (Avner and Heard, 2001). While this pattern is well defined for eutherians, several differences have been observed regarding this process in marsupials. The most relevant is that inactivation in both embryonic and extra-embryonic tissues always affects the X chromosome of paternal origin (Cooper, 1971; Sharman, 1971; Graves, 1996). Other differences are related to the mechanisms that trigger and maintain X chromosome inactivation. Strikingly, marsupials do not have an *XIST* gene (Davidow et al., 2007; Hore et al., 2007; Mikkelsen et al., 2007; Shevchenko et al., 2007), and although some chromatin modifications involved in the maintenance of the inactive state of the X chromosome are shared with eutherians (Wakefield et al., 1997; Mahadevaiah et al., 2009) some others have not been found in marsupials (Koina et al., 2009).

On these grounds, there is increasing support of the hypothesis that somatic XCI might have arisen in both eutherians and marsupials from the more general and conserved mechanism of MSCI (Cooper, 1971; Ayoub et al., 1997; McCarrey, 2001; Huynh and Lee, 2005; Hornecker et al., 2007; Namekawa et al., 2007). According to this hypothesis, MSCI during meiosis could result in the transmission of an imprinted X-inactivated chromosome from the sperm to the zygote. This inactivated X chromosome of paternal origin would remain so in marsupials, and also in the extra-embryonic tissues in eutherian mammals. The subsequent rising of

XIST in eutherian mammals could have allowed for the establishment of random X chromosome inactivation in the somatic tissues.

Although this is an appealing hypothesis, it is not exempt from obstacles. For example, it is not clear how the epigenetic markers on the paternal X chromosome could be transmitted. Histone or DNA modifications have not been observed to be transmitted through the sperm in mammals, although contradictory results have been reported (Huynh and Lee, 2003; Okamoto et al., 2005). Likewise, there are no data relative to the activation/inactivation of sex chromosomes in the marsupial zygote and early embryos that could shed light into this issue. Finally, the inactivation of sex chromosomes is currently known to be far more complex and incomplete than previously thought in both eutherians and marsupials (Carrel and Willard, 2005; Chow and Heard, 2009). Moreover, the recent finding that the X chromosome in marsupials can be reactivated in spermatids (Mahadevaiah et al., 2009) severely compromises the hypothesis of an meiosis-inactivated X chromosome being transmitted through the sperm. Once again, more experimental evidences are necessary before this idea can be confirmed or refuted.

10.6 Future Prospects

Despite recent advances in our understanding of marsupial sex chromosome behaviour during meiosis, several questions remain to be addressed. For example, the causes of SC absence between sex chromosomes seem to be more complex that previously thought. While it is assumed that in most marsupials species X and Y chromosomes do not possess a homologous region (Hayman, 1990; Graves and Watson, 1991), this inference was made mostly from meiotic studies that described the absence of synapsis between sex chromosomes in meiosis (Solari and Bianchi, 1975; Sharp, 1982). However, when more specific studies have been developed to assess the extent of homology between sex chromosomes, it was found that sex chromosomes in several species bear NORs in both sex chromosomes (Sharp, 1982; Hayman, 1990). In fact, sex chromosomes cross-hybridise using chromosome probes in some of these species (Toder et al., 1997). Nevertheless, these homologous regions are not able to promote sex chromosome synapsis, which is only observed in those species with complex and large sex chromosomes, such as XY_1Y_2 or X_1X_2Y, or in those in which both sex chromosomes have been translocated to autosomes (Sharp, 1982). Therefore, it is likely that the program of sex chromosome behaviour includes the proscription of SC formation in the absence of an extensive and/or functional homologous region (Hayman, 1990). As such, it would be relevant to study the process of sex chromosome pairing in a wider range of species with translocated sex chromosomes in order to test if these translocated segments allow sex chromosome synapsis, and also to explore if sex chromosomes in these cases develop a DP or any other modification of their AEs.

A second issue that deserves further attention is the evolutionary origin of SC modifications that leads to the formation of the DP. The fact that many eutherian

species also lack PARs, and develop SC-derived structures to associate and segregate achiasmatic sex chromosomes (de la Fuente et al., 2007), indicates that the features displayed during marsupial meiosis might not be exclusive to this group. For example, it is possible that the absence of a PAR and the presence of SC complex-derived structures associated with sex chromosomes stem from the therian ancestor of marsupials and eutherians. Under this scenario, marsupials would have maintained the ancestral mechanism of sex chromosome pairing and segregation. This would be in agreement with other cytogenetic features of the group, such as the extreme conservation of their karyotypes and the perpetuation of the Y chromosome, even though this chromosome is extremely reduced in size and gene number. This could be considered as proof that marsupials represent a frozen pathway in mammalian sex chromosome evolution, but conversely could be interpreted in terms of the "6% solution" proposed by Kirsch (Kirsch, 1977). This hypothesis states that some of the supposed primitive features of marsupials have been conserved simply because they represent extraordinary adaptive solutions to specific biological problems. On the other hand, eutherians would have followed a different evolutionary pathway that has resulted in a more dynamic karyotypic evolution, recurrent loss and addition of PARs to the sex chromosomes, and eventual loss of Y chromosomes in some species. Further studies on the nature of sex chromosomes and sex chromosome behaviour in eutherians lacking PARs could bring interesting information to test if modifications of the SC act as an ancient backup system to segregate achiasmate chromosomes in mammals. Unfortunately, little information is expected from the study of monotremes, as their complex sex chromosome system makes it very unlikely that they could represent the model from which therian sex chromosomes evolved. The study of other groups, such as some reptiles with XY determination systems (obviously not birds), could provide data relevant to this issue.

Finally, research concerning the mechanisms and origin of sex chromosome inactivation in mammals, both during male meiosis and in the female somatic line, is now being conducted and is sure to bring exciting results in the future. The study of marsupials could be relevant to understand the universality of the mechanisms described in eutherians. Likewise, the recent characterisation of some features of somatic X chromosome inactivation in monotremes (Deakin et al., 2008; Daish and Grützner, 2009) opens new lines of interpretation about the origin of dosage compensation in mammals. The extent of MSCI has not been investigated in monotremes, but the recent discovery of MSCI in birds (Schoenmakers et al., 2009) suggests that it could be more conserved than was previously recognised. Therefore, it seems likely that this mechanism could be related to the origin of gene dosage compensation in mammals.

Research into marsupial genetics, genomics, and cell biology in coming years will aid in the understanding of the general processes that to date have only been studied in eutherians. Sequencing of marsupial genomes has already fulfilled an extraordinary goal in this effort to put marsupials at the centre of mammalian genetics (Mikkelsen et al., 2007). The study of marsupial meiosis will also contribute relevant information concerning our general understanding of meiosis and mammalian genetics and cell biology.

Acknowledgments This work was supported by FONDECYT grant 1080090 (Chile), grant BFU2009-10987 from Ministerio de Ciencia e Innovación and grant A/017762/08 from Agencia Española de Cooperación Internacional para el Desarrollo (Spain).

References

Amrine-Madsen H, Scally M, Westerman M, Stanhope MJ, Krajewski C, Springer MS (2003) Nuclear gene sequences provide evidence for the monophyly of australidelphian marsupials. Mol Phylogenet Evol 28:186–196.

Ashley T (1994) Mammalian meiotic recombination: a reexamination. Hum Genet 94:587–593.

Ashley T, Moses MJ (1980) End association and segregation of the achiasmatic X and Y chromosomes of the sand rat, *Psammomys obesus*. Chromosoma 78:203–210.

Ashley T, Jaarola M, Fredga K (1989) Absence of synapsis during pachynema of the normal sized sex chromosomes of *Microtus arvalis*. Hereditas 111:295–304.

Avner P, Heard E (2001) X-chromosome inactivation: counting, choice and initiation. Nat Rev Genet 2:59–67.

Ayoub N, Richler C, Wahrman J (1997) Xist RNA is associated with the transcriptionally inactive XY body in mammalian male meiosis. Chromosoma 106:1–10.

Baarends WM, Wassenaar E, van der Laan R, et al. (2005) Silencing of unpaired chromatin and histone H2A ubiquitination in mammalian meiosis. Mol Cell Biol 25:1041–1053.

Berrios S, Fernandez-Donoso R (1990) Nuclear architecture of human pachytene spermatocytes: quantitative analysis of associations between nucleolar and XY bivalents. Hum Genet 86:103–116.

Borodin PM, Sablina OV, Rodionova MI (1995) Pattern of X-Y chromosome pairing in microtine rodents. Hereditas 123:17–23.

Carnero A, Jimenez R, Burgos M, Sanchez A, de la Guardia R (1991) Achiasmatic sex chromosomes in Pitymys duodecimcostatus: mechanisms of association and segregation. Cytogenet Cell Genet 56:78–81.

Carrel L, Willard HF (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434:400–404.

Cooper DW (1971) Directed genetic change model for X chromosome inactivation in eutherian mammals. Nature 230:292–294.

Costa Y, Speed R, Ollinger R, et al. (2005) Two novel proteins recruited by synaptonemal complex protein 1 (SYCP1) are at the centre of meiosis. J Cell Sci 118:2755–2762.

Costa Y, Speed RM, Gautier P, et al. (2006) Mouse MAELSTROM: the link between meiotic silencing of unsynapsed chromatin and microRNA pathway? Hum Mol Genet 15:2324–2334.

Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. Heredity 95:118–128.

Chow J, Heard E (2009) X inactivation and the complexities of silencing a sex chromosome. Curr Opin Cell Biol 21:359–366.

Daish T, Grützner F (2009) Location, location, location! Monotremes provide unique insights into the evolution of sex chromosome silencing in mammals. DNA Cell Biol 28:91–100.

Das NK, Siegel EP, Alfert M (1965) Synthetic activities during spermatogenesis in the locust. J Cell Biol 25:387–395.

Davidow LS, Breen M, Duke SE, Samollow PB, McCarrey JR, Lee JT (2007) The search for a marsupial XIC reveals a break with vertebrate synteny. Chromosome Res 15:137–146.

de la Fuente R, Parra MT, Viera A, et al. (2007) Meiotic pairing and segregation of achiasmate sex chromosomes in eutherian mammals: the role of SYCP3 protein. PLoS Genet 3:e198.

Deakin JE, Hore TA, Koina E, Graves JAM (2008) The status of dosage compensation in the multiple X chromosomes of the platypus. PLoS Genet 4:e1000140.

Escalier D, Garchon HJ (2000) XMR is associated with the asynapsed segments of sex chromosomes in the XY body of mouse primary spermatocytes. Chromosoma 109:259–265.

Fernandez-Capetillo O, Mahadevaiah SK, Celeste A, et al. (2003) H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. Dev Cell 4:497–508.
Fernandez-Donoso R, Berrios S, Pincheira J (1979) Position of the nucleolus within the nuclei of pachytene spermatocytes of *Dromiciops australis* and *Marmosa elegans* (Didelphoidea-Marsupialia). Experientia 35:1021–1023.
Franco MJ, Sciurano RB, Solari AJ (2007) Protein immunolocalization supports the presence of identical mechanisms of XY body formation in eutherians and marsupials. Chromosome Res 15:815–824.
Fredga K, Bulmer MG (1988) Aberrant Chromosomal Sex-Determining Mechanisms in Mammals, with Special Reference to Species with XY Females [and Discussion]. Phil Trans R Soc London B 322:83–95.
Gallardo MH, Paterson BD (1987) An additional 14-chromosome karyotype and sex-chromosome mosaicism in South American marsupials. Fieldiana Zool 39:111–116.
Graves JAM (1995) The origin and function of the mammalian Y chromosome and Y-borne genes – an evolving understanding. Bioessays 17:311–320.
Graves JAM (1996) Mammals that break the rules: genetics of marsupials and monotremes. Annu Rev Genet 30:233–260.
Graves JAM, Watson JM (1991) Mammalian sex chromosomes: evolution of organization and function. Chromosoma 101:63–68.
Graves JAM, Wakefield MJ, Toder R (1998) The origin and evolution of the pseudoautosomal regions of human sex chromosomes. Hum Mol Genet 7:1991–1996.
Grützner F, Rens W, Tsend-Ayush E, et al. (2004) In male platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. Nature 432:913–917.
Hamer G, Gell K, Kouznetsova A, Novak I, Benavente R, Hoog C (2006) Characterization of a novel meiosis-specific protein within the central element of the synaptonemal complex. J Cell Sci 119:4025–4032.
Handel MA (2004) The XY body: a specialized meiotic chromatin domain. Exp Cell Res 296:57–63.
Handel MA, Hunt PA (1992) Sex-chromosome pairing and activity during mammalian meiosis. Bioessays 14:817–822.
Hayman DL (1990) Marsupial cytogenetics. Aust J Zool 37:331–349.
Henderson SA (1963) Differential Ribonucleic Acid Synthesis of X and Autosomes During Meiosis. Nature 200:1235.
Heyting C (1996) Synaptonemal complexes: structure and function. Curr Opin Cell Biol 8:389–396.
Hore TA, Koina E, Wakefield MJ, Marshall Graves JAM (2007) The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. Chromosome Res 15:147–161.
Hornecker JL, Samollow PB, Robinson ES, Vandeberg JL, McCarrey JR (2007) Meiotic sex chromosome inactivation in the marsupial Monodelphis domestica. Genesis 45:696–708.
Hoyer-Fender S (2003) Molecular aspects of XY body formation. Cytogenet Genome Res 103:245–255.
Hoyer-Fender S, Costanzi C, Pehrson JR (2000) Histone macroH2A1.2 is concentrated in the XY-body by the early pachytene stage of spermatogenesis. Exp Cell Res 258:254–260.
Huynh KD, Lee JT (2003) Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. Nature 426:857–862.
Huynh KD, Lee JT (2005) X-chromosome inactivation: a hypothesis linking ontogeny and phylogeny. Nat Rev Genet 6:410–418.
Jimenez R, Carnero A, Burgos M, Sanchez A, de la Guardia R (1991) Achiasmatic giant sex chromosomes in the vole Microtus cabrerae (Rodentia, Microtidae). Cytogenet Cell Genet 57:56–58.
John B, King M (1985) Pseudoterminalisation, terminalisation and non-chiasmate modes of terminal association. Chromosoma 92:89–99.

Keeney S, Giroux CN, Kleckner N (1997) Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell 88:375–384.

Kelly WG, Aramayo R (2007) Meiotic silencing and the epigenetics of sex. Chromosome Res 15:633–651.

Khalil AM, Driscoll DJ (2006) Histone H3 lysine 4 dimethylation is enriched on the inactive sex chromosomes in male meiosis but absent on the inactive X in female somatic cells. Cytogenet Genome Res 112:11–15.

Khalil AM, Boyar FZ, Driscoll DJ (2004) Dynamic histone modifications mark sex chromosome inactivation and reactivation during mammalian spermatogenesis. Proc Natl Acad Sci USA 101:16583–16587.

Kirsch JA (1977) The six-percent solution: second thoughts on the adaptedness of the marsupialia. Am Sci 65:276–288.

Kirsch JA, Dickerman AW, Reig OA, Springer MS (1991) DNA hybridization evidence for the Australasian affinity of the American marsupial Dromiciops australis. Proc Natl Acad Sci USA 88:10465–10469.

Koina E, Chaumeil J, Greaves I, Tremethick D, Graves JAM (2009) Specific patterns of histone marks accompany X chromosome inactivation in a marsupial. Chromosome Res 17:115–126.

Koller PC (1936) The genetical and mechanical properties of chromosomes. II. Marsupials. J Genet 32:452–472.

Kralewski M, Benavente R (1997) XY body formation during rat spermatogenesis: an immunocytochemical study using antibodies against XY body-associated proteins. Chromosoma 106:304–307.

Lammers JHM, Offenberg HH, van Aalderen M, Vink ACG, Dietrich AJJ, Heyting C (1994) The gene encoding a major component of the lateral elements of synaptonemal complexes of the rat is related to X-linked lymphocyte-regulated genes. Mol Cell Biol 14:1137–1146.

Lyon MF (1961) Gene action in the X-chromosome of the mouse (Mus musculus L). Nature 190:372–373.

Mahadevaiah SK, Turner JM, Baudat F, et al. (2001) Recombinational DNA double-strand breaks in mice precede synapsis. Nat Genet 27:271–276.

Mahadevaiah SK, Royo H, Vandeberg JL, McCarrey JR, Mackay S, Turner JM (2009) Key features of the X inactivation process are conserved between marsupials and eutherians. Curr Biol 19:1478–1484.

Marcon E, Babak T, Chua G, Hughes T, Moens PB (2008) miRNA and piRNA localization in the male mammalian meiotic nucleus. Chromosome Res 16:243–260.

McCarrey JR (2001) X-chromosome inactivation during spermatogenesis: the original dosage compensation mechanism in mammals? In: Xue G, Xue Y, Xu Z, Holmes R, Hammond GL, Lim HA (eds) Gene Families: Studies of DNA, RNA, Enzymes and Proteins. World Scientific, River Edge.

McKee BD, Handel MA (1993) Sex chromosomes, recombination, and chromatin conformation. Chromosoma 102:71–80.

Meuwissen RLJ, Offenberg HH, Dietrich AJJ, Riesewijk A, van Iersel M, Heyting C (1992) A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. Embo J 11:5091–5100.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Monesi V (1965) Differential rate of ribonucleic acid synthesis in the autosomes and sex chromosomes during male meiosis in the mouse. Chromosoma 17:11–21.

Motzkus D, Singh PB, Hoyer-Fender S (1999) M31, a murine homolog of Drosophila HP1, is concentrated in the XY body during spermatogenesis. Cytogenet Cell Genet 86:83–88.

Namekawa SH, Park PJ, Zhang LF, et al. (2006) Postmeiotic sex chromatin in the male germline of mice. Curr Biol 16:660–667.

Namekawa SH, VandeBerg JL, McCarrey JR, Lee JT (2007) Sex chromosome silencing in the marsupial male germ line. Proc Natl Acad Sci USA 104:9730–9735.

Nilsson MA, Gullberg A, Spotorno AE, Arnason U, Janke A (2003) Radiation of extant marsupials after the K/T boundary: evidence from complete mitochondrial genomes. J Mol Evol 57(Suppl 1):S3–S12.

Offenberg HH, Schalk JA, Meuwissen RL, et al. (1998) SCP2: a major protein component of the axial elements of synaptonemal complexes of the rat. Nucleic Acids Res 26:2572–2579.

Ohno S (1967) Sex Chromosomes and Sex Linked Genes. Springer, Berlin.

Okamoto I, Arnaud D, Le Baccon P, et al. (2005) Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. Nature 438:369–373.

Page J, Berrios S, Rufas JS, et al. (2003) The pairing of X and Y chromosomes during meiotic prophase in the marsupial species Thylamys elegans is maintained by a dense plate developed from their axial elements. J Cell Sci 116:551–560.

Page J, Berrios S, Parra MT, et al. (2005) The program of sex chromosome pairing in meiosis is highly conserved across marsupial species: implications for sex chromosome evolution. Genetics 170:793–799.

Page J, de la Fuente R, Gomez R, et al. (2006a). Sex chromosomes, synapsis, and cohesins: a complex affair. Chromosoma 115:250–259.

Page J, Viera A, Parra MT, et al. (2006b). Involvement of synaptonemal complex proteins in sex chromosome segregation during marsupial male meiosis. PLoS Genet 2:e136.

Page SL, Hawley RS (2004) The genetics and molecular biology of the synaptonemal complex. Annu Rev Cell Dev Biol 20:525–558.

Palma RE (1995) The karyotypes of two South American mouse opossums of the genus *Thylamys* (Marsupialia: Didelphidae), from the Andes and eastern Paraguay. Proc Biol Soc Wash 108:1–5.

Palma RE, Spotorno AE (1999) Molecular systematics of marsupials based on the rRNA 12S mitochondrial gene: the phylogeny of didelphimorphia and of the living fossil microbiotheriid Dromiciops gliroides thomas. Mol Phylogenet Evol 13:525–535.

Pathak S, Elder FF, Maxwell BL (1980) Asynaptic behavior of X and Y chromosomes in the Virginia oppossum and the southern pygmy mouse. Cytogenet Cell Genet 26:142–149.

Rasmussen SW (1977) The transformation of the Synaptonemal Complex into the ‘elimination chromatin’ in Bombyx mori oocytes. Chromosoma 60:205–221.

Ratomponirina C, Viegas-Pequignot E, Dutrillaux B, Petter F, Rumpler Y (1986) Synaptonemal complexes in Gerbillidae: probable role of intercalated heterochromatin in gonosome-autosome translocations. Cytogenet Cell Genet 43:161–167.

Rens W, Grützner F, O’Brien PCM, Fairclough H, Graves JAM, Ferguson-Smith MA (2004) Resolution and evolution of the duck-billed platypus karyotype with an X1Y1X2Y2X3Y3X4Y4X5Y5 male sex chromosome constitution. Proc Natl Acad Sci USA 101:16257–16261.

Rens W, O’Brien P, Grutzner F, et al. (2007) The multiple sex chromosomes of platypus and echidna are not completely identical and several share homology with the avian Z. Genome Biol 8:R243.

Roche L, Seluja G, Wettstein R (1986) The meiotic behaviour of the XY pair in *Lutreolina crassicaudata* (Marsupialia: Didelphoidea). Genetica 71:213–224.

Roeder GS (1997) Meiotic chromosomes: it takes two to tango. Genes Dev 11:2600–2621.

Rogers RS, Inselman A, Handel MA, Matunis MJ (2004) SUMO modified proteins localize to the XY body of pachytene spermatocytes. Chromosoma 113:233–243.

Romanienko PJ, Camerini-Otero RD (2000) The mouse Spo11 gene is required for meiotic chromosome synapsis. Mol Cell 6:975–987.

Schimenti J (2005) Synapsis or silence. Nat Genet 37:11–13.

Schoenmakers S, Wassenaar E, Hoogerbrugge JW, Laven JSE, Grootegoed JA, Baarends WM (2009) Female meiotic sex chromosome inactivation in chicken. PLoS Genet 5:e1000466.

Seluja G, Roche L, Solari AJ (1987) Male meiotic prophase in Didelphis albiventris. J Heredity 78:218–222.

Sharman GB (1971) Late DNA replication in the paternally derived X chromosome of female kangaroos. Nature 230:231–232.

Sharman GB (1973) The chromosomes of non-eutherian mammals. In: Chiarelli AB, Capanna E (eds) Cytotaxonomy and Vertebrate Evolution. Academic Press, London.

Sharp P (1982) Sex chromosome pairing during male meiosis in marsupials. Chromosoma 86: 27–47.

Shevchenko AI, Zakharova IS, Elisaphenko EA, et al. (2007) Genes flanking Xist in mouse and human are separated on the X chromosome in American marsupials. Chromosome Res 15: 127–136.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837.

Solari AJ (1969) The evolution of the ultrastructure of the sex chromosomes (sex vesicle) during meiotic prophase in mouse spermatocytes. J Ultrastruct Res 27:289–305.

Solari AJ (1970) The spatial relationship of the X and Y chromosomes during meiotic prophase in mouse spermatocytes. Chromosoma 29:217–236.

Solari AJ (1974) The behavior of the XY pair in mammals. Int Rev Cytol 38:273–317.

Solari AJ (1993) Sex Chromosomes and Sex Determination in Vertebrates. CRC Press, Boca Raton, FL.

Solari AJ, Ashley T (1977) Ultrastructure and behavior of the achiasmatic, telosynaptic XY pair of the sand rat (Psammomys obesus). Chromosoma 62:319–336.

Solari AJ, Bianchi NO (1975) The synaptic behaviour of the X and Y chromosomes in the marsupial Monodelphis dimidiata. Chromosoma 52:11–25.

Spotorno AE, Marin JC, Yévenes M, et al. (1997) Chromosome divergences among American marsupials and the Australian affinities of the American Dromiciops. J Mamm Evol 4:259–269.

Svartman M, Vianna-Morgante AM (1998) Karyotype evolution of marsupials: from higher to lower diploid numbers. Cytogenet Cell Genet 82:263–266.

Szalay FS (1994) Evolutionary History of Marsupials and an Analysis of Osteological Characters. Cambridge University Press, Cambridge, MA.

Toder R, Wienberg J, Voullaire L, O'Brien PC, Maccarone P, Graves JAM (1997) Shared DNA sequences between the X and Y chromosomes in the tammar wallaby – evidence for independent additions to eutherian and marsupial sex chromosomes. Chromosoma 106:94–98.

Toder R, Wakefield MJ, Graves JAM (2000) The minimal mammalian Y chromosome – the marsupial Y as a model system. Cytogenet Cell Genet 91:285–292.

Tres LL (2005) XY chromosomal bivalent: nucleolar attraction. Mol Reprod Dev 72:1–6.

Turner JM (2007) Meiotic sex chromosome inactivation. Development 134:1823–1831.

Turner JM, Aprelikova O, Xu X, et al. (2004) BRCA1, histone H2AX phosphorylation, and male meiotic sex chromosome inactivation. Curr Biol 14:2135–2142.

Turner JM, Mahadevaiah SK, Fernandez-Capetillo O, et al. (2005) Silencing of unsynapsed meiotic chromosomes in the mouse. Nat Genet 37:41–47.

Turner JM, Mahadevaiah SK, Ellis PJ, Mitchell MJ, Burgoyne PS (2006) Pachytene asynapsis drives meiotic sex chromosome inactivation and leads to substantial postmeiotic repression in spermatids. Dev Cell 10:521–529.

van der Heijden GW, Derijck AA, Posfai E, et al. (2007) Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. Nat Genet 39:251–258.

Veyrunes F, Waters PD, Miethke P, et al. (2008) Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res 18:965–973.

von Wettstein D, Rasmussen SW, Holm PB (1984) The synaptonemal complex in genetic segregation. Annu Rev Genet 18:331–413.

Wahrman J, Nezer R, Freund O (1973) Multiple chromosome mechanisms with 'segregation bodies'. Chromosome Today 4:434.

Wakefield MJ, Keohane AM, Turner BM, Graves JAM (1997) Histone underacetylation is an ancient component of mammalian X chromosome inactivation. Proc Natl Acad Sci USA 94:9665–9668.
Wallis M, Waters PD, Delbridge ML, et al. (2007) Sex determination in platypus and echidna: autosomal location of SOX3 confirms the absence of SRY from monotremes. Chromosome Res 15:949–959.
Wallis M, Waters PD, Graves JAM (2008) Sex determination in mammals – Before and after the evolution of SRY. Cell Mol Life Sci 65:3182–3195.
Watson CM, Margan SH, Johnston PG (1998) Sex-chromosome elimination in the bandicoot Isoodon macrourus using Y-linked markers. Cytogenet Cell Genet 81:54–59.
Wolf KW (1994) How meiotic cells deal with non-exchange chromosomes. Bioessays 16:107–114.
Wolf KW, Baumgart K, Winking H (1988) Meiotic association and segregation of the giant sex chromosomes in male field vole (Microtus agrestis). Chromosoma 97:124–133.
Zickler D, Kleckner N (1998) The leptotene-zygotene transition of meiosis. Annu Rev Genet 32:619–697.
Zickler D, Kleckner N (1999) Meiotic chromosomes: integrating structure and function. Annu Rev Genet 33:603–754.

Chapter 11
Compact but Complex – The Marsupial Y Chromosome

Veronica J. Murtagh, Paul D. Waters, and Jennifer A. Marshall Graves

Abstract The mammalian Y chromosome is a gene poor element with enormous responsibility. Just how the mammalian Y chromosome acquired the function of sex determination, then surrendered most of its other functions, remains one of the most intriguing questions in genomics. The basic marsupial Y chromosome represents a degraded relic of the original mammalian Y, which did not receive the large autosomal addition that augmented the X and Y of eutherian mammals. Like the Y of other mammals, the marsupial Y contains species-specific repetitive sequences. It also bears an interesting mixture of genes; ancient male-specific genes, such as *SRY* and *RBMY*, which evolved from X-borne genes before the eutherian-marsupial divergence, and are conserved across therian mammals. The marsupial Y also bears novel genes such as *ATRY*, that in eutherian mammals have been lost from the Y but maintain a copy on the X. The unique properties of the marsupial Y chromosome have provided insights into the evolution of the mammalian Y, enabling us to gain a better understanding of the selection processes that shaped it.

Keywords Comparative genomics · Evolution · Mammals · Marsupials · Y chromosome

11.1 Introduction

The mammalian Y chromosome has two distinguishing properties; the absence of recombination (or its restriction to a small region of the chromosome), and a highly reduced number of active genes in the remaining (non-recombining) regions of the chromosome (Rice, 1996). These unique properties are the result of millions of years of evolution which have produced a Y chromosome that is small, highly heterochromatic, and highly specialized for functions in male sex and reproduction

V.J. Murtagh (✉)
Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia
e-mail: Veronica.Murtagh@anu.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_11, © Springer Science+Business Media B.V. 2010

(Graves, 2006). Since its origin as one member of a pair of autosomes, the mammalian Y chromosome appears to have undergone a rapid degradation, a result of the acquisition of a dominant male-determining gene followed by suppression of recombination between the X and Y (Graves, 2006).

11.2 Heterogametic Sex Chromosomes

Sex determination in vertebrates is particularly fascinating because of the variety of signals that direct male or female development. Sex determination in many reptiles and some fish is directed by environmental signals (usually temperature, TSD), but in mammals, birds, amphibians, many reptiles and most fish, it is determined genetically (GSD). Amongst vertebrates with genetic sex determination, some display male heterogamety (XX females: XY males) and others female heterogamety (ZW females: ZZ males) (that is either males or females produce two types of gamete). In many species, chromosomes that bear the sex determining gene are morphologically distinct, although most fish, frogs and many reptiles have indistinguishable sex chromosomes.

H. J. Muller (Muller, 1914; Muller, 1918) first suggested that sex chromosomes evolved from a pair of autosomes, retaining the same complement of genes but no longer possessing the ability to cross over with each other in at least part of the chromosome. Muller proposed that this restriction of recombination would result in the accumulation of recessive lethal or deleterious mutations, ultimately resulting in the degeneration of the Y chromosome (Muller, 1914). Ohno (1967) later developed this idea to explain the differences in ZW differentiation in different families of snake. The Z and W in the *Boidae* are homomorphic, whereas the Z and W of members of the *Colubridae* may differ largely in rearrangement, and vipers have highly differentiated Z and W chromosomes (Ohno, 1967). Ohno suggested that the genetic deterioration displayed by the W chromosome is the consequence of the accumulation of genes that govern female development, with the Z chromosome retaining its original gene composition (Ohno, 1967). Ohno proposed that this "differential accumulation" could only occur through the establishment of reduced recombination, which acts to preserve the accumulated female genes on the W chromosome. Birds, too, have a ZZ male ZW female system, in which different families have identical Z chromosomes, but W chromosomes that differ in size from virtually identical to the Z in flightless ratites, to the much reduced W of chicken and other carinate birds.

The differentiation of the mammal X and Y chromosome is thought to have taken a parallel course (Charlesworth, 1996; Graves, 2006). In species with an XY system, male-advantage alleles accumulated around the sex-determining locus, and recombination was suppressed to preserve the sex specificity of this region. Over time, mutation and deletion in this non-recombining region lead to degradation of the sex-specific chromosome, ultimately leading to the dimorphic sex chromosomes of extant mammals (Fig. 11.1). Dramatic confirmation of this hypothesis is provided by

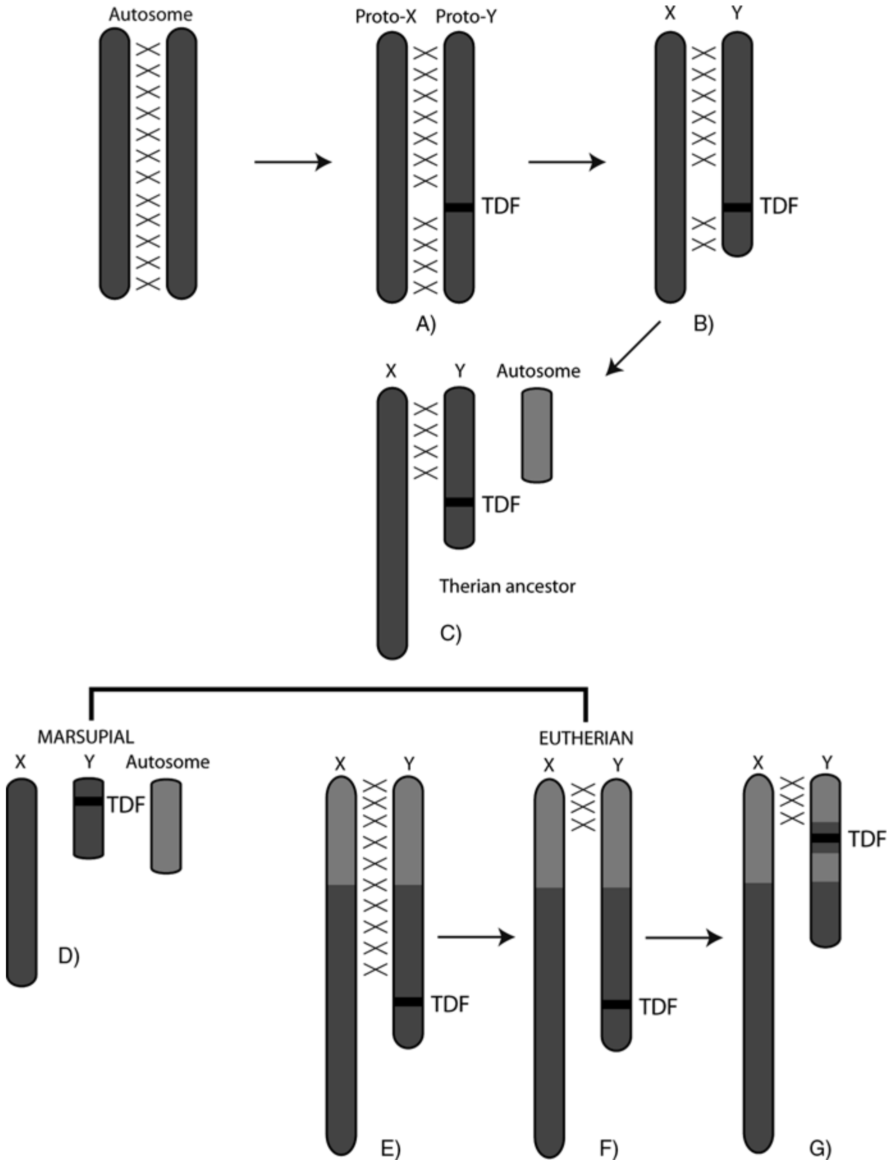


Fig. 11.1 Process of Y degradation in mammals. The X and Y began as a homologous pair of autosomes until one member of the pair acquired a testis determining locus (TDF; *black bar*). (**a**) Male advantage genes accumulate around the TDF, across which recombination was suppressed between the X and Y, leading to a region of the Y that was only inherited through males. (**b**) Recombination between the X and Y chromosomes was suppressed further, the Y chromosome began to lose functional genes and degrade. (**c**) In the therian ancestor >148 MYA the sex chromosomes were probably represented by a degraded Y only recombining with the X within a pseudoautosomal region (PAR). (**d**) In marsupials the Y chromosome further degraded from this ancestral therian state and all recombination between the X and Y chromosomes was eventually suppressed. (**e**) In eutherians, autosomal material (*light grey*) was translocated onto the sex chromosomes. (**f**) Recombination between the X and Y in this newly added region was suppressed, leaving a small PAR shared by the X and Y chromosomes. (**g**) The Y continued to rapidly degrade, with loss of functional genes and multiple rearrangements

the recent evidence that all the genes on the human and marsupial X are autosomal in platypus as well as chicken, implying that the therian mammal X and Y began their differentiation from an autosome pair a mere 166 million years ago (MYA) (Veyrunes et al., 2008; Smith et al., 2009).

The independent evolution of heteromorphic sex chromosomes in a wide range of taxa presents fascinating parallels in cytological and genetic characteristics across different orders (Charlesworth, 1996). In cases of extreme sex chromosome heteromorphism, the X and Y, or the Z and W, may differ in size and morphology, as well as content of genes (Bull, 1983). These characteristics also differ between the sex chromosomes and autosomes. In animal orders with male heterogametic systems, the X chromosome is typically larger than the Y, or at least differs in morphology (Bull, 1983) as a by-product of the degeneration of the sex-specific chromosome. In the absence of the genomic studies, which we take for granted these days, the best evidence cited by Bull for the absence of genes on the Y was the observation of a large number of X-linked traits, compared to very few known Y-linked traits. These days with new molecular techniques available, in particular sequencing technologies, we have a greater understanding of the genetic differences between the X and Y chromosomes, especially in their gene content.

The paucity of genes on the Y chromosome suggests that the Y represents a degraded form of the X chromosome (the corollary is that the X contains many functional genes that have been lost from the Y chromosome). Early evidence supporting this conclusion was based upon observations of XX/XO systems and viable XO (but not YO or YY) individuals. We now have much evidence for homology between the X and Y; for instance the pseudoautosomal regions (PAR) that are shared between them, and the homology of most Y-borne genes to an X-borne partner (Graves, 2006).

11.2.1 Sex Chromosomes of Birds and Mammals

There is now much evidence suggesting that the bird ZW system and mammalian XY system evolved independently from two different autosome pairs. Comparative gene mapping shows that the therian mammal XY pair and the bird ZW pair share no genes (Fridolfsson et al., 1998; Nanda et al., 1999; Graves and Shetty, 2001) and contain different sex determining genes. In mammals, male determination is triggered by the Y-borne *SRY* gene (mouse, Koopman et al., 1990; human, Sinclair et al., 1990), whereas in birds dosage of the Z-borne *DMRT1* gene (2x in males and 1x in females) is now known to be the trigger (Smith et al., 2009).

The basal monotreme mammals seem to be an exception. The absence of *SRY* in monotremes, and autosomal mapping of the gene from which it evolved (*SOX3*) (Wallis et al., 2007), implies that a different gene is sex determining in platypus and echidna. The mapping of *DMRT1* (which lies on the Z chromosome in chicken) to an X chromosome in platypus (one of a complex of ten sex chromosomes that form a chain at male meiosis; Grützner et al., 2004) was proposed to represent an evolutionary link between the mammalian and bird sex chromosomes.

However, orthologues of genes on the XY pair of therian mammals are autosomal in platypus (Veyrunes et al., 2008), which places the origin of the therian X and Y chromosomes somewhere between 166 and 148 MYA, after therians diverged from the monotremes, but before the marsupial/eutherian split. It is possible that the *SRY*-driven therian XY system took over from an ancient bird-like ZW system soon after the divergence of monotremes from therian mammals 166 MYA (Ezaz et al., 2006; Graves, 2008).

11.2.2 General Properties of Y Chromosomes

Y chromosomes in diverse taxa typically share three distinguishing properties; presence only in males, recombination with the X is often restricted to a small region of the chromosome, and highly reduced genetic activity in the remaining regions of the chromosome (Rice, 1996).

The independent evolution of a Y chromosome from an autosome in many different animal and plant lineages suggests that similar evolutionary forces are at work (Bull, 1983; Charlesworth, 1996). Comparisons between the Y chromosomes in different lineages have revealed different stages of Y degradation from a common beginning; the appearance of a male-determining locus or set of closely-linked loci. However, the extent of XY differentiation is not merely a reflection of the time since the acquisition of a male-determining gene by the proto-Y, as evidenced by the relatively undifferentiated sex chromosomes of ratite birds in contrast to the highly differentiated W chromosome in chicken and other carinate birds. Similarly, boid snake species have homomorphic Z and W chromosomes, whereas the Elapidae and Viperidae have highly degenerated, heterochromatic W chromosomes. Young male heterogametic systems are characterised by the absence of gross structural differences between the sex chromosomes (Charlesworth, 1996), with the Y chromosome in these systems possessing a sex determining factor. As time progresses, chromosomal rearrangements begin to prevent recombination in the sex-determining region and morphological differences between the sex chromosomes emerge. This sex determination stage is represented by the three-spine stickleback (*Gasterosteus aculeatus*) in which X and Y chromosomes are differentiated over only a tiny segment (Peichel et al., 2004), and in Medaka fish (*Oryzias latipes*), in which a retroposed *DMRT1* gene has established a new Y chromosome with a small differential segment corresponding to a region on the proto-X with reduced recombination frequencies (Kondo et al., 2006). Ultimately the suppression of recombination spreads to cover all or most of the length of the Y, which results in degradation and it becoming genetically very different from its partner – the X chromosome. The human Y chromosome (Skaletsky et al., 2003), gene-poor and riddled with non-coding repetitive elements, represents an advanced stage of Y chromosome evolution.

Numerous theories have been proposed to explain the mechanisms by which the Y chromosome evolves and degrades. One theory is that the suppression of recombination resulted in progressive degradation through higher variation, drift

and inefficient selection. The higher variation is predicted by the observation of male-biased mutation in humans, thought to result from the increased chances of mutation during the many mitoses of spermatogenesis and the oxidative environment and lack of DNA repair of the sperm (Aitken and Graves, 2002). The drift hypothesis ("Muller's ratchet") proposes that in a population of Y chromosomes, less mutated Ys that are lost by chance can never be regenerated through recombination. Inefficient selection occurs when a deleterious gene "hitchhikes" to fixation on the same Y as a male-advantage gene, or an advantageous gene is brought down by its association with deleterious genes.

The "sexually antagonistic genes hypothesis" is based on the theory that mutations conveying an advantage to the heterogametic sex could be unfavourable, even deleterious, in the homogametic sex (Rice, 1996). Genetic theory predicts the accumulation of such genes only when there is a greater benefit to one sex than there is disadvantage to the other. In the situation where there is tight linkage to the sex-inducing allele, sex-biased gene transmission can lead to accumulation of sexually antagonistic alleles without great disadvantage to the homogametic sex (Bull, 1983; Rice, 1987). Such a mechanism has been proposed as the major cause of the disproportionate accumulation of male-advantage genes in developing Y chromosomes.

11.3 Sex Chromosomes of Therian Mammals

The sex chromosomes of therian mammals are highly dimorphic, with a large, gene rich X and a small, heterochromatic Y (Graves, 2006). In eutherian mammals, the X chromosome is virtually identical in gene content between species, whereas the Y chromosome varies greatly in gene content, representing different (but overlapping) subsets of genes on the X (reviewed in Waters et al., 2007). There is great interest in the identification and characterisation of genes on the Y chromosome, but the highly-repetitive nature of the human Y chromosome has made sequencing difficult. In eutherian mammals a region of homology between the X and Y is still retained as the pseudo-autosomal region (PAR), which enables pairing and correct segregation at meiosis. The marsupial X chromosome is significantly smaller than the eutherian X, and the Y is often minute and no longer able to pair with the X, raising questions about the mechanism of segregation during male meiosis (see Fernández-Donoso et al., Chapter 10).

11.3.1 Sex Chromosomes of Eutherian Mammals

The Y chromosome of eutherian mammals ranges in size as illustrated by the characterised Ys of chimpanzee (24 Mb), human (60 Mb) and mouse (95 Mb). It is highly heterochromatic, and full of repetitive sequences that are poorly conserved, even between closely related species, so that numerous comparative

chromosome-painting studies resulted in no hybridization between Y chromosomes of different species (Goureau et al., 1996; Sherlock et al., 1996; Wienberg et al., 1997; Breen et al., 1999; Müller et al., 1999; Yang et al., 2000).

The advent of genome sequencing has not greatly advanced research on the Y chromosome because the highly repetitive nature of Ys makes sequence assembly extremely difficult. Virtually all genome sequencing studies have used female material, and detailed studies of the Y have been conducted only on human, mouse and chimpanzee Y chromosomes (Rozen et al., 2003; Skaletsky et al., 2003; Toure et al., 2005). Genes on the 94.7 Mb mouse Y are all concentrated in the tiny short arm, with the long arm being largely composed of repetitive sequences, although it does contain amplified genes (Burgoyne, 1992; Toure et al., 2004) and a terminal PAR. The 24 Mb chimpanzee Y has been explored in detail and compared with the human Y (Hughes et al., 2005; Kuroki et al., 2006), revealing similarities in structure and gene content.

11.3.1.1 Human Sex Chromosomes

Human sex chromosomes are highly differentiated. The 165 Mb human X (Ross et al., 2005) contains 1,529 genes (NCBI database, http://www.ncbi.nlm.nih.gov), in contrast to the much smaller 60 Mb human Y chromosome. Sequencing of the male-specific region of the human Y (Skaletsky et al., 2003) showed that it contains many simple repetitive elements, large stretches of complex repeats, and only 172 transcribed units. Many of these are untranslatable pseudogenes, and others are present in multiple copies, so that these sequences encode only 27 unique proteins. At least two of these genes are the result of transposition from other chromosomes, with the apparent insertion of autosomal copies of *DAZ* (Saxena et al., 1996) and retrotransposition of *CDY* (Lahn and Page, 1999) onto the Y.

The most dramatic finding from the BAC-based sequencing of the human Y chromosome (Skaletsky et al., 2003) was the revelation that the long arm of the human Y contains at least eight large “ampliconic” regions composed of highly symmetrical palindrome structures, which are apparently maintained by intra-chromosomal gene conversion. Comparisons with closely related primates (chimpanzee, bonobos and gorillas) indicate that most of these palindromic structures were already present in the common ancestor of humans and chimpanzees, also showing near uniformity of arm-to-arm sequences, suggesting that a steady-state balance exists between mutations in the palindrome and the gene-conversion events erase differences (Rozen et al., 2003). Although it was suggested by these authors that gene conversion could serve as a means of correcting mutations in one arm of a palindrome, it has been pointed out that conversion to the mutant form is equally probable (Graves, 2004), and may account for the tracts of pseudogenes such as those in the *RBMY* cluster.

The human X and Y chromosomes share a small region of homology, the pseudoautosomal region (PAR1) at the termini of the X and Y short arms, within which an obligatory recombination event occurs during male meiosis that mediates X and Y segregation. This PAR has a critical role in male fertility, since deletion of PAR1 in humans results in the failure of XY pairing (Mohandas et al., 1992). This critical

function seems not to depend on the gene content of the PAR in eutherians, which varies between species (Graves et al., 1998). There is also a second small PAR (PAR2) at the ends of the long arms of the human X and Y (Freije et al., 1992). PAR2 is not essential for X and Y segregation during male meiosis, and its absence from the sex chromosomes of chimpanzee and macaque implies that it was created more recently in human evolution (Bickmore and Cooke, 1987; Kvaloy et al., 1994; Li and Hamer, 1995; Kuhl et al., 2001).

11.4 Marsupial Y Chromosomes

The different gene contents of Y chromosomes are the result of independent degradation in different lineages. Marsupials (pouched mammals that give birth to altricial young) diverged from eutherian mammals 148 MYA, so there is special value in examining the differences between Y chromosomes that have been separated for such long evolutionary periods. Looking at the gene content of marsupial Y chromosomes will provide insight into the origin and fate of Y genes.

11.4.1 Marsupial Sex Determination

As for other species with male heterogamety, the marsupial sex chromosomes segregate at male meiosis into X and Y bearing sperm, and which sperm fertilizes the egg will determine whether the young develops as either male or female. However, there are significant differences between marsupial and eutherian meiosis and sex determination.

In contrast to eutherian mammals, the marsupial X and Y chromosomes do not share a PAR, so XY paring during male meiosis occurs in the absence of synaptonemal complex (SC) formation (see Fernández-Donoso et al., Chapter 10). In the absence of a PAR, the condensed X and Y remain separate until mid-pachytene of meiosis I, where they form end to end attachments without synapsis or crossing over (Solari and Bianchi, 1975; Sharp, 1982). Studies of male meiosis in American marsupials have revealed that sex chromosomes develop axial elements that associate by the formation of a dense plate, which attaches to the ends of the X and Y (Solari and Bianchi, 1975; Roche et al., 1986; Seluja et al., 1987). This dense plate in late prophase and metaphase I is composed of SCP3 and SCP1 proteins (Page et al., 2003; Page et al., 2005; Page et al., 2006), two of the constitutes of the tripartite synaptonemal complex in eutherian mammals. The X and Y chromosomes segregate in an organized way into female-(X bearing) and male-(Y bearing) determining sperm.

In eutherian mammals, the phenotypes of XO females and XXY males demonstrated that the presence of a Y chromosome determines maleness, which results from the presence of a dominant testis determining factor (TDF) on the Y that activates the testis-determining pathway. Hormones produced by the differentiated

eutherian testis control all other aspects of male development (Wilson et al., 1981). However, in marsupials, the role of the Y chromosome as the male-dominant sex determiner is somewhat equivocal. The Y chromosome also controls testis determination in marsupials since XXY animals have testes and XO animals lack testes. However, differentiation of the embryonic testis does not hormonally control all aspects of sexual differentiation, since XXY animals have a pouch with mammary glands and lack a scrotum, whereas XO animals lack a pouch and possess an empty scrotum (Sharman et al., 1990). Thus development of the mammary glands and scrotum is androgen-independent in marsupials (Shaw et al., 1988; Shaw et al., 1990), and must rely on either the dosage, or imprinting, of a gene on the X chromosome (Cooper et al., 1993). Refer to Pask and Renfree (Chapter 14) for more details on sex differentiation in marsupials.

In eutherian mammals the TDF has been identified as the *SRY* gene. The marsupial Y contains a homologue of *SRY*, but there is no direct evidence that it is sex determining, and, intriguingly, there are other candidates among the genes on the marsupial Y. These marsupial-specific differences are at the core of research into the content and function of genes on the marsupial Y, raising fundamental questions about the mechanisms behind sex determination and differentiation in marsupials, how they vary from eutherian mechanisms, and how both these systems evolved from a common ancestral system.

11.4.2 Genes on the Marsupial Y

Genes on the marsupial Y are particularly interesting as they provide us with an insight into the origin of the Y chromosome, as well as the process by which genes evolve novel functions (Graves, 1995). The marsupial Y chromosome bears only a small number of genes (Fig. 11.2), many of which have important roles in sex determination and differentiation. All of the known marsupial Y-borne genes have copies on the X chromosome, from which they differentiated.

Genes on the Y chromosome in mammals have been categorised as either disposable, as evidenced by their degradation compared to their homolog on the X, or selectable, having a function that provides a male-advantage, probably as a result of a direct or indirect role in sex determination or differentiation (Graves, 1995). Given the homology of the marsupial X chromosome to the ancient region of the X in eutherian mammals, genes on Y in both lineages had a common evolutionary origin. The retention of a gene on the Y in both marsupials and eutherians suggests that it has an important male-specific function, because it must have been independently retained over vast evolutionary time. Genes that are only on the eutherian Y are likely to be more recently differentiated from their X counterparts, and derived from the recent addition of an autosome to the eutherian sex chromosomes.

To date, just eight genes have been found to lie on the marsupial Y chromosome, five of which are conserved on the Y of at least some other therian mammals (Fig. 11.2). These five genes; *SRY*, *RBMY*, *KDM5D*, *RPS4Y* and *UBE1Y*, have been

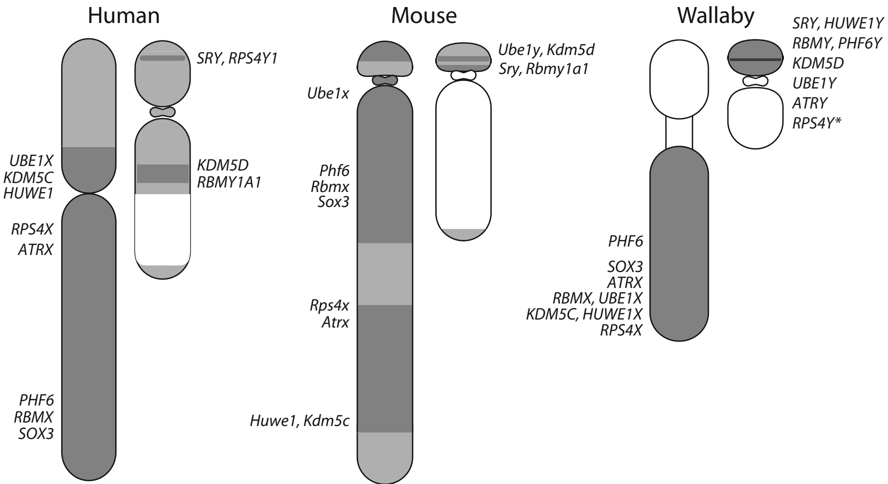


Fig. 11.2 Location of tammar wallaby (*Macropus eugenii*) Y genes and their X chromosome homologues, which are compared to their orthologues in human and mouse. The eight gene pairs shown here are: *ATRX/Y*, *HUWE1X/Y*, *KDM5C/D*, *PHF6X/Y*, *RBMX/Y*, *RPS4X/Y*, *SOX3/SRY* and *UBE1X/Y*. Shown in *grey* are the conserved regions of the X and Y chromosomes, with *light grey* representing the recently added region of the eutherian sex chromosomes. White indicates heterochromatin. The *asterix* next to *RPS4Y* indicates that there is no evidence it is present on the Y in Australian marsupials, (but there is for its location on the Y in Ameridelphia)

localised to the Y chromosome in a variety of eutherian mammals. The remaining three genes; *ATRY*, *HUWE1Y* and *PHF6Y* are located on the Y chromosome in marsupials but not eutherian mammals. X-borne paralogues for all eight genes are found on both the eutherian and marsupial X chromosomes.

11.4.2.1 *SRY*

The *SRY* gene (sex-determining region Y) was first discovered on the human Y chromosome (Sinclair et al., 1990), and subsequently a homologue was identified on the mouse Y (Gubbay et al., 1990). Mutation analysis in sex reversed humans, and transgenesis in mouse, proved it to be the testis determining factor in eutherian mammals (Berta et al., 1990; Koopman et al., 1990). Transcription of *SRY* in the genital ridge during early embryogenesis correlates with the initiation of Sertoli cell differentiation in a variety of eutherian species, including sheep (Payen et al., 1996), pigs (Daneau et al., 1996) and dogs (Meyers-Wallen, 2003). Expression of *Sry* is most specific in rodents, being confined to a narrow window prior to sexual differentiation (Koopman et al., 1990), whereas *SRY* expression in humans is maintained into adulthood (Hanley et al., 2000; reviewed by Wilhelm and Koopman, 2006).

The presence of *SRY* in marsupials (Foster et al., 1992) is further evidence of its important role in sex determination. However, the impracticality of mutation or transgenesis studies in marsupials means that there is no direct evidence that *SRY* is the marsupial testis determining gene.

At the core of *SRY* is an 80 amino acid motif, which shares homology with the High Mobility Group proteins, and hence, has been termed the HMG box (Sinclair et al., 1990). Paralogues of *SRY* have been grouped together into families of *SOX* (SRY-like HMG box containing) genes. The original *SRY* study in mouse identified four related genes (*Sox1–4*) present in males and females (Gubbay et al., 1990), which were shown to have homologues in birds (Griffiths, 1991). An *SRY*-related sequence was discovered on the marsupial X chromosome, and this gene (identified as a homologue to the mouse *Sox3*) was shown to be the SOX family member with the closest homology to *SRY* (Foster and Graves, 1994). Marsupial *SOX3* was therefore proposed to be the X chromosome partner of *SRY* and to represent the original gene from which *SRY* evolved. This was among the earliest evidence supporting the hypothesis that mammalian sex chromosomes evolved from a homologous autosome pair.

The pattern of conservation of the *SOX3/SRY* gene pair gives a particularly intriguing insight into Y chromosome degeneration. Cross-species comparisons between *SRY* sequences in eutherian mammals (human, mouse) and marsupials (tammar wallaby – *Macropus eugenii* and striped-face dunnart – *Sminthopsis macroura*) showed that sequence homology is conserved only within the HMG box, and homology drops down to less than 50% outside of the HMG box, even between the two marsupial species (Mitchell et al., 1992). Mouse *Sry* is poorly conserved with human *SRY*, and contains a rodent-specific 3' tail that codes for a polyglutamine region (Koopman et al., 1990). In contrast, *SOX3* is highly conserved across vertebrates. Expression of *SOX3* is detected in the immature ovaries of the toad *Xenopus laevis* (Koyano et al., 1997) and throughout gonadogenesis in male and female chicken (Smith et al., 1999), suggesting that expression of *SOX3* in the urogenital ridge is a conserved vertebrate trait.

Some exceptional rodents lack *SRY*. Among three species of mole vole, two have no Y; both sexes are XO in *Ellobius lutescens* and both are XX in *Ellobius tancrei* (Just et al., 1995). The XO Spiny Rat (*Tokudaia osimensis*), lacks *SRY* (Soullier et al., 1998), and ZOO-FISH with a mouse X chromosome probe suggests that the functional part of the Y chromosome has been translocated onto distal Xq in *Tokudaia* (Arakawa et al., 2002). These exceptions clearly show that male development can occur in the absence of a Y chromosome and *SRY*.

The mechanism by which *SRY* triggers testis differentiation has been long debated. Several genes have been identified in the human sex-determining pathway by their mutation in sex-reversed individuals. The first was *SOX9*, an autosomal relative of *SOX3* and *SRY*, whose deletion or translocation caused male to female sex reversal (Foster et al., 1994), and which appears to be key to Sertoli cell differentiation in all mammals. Its role as a main player in sexual development was confirmed by demonstrating that abnormal *Sox9* expression in XX gonads led to male development in the absence of *SRY* (Huang et al., 1999; Vidal et al., 2001). Most recently,

it has been proposed that *SRY* functions as a genetic switch, cooperating with *SF1* (steroidogenic factor 1) to up-regulate *Sox9* (Sekido and Lovell-Badge, 2008).

11.4.2.2 *RBMY*

RBMY was first suspected to play an important role in spermatogenesis when it was identified in a region of the long arm of the human Y chromosome deleted in men with azoospermia (Ma et al., 1993). *RBMY* is present in many copies on the human Y chromosome as part of palindromic expansions, with 15 copies clustered at Yq11.23 in humans (Chandley and Cooke, 1994). *RBMY* is expressed specifically in spermatogonia and early primary spermatocytes in adult males, as well as fetal and prepubertal germ cells, suggesting a role in germ cell development (Elliott et al., 1997). The discovery of a homologue of an *RBMY* on the Y chromosome in several marsupial species (Delbridge et al., 1997), suggested that it has an ancient and important male-specific role in all therian mammalians. The later discovery of an X-borne homologue of *RBMY* in marsupials, then in human (Delbridge et al., 1999) and mouse (Mazeyrat et al., 1999), challenged the theory that *RBMY*, and other male-specific genes on the Y, had autosomal origins. The evidence now implied that, like *SRY* and *SOX3*, *RBMY* and *RBMX* evolved from a gene on the mammalian proto-X and proto-Y pair (Delbridge et al., 1999). *RBMX* retained a widespread function and *RBMY* gained a male-specific function in spermatogenesis. These new findings lent weight to the theory that most Y-borne genes are remnants of homology between the X and Y chromosomes (Graves, 1995).

Like *SRY, RBMY* is poorly conserved between lineages, so that marsupial and human *RBMY* share more homology with *RBMX* than to each other. However, within the marsupial lineage *RBMY* is more highly conserved, with a cDNA from the tammar wallaby able to detect male-specific copies of *RBMY* in four other marsupial species (Brown Antechinus – *Antechinus stuartii*, South American opossum – *Monodelphis domestica*, Northern brown bandicoot – *Isoodon macrourus*, and brush-tail possum – *Trichosurus vulpecula*) (Delbridge et al., 1997). A remarkable observation was the apparent amplification of sequences in the bandicoot and brush-tailed possum, which contrasts with the single copy in antechinus, tammar and opossum. This implies that *RBMY* was originally present as a single copy on the Y chromosome of an ancestral therian mammal that was amplified and rearranged independently in at least three marsupial lineages, as well as in eutherian mammals, resulting in the observed interspecific *RBMY* variation.

11.4.2.3 *KDM5D*

KDM5D (lysine (K)-specific demethylase 5D) was first isolated from the deletion interval 2 (ΔSxr^b) of the mouse Y (Sutcliffe and Burgoyne, 1989; Agulnik et al., 1994a). Deletion of this ~250 kb region of the mouse Y, which also contains *Ube1y* and *Zfy*, affects spermatogenesis. A homologue on the X chromosome (*KDM5C*) was identified after a non-stringent cDNA library screen (Agulnik et al., 1994b). Originally named *Smcx* (*KDM5C*) and *Smcy* (*KDM5D*); expression of *KDM5C*

and *KDM5D* appears ubiquitous, with *KDM5C* detected even at the two-cell stage of pre-implantation mouse embryos, indicating an important housekeeping role (Agulnik et al., 1994a). Southern zoo blot analysis of the *KDM5C* cDNA showed male specific bands in human, horse, pig, dog and kangaroo, demonstrating its conservation on the Y throughout therian mammals. Although no male specific band has been detected in cattle or guinea pig, Southern blotting by Jegalian and Page detected a copy of *KDM5D* in rabbit, but reported that little or no *KDM5D* is expressed (Jegalian and Page, 1998).

Uniquely, the X-borne *KDM5C* appears to escape X-inactivation in most eutherians (Agulnik et al., 1994b; Jegalian and Page, 1998) except cattle, guinea and rabbit, the species in which the Y-borne *KDM5D* is absent or under-expressed, suggesting that the X and Y copies are functionally interchangeable (Agulnik et al., 1999). High homology is observed between *KDM5C* and *KDM5D*, aside from the last exon, coding for the C-terminal end of the protein, which shows only 32% protein identity (Agulnik et al., 1999). The presence of a zinc finger domain in both the X and Y copies indicates possible involvement in DNA binding and transcriptional regulation, with several splice variants of the transcripts detected in mouse and human (Agulnik et al., 1999). Recently, mutations in the X-borne *KDM5C* leading to X-linked mental retardation have been cited as evidence that *KDM5C* plays an important role in human brain function (Jensen et al., 2005).

Interestingly, *KDM5D* maps to the same region of the human Y chromosome as the locus controlling expression of the histocompatibility antigen H-Y, which was originally credited with a sex-determining role, but later eliminated by gene mapping (McLaren et al., 1988). The widespread expression of *KDM5C* and *KDM5D*, along with differences between the two, could produce the H-Y male-specific minor antigen system. *KDM5D* was later shown to encode several H-Y antigen epitopes (Agulnik et al., 1999).

Unlike other genes on the Y which are highly amplified and degenerate, *KDM5D* is present as a single active full-length copy on the mouse and human Y chromosomes (Agulnik et al., 1994a). In contrast, when *KDM5C* and *KDM5D* were isolated and FISH mapped in tammar wallaby, the *KDM5D* probe mapped to two regions on the wallaby Y chromosome, suggesting multiple Y borne copies (Waters et al., 2001).

11.4.2.4 *UBE1Y*

Ube1 (ubiquitin-activating enzyme E1), was also first isolated from the mouse ΔSxr^{b} deletion interval. *UBE1Y* has been identified in a range of eutherian mammals, excluding primates (Kay et al., 1991; Mitchell et al., 1991). Mouse *Ube1y* displays 85% nucleic acid identity to its X homologue (*Uba1*), orthologues of which have been identified in human and marsupials (Mitchell et al., 1991; Mitchell et al., 1992).

The testis-specific expression of *Ube1y* contrasts greatly to the ubiquitous expression of its X-borne homologue (*Uba1*). The essential role *Uba1* plays in eukaryotic cell proliferation made *Ube1y* a strong candidate for a spermatogenesis factor on

the Y chromosomein mice (Mitchell et al., 1991; Odorisio et al., 1996). In particular, it was a candidate for the spermatogonial proliferation factor (*Spy*) on the Y, this is until identification of *Eif2s3y* (a subunit of the eukaryotic translation initiation factor *Eif 2*) as the factor necessary to overcome the spermatogenic block that results from the ΔSxr^b deletion (Mazeyrat et al., 2001). In marsupials, *EIF2S3* was localised to tammar wallaby chromosome 5p (Delbridge and Graves, 2004), the region homologous to the added region of the eutherian sex chromosomes, which contains many genes with roles in sex determination and differentiation. The absence of a Y-borne homologue of *Eif2s3* in humans, as well as the absence of *Ube1y* in primates indicates that spermatogenic proliferation in primates is under the influence of a third, as yet unknown factor.

11.4.2.5 *RPS4Y*

Ribosomal protein small subunit 4 (*RPS4*) codes for a basic ribosomal protein involved in mRNA binding. Its protein product is located at the 40/60 s subunit interface of the small ribosomal subunit, responsible for accurate translation (Nygard and Nika, 1982). Uniquely amongst human ribosomal proteins, two homologous loci have been identified, *RPS4X* (located on Xq13) and *RPS4Y* (located on distal Yp,), encoding proteins with 93% amino acid identity (Fisher et al., 1990). A *RPS4* homologue in chicken, located on chromosome 4, is nearly identical to *RPS4X* (Zinn et al., 1994), suggesting a crucial role in development throughout vertebrates.

Deletions of a 90 kb region encompassing *RPS4Y* have been associated with the phenotypic features of Turner syndrome (Zinn et al., 1994). It may therefore be at least partially responsible for the gonadal dysgenesis and short stature associated with this syndrome, which is thought to result from monosomy of one or more genes shared between the X and Y chromosomes (Ferguson-Smith, 1965).

Although *RPS4X* is X-borne in all eutherian mammals (Bergen et al., 1998), *RPS4Y* has been described only in primates, including macaque (Omoe and Endo, 1996), gibbons and various hominoids (Samollow et al., 1996). Accordingly, *RPS4X* undergoes X inactivation in most eutherian mammals, escaping X-inactivation only in primates (Jegalian and Page, 1998). The isolation of full-length cDNA clones from two different *RPS4* genes in opossum (Jegalian and Page, 1998), corresponding to X and Y copies in marsupials, is evidence that *RPS4Y* is an ancient mammalian Y gene, and was not recently acquired on the Y chromosome in primates. However, no male-specific bands have been observed in any Australian marsupial, suggesting that *RPS4Y* might have been retained only in American marsupials. The observation that in humans *RPS4Y* is expressed at just one-tenth the level of *RPS4X* (Zinn et al., 1994), suggests that *RPS4Y* has been partially inactivated, and is some way along the path to elimination from the human Y chromosome.

11.4.2.6 *ATRY*

A search for the mutated gene that results in mental retardation associated with α-thalassaemia (ATR-X syndrome), led to the discovery of *ATRX* (alpha-thalassemia

and mental retardation on the X chromosome) in humans (Gibbons et al., 1995b). The observation of genital abnormalities in ATR-X patients pinpointed a function for *ATRX* in the urogenital development pathway (Gibbons et al., 1995a), particularly with male-to-female sex reversal reported in a pedigree with a deletion at an intron/exon boundary (Ion et al., 1996).

Examination of the *ATRX* protein assigned it to a helicase superfamily subgroup (SNF2-like family) containing genes involved in DNA recombination, repair and the regulation of transcription (Picketts et al., 1996). A highly similar (83% identity) gene *Atrx* was isolated on the mouse X, enabling identification of two functionally important regions; a catalytic domain and a zinc finger domain (Picketts et al., 1998). *ATRX* has no Y homologue in any eutherian mammal.

The search for marsupial *ATRX* led to the surprising discovery of an active homologue on the marsupial Y (*ATRY*). Southern blot, fluorescent in situ hybridisation and partial sequence analyses demonstrated the presence of an X-borne *ATRX* and Y-borne *ATRY* in both Australian (Pask et al., 2000) and South American (Carvalho-Silva et al., 2004) marsupials. *ATRY* is testis-specific, which contrasts to the ubiquitous expression of human and marsupial *ATRX* (Gibbons et al., 1995b; Villard et al., 1997), although marsupial *ATRX* is not expressed in the developing testis (Pask et al., 2000). The different expression patterns of *ATRX* and *ATRY* in marsupials suggests subfunctionalization, and indicates that *ATRY* plays a specific role in marsupial sex differentiation.

The ancestral core of *ATRX*, identified by sequence comparisons between the tammar and human *ATRX*, consists of a cysteine-rich region, SWI2/SNF2 helicase-like domains, and modular protein binding elements (Park et al., 2004). This core is conserved in *ATRY*, suggesting that *ATRX* and *ATRY* operate via similar chromatin remodeling mechanisms. However, their difference in function, particularly tissue location, could be the result of a difference in binding partners (Park et al., 2005), with alterations of *ATRY* across a number of protein motifs a result of its testis-specific role in marsupials.

The involvement of *ATRX* in the human sex determining pathway, and its testis-specific expression pattern in marsupials, means that *ATRY* must be considered to be a candidate testis-determining gene in marsupials. Whether *ATRY* functions in addition to *SRY* or replaces *SRY* as the marsupial testis-determining factor, is still unresolved.

11.4.2.7 *PHF6Y* and *HUWE1Y*

Full-sequencing of two tammar wallaby Y-specific BAC (bacterial artificial chromosome) clones containing *RBMY* and *SRY* unexpectedly revealed the presence of two novel genes on the tammar Y chromosome; *PHF6Y* (on the same BAC clone as *RBMY*) and *HUWE1Y* (on the same BAC clone as *SRY*) (Murtagh et al., in preparation; Sankovic, 2005). Both genes have X-borne partners in marsupials, as well as humans and mice, but no Y-borne homologues have been detected in any eutherian mammal.

PHF6 (plant homeodomain PhD-like finger) is located on Xq26.3 in humans, and has been implicated in Börjeson–Forssman–Lehmann syndrome (BFLS) (Lower

et al., 2002), which includes moderate to severe mental retardation, gynaecomastia and hypogonadism. *PHF6* is a nucleolar protein, encoding a 4.5 kb alternatively spliced transcript with 11 exons and an ORF (open reading frame) of 1,095 bp. The presence of two zinc finger domains in the *PHF6* protein indicates a role in transcription regulation, and highly conserved orthologues have been identified in divergent vertebrate species such as mouse, zebrafish and African clawed frog (Lower et al., 2002). *PHF6* is ubiquitously expressed in eutherian mammals, and highly expressed in the embryonic central nervous system of mouse (Lower et al., 2002). An opossum copy of *PHF6* has also been identified on the X chromosome (Shevchenko et al., 2007).

HUWE1 (HECT, UBA and WWE domain containing 1) is a ubiquitin ligase belonging to the HECT (homologue of E6AP) family of E3 enzymes, and is thought to target Mcl1 (which inhibits apoptosis) and p53 (a tumour suppressor) for polyubiquitination, triggering degradation of these proteins in the proteasome (Chen et al., 2005; Zhong et al., 2005). Most significantly, recent studies in humans have revealed that a linear pathway involving *HUWE1* allows neuronal differentiation and cell-cycle arrest of stem cells and cortical progenitors (Zhao et al., 2008). An orthologue of *HUWE1* in *Caenorhabditis elegans* has also been found to be essential for synaptogenesis (Sieburth et al., 2005). In rats, the X-borne *Huwe1* ubiquitinates multiple core histones (Liu et al., 2005) and shows strong expression in spermatogonia and early spermatocytes, but little to no expression in spermatids. Although *Huwe1* is also expressed in other tissues, this expression is mainly cytoplasmic except in the neuronal cells of the brain, indicating that *Huwe1* has a specific function in the testis compared to other tissues. These observations have been taken to suggest that *Huwe1* is involved in chromatin modification in early germs cells of the testis (Liu et al., 2007).

11.4.3 Summary of Genes on the Marsupial Y

The eight genes so far identified on the Y chromosome of marsupials share three characteristics: (1) There is a clear bias of genes with important roles in sex determination and differentiation, as shown by the presence of *RBMY* and *SRY* on the marsupial, as well as the eutherian, Y chromosome. There is also evidence of sexual development phenotypes for the X orthologues of six of the eight genes. (2) All eight of the known marsupial Y-borne genes have copies on the X chromosome, a relic of the ancient homology between the X and Y chromosomes. (3) Not all of the Y-borne genes in marsupials are found on the Y in eutherians, with three of the eight genes (*ATRY*, *PHF6Y*, *HUWE1Y*) lacking a homologue on the eutherian Y chromosome.

11.5 Conclusion

Generally, the rules of gene content and function of the Y chromosome, established for human and mouse, apply also to marsupials. However, the marsupial

Y chromosome, although it is tiny in many species, contains more genes (eight) than the conserved region of the Y in human (four) or mouse (five). This suggests that gene loss from the Y in marsupials has not proceeded as far as for the comparable region of the human and mouse Y.

As for the human Y, genes located on the marsupial Y chromosome, and their respective partners on the X chromosome, show a clear involvement in sexual differentiation and/or brain development. It is interesting that these "brains and balls" genes on the X chromosome (Graves et al., 2002) are the ones whose Y homologues have survived, presumably because the original autosomal gene already had a selectable function in male reproduction.

The presence of *SOX3* on the X and *SRY* on the Y in marsupials as well as eutherian mammals suggests that this gene differentiated from an original brains-and-balls gene before marsupials diverged from eutherian mammals. However, there is no direct evidence that *SRY* is the sex-determining gene in marsupials, and it remains possible that another gene (*ATRY*? *PHF6Y*?) acquired the sex determining function that led to the origin of the mammalian X and Y chromosomes, and was supplanted by *SRY*.

Our knowledge of the marsupial Y therefore plays an important role in understanding the evolution of mammalian sex chromosomes, and is a shining example of the value of non-model species for investigating genome organization and function.

References

Agulnik AI, Mitchell MJ, Lerner JL, Woods DR, Bishop CE (1994a). A mouse Y chromosome gene encoded by a region essential for spermatogenesis and expression of male-specific minor histocompatibility antigens. Hum Mol Genet 3:873–878.

Agulnik AI, Mitchell MJ, Mattei M-G, et al. (1994b). A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation in mouse and human. Hum Mol Genet 3:879–884.

Agulnik AI, Longepied G, Ty MT, Bishop CE, Mitchell M (1999) Mouse H-Y encoding Smcy gene and its X Chromosomal homolog Smcx. Mamm Genome 10:926–929.

Aitken RJ, Graves JAM (2002) Human spermatozoa: the future of sex. Nature 415:963.

Arakawa Y, Nishida-Umehara C, Matsuda Y, Sutou S, Suzuki H (2002) X-chromosomal localization of mammalian Y-linked genes in two XO species of the Ryukyu spiny rat. Cytogenet Genome Res 99:303–309.

Bergen A, Pratt M, Mehlman P, Goldman D (1998) Evolution of RPS4Y. Mol Biol Evol 15: 1412–1419.

Berta P, Hawkins JB, Sinclair AH, et al. (1990) Genetic evidence equating SRY and the testis-determining factor. Nature 348:448–450.

Bickmore WA, Cooke HJ (1987) Evolution of homologous sequences on the human X and Y chromosomes, outside of the meiotic pairing segment. Nucleic Acids Res 15:6261–6271.

Breen M, Thomas R, Binns MM, Carter NP, Langford CF (1999) Reciprocal Chromosome Painting Reveals Detailed Regions of Conserved Synteny between the Karyotypes of the Domestic Dog (*Canis familiaris*) and Human. Genomics 61:145–155.

Bull JJ (1983) Evolution of Sex Determining Mechanisms. Benjamin/Cummings Publishing Company Inc, Menlo Park.

Burgoyne PS (1992) Fertility in mice requires X-Y pairing and a Y-chromosomal "spermatogenesis" gene mapping to the long arm. Cell 71:391–398.

Carvalho-Silva DR, O'Neill R, Brown JD, et al. (2004) Molecular characterization and evolution of X and Y-borne ATRX homologues in American marsupials. Chromosome Res 12:795–804.

Chandley A, Cooke HJ (1994) Human male fertility – Y-linked genes and spermatogenesis. Hum Mol Genet 3:1449–1452.

Charlesworth B (1996) The evolution of chromosomal sex determination and dosage compensation. Curr Biol 6:149–162.

Chen D, Kon N, Li M, Zhang W, Qin J, Gu W (2005) ARF-BP1/Mule Is a Critical Mediator of the ARF Tumor Suppressor. Cell 121:1071–1083.

Cooper DW, Johnston PG, Watson JM, Graves JAM (1993) X-inactivation in marsupials and monotremes. Sem Dev Biol 4:117–128.

Daneau I, Ethier J-F, Lussier JG, Silversides DW (1996) Porcine SRY gene locus and genital ridge expression. Biol Reprod 55:47–53.

Delbridge ML, Graves JAM (2004) Assignment of the eukaryotic translation initiation factor (EIF2S3) to tammar wallaby chromosome 5p by in situ hybridization. Cytogenet Genome Res 107:139B.

Delbridge M, Harry J, Toder R, et al. (1997) A human candidate spermatogenesis gene, RBM1, is conserved and amplified on the marsupial Y chromosome. Nat Genet 15:131–136.

Delbridge M, Lingenfelter PA, Disteche CM, Graves JAM (1999) The candidate spermatogenesis gene RBMY has a homologue on the human X chromosome. Nat Genet 22:223–224.

Elliott DJ, Millar MR, Oghene K, et al. (1997) Expression of RBM in the nuclei of human germ cells is dependent on a critical region of the Y chromosome long arm. Proc Natl Acad Sci USA 94:3848–3853.

Ezaz T, Stiglec R, Veyrunes F, Marshall Graves JA (2006) Relationships between vertebrate ZW and XY sex chromosome systems. Curr Biol 16:R736–R743.

Ferguson-Smith MA (1965) Karyotype-phenotype correlations in gonadal dysgenesis and their bearing on the pathogenesis of malformations. J Med Genet 39:142–155.

Fisher EMC, Beer-Romero P, Brown LG, et al. (1990) Homologous ribosomal protein genes on the human X and Y chromosomes: escape from X inactivation and possible implications for turner syndrome. Cell 63:1205–1218.

Foster JW, Graves JAM (1994) An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. Proc Natl Acad Sci USA 91:1927–1931.

Foster JW, Brennan FE, Hampikian GK, et al. (1992) Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. Nature 359:531–533.

Foster JW, Dominguez-Steglich MA, Guioli S, et al. (1994) Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. Nature 372:525–530.

Freije D, Helms C, Watson MS, Donis-Keller H (1992) Identification of a Second Pseudoautosomal Region Near the Xq and Yq Telomeres. Science 258:1784–1787.

Fridolfsson AK, Cheng H, Copeland NG, et al. (1998) Evolution of the avian sex chromosomes from an ancestral pair of autosomes. Proc Natl Acad Sci USA 95:8147–8152.

Gibbons RJ, Brueton L, Buckle VJ, et al. (1995a). Clinical and hematologic aspects of the X-linked alpha-thalassemia/mental retardation syndrome (ATR-X). Am J Med Genet 55:288–299.

Gibbons RJ, Picketts DJ, Villard L, Higgs DR (1995b). Mutations in a Putative Global Transcriptional Regulator Cause X-Linked Mental Retardation with □-Thalassemia (ATR-X Syndrome). Cell 80:837–845.

Goureau A, Yerle M, Schmitz A, et al. (1996) Human and Porcine Correspondence of Chromosome Segments Using Bidirectional Chromosome Painting. Genomics 36:252–262.

Graves JAM (1995) The origin and function of the mammalian Y chromosome and Y-borne genes – an evolving understanding. BioEssays 17:311–320.

Graves JAM (2004) The degenerate Y chromosome – can conversion save it? Reprod Fert Develop 16:527–534.

Graves JAM (2006) Sex chromosome specialiation and degeneration in mammals. Cell 124: 901–914.

Graves JAM (2008) Weird Animal Genomes and the Evolution of Vertebrate Sex and Sex Chromosomes. Annu Rev Genet 42:565–586.

Graves JAM, Shetty S (2001) Sex from W to Z: evolution of vertebrate sex chromosomes and sex determining genes. J Exp Zool 290:449–462.
Graves JAM, Wakefield MJ, Toder R (1998) The origin and evolution of the pseudoautosomal regions of human sex chromosomes. Hum Mol Genet 7:1991–1996.
Graves JAM, Gecz J, Hameister H (2002) Evolution of the human X – a smart and sexy chromosome that controls speciation and development. Cytogenet Genome Res 99: 141–145.
Griffiths R (1991) The Isolation of Conserved DNA Sequences Related the Human Sex-Determining Region Y Gene from the Lesser Black-Backed Gull (*Larus fuscus*). P Roy Soc B Biol Sci 244:123–128.
Grützner F, Rens W, Tsend-Ayush E, et al. (2004) In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. Nature 432:913–917.
Gubbay J, Collignon J, Koopman P, et al. (1990) A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature 346:245–250.
Hanley NA, Hagan DM, Clement-Jones M, et al. (2000) SRY, SOX9, and DAX1 expression patterns during human sex determination and gonadal development. Mech Develop 91:403–407.
Huang B, Wang S, Ning Y, Lamb AN, Bartley J (1999) Autosomal XX sex reversal caused by duplication of SOX9. Am J Med Genet 87:349–353.
Hughes JF, Skaletsky H, Pyntikova T, et al. (2005) Conservation of Y-linked genes during human evolution revealed by comparative sequencing in chimpanzee. Nature 437:101–104.
Ion A, Telvi L, Chaussain JL, et al. (1996) A novel mutation in the putative DNA helicase XH2 is responsible for male-to-female sex reversal associated with an atypical form of the ATR-X syndrome. Am J Med Genet 58:1185–1191.
Jegalian K, Page DC (1998) A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. Nature 394:776–780.
Jensen LR, Amende M, Gurok U, et al. (2005) Mutations in the JARID1C Gene, Which Is Involved in Transcriptional Regulation and Chromatin Remodeling, Cause X-Linked Mental Retardation. Am J Med Genet 76:227–236.
Just W, Rau W, Vogel W, et al. (1995) Absence of Sry in species of the vole Ellobius. Nat Genet 11:117–118.
Kay GF, Ashworth A, Penny GD, et al. (1991) A candidate spermatogenesis gene on the mouse Y chromosome is homologous to ubiquitin-activating enzyme E1. Nature 354:486–489.
Kondo M, Hornung U, Nanda I, et al. (2006) Genomic organization of the sex-determining and adjacent regions of the sex chromosomes of medaka. Genome Res 16:815–826.
Koopman P, Munsterberg A, Capel B, Vivian N, Lovell-Badge R (1990) Expression of a candidate sex-determining gene during mouse testis differentiation. Nature 348:450–452.
Koyano S, Ito M, Takamatsu N, Takiguchi S, Shiba T (1997) The *Xenopus* Sox3 gene expressed in oocytes of early stages. Gene 188:101–107.
Kuhl H, Rottger S, Heilbronner H, Enders H, Schempp W (2001) Loss of the Y chromosomal PAR2-region in four familial cases of satellited Y chromosomes (Yqs). Chromosome Res 9:215–222.
Kuroki Y, Toyoda A, Noguchi H, et al. (2006) Comparative analysis of chimpanzee and human Y chromosomes unveils complex evolutionary pathway. Nat Genet 38:158–167.
Kvaloy K, Galvagni F, Brown WRA (1994) The sequence organization of the long arm pseudoautosomal region of the human sex chromosomes. Hum Mol Genet 3:771–778.
Lahn BT, Page DC (1999) Four evolutionary strata on the human X chromosome. Science 286:964–967.
Li L, Hamer DH (1995) Recombination and allelic association in the Xq/Yq homology region. Hum Mol Genet 4:2013–2016.
Liu Z, Oughtred R, Wing SS (2005) Characterization of E3Histone, a novel testis ubiquitin protein ligase which ubiquitinates histones. Mol Cell Biol 25:2819–2831.

Liu Z, Miao D, Xia Q, Hermo L, Wing SS (2007) Regulated expression of the ubiquitin protein ligase, E3(Histone)/LASU1/Mule/ARF-BP1/HUWE1, during spermatogenesis. Dev Dyn 236:2889–2898.

Lower KM, Turner G, Kerr BA, et al. (2002) Mutations in PHF6 are associated with Borjeson-Forssman-Lehmann syndrome. Nat Genet 32:661–665.

Ma K, Inglis JD, Sharkey A, et al. (1993) A Y chromosome gene family with RNA-binding protein homology: candidates for the azoospermia factor AZF controlling human spermatogenesis. Cell 75:1287–1295.

Mazeyrat S, Saut N, Mattei M-G, Mitchell MJ (1999) RBMY evolved on the Y chromosome from a ubiquitously transcribed X-Y identical gene. Nat Genet 22:224–226.

Mazeyrat S, Saut N, Grigoriev V, et al. (2001) A Y-encoded subunit of the translation initiation factor Eif2 is essential for mouse spermatogenesis. Nat Genet 29:49–53.

McLaren A, Simpson E, Epplen JT, et al. (1988) Location of the genes controlling H-Y antigen expression and testis determination on the mouse Y chromosome. Proc Natl Acad Sci USA 85:6442–6445.

Meyers-Wallen VN (2003) Sry and Sox9 expression during canine gonadal sex determination assayed by quantitative reverse transcription-polymerase chain reaction. Mol Reprod Dev 65:373–381.

Mitchell MJ, Woods DR, Tucker PK, Opp JS, Bishop CE (1991) Homology of a candidate spermatogenic gene from the mouse Y chromosome to the ubiquitin-activating enzyme El. Nature 354:483–486.

Mitchell MJ, Woods DR, Wilcox SA, Marshall Graves JA, Bishop CE (1992) Marsupial Y chromosome encodes a homologue of the mouse Y-linked candidate spermatogenesis gene Ube1y. Nature 359:528–531.

Mohandas TK, Speed RM, Passage MB, Yen PH, Chandley A, Shapiro LJ (1992) Role of the pseudoautosomal region in sex-chromosome pairing during male meiosis: meiotic studies in a man with a deletion of distal Xp. Am J Hum Genet 51:526–533.

Muller HJ (1914) A gene for the fourth chromosome of Drosophila. J Exp Zool 17:325–336.

Muller HJ (1918) Genetic Variability, Twin Hybrids and Constant Hybrids, in a Case of Balanced Lethal Factors. Genetics 3:422–499.

Müller S, Stanyon R, O'Brien P, Ferguson-Smith MA, Plesker R, Wienberg J (1999) Defining the ancestral karyotype of all primates by multidirectional chromosome painting between tree shrews, lemurs and humans. Chromosoma 108:393–400.

Nanda I, Shan Z, Schartl M, et al. (1999) 300 million years of conserved synteny between chicken Z and human chromosome 9. Nat Genet 21:258–259.

Nygard O, Nika H (1982) Identification by RNA-protein cross-linking of ribosomal proteins located at the interface between the small and the large subunits of mammalian ribosomes. EMBO J 1:357–362.

Odorisio T, Mahadevaiah SK, Mccarrey JR, Burgoyne PS (1996) Transcriptional analysis of the candidate spermatogenesis gene Ube1y and of the closely related Ube1x shows that they are coexpressed in spermatogonia and spermatids but are repressed in pachytene spermatocytes. Dev Biol 180:336–343.

Ohno S (1967) Sex Chromosomes and Sex-Linked Genes. Springer, New York, NY.

Omoe K, Endo A (1996) Relationship between the monosomy X phenotype and Y-linked ribosomal protein S4 (Rps4) in several species of mammals: a molecular evolutionary analysis of Rps4 homologs. Genomics 31:44–50.

Page J, Berrios S, Rufas JS, et al. (2003) The pairing of X and Y chromosomes during meiotic prophase in the marsupial species Thylamys elegans is maintained by a dense plate developed from their axial elements. J Cell Sci 116:551–560.

Page J, Berrios S, Parra MT, et al. (2005) The program of sex chromosome pairing in meiosis is highly conserved across marsupial species: implications for sex chromosome evolution. Genetics 170:793–799.

Page J, Viera A, Parra MT, et al. (2006) Involvement of synaptonemal complex proteins in sex chromosome segregation during marsupial male meiosis. PLoS Genet 2:e136.

Park DJ, Pask AJ, Huynh K, Renfree MB, Harley VR, Marshall Graves JAM (2004) Comparative analysis of ATRX, a chromatin remodeling protein. Gene 339:39–48.
Park DJ, Pask AJ, Huynh K, Harley VR, Renfree MB, Graves JAM (2005) Characterisation of the marsupial-specific ATRY gene: implications for the evolution of male-specific function. Gene 362:29–36.
Pask AJ, Renfree MB, Graves JAM (2000) The human sex-reversing ATRX gene has a homologue on the marsupial Y chromosome, ATRY: implications for the evolution of mammalian sex determination. Proc Natl Acad Sci USA 97:13198–13202.
Payen E, Pailhoux E, Merhi Ra, et al. (1996) Characterization of ovine SRY transcript and developmental expression of genes involved in sexual differentiation. Int J Dev Biol 40:567–575.
Peichel CL, Ross A, Matson CK, et al. (2004) The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. Curr Biol 14:1416–1424.
Picketts DJ, Higgs DR, Bachoo S, Blake DJ, Quarrell OWJ, Gibbons RJ (1996) ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. Hum Mol Genet 5:1899–1907.
Picketts DJ, Tastan AO, Higgs DR, Gibbons RJ (1998) Comparison of the human and murine ATRX gene identifies highly conserved, functionally important domains. Mamm Genome 9:400–403.
Rice WR (1987) Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. Genetics 116:161–167.
Rice WR (1996) Evolution of the Y sex chromosome in animals. Bioscience 46:331–343.
Roche L, Seluja G, Wettstein R (1986) The meiotic behaviour of the XY pair in Lutreolina crassicaudata (Marsupialia: Didelphoidea). Genetica 71:213–224.
Ross MT, Grafham DV, Coffey AJ, et al. (2005) The DNA sequence of the human X chromosome. Nature 434:325–337.
Rozen S, Skaletsky H, Marszalek JD, et al. (2003) Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature 423:873–876.
Samollow PB, Cherry LM, Witte SM, Rogers J (1996) Interspecific variation at the Y-linked RPS4Y locus in hominoids: implications for phylogeny. Am J Phys Anthropol 101: 333–343.
Sankovic N (2005) Molecular Characterisation of a Marsupial Y Chromosome. Research School of Biological Sciences, The Australian National University, Canberra.
Saxena R, Brown LG, Hawkins T, et al. (1996) The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. Nat Genet 14:292–299.
Sekido R, Lovell-Badge R (2008) Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature 453:930–934.
Seluja GA, Roche L, Solari AJ (1987) Male meiotic prophase in *Didelphis albiventris*. J Hered 78:218–222.
Sharman G, Hughes R, Cooper D (1990) The Chromosomal Basis of Sex-Differentiation in Marsupials. Aust J Zool 37:451–466.
Sharp P (1982) Sex chromosome pairing during male meiosis in marsupials. Chromosoma 86: 27–47.
Shaw G, Renfree M, Short R, O WS (1988) Experimental manipulation of sexual differentiation in wallaby pouch young treated with exogenous steroids. Development 104:689–701.
Shaw G, Renfree MB, Short R (1990) Primary genetic control of sexual differentiation in marsupials. Aust J Zool 37:443–450.
Sherlock JK, Griffin DK, Delhanty JDA, Parrington JM (1996) Homologies between human and marmoset (*Callithrix jacchus*) chromosomes revealed by comparative chromosome painting. Genomics 33:214–219.
Shevchenko A, Zakharova I, Elisaphenko E, et al. (2007) Genes flanking Xist in mouse and human are separated on the X chromosome in American marsupials. Chromosome Res 15:127–136.
Sieburth D, Ch'ng Q, Dybbs M, et al. (2005) Systematic analysis of genes required for synapse structure and function. Nature 436:510–517.

Sinclair AH, Berta P, Palmer MS, et al. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature 346:240–244.
Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837.
Smith CA, Roeszler KN, Ohnesorg T, et al. (2009) The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature 461:267–271.
Smith CA, Smith MJ, Sinclair AH (1999) Gene expression during gonadogenesis in the chicken embryo. Gene 234:395–402.
Solari AJ, Bianchi NO (1975) The synaptic behaviour of the X and Y chromosomes in the marsupial Monodelphis dimidiata. Chromosoma 52:11–25.
Soullier S, Hanni C, Catzeflis F, Berta P, Laudet V (1998) Male sex determination in the spiny rat Tokudaia osimensis (Rodentia: Muridae) is not Sry dependent. Mamm Genome 9:590–592.
Sutcliffe MJ, Burgoyne PS (1989) Analysis of the testes of H-Y negative XOSxrb mice suggests that the spermatogenesis gene (Spy) acts during the differentiation of the A spermatogonia. Development 107:373–380.
Toure A, Grigoriev V, Mahadevaiah S, Rattigan Á, Ojarikre O, Burgoyne PS (2004) A protein encoded by a member of the multicopy SSty gene family located on the long arm of the mouse Y chromosome is expressed during sperm development. Genomics 83:140–147.
Toure A, Clemente E, Ellis P, et al. (2005) Identification of novel Y chromosome encoded transcripts by testis transcriptome analysis of mice with deletions of the Y chromosome long arm. Genome Biol 6:R102.
Veyrunes F, Waters PD, Miethke P, et al. (2008) Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res 18:965–973.
Vidal VPI, Chaboissier M-C, de Rooij DG, Schedl A (2001) Sox9 induces testis development in XX transgenic mice. Nat Genet 28:216–217.
Villard L, Lossi A-M, Cardoso C, et al. (1997) Determination of the Genomic Structure of the XNP/ATRX Gene Encoding a Potential Zinc Finger Helicase. Genomics 43:149–155.
Wallis M, Waters P, Delbridge M, et al. (2007) Sex determination in platypus and echidna: autosomal location of SOX3 confirms the absence of SRY from monotremes. Chromosome Res 15:949–959.
Waters P, Duffy B, Frost C, Delbridge M, Graves JAM (2001) The human Y chromosome derives largely from a single autosomal region added to the sex chromosomes 80–130 million years ago. Cytogenet Cell Genet 92:74–79.
Waters P, Wallis MC, Graves JAM (2007) Mammalian sex – origin and evolution of the Y chromosome and SRY. Semin Cell Dev Biol 18:389–400.
Wienberg J, Stanyon R, Nash W, et al. (1997) Conservation of human vs. feline genome organization revealed by reciprocal chromosome painting. Cytogenet Cell Genet 77:211–217.
Wilhelm D, Koopman P (2006) The makings of maleness: towards an integrated view of male sexual development. Nat Rev Genet 7:620–631.
Wilson J, George F, Griffin J (1981) The hormonal control of sexual development. Science 211:1278–1284.
Yang F, O'Brien P, Ferguson-Smith M (2000) Comparative chromosome map of the laboratory mouse and chinese hamster defined by reciprocal chromosome painting. Chromosome Res 8:219–227.
Zhao X, Heng JI-T, Guardavaccaro D, et al. (2008) The HECT-domain ubiquitin ligase Huwe1 controls neural differentiation and proliferation by destabilizing the N-Myc oncoprotein. Nat Cell Biol 10:643–653.
Zhong Q, Gao W, Du F, Wang X (2005) Mule/ARF-BP1, a BH3-only E3 ubiquitin ligase, catalyzes the polyubiquitination of Mcl-1 and regulates apoptosis. Cell 121:1085–1095.
Zinn AR, Alagappan RK, Brown LG, Wool I, Page DC (1994) Structure and function of ribosomal protein S4 genes on the human and mouse sex chromosomes. Mol Cell Biol 14:2485–2492.

Part V
Marsupial Epigenetics

Paul D. Waters

The term epigenetics was originally coined to explain phenotypic variation that was independent of genotype; hence the name epi-(above) genetics. The term is now widely used to explain control of gene expression that is independent of underlying DNA sequence. Epigenetic regulation has been (and still is) extensively studied in mouse and human. But outside of these model species much less is known about eutherian epigenetic regulation, and even less is known about marsupial epigenetic regulation, leaving a large gap in our knowledge of even the most stunning examples of mammalian epigenetics.

In this part two classic examples of epigenetic regulation are reviewed. Hore and colleagues in Chapter 12 review a biologically and medically important example of epigenetic regulation; genomic imprinting, where one allele is preferentially silenced depending on parent of origin. The first example ever observed of imprinting was in marsupials as paternal X chromosome inactivation (Cooper et al., 1971; Richardson et al., 1971; Sharman, 1971; see Al Nadaf et al., this part). Hore et al. discuss classic imprinted eutherian loci in marsupials, and the molecular mechanisms that regulate them. They use data collected from various vertebrate representatives to decipher when these mechanisms might have arisen during mammalian evolution. The authors also highlight that the distribution of imprinting throughout the mammalian tree correlates to the evolution of viviparity and the placenta. There is an estimated 80 imprinted loci in eutherians (with a well developed placenta) (Morison et al., 2005), whereas in marsupials (with an under developed placenta) there are considerably fewer, and imprinting appears to be absent in monotremes (egg laying mammals) (reviewed in Renfree et al., 2008).

The evolution of the placenta and genomic imprinting occurring at a similar time, and the observation that many imprinted genes are expressed and imprinted (often exclusively) in the placenta and yoke sac (influencing placenta size and function), has led to the parental-conflict hypothesis (reviewed in Wilkins and Haig, 2003). This hypothesis states that imprinting results from conflict between the maternal and paternal genomes for provisioning of the mother's resources to the fetus. The father wants the offspring to receive as much of the maternal resources as possible, regardless of the mother's future fitness. In contrast, the mother wants to share resources evenly to all her young, irrespective of which male fathers them. The predicted result is parent specific gene expression that alters the supply of resources to the

fetus. Indeed, many paternally expressed loci stimulate placental and fetal growth, whereas maternally expressed loci suppress placental and fetal growth. However, the parental-conflict hypothesis does not explain imprinted loci that do not appear to function in placenta. Hore et al. outline alternative hypotheses that have been put forward to explain the evolution of imprinting, and highlight how understanding imprinting of different loci in distantly related mammals can help distinguish which hypothesis is more correct.

Al Nadaf and colleagues (Chapter 13) review a second classic example of epigenetic regulation, X chromosome inactivation (XCI), a process in which one of the X chromosomes is transcriptionally silenced in the somatic cells of females. This phenomenon is arguably the most spectacular example of epigenetic regulation in therian mammals. It arose because the Y chromosome was once genetically identical to the X, but has since lost most of its functional genes (see Murtagh et al., previous part). Loss of Y gene function resulted in double the expression of X genes in females (with two X chromosomes) compared to males (with one X- and one Y-chromosome), and a dosage compensation mechanism was required to restore equal expression of X gene between the sexes.

In human and mouse XCI is proposed to be complete along the X, stable, somatically heritable, and random with respect to which X is chosen for inactivation. In contrast, the limited information available on marsupial XCI indicates that it is incomplete, unstable, tissue specific and non-random (Cooper et al., 1993) Al Nadaf et al. contrasts what is known about the cytogenetic and molecular manifestations of XCI in eutherian and marsupial mammals. This review extends beyond therian mammals to compare XCI with monotreme and bird dosage compensation, which in both groups appears to be incomplete and locus specific (Itoh et al., 2007; Melamed and Arnold, 2007; Deakin et al., 2008). Because birds, monotremes and therian mammals have independently evolved dosage compensation systems, any common epigenetic signatures must have been independently recruited from a universal epigenetic "toolbox".

Marsupial genome projects and next generation sequencing are allowing research in this field to progress at a much faster rate than has been permitted in the past. It is now possible to move from looking at single genes to studies that encompass the entire X for XCI, and the entire genome for genomic imprinting. These global studies will permit broader comparisons between not just mammalian lineages, but vertebrates as a whole.

References

Cooper DW, Johnston PG, Watson JM, Graves JAM (1993) X inactivation in marsupials and monotremes. Semin Dev Biol 4:117–128.

Cooper DW, VandeBerg JL, Sharman GB, Poole WE (1971) Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. Nat New Biol 230:155–157.

Deakin JE, Hore TA, Koina E, Marshall Graves JA (2008) The status of dosage compensation in the multiple X chromosomes of the platypus. PLoS Genet 4:e1000140.

Itoh Y, Melamed E, Yang X, et al. (2007) Dosage compensation is less effective in birds than in mammals. J Biol 6:2.

Melamed E, Arnold AP (2007) Regional differences in dosage compensation on the chicken Z chromosome. Genome Biol 8:R202.

Morison IM, Ramsay JP, Spencer HG (2005) A census of mammalian imprinting. Trends Genet 21:457–465.

Renfree MB, Ager EI, Shaw G, Pask AJ (2008) Genomic imprinting in marsupial placentation. Reproduction 136:523–531.

Richardson BJ, Czuppon AB, Sharman GB (1971) Inheritance of glucose-6-phosphate dehydrogenase variation in kangaroos. Nat New Biol 230:154–155.

Sharman GB (1971) Late DNA replication in the paternally derived X chromosome of female kangaroos. Nature 230:231–232.

Wilkins JF, Haig D (2003) What good is genomic imprinting: the function of parent-specific gene expression. Nat Rev Genet 4:359–368.

Chapter 12
The Evolution of Genomic Imprinting – A Marsupial Perspective

Timothy A. Hore, Marilyn B. Renfree, Andrew J. Pask, and Jennifer A. Marshall Graves

Abstract Genomic imprinting is a medically significant epigenetic trait in which genes are expressed from only one of the two alleles, according to their parent of origin. Among vertebrates, imprinted gene expression has been found only in the live-bearing therian mammals – eutherians (placental mammals) and marsupials. Because marsupials are so distantly related to eutherians, comparisons of imprinting between the two mammalian infraclasses are particularly valuable. However, the popular mammalian model organisms (humans, mice and domestic mammals) are all eutherian, so imprinting in marsupials has not received the attention it deserves. Research gathered over some years shows that marsupial orthologues of many imprinted domains in eutherians, such as the well characterised *IGF2/H19* and *PEG10* domains, comprise fewer imprinted genes. Other eutherian imprinted domains, such as the X-inactivation centre, Callipyge and Prader-Willi/Angleman syndrome domains, are either completely absent in marsupials or are not imprinted. The occurrence of imprinting in marsupials (with some imprinted genes) and eutherians (with many imprinted genes) contrasts with the likely absence of imprinting in egg-laying vertebrates (monotreme mammals, birds and reptiles). The acquisition of imprinting by therian mammals correlates with increased dependence on placentation for early development and coincides with the acquisition of a unique repeat content within the genome and germline expression of *BORIS*. Analysis of the evolutionary trajectory of these traits offers us insights into how and why genomic imprinting evolved in mammals.

Keywords Evolution · Genomic imprinting · Marsupials · Parental conflict · Regulators of imprinting

---

T.A. Hore (✉)
Laboratory of Developmental Genetics and Imprinting, The Babraham Institute, Cambridge CB22 3AT, UK
e-mail: tim.hore@bbsrc.ac.uk

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_12, © Springer Science+Business Media B.V. 2010

12.1 Introduction

12.1.1 Genomic Imprinting: Definition and Relevance

Mammals, like all diploid organisms, have two copies of their autosomal genes – one copy (or allele) from each parent. Usually, both of these alleles are expressed at equal levels, thus maximising the protective benefit of diploidy should one allele fail due to mutation (Otto and Goldstein, 1992). Despite this, about 80 genes in mammals are expressed from only one of the two alleles, depending on their parental origin, a phenomenon known as genomic imprinting.

Genomic imprinting is important because it represents a striking example of epigenetic regulation, now recognised to be critical in animal development. The expression of imprinted genes from only one allele (monoallelic) is determined on the basis of the sex of the parent from which alleles were inherited, even though the maternal and paternal alleles might have identical DNA sequences. Thus, the control of genomic imprinting must occur at a level which is above or outside DNA sequence (hence, epi-genetic – see Russo et al., 1996).

It is now known that an epigenetic mark (or imprint) is imparted in the previous generation during spermatogenesis in the father or oogenesis in the mother, to distinguish one parental allele from the other. It is thought that this original parent-distinguishing imprint is in the form of DNA methylation at CpG nucleotides, which is then "read" and interpreted during development and adult life to establish parent-specific expression of nearby genes. Imprinted genes are often clustered and regulated co-ordinately by a single region of differential methylation known as an imprint control region (ICR).

As well as their significance as an epigenetic model, imprinted genes are important medically. Although they number fewer than a hundred, imprinted genes are disproportionately associated with human disease (Morison et al., 2005). This is because the mono-allelic expression of imprinted genes means that any mutation in the single active copy of an imprinted gene is always fully penetrant, as no backup copy is expressed that can compensate. Moreover, without even an alteration in DNA sequence, imprinted genes can cause disease. For example, if two sister chromosomes are inherited from only one parent (uniparental disomy), imprinted genes on that chromosome show twice their normal expression or no expression at all. Loss of imprinting (i.e. reactivation of a silenced allele) can also cause diseases such as cancer, as was discovered during initial characterisation of the archetypal imprinted gene, insulin-like growth factor 2 (*IGF2*) (Ogawa et al., 1993; Rainier et al., 1993).

Because of its relevance to human biology, agriculture and medicine, genomic imprinting has been examined in detail in humans and mice, and sporadically in cattle and other domestic mammals. However, little work has been done on imprinting in marsupials (Hore et al., 2007b; Renfree et al., 2008; Renfree et al., 2009). This is perhaps surprising, given that the first discovery of imprinting in a mammal was paternal-specific silencing of the X-chromosome in female kangaroos (Cooper et al., 1971; Richardson et al., 1971). In this chapter we examine observations of

imprinted clusters and their regulators in marsupials, comparing this to observations of imprinting in eutherians. We also highlight what this tells us about how and why genomic imprinting evolved.

12.1.2 Evolution of Genomic Imprinting in Mammals: Why Marsupials Are Important

Given the high disease burden associated with genomic imprinting, it seems bizarre that it would ever evolve in the first place. Indeed, imprinting has been discovered in only a few taxonomic groups – mammals, flowering plants and some insects – perhaps reflecting selection against imprinting in most diploids. Moreover, the nature and mechanism of imprinting in these distantly-related groups is so dissimilar that it probably evolved independently in each lineage (Scott and Spielman, 2006).

Currently, the expectation is that mammalian imprinting evolved some time after the divergence of mammals from birds, reptiles and other vertebrates (Fig. 12.1). The most robust evidence for this proposition lies, not in the failure to discover genomic imprinting in non-mammalian vertebrates (despite numerous attempts – O'Neill et al., 2000; Nolan et al., 2001; Yokomine et al., 2001; Lawton et al., 2005; Colosi et al., 2006), but rather the ability of non-mammalian vertebrates to undergo parthenogenesis. This contrasts with the inability of mammals to form parthenogenetic individuals (Mann and Lovell-Badge, 1984; McGrath and Solter, 1984; Surani et al., 1984). Parthenogenotes receive both haploid copies of their genome from their mother, and require no genetic input from a father. Therefore, imprinted genes in a parthenogenote would have twice the normal expression of maternally expressed if imprinted genes are dosage sensitive or essential, genomic imprinting will provide a barrier to parthenogenesis. All non-mammalian vertebrate groups, including fish, amphibians, reptiles and even birds can tolerate parthenogenesis and other forms of unequal parental inheritance in nature and the laboratory (reviewed Hore, 2008). In contrast, mammalian parthenogenesis cannot occur without genetic modification of imprinted loci and oocyte manipulation techniques that specifically negate the effects of imprinted genes (Kono et al., 2004; Kawahara et al., 2007).

Since mammalian imprinting probably evolved after mammals diverged from other vertebrates, distantly related mammalian taxa such as marsupials and monotremes are likely to hold the key to understanding how and why imprinting evolved (Graves, 1996; Wilkins and Haig, 2003; Hore et al., 2007b; Renfree et al., 2009). This is because they diverged from other vertebrates at about the same time that imprinting evolved (Fig. 12.1), and may still carry with them traces of the earliest innovations that were required for this specialised form of gene expression. Furthermore, marsupials represent the most distantly related mammalian taxa that give birth to live young, being nurtured *in-utero* with a placenta, a feature that is thought to be intimately linked with the evolution of genomic imprinting (discussed later).

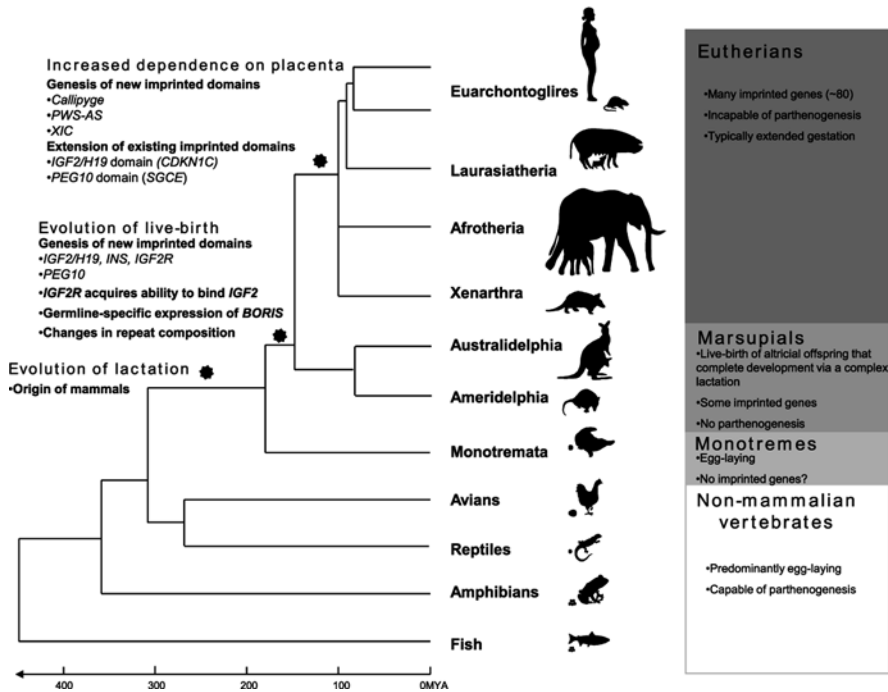


Fig. 12.1 The evolution of genomic imprinting in mammals. Eutherian mammals have many imprinted genes (~80) some of which are imprinted in marsupials, and some which are not (see *right column*). Monotremes and non-mammalian vertebrates are thought not to possess imprinting. This implies that genomic imprinting first emerged between 148 and 166 million years ago (MYA) (Bininda-Emonds et al., 2007), at a similar time to the evolution of viviparity, germline-specific expression of *BORIS* and significant changes in repeat composition in the genome. Later, in the eutherian lineage alone, imprinting apparently became more wide-spread through the genesis of new imprinted domains, or the expansion of existing ones

12.2 Imprinted Clusters in Marsupials

As nearly all research into mammalian imprinted genes has been conducted on eutherian mammals (mice in particular), the experimental approach for researchers studying marsupial imprinting has often been to choose imprinted loci well-characterised in humans and mice, and then study their orthologues in marsupials. Although this approach may miss marsupial-specific imprinted loci, it has certainly been the most efficient method to generate data on imprinting in marsupials, for which, until recently, there was little sequence information or other genomic tools that researchers of humans and mice take for granted. As a consequence of this research, there is now a greater understanding of the construction and evolution of the domains that are imprinted in humans and mice, only a subset of which are imprinted in marsupials. Imprinted loci can be assembled over time in unexpected ways, including genome rearrangement, insertion of mobile elements,

genesis of large non-coding RNAs and insertion and expansion of families of small RNAs. Here we review our understanding of the construction and evolution of major imprinted loci and their regulators, focussing on data gathered from marsupials.

12.2.1 IGF2/H19 and CDKN1C Locus

The *IGF2/H19* and *CDKN1C* cluster is the best characterised imprinted gene cluster in mammals. IGF2 is an essential positive regulator of fetal and postnatal growth that is maternally expressed (that is, paternally silenced) in eutherian mammals (DeChiara et al., 1991). *H19* lies downstream of *IGF2* and is paternally expressed (maternally silenced) (Bartolomei et al., 1991). *H19* produces a long non-coding RNA molecule that harbours the micro RNA *miR-675* (Cai and Cullen, 2007). The ICR of the *IGF2/H19* region is located just upstream of *H19* and is methylated on the paternally derived chromosome (Tremblay et al., 1997). The methylation differences between parental alleles at the ICR results in imprinted expression of *IGF2* and *H19* through differential binding of an insulator protein (Bell and Felsenfeld, 2000; Hark et al., 2000; Han et al., 2008).

The nearby *CDKN1C* gene is oppositely imprinted to *IGF2* (paternally expressed) and, more interestingly, opposes effects on growth of *IGF2* (Ager et al., 2008a) (Fig. 12.2). It was hypothesised that imprinting may have co-evolved between *IGF2* and *CDKN1C* to balance out the effects on fetal growth and nutrition of halpo-insufficency of either gene. Examination of the cluster in marsupials, however, reveals a stepwise evolution of complex imprinted regions in mammals.

IGF2 was found to be highly conserved in marsupials, differentially methylated and expressed from the paternal chromosome in somatic tissues of pouch young and the marsupial placenta (O'Neill et al., 2000; Suzuki et al., 2005; Lawton et al., 2008; Ager et al., 2008b). Until recently, marsupials were thought to lack an *H19* orthologue. However, the availability of the tammar wallaby genome sequence allowed more thorough sequence searches to identify an *H19* orthologue containing the microRNA, *miR-675* (Smits et al., 2008). *H19* was also found to be paternally expressed in marsupials and associated with a region of differential methylation containing insulator binding sites that is likely to be the ICR. Thus, despite initial expectations, marsupials appear to regulate imprinting at the *IGF2/H19* domain in an identical way to eutherians.

The nearby *CDKN1C* gene, on the other hand, was found to be bi-alellically expressed in marsupials and not imprinted (Ager et al., 2008a). Imprinting of this gene in mice is mediated through the transcription of an antisense RNA *Kcnqot1* derived from the *Kcnq1* gene. *KCNQ1* was also found to be conserved in marsupials and produced an antisense transcript despite the lack of *CDKN1C* imprinting. *CDKN1C* was expressed in the marsupial placenta suggesting a function antagonistic to *IGF2* preceded its acquisition of imprinting (Ager et al., 2008a). Imprinting of the two genes evidently did not occur concurrently to balance maternal and paternal influences on the growth of the placenta. Furthermore, the expression of *KCNQ1OT1* in the absence of *CDKN1C* imprinting suggests that antisense

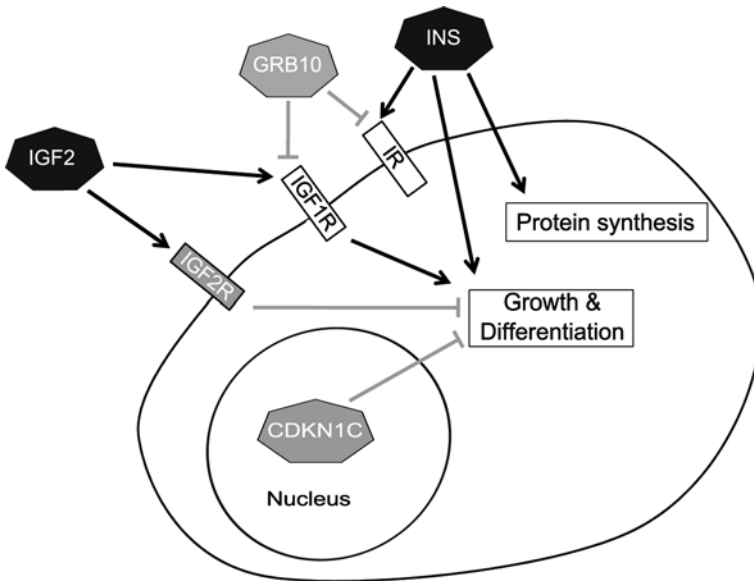


Fig. 12.2 Imprinted gene interactions, showing their antagonistic (*grey lines*) and complementary (*black arrows*) roles in growth and differentiation within a cell. Imprinted proteins are indicated *black* (paternally expressed) and *grey* (maternally expressed), whereas non-imprinted proteins are shown as white boxes. Note, *CDKN1C* is maternally expressed in mice and humans, but expressed from both alleles in the marsupials. Figure adapted from Ager et al. (2008a)

transcription at this loci preceded imprinting of this domain. It would appear that imprinting of the region has spread to encompass more neighbouring genes, and become more complex in the eutherian lineage.

12.2.2 INS

INS provides a special case of imprinting because it is imprinted only in the yolk sac. Most eutherians form both a chorio-vitelline (yolk sac) and a chorioallantoic placenta to support fetal development, whereas most marsupials rely exclusively on a yolk sac placenta (Renfree, 1972, 1982). The marsupial yolk sac placenta actively synthesises material needed for its own maintenance and for fetal growth. Similarly, the eutherian yolk sac is essential for fetal-maternal exchange, hematopoesis, and synthesis of several biologically active molecules, and a functional yolk sac is essential for successful pregnancy, even in humans (Livesey and Williams, 1981; Freeman, 1990; Jones and Jauniaux, 1995; Gulbis et al., 1998; McGrath and Palis, 2005). As is well known, *INS* encodes the polypeptide hormone insulin that regulates carbohydrate metabolism. However, it also has a role in cell growth, through influencing amino acid and fatty acid transporters as well as protein synthesis (Fehlmann et al., 1979; Chabowski et al., 2004; Hernandez-Sanchez et al., 2006) (Fig. 12.2). *INS* lies on human chromosome 11p15.5, located 5' of *IGF2,* and is

paternally expressed, but apparently only in the yolk sac (Moore et al., 2001). In the mouse, there are two copies of the gene: *Ins1* and *Ins2*, the latter of which is the orthologue of *INS* and lies downstream of *IGF2.* Whereas the maternal *Ins2* allele is gradually silenced in the mouse yolk sac, there is no silencing of *Ins1* and there is equal expression of *Ins2* alleles in embryonic body and head. Thus, as in the human, the mouse *Ins2* is also imprinted only in the yolk sac placenta. The silencing of the maternal allele is complete by day 14.5 of gestation (Deltour et al., 1995). The imprint status of *INS/Ins2* has been determined only in the pancreas (not imprinted) and the yolk sac (imprinted) of mice and humans (Giddings et al., 1994; Deltour et al., 1995; Moore et al., 2001). Differential allelic expression of *INS* has been identified in the thymus, but its imprinting in this tissue has yet to be confirmed (Pugliese et al., 1997; Vafiadis et al., 1997). As yet, no other tissues have been examined for imprinting of *INS*.

In the tammar wallaby, *INS* is paternally expressed in the placenta, as it is in mouse and human (Ager et al., 2007). INS protein is found in the endoderm of the yolk sac placenta, and increases during late gestation up to 2 days before birth. The conservation of paternally-biased *INS* expression in the yolk sac placenta in both eutherian and marsupial mammals suggests that this imprint was acquired at least 130–148 million years ago, before these mammalian lineages diverged (Luo et al., 2003; Bininda-Emonds et al., 2007) (Fig. 12.1). Four other genes so far examined (*PEG1/MEST, IGF2* and *PEG10* and *H19*) are also imprinted in the marsupial placenta (Suzuki et al., 2005; Suzuki et al., 2007; Smits et al., 2008). However, *INS* is the first example of placental-specific imprinting in marsupials, confirming that, as in eutherians, imprinting in the placenta can be independent of imprinting in other fetal tissues (Ager et al., 2007). Importantly, these findings indicate that placental expression was sufficient in the common ancestor to have driven the evolution of genomic imprinting for this gene.

12.2.3 IGF2R

IGF2R is known as the IGF2-receptor because in humans and mice its gene product acts to internalise IGF2 (as well as other extracellular ligands) to the lysozymes for degradation (Kornfeld, 1992). *IGF2R* is maternally expressed in the American opossum, *Didelphis virginiana* (Killian et al., 2000). Significantly, non-imprinted expression of *IGF2R* was found in monotremes and chicken, implying that, like *IGF2*, *IGF2R* acquired imprinting after therian mammals (eutherians and marsupials) diverged from monotremes (Killian et al., 2000; Nolan et al., 2001). Interestingly, IGF2R does not bind IGF2 in these non-imprinted species presumably because it is missing key residues in its attachment site (Killian et al., 2000). Thus, not only did imprinted expression of *IGF2R* evolve at a similar time to *IGF2*, but so did its antagonistic relationship with *IGF2* (Fig. 12.1).

In mice, *Igf2r* imprinting, as well as imprinted expression of two neighbouring genes *Slc22a2* and *Slc22a3*, is under the control of a differentially methylated region in the second intron of *Igf2r* (Lyle et al., 2000) and a paternally expressed

non-coding RNA (*Air*) that is antisense to *Igf2r* (Sleutels et al., 2002). It is thought that marsupial *IGF2R* does not require this region of differential methylation or the *AIR* transcript for its imprinted expression, as neither have been found in opossum (Killian et al., 2000; Weidman et al., 2006a). The same also appears to be true of canine *IGF2R* (O'Sullivan et al., 2007), casting doubt upon the applicability of mouse as a model for all mammalian imprinting at the *IGF2R* locus. Despite previous doubts of its existence (Oudejans et al., 2001) an *AIR* transcript has recently been identified in human, and along with differential methylation of intron 2, may contribute to the regulation of polymorphic imprinting at human *IGF2R* (Yotova et al., 2008). Thus, it is possible that *AIR* is a recent addition to the control of imprinting in the order Euarchontoglires (Supraprimates), which diverged from other eutherians about 95 MYA (Bininda-Emonds et al., 2007).

12.2.4 PEG10 Locus

The Paternally Expressed Gene 10 (*PEG10*) locus provides a unique example of how an imprinted region was constructed during mammalian evolution. Genomic imprinting is thought to be derived from methylation mechanisms already existing in the cell to silence foreign inserting DNA (Barlow, 1993). This occurs in retrotransposons and often in transgenic inserted DNA fragments (Chaillet, 1994). *PEG10* is itself derived from a retrotransposon (of the sushi-ichi class; Ono et al., 2001). Comparisons of the genomes of the three extant mammalian lineages showed that *PEG10* is absent from the platypus genome, but present in the marsupial genome, implying that it was inserted after the monotreme-therian divergence around 170 million years ago (Suzuki et al., 2007). Its insertion into the genome coincided with its acquisition of imprinting. Furthermore, this imprint is controlled by a differentially methylated region, as in eutherians, and confers complete silencing of the maternal allele. Differentially methylated regions associated with imprinting have so far been identified only at two marsupial domains (*IGF2/H19* and *PEG10*) (Suzuki et al., 2007; Lawton et al., 2008; Smits et al., 2008). Other marsupial imprinted genes, such as *PEG1/MEST* lack regions of differential methylation and, despite showing allelic bias, they do not completely silence the attenuated allele (Suzuki et al., 2005). However, monoallelic expression need not be absolute or complete even in eutherians, with low expression from the "silent" parental allele occurring for several imprinted genes (Matsuoka et al., 1996). This may suggest, that at least in some imprinted clusters, differential methylation evolved after imprinting was established, possibly to *strengthen* the imprint.

In eutherian mammals, imprinting in the *PEG10* region encompasses its upstream neighbour *SGCE* (Ono et al., 2001). *SGCE* is also adjacent to *PEG10* in marsupials, but imprinting is restricted to the *PEG10* gene and *SGCE* is bialleleic (Suzuki et al., 2007). Thus, the *PEG10* locus provides a clear picture of the genesis and evolution of a new imprinted domain. Initially, insertion of the retrotransposed *PEG10* gene into the genome attracted silencing. As *PEG10* has a role in placental development, this silencing could have conferred an evolutionary advantage, so *PEG10* imprinting was selected for and maintained in therian mammals. In

eutherians, imprinting spread to encompass the neighbouring *SGCE* gene. These findings add strong support to the host defence hypothesis (Barlow, 1993), as does examination of the genomes of non-imprinted species such as platypus (see below), potentially explaining how silencing mechanisms were recruited to imprinted loci.

12.3 Clusters Not Imprinted in Marsupials

Proving that imprinting does not occur at a given locus is more problematic than proving that it does, as it is difficult to exclude monoallelic expression of a gene in all the tissues and possible developmental time-frames. However, a combination of gene expression and genomic analyses has recently provided very compelling evidence for the absence of imprinting at the marsupial orthologues of the Prader-Willi and Angleman's syndrome (PWS-AS) locus and the Callipyge locus (Rapkins et al., 2006; Weidman et al., 2006b; Edwards et al., 2008). Significantly, this work demonstrates that not all domains evolved imprinting in the ancestor of therian mammals; rather they acquired parent-specific expression only in eutherians, perhaps in response to new or changing selective pressures (Hore et al., 2007b).

12.3.1 Prader-Willi and Angelman's Syndrome Locus

Deletions or malfunction of imprinted genes on human chromosome 15q11–13 are responsible for the neurological disorders Prader-Willi and Angelman's syndromes (PWS and AS – Nicholls and Knepper, 2001). This locus contains multiple paternally- and maternally-expressed genes, all of which are under the control of a single ICR. This ICR is located in the promoter of a large (>100 kb), paternally expressed transcript that is responsible for the expression of two distinct proteins (SNURF and SmN – see Gray et al., 1999a) and several classes of small nucleolar RNAs (snoRNAs) (Runte et al., 2001). This remarkable transcript (called *SNURF-SNRPN* after the two protein-encoding genes it contains) also overlaps a neighbouring maternally expressed gene (*UBE3A*) in an antisense fashion, presumably providing the regulatory mechanism linking imprinted expression of *UBE3A* to the ICR (Runte et al., 2004). Located distal to *UBE3A* is another maternally expressed gene (*ATP10A*), and there are three intronless, paternally expressed genes located proximal to *SNRPN* (*NDN*, *MAGEL2* and *MKRN3*) (Nicholls and Knepper, 2001).

Although genes from the PWS-AS region are conserved in number and arrangement throughout eutherians, most appear not to exist outside this mammal infraclass. Attempts to find *NDN*, *MAGEL2* and *MKRN3* in marsupials resulted only in the discovery of their intron-bearing progenitors (implying that they arose by retrotransposition), whereas no sequence with similarity to the *SNURF* gene could be found (Gray et al., 2000; Rapkins et al., 2006). An orthologue of *SNRPN* was discovered in marsupials adjacent to a very similar gene *SNRPB*, but could not be found in non-therians or other vertebrates. The position of *SNRPN* in marsupials and its

absence from the genomes of monotremes and non-mammalian vertebrates indicates that *SNRPN* arose from *SNRPB* by tandem duplication during early eutherian radiation (Gray et al., 1999b; Rapkins et al., 2006).

In marsupials, *UBE3A* (the gene that lies next to *SNRPN* in eutherians) is located on a different chromosome to *SNRPN*, lying next to a gene called *CNGA3*. This unexpected *UBE3A-CNGA3* conformation was shared with the platypus, chicken and fish genomes, implying that it was the ancestral vertebrate arrangement. A rearrangement that brought *UBE3A* and *SNRPN* together must have occurred in eutherians soon after their divergence from the marsupials. A single snoRNA (*SNORD119*) was found within marsupial *SNRPN* and is thought to have seeded at least one of the large arrays of snoRNAs characteristic of the eutherian locus, presumably by tandem duplication (Nahkuri et al., 2008). No trace of the regulatory elements, including the ICR of the PWS-AS region, could be found in marsupials or any other non-eutherian. In line with the absence of regulatory elements and the dislocation of the PWS-AS domain, expression of marsupial and monotreme *SNRPN* and *UBE3A* orthologues was found to be non-imprinted (Rapkins et al., 2006). Thus, the PWS-AS region was assembled from a range of components found in disparate regions of the genome, and acquired imprinting only in the eutherian lineage.

12.3.2 Callipyge Locus

The Callipyge locus on human chromosome 14q32 contains an diverse collection of imprinted genes and is named after a striking posterior muscle trait for which its orthologue in sheep is responsible (Callipyge means beautiful buttock) (Cockett et al., 1996). This imprinted cluster contains three paternally expressed protein-encoding genes; the delta-like homologue which causes the callipyge phenotype when overexpressed (*DLK1*), the retrotransposon-derived *RTL1* gene and the type 3 deiodinase, *DIO3* (reviewed in da Rocha et al., 2008). The Callipyge domain also contains maternally expressed sequences, including a large non-coding transcript (*GTL2*) and several classes of micro RNAs (miRNAs) and snoRNAs (Charlier et al., 2001; Cavaille et al., 2002; Seitz et al., 2004). Interestingly, seven of these miRNAs lie within the *RTL1* gene, but on the anti-sense strand, and have been shown to repress *RTL1* expression on the maternally derived chromosome through RNA-interference (Davis et al., 2005).

In vertebrates, the *DLK1* and *DIO3* genes lie in close proximity (Dunzinger et al., 2005). Thorough searches for homology between these genes has uncovered around 50 regions that are conserved between mammalian groups, but none that overlap with the ICR or any of the large imprinted non-coding RNAs, miRNAs or snoRNAs. *DLK1* and *DIO3*, the only identifiable marsupial orthologues from the Callipyge domain, were found to have bi-allelic expression in wallaby and platypus (Weidman et al., 2006b; Edwards et al., 2008). Thus, in parallel to the PWS-AS region (Rapkins et al., 2006), many of the genes from the Callipyge domain do not exist in marsupials and other non-eutherian species, and of the ones that are present, none are imprinted (Weidman et al., 2006b; Edwards et al., 2008).

Although Callipgye is clearly a eutherian-specific imprinted domain, traces of a degraded *RTL1* sequence have been found in marsupials, but not platypus or any non-mammals (Edwards et al., 2008). Thus, it is proposed that *RTL1* was inserted by retrotransposition sometime during the early evolution of therian mammals. In eutherians this inserted sequence apparently gained functionality, as *RTL1* is required for placental development in mice, but the gene appears to have succumbed to pseudogenisation in marsupials. It is interesting to compare this initial insertion event with that of the *PEG10* insertion, which is thought to have seeded imprinting at the *PEG10* locus. A common thread between these domains may be that repeat insertion provided a mechanism for imprinting to evolve, perhaps through host-defence mechanisms (Barlow, 1993; Pask et al., 2009).

12.4 Distribution of Repeats in Mammalian Genomes and the Host Defence Hypothesis

The recent sequencing of the platypus genome (Warren et al., 2008) made it possible to perform genome wide comparisons across all three extant lineages of mammals. This was of particular importance to understanding the genesis of genomic imprinting, as imprinting is thought not to occur in the monotreme genome, but be restricted to marsupials and eutherians.

The distribution of repeat elements known to attract epigenetic silencing was therefore examined in the genomic regions orthologous to eutherian imprinted domains from representatives of each mammalian lineage (Pask et al., 2009). A significant accumulation of long terminal repeats and DNA elements was identified in therian mammals when compared to platypus, implying that they expanded in therians after their divergence from monotremes. Repeat insertion can trigger epigenetic silencing of surrounding alleles. If this silencing conferred an evolutionary advantage when applied in a parent-specific manner (such as in the *PEG10* gene, for example) it would be selected for and maintained, evolving into an imprinted locus.

Significantly, repeat accumulation is not a feature just of the imprinted domains in therian mammals, but of the entire genome of therian mammals (Warren et al., 2008; Pask et al., 2009). Thus, it is possible that repeat accumulation in nearby regions was not critical for the evolution of imprinted genes per se. Rather; it may have been responsible for the evolution of specialised epigenetic machinery which was primarily used to silence the new wave of therian repeats, then secondarily recruited to the regulation of imprinting.

12.5 The X-Chromosome

One spectacular, yet often ignored, example of mammalian imprinting is the inactivation of the paternally derived X chromosome in female marsupials (Cooper et al., 1971; Richardson et al., 1971). X inactivation balances gene dosage in male marsupials that have only one X chromosome and a small gene poor Y chromosome.

X-chromosome inactivation also occurs in humans and mice, but affects X chromosomes originating from both parents by inactivating one or the other X randomly during early development. This leads to a mosaic phenotype that is observable in female mice and cats that possess X-linked coat colour genes (Lyon, 1961) and in women suffering from X-linked skin diseases that display developmental "Lines of Blascko" (Happle, 1985). Interestingly, in the extraembryonic tissue (including the yolk sac and allantoic placentae) of rodents (Takagi and Sasaki, 1975; Wake et al., 1976; West et al., 1977) and cattle (Dindot et al., 2004) (but apparently not humans; Migeon et al., 1985), it appears that the inactive X is paternally derived, perhaps indicating an evolutionary link with marsupial X inactivation.

Although X inactivation deserves discrete review in its own right (see Al Nadaf et al., Chapter 13), it is important to recognise that the X does possess *bona fide* imprinted genes worthy of evolutionary analysis. Perhaps the most interesting of these belong to the X-inactivation centre that orchestrates silencing of the entire eutherian X chromosome, largely through the action of a large non-coding RNA called the X-inactive specific transcript (*XIST*) (reviewed Heard et al., 2004; Payer and Lee, 2008).

Marsupials lack *XIST* and the other non-coding RNAs from the eutherian X-inactivation centre (Duret et al., 2006; Davidow et al., 2007; Hore et al., 2007a; Shevchenko et al., 2007), so this locus must have evolved more recently in the eutherian lineage. It appears that *XIST* and at least some of the other imprinted non-coding RNAs from this locus have been derived from ancestral protein-coding genes that succumbed to pseudogenistation during mammalian evolution (Duret et al., 2006; Elisaphenko et al., 2008). Since marsupials possess these protein coding genes (albeit somewhat rearranged compared to chicken and frog; Davidow et al., 2007; Hore et al., 2007a; Shevchenko et al., 2007), it is thought that these pseudogenisation events occurred after the divergence of marsupials from eutherians. Concomitant with pseudogenisation and rearrangement was invasion of retroposed genes, repeats and small non-coding RNAs. Thus the X-inactivation centre has been constructed in a manner surprisingly similar to that of other imprinted domains, including PWS-AS and Callipyge (Hore et al., 2007b).

Whether or not marsupial X inactivation is under the control of a non-homologous imprinted domain analogous to the eutherian X-inactivation centre, or depends upon some other mechanism (such as meiotic sex-chromosome inactivation in males – see Al Nadaf et al., Chapter 13) is currently a field of active research and debate.

12.6 Regulators of Imprinting in Marsupials

12.6.1 Evolution of the DNA Methyltransferase 3 Family

In the same way that comparative genomics can shed light on the evolution of imprinted gene clusters, similar approaches have recently elucidated the evolution of genes that regulate imprinting.

Imprinted genes undergo a generational "life-cycle". In the germ cells there is removal of imprint marks received from mother and father, and establishment of new sex-specific marks ready to be passed on to the offspring (Reik et al., 2001). To establish DNA methylation at non-methylated sites, mammals use the DNA methyltransferase 3 family members *DNMT3A* and *DNMT3B* (Goll and Bestor, 2005). There is a third member of the DNA methyltransferase 3 family in mammals called DNMT3L, which lacks the critical domain for DNA methyltransferase ability, but apparently enhances the function of the other DNMT3 members at a subset of methylated sites by connecting them to histone H3 tails that are unmethylated at lysine 4 (Jia et al., 2007; Ooi et al., 2007). Male mice with *Dnmt3L* knocked-out are viable, but are sterile because of defects in meiosis, presumably due to repeat activation in the absence of methylation (Bourc'his et al., 2001; Hata et al., 2002; Kaneda et al., 2004; Webster et al., 2005). Interestingly, global methylation patterns in female knockouts appear relatively unaffected, but they are also sterile and are unable to impart maternal specific methylation at ICRs in growing oocytes (Bourc'his et al., 2001; Hata et al., 2002; Kaneda et al., 2004).

Recently, the existence of the DNMT3 family was investigated in marsupials and non-mammal vertebrates (Yokomine et al., 2006). *DNMT3A* and *DNMT3B* orthologues were discovered in all species examined, which is perhaps not surprising given that they are the only de-novo methyltransferases detected in vertebrates. *DNMT3L* showed a reduced phylogenetic presence, being found only in eutherians and marsupials, but not birds, amphibians or fish. The presence of *DNMT3L* in marsupials as well as eutherians suggests that its role in establishing de-novo methylation at maternally imprinted clusters is conserved in therians (Yokomine et al., 2006; Smits et al., 2008). As *DNMT3L* was not detected in birds, amphibians or fish, it was originally proposed that *DNMT3L* arose in early therian mammals prior to the divergence of marsupials and eutherians, around the same time as imprinting. However, searches in the lizard genome (*Anolis carolinensis*) yield a clear orthologue of *DNMT3L* (assembly AnoCar 1.0, scaffold118), implying that *DNMT3L* is much more ancient than predicted, having arisen sometime before the divergence of mammals from birds and reptiles.

12.6.2 Evolution of CTCF and BORIS

CCCTC-binding factor (*CTCF*) and its paralogue, Brother of Regulator of Imprinted Sites (*BORIS*) are epigenetic regulators that share the same DNA-binding zinc-finger domain. However, outside the DNA binding domain they share almost no similarity and have completely different expression patterns – *CTCF* is expressed ubiquitously, whereas *BORIS* is expressed only in the germline (Loukinov et al., 2002; Jelinic et al., 2006; Kholmanskikh et al., 2008; Monk et al., 2008). Furthermore, the products of these genes appear to have quite different functions at the *IGF2/H19* ICR; BORIS is thought to *establish* paternal-specific methylation of this site through interaction with members of the DNMT3 family and a histone modifier called PRMT7 (Jelinic et al., 2006), whereas CTCF *interprets* this imprint later in

development, giving rise to parent-specific expression. CTCF does this by binding the ICR in a methylation-sensitive fashion. On the maternally derived, unmethylated ICR, CTCF insulates the nearby *IGF2* gene from downstream enhancers, silencing its expression (Bell and Felsenfeld, 2000; Hark et al., 2000). On the paternally derived chromosome, CTCF cannot bind the methylated ICR, leaving *IGF2* open for expression.

Recently, researchers have endeavoured to find out if *CTCF* and *BORIS* are present in marsupial genomes, and if there is anything related to their history that can shed light upon the evolution of imprinting. In line with its essential and varied functions within the vertebrate cell (reviewed Filippova, 2008) *CTCF* was found to be highly conserved in sequence and expressed ubiquitously in all vertebrates (Klenova et al., 1993; Klenova et al., 1998; Burke et al., 2002; Moon et al., 2005; Pugacheva et al., 2006), including marsupials (Hore et al., 2008). Thus, there is nothing striking about the history of *CTCF* that correlates with the evolution of imprinting.

However, *BORIS* shows a much reduced phylogenetic presence. Originally, it was thought that it arose early in the mammal lineage, as *BORIS* orthologues were discovered in humans and mice, but not chicken (Loukinov et al., 2002). However, a more recent analysis has revealed *BORIS* orthologues in all mammalian groups and two lizard species (*A. carolinensis* and *Pogona vitticeps*), as well as traces of a degraded *BORIS* sequence in bird genomes (Hore et al., 2008). Thus, in contrast to what was originally proposed, the genesis of *BORIS* is not correlated with the evolution of imprinting.

However, examination of *BORIS* expression showed that it is actively transcribed in somatic tissues of monotremes (platypus) and reptiles (bearded dragon lizard), yet in marsupials (tammar wallaby) and cattle, it is expressed only in the germline, just as in humans and mice. Thus, *BORIS* is specific to the germ-line in mammal groups that undergo imprinting, but shows wider expression in species without imprinting such as monotremes and non-mammalian vertebrates. This implies that, although *BORIS* arose well before imprinting, restriction of its expression to the germline was associated with its recruitment to the regulation of imprinted genes.

12.7 Why Did Genomic Imprinting Evolve?

So far we have reviewed what is known about marsupial orthologues of human and mouse imprinted genes and their regulators. In doing so, we gather an understanding of *how* they have been constructed over evolutionary time; a process apparently involving considerable genomic rearrangement, changes in repeat composition, insertion of retroposed genes, evolution of epigenetic regulators and the genesis of large and small non-coding RNAs (Fig. 12.3). However, it is of equal interest to understand *why* imprinted genes have discarded biallelic expression, despite exposure to considerable mutational load. Although this topic is fiercely debated, most agree that imprinting is related to the evolution of viviparity and the placenta.

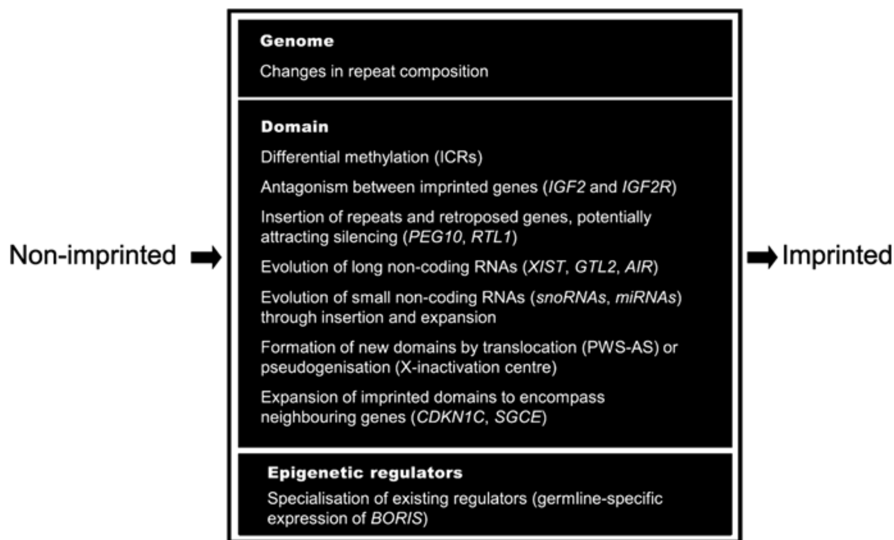


Fig. 12.3 Genomic changes associated with the evolution of imprinting. Changes occurring at the genomic level, domain level and at epigenetic regulators are listed, with specific examples given in *brackets*

12.7.1 Genomic Imprinting and the Evolution of the Placenta

There is an intimate link between imprinting and the mammalian placenta. Nearly all of the imprinted genes so far tested in mice are expressed and imprinted (often exclusively) in the placenta or yolk sac (Reik et al., 2003). Furthermore, many imprinted genes affect the size and function of the placenta – when paternally expressed genes such as *Igf2*, *Peg1*, *Peg3*, *Rtl1* and *Slc38a4* are knocked out, the size of the placenta and fetus is reduced, whereas ablation of maternally expressed *Igf2r*, *Phlda2*, *Grb10*, *Cdkn1c* and *H19* causes placental and fetal overgrowth (reviewed Angiolini et al., 2006). Some imprinted genes can simultaneously affect the supply of nutrients through the placenta and the demand of nutrients from the fetus, suggesting that imprinted genes play an important role in fine tuning both sides of the materno-fetal interface (Constancia et al., 2002; Kaneko-Ishino et al., 2003; Reik et al., 2003; Angiolini et al., 2006). Because of the close relationship between imprinted genes and the mammalian placenta, it has often been speculated that their evolution is also somehow related.

Marsupials represent the most distantly related mammalian clade that transfers sufficient nutrients to their offspring *in-utero* to make possible live-birth, albeit to altricial young. All genes that have been identified as imprinted in marsupials show imprinted expression in the placenta where tested [*IGF2, PEG1/MEST* (Suzuki et al., 2005), *PEG10* (Suzuki et al., 2007) and *INS* (Ager et al., 2007)]. In the case of *INS*, the placenta is the only tissue displaying imprinted expression in marsupials.

These results suggest that in marsupials, as in eutherians, the placenta is a key tissue for imprinted gene expression (Renfree et al., 2008).

In contrast to live-bearing therian mammals, no parental bias has so far been discovered in the expression of genes from the egg-laying monotremes, perhaps indicating that like other egg-laying vertebrates, they lack genomic imprinting (Killian et al., 2000; O'Neill et al., 2000; Killian et al., 2001b; Killian et al., 2001a; Nolan et al., 2001; Rapkins et al., 2006; Deakin et al., 2008; Edwards et al., 2008). Thus, the change in the mode of reproduction, as the livebearing therian mammals diverged from egg-laying progenitors, may have begun to apply the evolutionary pressure from which imprinting evolved.

What aspects of live-birth and the placenta promote the evolution of parent-specific expression? Currently, there is no accepted answer to this question, but several hypotheses have been discussed extensively in the literature. Perhaps the most widely discussed, and with the most relevance to marsupials, is the kinship hypothesis.

12.7.2 The Kinship Hypothesis of Genomic Imprinting

The kinship hypothesis (also known as the parental-conflict, or parental tug-of-war hypothesis) states that imprinting evolved out of a conflict between the mother's and father's genomes over the provision of maternal resources (Haig and Westoby, 1989; Moore and Haig, 1991; Haig, 2000; Wilkins and Haig, 2003). Specifically, a father can maximise his reproductive output by ensuring that his offspring receive as much maternal resources as possible, regardless of the expense to the mother and her offspring fathered by other males. Conversely, a mother maximises her reproductive output by sharing her maternal resources equally to all her offspring, regardless of the father. The disparity between these two selection pressures is expected to result in parent-specific expression of genes which can modify the supply of maternal resources during development.

As the placenta is entirely fetally-derived, often invasive and can affect the supply and demand of nutrients to offspring during pregnancy, it offers an ideal arena for parental conflict. As previously discussed, many paternally expressed genes promote growth of the placenta and offspring, whereas maternally expressed genes act to suppress growth of the placenta and offspring (Angiolini et al., 2006). Thus, the function of imprinted genes in the placenta provides considerable support for the kinship hypothesis. Some imprinted genes with effects outside of the placenta also provide support for the kinship hypothesis. For instance, mutations in the paternally expressed XLαs protein cause growth retardation and a failure to suckle in newborn humans and mice (Plagge et al., 2004; Genevieve et al., 2005), indicating that parental conflict acting during lactation may also give rise to parent-specific expression.

The phylogenetic distribution of imprinting provides solid support for the kinship hypothesis. Imprinting does not occur in birds, amphibians, reptiles and fish which are predominately egg-laying, so that paternally derived genes have

little or no chance of affecting the supply of maternal resources. Paradoxically, non-mammalian vertebrates that possess placenta-like structures also support the kinship hypothesis. Live-bearing killifish *Heteroandria formosa* and *Poeciliopsis prolifica* exhibit accelerated positive selection of *IGF2* sequence compared with their egg-laying relatives (O'Neill et al., 2007). Thus, although these fish are without imprinting, accelerated positive selection is thought to provide evidence for parental-conflict due to viviparity, being caused by successive selective sweeps of antagonistic mutations that favour either the paternal or maternal lineage.

Marsupials represent an interesting case for the kinship hypothesis, as they are live-bearing, but give birth to tiny altricial young. One would hypothesise that parental-conflict would be a weaker selective force than in eutherian mammals that typically give birth to offspring that are much larger relative to the mother (Renfree et al., 2009). Perhaps this is why imprinting in marsupials is less pronounced: often there is some degree of expression from silenced alleles in marsupials, and their imprinted regions are smaller or, in some cases, absent. Moreover, the evolution of eutherian-specific imprinted loci (e.g. Callipyge and PWS-AS) or the expansion of existing imprinted loci in eutherians (e.g., *CDKN1C* and *PEG10*) may be the result of stronger selection for parent-specific expression due to greater parental conflict *in-utero*.

12.7.3 Alternative Theories: Ovarian Time-Bomb and Co-adaptation

Although successful in explaining imprinting of genes involved in growth, the kinship hypothesis has trouble accounting for the function of all imprinted genes. For instance, the maternally expressed *Slc22a2* and *Slc22a3* are small solute carriers that appear to have no impact on growth or acquisition of maternal resources, so are hard to explain by kinship (Haig, 2004). Furthermore, although interesting explanations have been offered (Haig and Wharton, 2003; Ubeda, 2008), the largely neurological disorders that arise from perturbation of the imprinted PWS-AS domain also do not neatly fit with the kinship hypothesis. Thus, other selective advantages which are thought to outweigh the costs of genomic imprinting have been proposed; including intralocus sexual conflict (Day and Bonduriansky, 2004), protection against invasive trophoblastic disease (Varmuza and Mann, 1994), and maternal-infant co-adaptation (Keverne and Curley, 2008).

Of these, marsupial imprinting is particularly relevant to the latter two, as they both require placentation and viviparity for their application. For instance, the hypothesis that imprinting protects against invasive trophoblastic disease (dubbed the "Ovarian time-bomb" hypothesis) predicts that imprinting prevents eggs from spontaneously developing by ensuring that essential growth factors are paternally-expressed and are therefore only active in the zygote post-fertilisation (Varmuza and Mann, 1994). The risk that unfertilised eggs could spontaneously develop into a cancerous growth capable of harming the mother is presumably much less in species without a placenta, potentially explaining why egg-laying species do not

show imprinting. However, the ovarian time-bomb hypothesis falters in that it fails to explain why so many genes are imprinted, and why their mono-allelic expression often persists into adulthood, both of which are presumably unnecessary for the prevention of trophoblastic disease (Wilkins and Haig, 2003).

The maternal co-adaptation hypothesis for the evolution of imprinting bases its arguments on the maternal-young interactions from fetus to neonate, so may apply to marsupials. A subset of imprinted genes are active in both the placenta (tissues of fetal origin), and the brain (of offspring and mother). These genes can alter physiological and behavioural traits associated with in utero development, mothering and feeding (Li et al., 1999; Curley et al., 2004). The developing brain is as equally susceptible to imprinting as the placenta. Since both the placenta and the hypothalamus are responsible for the regulation of gases, water and nutrients, these effects overlap in function and developmental timing. Thus they may be co-adapted to ensure complementary functions between mother and offspring, rather than conflict over maternal resources (Keverne and Curley, 2008). Despite the attractiveness of this hypothesis, maternal-infant co-adaptation cannot provide evolutionary reasons for all imprinted genes to have acquired monoallelic expression, because not all imprinted genes have overlapping functions in the mother and infant. Therefore, like all other theories regarding the evolution of imprinting, this area will require further research.

12.8 Conclusions/Future Research

The hazy understanding of what selective pressures caused the evolution of genomic imprinting underlines the importance of studying imprinted genes in marsupials. All major theories to explain the evolution of imprinting are levered on the same fulcrum of mammalian history – the evolution of placentation sufficient to support live-birth. Yet only the major imprinted loci have been investigated in marsupials, and often to a lower level than their human and mouse orthologues. Despite this, it is now clear that imprinting in marsupials is less common and less penetrant than eutherians at the well characterised loci. Perhaps this is a reflection of the lower dependence of marsupials on placentation to nurture their young, so they do not experience the same selective pressures that eutherians do.

Alternatively, and perhaps more excitingly, we may soon discover that marsupials possess a suite of imprinted genes that are specific to their taxa and reproductive lifecycles. For instance, since lactation is relatively more important than placental nutrition in marsupials, additional imprinted genes may be found that affect the supply and demand of milk. Moreover, there is potential for exaggerated imprinting in the dasyurid and didelphid marsupials, certain members of which are the only mammals to achieve near semelparity (i.e. single birth) and experience mass die-off after breeding (Cockburn, 1997). Perhaps the best studied of these is the broad-footed marsupial mouse (*Antechinus spp*). Males of this species participate in only one breeding season, and only 7% of females survive for two breeding seasons (Holleley et al., 2006). However, they make up for

this by producing large litters that are sired by up to four fathers (Fisher et al., 2006). Fierce selection between half-sibs competing for the same maternal resources might accelerate the evolution of imprinting to levels not yet observed in other mammals.

In either case, the study of marsupials is paramount if our understanding of the evolution and function of genomic imprinting is to be advanced and extended in the forthcoming era of massively-parallel sequencing and epigenomics. Given the wealth of significant information already gleaned from the study of marsupial epigenetics, it is almost certain these “alternative mammals” will yield many more fundamental insights critical to mammalian biology.

References

Ager E, Suzuki S, Pask AJ, et al. (2007) Insulin is imprinted in the placenta of the marsupial, *Macropus eugenii*. Dev Biol 309:317–328.

Ager EI, Pask AJ, Gehring HM, Shaw G, Renfree MB (2008a). Evolution of the CDKN1C-KCNQ1 imprinted domain. BMC Evol Biol 8:163.

Ager EI, Pask AJ, Shaw G, Renfree MB (2008b). Expression and protein localisation of IGF2 in the marsupial placenta. BMC Dev Biol 8:17.

Angiolini E, Fowden A, Coan P, et al. (2006) Regulation of placental efficiency for nutrient transport by imprinted genes. Placenta 27(Suppl A):S98–S102.

Barlow DP (1993) Methylation and imprinting: from host defense to gene regulation? Science 260:309–310.

Bartolomei MS, Zemel S, Tilghman SM (1991) Parental imprinting of the mouse H19 gene. Nature 351:153–155.

Bell AC, Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 405:482–485.

Bininda-Emonds OR, Cardillo M, Jones KE, et al. (2007) The delayed rise of present-day mammals. Nature 446:507–512.

Bourc’his D, Xu GL, Lin CS, Bollman B, Bestor TH (2001) Dnmt3L and the establishment of maternal genomic imprints. Science 294:2536–2539.

Burke LJ, Hollemann T, Pieler T, Renkawitz R (2002) Molecular cloning and expression of the chromatin insulator protein CTCF in *Xenopus laevis*. Mech Dev 113:95–98.

Cai X, Cullen BR (2007) The imprinted H19 noncoding RNA is a primary microRNA precursor. RNA 13:313–316.

Cavaille J, Seitz H, Paulsen M, Ferguson-Smith AC, Bachellerie JP (2002) Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. Hum Mol Genet 11:1527–1538.

Chabowski A, Coort SLM, Calles-Escandon J, et al. (2004) Insulin stimulates fatty acid transport by regulating expression of FAT/CD36 but not FABPpm. Am J Physiol Endocrinol Metab 287:781–789.

Chaillet JR (1994) Genomic imprinting: lessons from mouse transgenes. Mutat Res 307:441–449.

Charlier C, Segers K, Wagenaar D, et al. (2001) Human-ovine comparative sequencing of a 250-kb imprinted domain encompassing the callipyge (clpg) locus and identification of six imprinted transcripts: DLK1, DAT, GTL2, PEG11, antiPEG11, and MEG8. Genome Res 11:850–862.

Cockburn A (1997) Living slow and dying young: senescence in marsupials. In: Saunders N, Hinds L (eds) Marsupial Biology: Recent Research, New Perspectives. University of New South Wales Press, Sydney.

Cockett NE, Jackson SP, Shay TL, et al. (1996) Polar overdominance at the ovine callipyge locus. Science 273:236–238.

Colosi DC, Martin D, More K, Lalande M (2006) Genomic organization and allelic expression of UBE3A in chicken. Gene 383:93–98.

Constancia M, Hemberger M, Hughes J, et al. (2002) Placental-specific IGF-II is a major modulator of placental and fetal growth. Nature 417:945–948.

Cooper DW, VandeBerg JL, Sharman GB, Poole WE (1971) Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. Nat New Biol 230:155–157.

Curley JP, Barton S, Surani A, Keverne EB (2004) Coadaptation in mother and infant regulated by a paternally expressed imprinted gene. Proc Biol Sci 271:1303–1309.

da Rocha ST, Edwards CA, Ito M, Ogata T, Ferguson-Smith AC (2008) Genomic imprinting at the mammalian Dlk1-Dio3 domain. Trends Genet 24:306–316.

Davidow LS, Breen M, Duke SE, et al. (2007) The search for a marsupial XIC reveals a break with vertebrate synteny. Chromosome Res 15:137–146.

Davis E, Caiment F, Tordoir X, et al. (2005) RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. Curr Biol 15:743–749.

Day T, Bonduriansky R (2004) Intralocus sexual conflict can drive the evolution of genomic imprinting. Genetics 167:1537–1546.

Deakin JE, Hore TA, Koina E, Graves JAM (2008) The status of dosage compensation in the multiple X chromosomes of the platypus. PLoS Genet 4:e1000140.

DeChiara TM, Robertson EJ, Efstratiadis A (1991) Parental imprinting of the mouse insulin-like growth factor II gene. Cell 64:849–859.

Deltour L, Montagutelli X, Guenet JL, Jami J, Paldi A (1995) Tissue- and developmental stage-specific imprinting of the mouse proinsulin gene, Ins2. Dev Biol 168:686–688.

Dindot SV, Farin PW, Farin CE, et al. (2004) Epigenetic and genomic imprinting analysis in nuclear transfer derived *Bos gaurus/Bos taurus* hybrid fetuses. Biol Reprod 71: 470–478.

Dunzinger U, Nanda I, Schmid M, Haaf T, Zechner U (2005) Chicken orthologues of mammalian imprinted genes are clustered on macrochromosomes and replicate asynchronously. Trends Genet 21:488–492.

Duret L, Chureau C, Samain S, Weissenbach J, Avner P (2006) The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. Science 312:1653–1655.

Edwards CA, Mungall AJ, Matthews L, et al. (2008) The evolution of the DLK1-DIO3 imprinted domain in mammals. PLoS Biol 6:e135.

Elisaphenko EA, Kolesnikov NN, Shevchenko AI, et al. (2008) A dual origin of the Xist gene from a protein-coding gene and a set of transposable elements. PLoS ONE 3:e2521.

Fehlmann M, Le Cam A, Freychet P (1979) Insulin and glucagon stimulation of amino acid transport in isolated rat hepatocytes. Synthesis of a high affinity component of transport. J Biol Chem 254:10431–10437.

Filippova GN (2008) Genetics and epigenetics of the multifunctional protein CTCF. Curr Top Dev Biol 80:337–360.

Fisher DO, Double MC, Blomberg SP, Jennions MD, Cockburn A (2006) Post-mating sexual selection increases lifetime fitness of polyandrous females in the wild. Nature 444:89–92.

Freeman SJ (1990) Functions of extraembryonic membranes. In: Postimplantation Mammalian Embryos: A Practical Approach. Oxford University Press, New York.

Genevieve D, Sanlaville D, Faivre L, et al. (2005) Paternal deletion of the GNAS imprinted locus (including Gnasxl) in two girls presenting with severe pre- and post-natal growth retardation and intractable feeding difficulties. Eur J Hum Genet 13:1033–1039.

Giddings SJ, King CD, Harman KW, Flood JF, Carnaghi LR (1994) Allele specific inactivation of insulin 1 and 2, in the mouse yolk sac, indicates imprinting. Nat Genet 6:310–313.

Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. Annu Rev Biochem 74: 481–514.

Graves JAM (1996) Mammals that break the rules: genetics of marsupials and monotremes. Annu Rev Genet 30:233–260.

Gray TA, Saitoh S, Nicholls RD (1999a). An imprinted, mammalian bicistronic transcript encodes two independent proteins. Proc Natl Acad Sci USA 96:5616–5621.

Gray TA, Smithwick MJ, Schaldach MA, et al. (1999b). Concerted regulation and molecular evolution of the duplicated SNRPB'/B and SNRPN loci. Nucleic Acids Res 27: 4577–4584.

Gray TA, Hernandez L, Carey AH, et al. (2000) The ancient source of a distinct gene family encoding proteins featuring RING and C(3)H zinc-finger motifs with abundant expression in developing brain and nervous system. Genomics 66:76–86.

Gulbis B, Jauniaux E, Cotton F, Stordeur P (1998) Protein and enzyme patterns in the fluid cavities of the first trimester gestational sac: relevance to the absorptive role of secondary yolk sac. Mol Hum Reprod 4:857–862.

Haig D, Westoby M (1989) Parent-specific gene-expression and the triploid endosperm. American Naturalist 134:147–155.

Haig D (2000) The kinship theory of genomic imprinting. Ann Rev Ecol Syst 31:9–32.

Haig D, Wharton R (2003) Prader-Willi syndrome and the evolution of human childhood. Am J Hum Biol 15:320–329.

Haig D (2004) Genomic imprinting and kinship: how good is the evidence? Annu Rev Genet 38:553–585.

Han L, Lee DH, Szabo PE (2008) CTCF is the master organizer of domain-wide allele-specific chromatin at the H19/Igf2 imprinted region. Mol Cell Biol 28:1124–1135.

Happle R (1985) Lyonization and the lines of Blaschko. Hum Genet 70:200–206.

Hark AT, Schoenherr CJ, Katz DJ, et al. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature 405:486–489.

Hata K, Okano M, Lei H, Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 129:1983–1993.

Heard E, Chaumeil J, Masui O, Okamoto I (2004) Mammalian X-chromosome inactivation: an epigenetics paradigm. Cold Spring Harb Symp Quant Biol 69:89–102.

Hernandez-Sanchez C, Mansilla A, de la Rosa EJ, de Pablo F (2006) Proinsulin in development: new roles for an ancient prohormone. Diabetologia 49:1142–1150.

Holleley CE, Dickman CR, Crowther MS, Oldroyd BP (2006) Size breeds success: multiple paternity, multivariate selection and male semelparity in a small marsupial, Antechinus stuartii. Mol Ecol 15:3439–3448.

Hore TA, Koina E, Wakefield MJ, Graves JAM (2007a). The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. Chromosome Res 15:147–161.

Hore TA, Rapkins RW, Graves JAM (2007b). Construction and evolution of imprinted loci in mammals. Trends Genet 23:440–448.

Hore TA (2008) The Evolution of Genomic Imprinting and X-Chromosome Inactivation in Mammals. PhD thesis, Research School of Biological Sciences, Canberra.

Hore TA, Deakin JE, Graves JAM (2008) The evolution of epigenetic regulators CTCF and BORIS/CTCFL in amniotes. PLoS Genet 4:e1000169.

Jelinic P, Stehle JC, Shaw P (2006) The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation. PLoS Biol 4:e355.

Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature 449:248–251.

Jones CJ, Jauniaux E (1995) Ultrastructure of the materno-embryonic interface in the first trimester of pregnancy. Micron 26:145–173.

Kaneda M, Okano M, Hata K, et al. (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature 429:900–903.

Kaneko-Ishino T, Kohda T, Ishino F (2003) The regulation and biological significance of genomic imprinting in mammals. J Biochem 133:699–711.

Kawahara M, Wu Q, Takahashi N, et al. (2007) High-frequency generation of viable mice from engineered bi-maternal embryos. Nat Biotechnol 25:1045–1050.

Keverne EB, Curley JP (2008) Epigenetics, brain evolution and behaviour. Front Neuroendocrinol 29:398–412.

Kholmanskikh O, Loriot A, Brasseur F, De Plaen E, De Smet C (2008) Expression of BORIS in melanoma: lack of association with MAGE-A1 activation. Int J Cancer 122: 777–784.

Killian JK, Byrd JC, Jirtle JV, et al. (2000) M6P/IGF2R imprinting evolution in mammals. Mol Cell 5:707–716.

Killian JK, Nolan CM, Wylie AA, et al. (2001a). Divergent evolution in M6P/IGF2R imprinting from the Jurassic to the Quaternary. Hum Mol Genet 10:1721–1728.

Killian JK, Nolan CM, Stewart N, et al. (2001b). Monotreme IGF2 expression and ancestral origin of genomic imprinting. J Exp Zool 291:205–212.

Klenova EM, Nicolas RH, Paterson HF, et al. (1993) CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. Mol Cell Biol 13:7612–7624.

Klenova EM, Fagerlie S, Filippova GN, et al. (1998) Characterization of the chicken CTCF genomic locus, and initial study of the cell cycle-regulated promoter of the gene. J Biol Chem 273:26571–26579.

Kono T, Obata Y, Wu Q, et al. (2004) Birth of parthenogenetic mice that can develop to adulthood. Nature 428:860–864.

Kornfeld S (1992) Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors. Annu Rev Biochem 61:307–330.

Lawton BR, Sevigny L, Obergfell C, et al. (2005) Allelic expression of IGF2 in live-bearing, matrotrophic fishes. Dev Genes Evol 215:207–212.

Lawton BR, Carone BR, Obergfell CJ, et al. (2008) Genomic imprinting of IGF2 in marsupials is methylation dependent. BMC Genomics 9:205.

Li L, Keverne EB, Aparicio SA, et al. (1999) Regulation of maternal behavior and offspring growth by paternally expressed Peg3. Science 284:330–333.

Livesey G, Williams KE (1981) Rates of pinocytic capture of simple proteins by rat yolk sacs incubated in vitro. Biochem J 198:581–586.

Loukinov DI, Pugacheva E, Vatolin S, et al. (2002) BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. Proc Natl Acad Sci USA 99:6806–6811.

Luo ZX, Ji Q, Wible JR, Yuan CX (2003) An Early Cretaceous tribosphenic mammal and metatherian evolution. Science 302:1934–1940.

Lyle R, Watanabe D, Vruchte D, et al. (2000) The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. Nat Genet 25:19–21.

Lyon MF (1961) Gene action in the X-chromosome of the mouse (*Mus musculus L*). Nature 190:372–373.

Mann JR, Lovell-Badge RH (1984) Inviability of parthenogenones is determined by pronuclei, not egg cytoplasm. Nature 310:66–67.

Matsuoka S, Thompson JS, Edwards MC, et al. (1996) Imprinting of the gene encoding a human cyclin-dependent kinase inhibitor, p57KIP2, on chromosome 11p15. Proc Natl Acad Sci USA 93:3026–3030.

McGrath J, Solter D (1984) Completion of mouse embryogenesis requires both the maternal and paternal genomes. Cell 37:179–183.

McGrath KE, Palis J (2005) Hematopoiesis in the yolk sac:more than meets the eye. Exp Hematol 33:1021–1028.

Migeon BR, Wolf SF, Axelman J, Kaslow DC, Schmidt M (1985) Incomplete X chromosome dosage compensation in chorionic villi of human placenta. Proc Natl Acad Sci USA 82: 3390–3394.

Monk M, Hitchins M, Hawes S (2008) Differential expression of the embryo/cancer gene ECSA(DPPA2), the cancer/testis gene BORIS and the pluripotency structural gene OCT4, in human preimplantation development. Mol Hum Reprod 14:347–355.

Moon H, Filippova G, Loukinov D, et al. (2005) CTCF is conserved from Drosophila to humans and confers enhancer blocking of the Fab-8 insulator. EMBO Rep 6:165–170.
Moore GE, Abu-Amero SN, Bell G, et al. (2001) Evidence that insulin is imprinted in the human yolk sac. Diabetes 50:199–203.
Moore T, Haig D (1991) Genomic imprinting in mammalian development: a parental tug-of-war. Trends Genet 7:45–49.
Morison IM, Ramsay JP, Spencer HG (2005) A census of mammalian imprinting. Trends Genet 21:457–465.
Nahkuri S, Taft RJ, Korbie DJ, Mattick JS (2008) Molecular Evolution of the HBII-52 snoRNA Cluster. J Mol Biol 381:810–815.
Nicholls RD, Knepper JL (2001) Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. Annu Rev Genomics Hum Genet 2:153–175.
Nolan CM, Killian JK, Petitte JN, Jirtle RL (2001) Imprint status of M6P/IGF2R and IGF2 in chickens. Dev Genes Evol 211:179–183.
O'Neill MJ, Ingram RS, Vrana PB, Tilghman SM (2000) Allelic expression of IGF2 in marsupials and birds. Dev Genes Evol 210:18–20.
O'Neill MJ, Lawton BR, Mateos M, et al. (2007) Ancient and continuing Darwinian selection on insulin-like growth factor II in placental fishes. Proc Natl Acad Sci USA 104:12404–12409.
O'Sullivan FM, Murphy SK, Simel LR, et al. (2007) Imprinted expression of the canine IGF2R, in the absence of an anti-sense transcript or promoter methylation. Evol Dev 9:579–589.
Ogawa O, Eccles MR, Szeto J, et al. (1993) Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. Nature 362:749–751.
Ono R, Kobayashi S, Wagatsuma H, et al. (2001) A retrotransposon-derived gene, PEG10, is a novel imprinted gene located on human chromosome 7q21. Genomics 73:232–237.
Ooi SK, Qiu C, Bernstein E, et al. (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature 448:714–717.
Otto SP, Goldstein DB (1992) Recombination and the evolution of diploidy. Genetics 131: 745–751.
Oudejans CB, Westerman B, Wouters D, et al. (2001) Allelic IGF2R repression does not correlate with expression of antisense RNA in human extraembryonic tissues. Genomics 73:331–337.
Pask AJ, Papenfuss AT, Ager EI, et al. (2009) Analysis of the platypus genome suggests a transposon origin for mammalian imprinting. Genome Biol 10:R1.
Payer B, Lee JT (2008) X chromosome dosage compensation: how mammals keep the balance. Annu Rev Genet 42:733–772.
Plagge A, Gordon E, Dean W, et al. (2004) The imprinted signaling protein XL alpha s is required for postnatal adaptation to feeding. Nat Genet 36:818–826.
Pugacheva EM, Kwon YW, Hukriede NA, et al. (2006) Cloning and characterization of zebrafish CTCF: developmental expression patterns, regulation of the promoter region, and evolutionary aspects of gene organization. Gene 375:26–36.
Pugliese A, Zeller M, Fernandez A Jr., et al. (1997) The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. Nat Genet 15:293–297.
Rainier S, Johnson LA, Dobry CJ, et al. (1993) Relaxation of imprinted genes in human cancer. Nature 362:747–749.
Rapkins RW, Hore T, Smithwick M, et al. (2006) Recent assembly of an imprinted domain from non-imprinted components. PLoS Genet 2:e182.
Reik W, Dean W, Walter J (2001) Epigenetic reprogramming in mammalian development. Science 293:1089–1093.
Reik W, Constancia M, Fowden A, et al. (2003) Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. J Physiol 547:35–44.
Renfree MB (1972) Influence of the embryo on the marsupial uterus. Nature 240:475–477.
Renfree MB (1982) Implantation and placentation. Reprod Mamm 2:26–69.
Renfree MB, Ager EI, Shaw G, Pask AJ (2008) Genomic imprinting in marsupial placentation. Reproduction 136:523–531.

Renfree MB, Hore TA, Shaw G, Marshall Graves JAM, Pask AJ (2009) Evolution of genomic imprinting: insights from marsupials and monotremes. Annu Rev Genomics Hum Genet. 10:241–262.

Richardson BJ, Czuppon AB, Sharman GB (1971) Inheritance of glucose-6-phosphate dehydrogenase variation in kangaroos. Nat New Biol 230:154–155.

Runte M, Huttenhofer A, Gross S, et al. (2001) The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. Hum Mol Genet 10:2687–2700.

Runte M, Kroisel PM, Gillessen-Kaesbach G, et al. (2004) SNURF-SNRPN and UBE3A transcript levels in patients with Angelman syndrome. Hum Genet 114:553–561.

Russo VEA, Martienssen RA, Riggs AD (1996) Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor Laboratory Press, Plainview.

Scott RJ, Spielman M (2006) Genomic imprinting in plants and mammals: how life history constrains convergence. Cytogenet Genome Res 113:53–67.

Seitz H, Royo H, Bortolin ML, et al. (2004) A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Res 14:1741–1748.

Shevchenko AI, Zakharova IS, Elisaphenko EA, et al. (2007) Genes flanking Xist in mouse and human are separated on the X chromosome in American marsupials. Chromosome Res 15: 127–136.

Sleutels F, Zwart R, Barlow DP (2002) The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature 415:810–813.

Smits G, Mungall AJ, Griffiths-Jones S, et al. (2008) Conservation of the H19 noncoding RNA and H19-IGF2 imprinting mechanism in therians. Nat Genet 40:971–976.

Surani MA, Barton SC, Norris ML (1984) Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature 308:548–550.

Suzuki S, Renfree MB, Pask AJ, et al. (2005) Genomic imprinting of IGF2, p57(KIP2) and PEG1/MEST in a marsupial, the tammar wallaby. Mech Dev 122:213–222.

Suzuki S, Ono R, Narita T, et al. (2007) Retrotransposon silencing by DNA methylation can drive mammalian genomic imprinting. PLoS Genet 3:e55.

Takagi N, Sasaki M (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature 256:640–642.

Tremblay KD, Duran KL, Bartolomei MS (1997) A 5′ 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. Mol Cell Biol 17:4322–4329.

Ubeda F (2008) Evolution of genomic imprinting with biparental care: implications for Prader-Willi and Angelman syndromes. PLoS Biol 6:e208.

Vafiadis P, Bennett ST, Todd JA, et al. (1997) Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus. Nat Genet 15:289–292.

Varmuza S, Mann M (1994) Genomic imprinting–defusing the ovarian time bomb. Trends Genet 10:118–123.

Wake N, Takagi N, Sasaki M (1976) Non-random inactivation of X chromosome in the rat yolk sac. Nature 262:580–581.

Warren WC, Hillier LW, Graves JAM, et al. (2008) Genome analysis of the platypus reveals unique signatures of evolution. Nature 453:175–183.

Webster KE, O'Bryan MK, Fletcher S, et al. (2005) Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. Proc Natl Acad Sci USA 102:4068–4073.

Weidman JR, Dolinoy DC, Maloney KA, Cheng JF, Jirtle RL (2006a). Imprinting of Opossum Igf2r in the absence of differential methylation and air. Epigenetics 1:49–54.

Weidman JR, Maloney KA, Jirtle RL (2006b). Comparative phylogenetic analysis reveals multiple non-imprinted isoforms of opossum Dlk1. Mamm Genome 17:157–167.

West JD, Frels WI, Chapman VM, Papaioannou VE (1977) Preferential expression of the maternally derived X chromosome in the mouse yolk sac. Cell 12:873–882.

Wilkins JF, Haig D (2003) What good is genomic imprinting: the function of parent-specific gene expression. Nat Rev Genet 4:359–368.

Yokomine T, Kuroiwa A, Tanaka K, et al. (2001) Sequence polymorphisms, allelic expression status and chromosome locations of the chicken IGF2 and MPR1 genes. Cytogenet Cell Genet 93:109–113.

Yokomine T, Hata K, Tsudzuki M, Sasaki H (2006) Evolution of the vertebrate DNMT3 gene family: a possible link between existence of DNMT3L and genomic imprinting. Cytogenet Genome Res 113:75–80.

Yotova IY, Vlatkovic IM, Pauler FM, et al. (2008) Identification of the human homolog of the imprinted mouse Air non-coding RNA. Genomics 92:464–473.

Chapter 13
Marsupial Genetics Reveals Insights into Evolution of Mammalian X Chromosome Inactivation

Shafagh Al Nadaf, Paul D. Waters, Janine E. Deakin, and Jennifer A. Marshall Graves

Abstract X chromosome inactivation (X-inactivation; XCI) in mammals equalizes X gene dosage between XX females and XY males. It is the prime example of epigenetic repression on a grand scale. This regulatory process differentially treats homologous chromosomes within the same nucleus to ensure that only a single X chromosome remains active in a diploid female cell. The best-studied models of X-inactivation, humans and mice, represent only one of four clades of placental (eutherian) mammals. Comparisons of dosage compensation mechanisms in distantly related eutherian mammals, marsupial and monotreme mammals, and even birds, will offer entirely new insights to the mechanisms and evolution of dosage compensation. In order to reconstruct what dosage compensation mechanisms might have been functioning in the mammalian ancestor, we highlight the molecular similarities and differences between X-inactivation in marsupials and eutherians, and compare them with the partial dosage compensation system observed in monotreme mammals, which appears more similar to bird dosage compensation. We draw parallels between these mechanisms, which may well have evolved independently by drawing from a common epigenetic "toolbox", and therefore utilize similar molecular mechanisms to down-regulate gene expression.

Keywords Dosage compensation · Epigenetics · Evolution · Marsupials · X inactivation

---

S. Al Nadaf (✉)
Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 2601, Australia
e-mail: shafagh.alnadaf@anu.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_13, © Springer Science+Business Media B.V. 2010

13.1 Introduction

One of the most intriguing discoveries in mammalian genetics during the last 50 years is an epigenetic gene regulation event referred to as X chromosome inactivation (X-inactivation; XCI).

The evolution of X-inactivation is inextricably linked to the emergence of heterogamety and the evolution of sex chromosomes. In mammals, sex is normally determined by an XX/female XY/male sex chromosome system in which the Y bears the dominant male determining gene (*SRY*) that sets off a cascade of events leading to testis development. Suppression of recombination with the X to preserve the male specificity of nearby male beneficial loci resulted in Y chromosome degradation. Thus, in many species sex chromosomes are morphologically distinct. Outside of mammals, vertebrates have a variety of genetic and environmental sex determining systems. Birds have female heterogamety, with a ZZ male: ZW female system. The ZW sex pair is completely non-homologous to the mammal XY pair (Fridolfsson et al., 1998; Nanda et al., 1999; Graves, 2001), but there are strong parallels in that the small and gene poor heterochromatic Y/W is a degraded relic of the gene-rich highly conserved X/Z chromosome (Charlesworth, 1996; Graves, 2006).

Since the sex determining chromosome is often small and gene-poor, males and females have different gene dosage from the X (in XY systems) or Z (in ZW systems). A dosage compensation mechanism equalizes this imbalance of gene expression between the sexes. In female therian (marsupial and eutherian) mammals, one X chromosome becomes transcriptionally inactive at an early stage of embryogenesis, a stable and somatically heritable state. This mechanism, which results in differential activity of the two homologous chromosomes in the same nucleus, is an important model system for epigenetic silencing.

13.2 Evolution of Mammalians Sex Chromosomes

Therian mammals have an XX female XY male sex determining system in which the X is large and bears many genes, and the Y chromosome is small, gene-poor and bears the male determining gene *SRY*. Comparative gene mapping between the three extant lineages of mammals shows that the most basal lineage, subclass Prototheria (monotremes), possess a complex of serially translocated sex chromosomes (Grützner et al., 2004; Rens et al., 2004) that share no homology to the sex chromosomes of therian mammals. Monotreme sex chromosomes do not bear the *SRY* gene, but share extensive sequence homology with the bird ZW sex chromosome system (Veyrunes et al., 2008). This dates the emergence of the therian XY pair, and the *SRY* gene, to after the divergence of monotremes from therian mammals 166MYA (million years ago) (Ezaz et al., 2006; Veyrunes et al., 2008), but before the marsupial/eutherian split.

Mapping between the two infraclass of the therian mammals, Eutheria (placental mammals) and Metatheria (marsupials), has revealed extensive homology of their

sex chromosomes (reviewed in Graves, 1995). Genes on the long arm (Xq) and proximal short arm (Xp) of human X chromosome map to the X in marsupials, implying that this X conserved region (XCR) has been part of the X chromosome since before the marsupial/eutherian divergence. However, a large region of the eutherian X chromosome (represented by much of human Xp) is orthologous to marsupial autosomes (Graves et al., 1998; Sinclair et al., 1988; Spencer et al., 1991). This region was added to the eutherian X (therefore named the X added region; XAR) after the divergence of eutherians from marsupials (Graves, 1995), but prior to eutherian radiation (Rodriguez Delgado et al., 2009).

Although therian X and Y chromosomes arose from an ordinary autosomal pair, they have undergone a lengthy evolutionary process that has resulted in strikingly different morphology and gene content. Initially, one member of the autosomal pair (the proto-Y) acquired a testis-determining locus (*SRY*). Genes that were advantageous to males then accumulated nearby this new sex-determining locus, either by transposition from autosomal sites or by mutations in existing genes. Subsequent selection to keep this male-advantageous gene cassette together resulted in the suppression of recombination with the X. In the absence of recombination, the male specific region of the Y (MSY) was vulnerable to drift and inefficient selection, leading to accumulation of slightly deleterious mutations (reviewed in Graves, 2006). Ultimately, the Y chromosome degraded over time to become small and gene poor (Charlesworth and Charlesworth, 2000).

The human X chromosome is large (165 Mb) (Ross et al., 2005) and bears 1,529 genes (NCBI database, http://www.ncbi.nlm.nih.gov). Gene content and order are highly conserved between even distantly related eutherian lineages (Murphy et al., 2007; Quilter et al., 2002; Raudsepp et al., 2004); with the exception of the rodent X, which has undergone several rearrangements (Amar et al., 1988; Waterston et al., 2002). In contrast, the human Y chromosome is 65 Mb in size, about half of which is heterochromatic, and codes for only 27 distinct proteins (Skaletsky et al., 2003). The Y varies in size and gene content between mammalian species, consistent with independent degradation and gene loss in different lineages (Graves, 2006; Waters et al., 2007).

This progressive loss of gene activity from the newly minted mammalian proto-Y resulted in dosage imbalance of X-linked genes between XX and XY individuals, for which a dosage compensation mechanism evolved to equalize.

13.3 Evolution of Dosage Compensation

The evolution of heteromorphic sex chromosomes in a wide range of taxa presents a similar problem, regardless of the sex chromosome system in question. All animals with differentiated sex chromosomes will exhibit different levels of X or Z gene dosage between the sexes. It is expected that such differences must be compensated for by a series of genetic and epigenetic regulatory mechanism that equalizes gene dosage. Dosage compensation is a common feature

of heteromorphic sex chromosomes in many diverse taxa, such as the fruitfly *Drosophila melanogaster,* the roundworm *Caenorhabditis elegans,* and most mammals. In mammals, X-inactivation effectively balances gene expression from the X chromosome between the sexes by means of a complex epigenetic pathway that leads to the transcriptional silencing of one X chromosome in the somatic cells of females.

13.4 X Chromosome Inactivation

Fifty years ago the observation of mosaic coat colour in female mice, heterozygous for the X-linked coat colour gene, lead to the hypothesis that one of two X chromosomes in females is inactivated early in embryogenesis (Lyon, 1961). It was proposed that inactivation was stable, and that both the active X (Xa) and inactive X (Xi) maintained their states through successive cell generations. Thus, this somatically heritable epigenetic change (Lyon, 1961) produces a mosaic of cells with either one or the other X chromosome transcriptionally silenced.

Several manifestations of XCI were observed at the cytogenetic level. These include the observation of a heterochromatic sex chromatin body (a Barr body; Barr and Carr, 1962) in female, but not male, interphase cells. In females with more then two X chromosomes extra Barr bodies were observed (Chu et al., 1960). It was also discovered that the two X chromosomes replicate asynchronously in female cells during mitosis (Taylor, 1960). Delayed replication timing is perhaps the most universal characteristic of the Xi chromosome, and is still one of the most reliable markers of inactivation. The Xi as a whole, and individual genes on the inactive chromosome, replicate late in S phase (Hansen et al., 1996; Priest et al., 1967; Schmidt and Migeon, 1990; Taylor, 1960; Torchia et al., 1994).

Not surprisingly, most information about X-inactivation comes from studies on humans and mice. Although these mammals belong to the same eutherian clade, their mechanism of X-inactivation is different, so it is hard to generalize to all mammals. Sporadic information available from other eutherians, such as cattle, horse and elephant, suggests that similar modes of X-inactivation operate in all eutherians. However, comparisons between eutherians and their most distantly related cousins that display X-inactivation, the marsupials, reveal marked differences that help to answer many fundamental questions about sex chromosome organization, function and evolution, and allows us to deduce how the X-inactivation system first evolved. Significantly, parallels of therian X-inactivation can be drawn with dosage compensation systems of other vertebrates.

13.4.1 Eutherian X-Inactivation

A master regulator locus, called the X-inactivation centre (XIC), controls the eutherian X-inactivation process (Avner and Heard, 2001). The XIC contains several transcribed non-coding RNA genes, whose sequences are poorly conserved between

species. As such, multiple sequence comparisons of the XIC in human, mouse, vole and cow have revealed little sequence conservation of non-coding elements (Chureau et al., 2002; Hendrich et al., 1993; Nesterova et al., 2001). Only 3.4% of the intergenic region within the XIC shows sequence conservation between mouse and human, yet the intron-exon structure and the exonic repeats of a non-coding gene, *XIST* (X-Inactive Specific Transcript) are maintained between human, mouse and cow (Chureau et al., 2002; Hendrich et al., 1993; Nesterova et al., 2001). It has been demonstrated that even the most distantly related eutherian lineage (Afrotheria; represented by the elephant) has an *XIST* orthologue, although the integrity of the elephant XIC is not known (Duret et al., 2006).

13.4.1.1 X-Inactivation in Eutherian Development

The original hypothesis put forward by Lyon suggested that X-inactivation was random, giving rise to a mosaic of cells expressing either the paternally or maternally derived X chromosome (Lyon, 1961). However, during the earliest stages of female embryogenesis, both X chromosomes remain active. X-inactivation is first observed in the trophectodermal tissue of the early blastocyst, which gives rise to early differentiating extra-embryonic tissues. In these cells the choice of which parentally derived X chromosome becomes inactive is parent specific (imprinted), whereby the paternally derived X chromosome is always inactivated in some eutherian species. The earliest sign of an Xi chromosome in the cells that give rise to the embryo proper is during the pre-implantation blastocyst, where either the maternally or paternally derived X is chosen randomly for inactivation.

13.4.1.2 Imprinted X-Inactivation

Paternal X-inactivation has been demonstrated to occur in the extraembryonic tissue of both mouse and cow (Harper et al., 1982; Dindot et al., 2004; Takagi and Sasaki, 1975; Wake et al., 1976; West et al., 1977). However, most studies report no preferential expression from the paternal X in human extraembryonic material (Migeon et al., 1985; Mohandas et al., 1989), but there have been some suggestions of an imprint in an early extraembryonic lineage (Goto et al., 1997; Harrison, 1989).

Imprinted X-inactivation of mice and cattle requires expression of *Xist* from the paternal X early in the preimplantation embryo, possibly as early as the 2-cell stage (Dindot et al., 2004; Kay et al., 1993). Interestingly, it has recently been shown that initiation of imprinted X-inactivation in mice is not *Xist* dependent but that the inactivation signal will not stabilize in the absence of *Xist* (Kalantry et al., 2009). By the 8-cell stage an *Xist* coated repressive domain has formed, from which gene expression is excluded (Okamoto et al., 2005). The nature of the paternal specific imprint is currently unknown. There are two alternative possibilities; either the maternal X is marked to remain active (Lyon and Rastan, 1984), or the paternal X is predisposed to inactivation (McLaren and Monk, 1981). The latter is supported by demonstrations that the paternal X is inactivated in cleavage stage embryos (Okamoto et al., 2004). This led to the suggestion that the paternal X is inherited by the zygote

in a pre-inactivated state, which is acquired at male meiosis during meiotic sex chromosome inactivation (MSCI) (reviewed in Turner, 2007). However, the observation of post meiotic transcription of several X-linked genes indicates reversal of MSCI before fertilization (Mahahevaiah et al., 2009; Okamoto et al., 2004).

This imprinted X-inactivation appears to be erased or become ineffective, at the time of implantation, in the inner cell mass that gives rise to embryonic tissue (Okamoto et al., 2004). This is followed by the random choice to inactivate either the paternal or maternal X chromosome (Reik and Lewis, 2005).

13.4.1.3 Random X-Inactivation

In contrast to extra embryonic tissue, X-inactivation is random in cells of the embryo proper and always occurs in a mutually exclusive manner. Random X-inactivation is also controlled by the XIC. *Xist* deletion studies demonstrated that it is essential for initiation of silencing, and that it plays a role in the choice of which X is inactivated (Marahrens et al., 1998, 1997; Penny et al., 1996; Wutz and Jaenisch, 2000). Furthermore, a point mutation in the human *XIST* promoter (in two unrelated families) is associated with the primary skewing of X-inactivation choice (Plenge et al., 1997), suggesting that in humans *XIST* itself participates in choice. Consequently, genomic regions that regulate *Xist* expression also influence the fate of chromosome X-inactivation status. For instance, deletion of the *Tsix* (the antisense of *Xist*) transcription start site was shown to abolish choice, in that the mutated chromosome was always inactivated (Lee and Lu, 1999). Moreover, deletion of an element at the 5′ end of the *Tsix* gene (Vigneau et al., 2006) named *DXPas34*, resulted in non-random X-inactivation due to disruption of choice (Migeon, 2003).

In mouse, a locus called X controlling element (*XCE*), which affects choice, was also mapped to the XIC (Johnston and Cattanach, 1981), but it remains possible that more control sequences lie in this region. One such element is a cluster of binding sites at the *DXPas34* locus and the *Tsix* promoter region for CTCF, a transcription factor that possibly regulates *Tsix* and /or *Xist* (Chao et al., 2002; Clerc and Avner, 2003). It is still debated whether the human orthologue of *Tsix* shares the same functions (Migeon, 2003).

13.4.1.4 Transcriptional Silencing of the Inactive X Chromosome

A fundamental assumption of X-inactivation was that genes on the Xi were silenced by transcriptional repression. This was eventually demonstrated by finding that the inactive wild-type *HPRT* allele was not expressed in cells cloned from a woman heterozygous for *HPRT* deficiency (Graves and Gartler, 1986). This has since been repeatedly confirmed by the finding that genes on an inactive human X in a rodent-human cell hybrid are not transcribed (Carrel et al., 1999), and is consistent with the finding that DNA polymerase II is excluded from the Xi domain (Chaumeil et al., 2006).

The earliest step of initiation of X-inactivation is thought to be a counting event, by which the ratio of X chromosomes to autosomes is measured, ensuring that

the single X chromosome in XY males is not inactivated, and that in female cells all X chromosomes in excess of one are silenced. At the onset of X-inactivation, *Xist* RNA expression is increased and coats the future Xi. However, *Xist* induced silencing is reversible (Chadwick and Willard, 2003); and in order to establish an inactive state, *Xist* RNA is required to spread along the entire length of the X chromosome. Spreading only initiates from the XIC and is strictly limited *in cis* (Avner and Heard, 2001). The attenuation of inactivation spreading, when *Xist* was moved to and expressed from some (but not all) autosomes, indicates that although not essential for spreading, X chromatin is somehow more favourable to the X-inactivation machinery (Lyon, 1998b). It has been proposed that long interspersed repeat elements (LINEs) could act as booster elements for *Xist* RNA to bind to, thereby spreading along the chromosome (Hansen, 2003; Lyon, 1998a). Evidence for the role of LINEs in spreading of the inactivation signal include the observation that these elements are concentrated on the human and mouse X (Korenberg and Rykowski, 1988; Boyle et al., 1990; Bailey et al., 2000), and are correlate with susceptibility to and escape from inactivation (Carrel et al., 2006; Wang et al., 2006).

Once spreading of *Xist* RNA along the Xi has been achieved, several epigenetic modifications occur to maintain a heterochromatic and transcriptionally silent state throughout the cell cycle. The first changes in histone variants and modifications to Xi include the loss of histones associated with transcriptional activity, such as H3K4me and H3K9ac (reviewed in Heard, 2005). Next, the Xi gains specific repressive marks, H3K9me2, H3K27me3, H4K20me1, H2AK119ub (reviewed in Heard, 2005). Finally, concentration of variant histones such as macro-H2A1, and hypoacetylation on histone H4 appears (reviewed in Heard, 2005). Experiments using 5-azacytidine to reactivate repressed genes provided the first evidence for the role that methylation plays in X-inactivation (Graves, 1982; Mohandas et al., 1981). It was later demonstrated that it was preferential methylation of CpG islands upstream of inactive genes on the inactive X, when compared to active X (Lock et al., 1986). DNA methylation seems to occur at a late stage of the XCI process, at least in teratocarcinoma and embryonic stem cells, so is thought to be involved in stabilizing the inactive eutherian X (Lock et al., 1987).

13.4.1.5 Escape from Inactivation

It was originally proposed that all genes were coordinately silenced on the Xi. Inactivation of X-specific genes appears to be nearly ubiquitous for mouse, with just a few exceptions. However, research on different mammalian species has shown that inactivation is not complete. Approximately 15% of genes on the human X escape inactivation in fibroblasts (Carrel and Willard, 2005) and somewhat fewer in lymphoblastoid cell lines (Johnston et al., 2008). These “escapee” genes lack the chromatin modifications characteristic of the Xi, and can be either clustered together, or interspersed with genes known to be subject to inactivation. The mechanism that mediates escape from X-inactivation remains unclear, but non-inactivated

domains appear to be protected from the inactivation signal by CTCF-containing insulator elements (Disteche, 1995).

The distribution of XCI for escapee genes, with genes located largely on the short arm of the human X, is a reflection of the evolutionary history of the X chromosome. This region was recently added to the eutherian X, so genes within it did not require dosage compensate prior to marsupial/eutherian divergence. The addition of this region to the X and Y, followed by degradation of Y homologues, resulted in dosage differences of X-linked genes between the sexes and selection for recruitment into the X-inactivation system (Graves et al., 1998; Graves and Schmidt, 1992). Thus, an understanding of X chromosome evolution presents an important opportunity in understanding the origin and mechanism of X-inactivation.

Little is known about the nature and molecular mechanisms of X-inactivation in eutherian mammals outside of humans and mice. This is unfortunate because mice and human X-inactivation differs in several important respects, including the extent to which the X is silenced, and the mechanism by which *Xist* functions. Furthermore, rodents and primates belong to the same eutherian superordinal clade, and as a result our understanding of X-inactivation lacks evolutionary depth. Therefore, it is not clear whether the mouse model of X-inactivation is applicable to other eutherian species, so the study of marsupial dosage compensation is important in understanding the evolution of X-inactivation mechanisms.

13.4.2 Marsupial X-Inactivation

Like eutherian mammals, dosage compensation in marsupials is achieved via transcriptional silencing of one X in female somatic cells. Early work on marsupial lymphocytes showed that the two X chromosomes replicate asynchronously in female somatic cells, with the Xa replicating earlier in S phase than the Xi (Graves, 1967). Barr bodies are more difficult to detect in marsupials than in eutherian mammals, and have only been observed in some species and tissues. Barr bodies were absent in adult brush-tail possum (*Trichosurus vulpecula*) (McKay et al., 1987), although they were observed in bilaminar and trilaminar blastocysts of a marsupial mouse (*Antechinus stuartii*) (Johnston and Robinson, 1987). Barr bodies have also been observed in *Monodelphis* embryo cells, and at a low frequency in nuclei of *T. vulpecula* corneal epitheium cells (Robinson et al., 1994; McKay et al., 1987).

13.4.2.1 X-Inactivation in Marsupial Development

Studies of the expression of isozyme variants in female kangaroo heterozygous for X-linked markers confirmed that genes on one X chromosome were inactivated (reviewed in Cooper et al., 1993). For instance, it was demonstrated that the classic X-borne genes *PGK* and *G6PD* were inactivated on the marsupial X chromosome (Richardson et al., 1971). However, these studies also revealed striking differences between eutherian and marsupial X-inactivation.

Possibly, the biggest difference between marsupial and eutherian X-inactivation is the control of Xi choice. Using tritiated thymidine, the paternal X was identified as late replicating in lymphocytes of female hybrid kangaroos with heteromorphic X chromosomes (Sharman, 1971), suggesting that the paternally derived X chromosome is preferentially inactivated and marsupial X-inactivation is non-random.

The observation that some X-linked genes on the paternal X retain activity and are expressed from both alleles in some tissues of kangaroos, lead to the conclusion that marsupial X-inactivation is incomplete and tissue specific. These results were inconsistent in different species; for instance in the Virginia opossum (*Didelphis virginiana) G6PD* was shown to be fully inactivated in lymphocytes and cultured fibroblasts, but active from both Xs in cultured firbroblasts from two kangaroo species (reviewed in Cooper et al., 1993). A more recent study by Hornecker et al. demonstrated that the maternal allele of the *G6PD* gene was expressed predominantly in seven adult female tissues examined, whereas expression of the paternal allele was not detected. Yet, equal or lower expression levels of the paternal *PGK-1* allele were consistently observed in most tissues tested (Hornecker et al., 2007). Although this is consistent with previous demonstrations indicating that the degree of repression of the paternal X varies among marsupials in a locus-, tissue-, and species-specific manner, it is difficult to generalize the findings of isozyme studies to the whole X chromosome because they could only be performed for five genes (Table 13.1).

Results of two recent studies have confirmed the incomplete and locus specific nature of X-inactivation in marsupials. The frequency of mono-allelic transcription of X-linked genes on the marsupial X, in female fibroblasts of tammar wallaby was tested at the cellular level by RNA-FISH, which targets primary transcripts at the point of transcription. Surprisingly, the result of both studies indicates that the inactivation status of X-linked genes in marsupial fibroblasts was different between nuclei from the same individual. None of the X-linked genes tested displayed 100% mono-allelic expression, instead different genes show different proportions of cells with either one or both alleles active (Koina et al., 2005; Al Nadaf et al., in preparation). It then seems possible that locus-specific regulatory factors function within a chromosomal wide inactivation mechanism by which the paternal X is invariably inactivated.

Although the incomplete X-inactivation system of marsupials has not been well explored, these RNA-FISH results imply that many genes may not absolutely require the exact balancing of X gene dosage between the sexes. If one was to assume that partial compensation is sufficient to balance fitness between sexes, then maybe the dosage compensation in marsupials evolved to maintain dosage of genes critical for sexual differences.

13.4.2.2 Molecular Mechanism of Marsupial X-Inactivation

Much less is known of the molecular mechanism of X-inactivation in marsupials than in eutherians. Only very recently has it been possible to test the fundamental hypothesis that silencing of the marsupial X is at the transcriptional level. Both the

Table 13.1 Inactivation status of genes on the marsupial X (adapted from Cooper et al., 1993; Deakin et al., 2009)

Gene	Species	Method	Inactivation status: Somatic tissues	Inactivation status: Cultured fibroblasts	References
G6PD	*Macropus robustus* (Wallaroo)	Isozyme[1]/ SNuPE[2]	Complete	Incomplete	Johnston and Robinson (1986), Johnston et al., (1985), Richardson et al., (1971)
	Macropus rufogriseus (Red-necked wallaby)	Isozyme	Complete	Incomplete	Johnston and Robinson (1987)
	Didelphis virginiana (Virginia opossum)	Isozyme	Incomplete	Incomplete	Samollow et al., (1987, 1989, 1995)
	Monodelphis domestica (Short grey-tailed opossum)	Allele-specific PCR	Mostly complete	–	Hornecker et al., (2007)
GLA	*Antechinus stuarttii* (Brown antechninus)	Isozyme	Complete	Complete	Johnston and Robinson (1987)
	Antechinus rosamondae (Little red marsupial mouse)	Isozyme	Complete	Complete	Cooper et al., (1983)
	Kangaroo hybrids	Isozyme	Complete	Complete	Cooper et al., (1983)
HPRT	*Didelphis virginiana*	Isozyme	–	Incomplete	Migeon et al., (1989)
PGK1	*Macropus giganteus* (Eastern grey kangaroo)	Isozyme	Tissue specific	Incomplete	Cooper et al. (1971), VandeBerg et al., (1977)
	Macropus parryi (Whiptail wallaby)	Isozyme	Tissue specific	Incomplete	VandeBerg et al., (1977)
	Trichosurus vulpecula (Brushtail possum)	Isozyme	Tissue specific	–	VandeBerg (1979)
	Didelphis virginiana	Isozyme	Complete	Complete	Samollow et al., (1987, Samollow et al.,1989)
	Monodelphis domestica	SNuPE	Incomplete	–	Hornecker et al., (2007)
SLC16A2	*Macropus eugenii* (Tammar wallaby)	RNA-FISH	–	76% cells inactivated	Koina et al. (2005)

[1]Isoenzymes
[2]Single Nucleotide Primer Extension

absence of RNA polymerase from one X and the presence of RNA-FISH signal on only one X confirm that silencing is at the transcriptional level (Koina et al., 2005; Al Nadaf et al., in preparation; Chaumeil et al., in preparation).

A prolonged hunt for a marsupial *Xist* was unsuccessful, and it was not possible to demonstrate that there was no marsupial *Xist* until sequencing of the opossum genome was complete (Mikkelsen et al., 2007). Although genes flanking the XIC are well conserved across vertebrates, no *Xist* orthologue could be identified in any marsupial in the region that juxtapose flanking markers (Davidow et al., 2007; Duret et al., 2006; Hore et al., 2007). Comparative mapping data revealed that during evolution of the marsupial X chromosome, one or more rearrangements broke up the conserved block of vertebrate genes that makes up the current eutherian XIC. Exhaustive scanning of sequence on either side of the breakpoint revealed no sign of *XIST*; thus, it was concluded that marsupials lack *XIST* (Hore et al., 2007; Davidow et al., 2007).

In chicken and frogs there are five protein-coding genes that occupy the region between the XIC flanking markers. None of these are present in the human or mouse XIC, having become inactivated pseudogenes that were lost in eutherians. However, one of these genes (*LNX2*) remains intact and active in the opossum. Duret et al. claimed that this gene has homology to *XIST*, and suggested that the eutherian *XIST* evolved from pseudogenization of this protein-coding gene (Duret et al., 2006). However, Hore et al. (2007) considered that the homology was very weak and proposed that this gene may have donated only its transcription start site to *XIST*.

Consequently, imprinted X-inactivation in marsupials is achieved by an *XIST* independent method. Although not identified, it is possible that another noncoding RNA with a comparable function to *XIST* acts in marsupial X-inactivation. Since marsupial X-inactivation is *XIST* independent and initiation of inactivation does not seem to begin and spread from an inactivation centre, it is unlikely that the marsupial process of inactivation completely mirrors that of eutherians.

Several histone modifications have been observed to accumulate on the inactive marsupial X, but not all of these histone marks are the same as for the inactive X of eutherian mammals. Antibodies to acetylated histone H4 bound to only one of the two X chromosomes in cells from female kangaroos, revealing a difference in histone modification between Xa and Xi (Wakefield et al., 1997). More recently, immunofluorescence on metaphase chromosomes of tammar wallaby (*Macropus eugenii*) has identified a more detailed pattern of histone modifications. It appears that all active marks (H3K4me2, H3K9ac) are excluded from the Xi in marsupials, as in eutherian mammals, representing a common feature of X-inactivation throughout mammals.

No DNA methylation differences have been observed between alleles on the active and inactive marsupial X. Early studies using methylation-sensitive enzymes on DNA from the Virginia opossum found no methylation differences that correlated with activity of the *HPRT* and *G6PD* loci (Kaslow and Migeon, 1987). The CpG islands associated with these genes were hypomethylated on both Xa and Xi

chromosome. More recently, bisulfite sequencing detected no methylation difference at the kangaroo *G6PD* locus (Loebel and Johnston, 1996) or the opossum *G6PD* and *PGK1* loci (Hornecker et al., 2007). Intriguingly, a chromosome-wide examination of methylation by in situ nick translation detected considerable methylation of the maternal X chromosome (Loebel and Johnston., 1993). Hence, it is too early to rule out a role for methylation in marsupial X-inactivation and further studies are required to probe deeper into the methylation status of marsupial X-borne genes.

13.5 Evolution of Dosage Compensation and X-Inactivation

Due to the molecular and cytological similarities, many have speculated that marsupial and eutherian X-inactivation share a common evolutionary history. Studying dosage compensation in distantly related mammalian groups helps us understand how and when therian X-inactivation evolved. Mechanisms that are shared by different species are likely to have been present in a common ancestor, whereas features that are lineage specific were probably acquired after species divergence. The more distantly related the species that are compared, the farther back in evolutionary time we can look. Thus, we can date the acquisition of different elements of this complex silencing mechanism.

13.5.1 Evolution of Dosage Compensation in Monotremes

Monotremes, with their mix of reptilian and mammalian features, might provide new insight into the origin of X chromosome inactivation. In particular, the homology of the sex chromosomes with the bird ZW, rather than the therian XY system, could offers insights about dosage compensation in the last common ancestor of the three extant mammalian groups. The independent evolution of monotreme and therian sex chromosomes, therefore, makes them a most interesting model for exploring the origins of mammalian dosage compensation.

The need for a dosage compensation mechanism would seem to be acute in monotremes because of the very large portion of the genome (approx. 12%) that is X specific, and therefore represented unequally in males and females. However, early studies into the replication timing of platypus X_1 revealed no asynchronous replication of the unpaired region of this chromosome (Wrigley and Graves, 1988). This suggested that if the platypus does dosage compensate, it is unlikely to do so by X-inactivation.

A recent study by Deakin et al. (2008a), found that genes on the multiple platypus Xs show partial and variable dosage compensation. In heterozygotes for transcribed sequence polymorphisms, transcripts of both alleles were detected, eliminating paternal X-inactivation as a possibility. RNA-FISH showed a mixture of nuclei, some with one signal, and some with two. Different genes had different and reproducible frequencies of one X-active and two X-active cells. Moreover, quantitative

PCR showed that the male:female ratio differed for different genes, ranging between 1:1 (full compensation) and 1:2 (no dosage compensation). This surprising result suggested that the partial compensation results, not from partial expression in each cell, but from a stochastic regulation of expression, such that different genes have different probabilities of expression from one or both alleles.

This partial and variable dosage compensation is similar to that observed for genes on the bird Z chromosome, which also display incomplete and locus-specific dosage compensation (Itoh et al., 2007). Measurement of male:female expression ratios for all Z-borne genes on microarrays produced a distribution of ratios varying between complete compensation (1:1) and no compensation (1:2). Dosage compensation in birds does not appear to be achieved by Z inactivation in ZZ males, as biallelic expression has been shown for six Z-borne genes by either RNA-FISH or SNP analysis (Kuroda et al., 2001; Kuroiwa et al., 2002). However, four of these genes showing biallelic expression do not appear to be subject to dosage compensation based on quantitative PCR (Kuroiwa et al., 2002) or microarray data (Itoh et al., 2007), with further work required to determine if Z inactivation exists in birds.

Thus, the platypus and bird sex chromosomes not only share homology, they seem to share at least some features of dosage compensation, which is distinct to the X-inactivation system of therian mammals. This effectively dates the origin of the therian X-inactivation system to after 165MYA, when the therian mammal XY chromosome pair began its differentiation.

13.5.2 Evolution of X-Inactivation in Therian Mammals

While the precise routes by which X-inactivation has evolved are not known, plausible models have been put forward. One evolutionary scenario considers the possibility of independent genesis of X-inactivation systems in eutherian and marsupials after the two groups diverged. In this scenario the sex chromosomes of the therian ancestor are not sufficiently degraded to require a chromosome wide dosage compensation system (Duret et al., 2006). After separation, the marsupial and eutherian sex chromosomes degraded independently, thus requiring independent genesis of dosage compensation systems. This hypothesis is supported by the differences observed between marsupials and eutherians X-inactivation, including the absence of a marsupial *XIST* and some differences in molecular machinery, including lack of DNA methylation and accumulation of inactive histone marks in marsupials (Table 13.2).

Additional support for this hypothesis came recently from a bioinformatic analysis examining the effects of MSCI upon gene movements. This study reported significantly elevated rates of gene retroposition off the X chromosome to autosomes, in both marsupials and eutherians, relative to the expected retroposition rate (Potrzebowski et al., 2008). However, not one of the 46 X to autosome retroposition events identified by Potrzebowski et al. (2008) occurred in the ancestor of all therian mammals, instead occurring independently in marsupials and eutherians.

Table 13.2 Comparison of the features of X chromosome inactivation in marsupial and eutherian mammals

XCI features	Eutherians	Marsupials
Asynchronous replication	Yes	Yes
Barr bodies	Consistent	Inconsistent
Stability	Stable 15% escapees	Leaky
Control of Xi choice	Random (soma) Imprinted (placenta)	Imprinted – Stochastic
XIST locus	Yes	No
Inactivation centre	Yes	NE
LINE1 accumulation	Yes	No
Active histone marks	Yes	Yes
Repressive histone marks	Yes	CD
CpG island methylation	Yes	No (PD)

CD, Conflicting Data; NE, No Evidence for it; PD, Provisional Data

This implies that the majority of X and Y divergence (and therefore selection for X-inactivation) occurred after the divergence of marsupials and eutherians.

Nonetheless the possibility exists that the X-inactivation systems of eutherian and marsupial mammals share a common ancestry. Marsupials, which last shared a common ancestor with eutherian mammals 148MYA, have been proposed to possess a similar but simpler X-inactivation system than eutherians. Its unique features (paternal imprinting, lower stability, the lack of consistent Barr body formation, and incomplete and tissue-specific inactivation) may therefore have been present in a common therian ancestor. The imprinted X-inactivation system of marsupials has therefore been proposed to be more similar to the ancestral therian XCI system, upon which layers of molecular complexity were added during eutherian evolution (Table 13.2).

Indeed, some of the features that were thought to be specific to marsupial X-inactivation now appear to be common in humans. For instance, the incomplete inactivation of the marsupial X is mirrored in the discovery of many genes on the human X that escape X-inactivation, and the observation of variable inactivation of some genes between different females (Carrel and Willard, 2005). There is also some evidence that marsupials are not alone in displaying tissue-specific inactivation. Although this has been little studied in humans or mice, inactivation of at least some genes is tissue specific; for instance, variable patterns of dystrophin gene (*mdx*) expression have been observed in skeletal versus cardiac muscle of adult mice heterozygous for *mdx* (Bittner et al., 1997).

Perhaps the most dramatic support for this hypothesis was the discovery that X-inactivation is paternal, not only in marsupials, but also in the extra-embryonic tissues of the rodent embryo. This immediately suggested that paternal inactivation

was ancestral in both groups of mammals, and randomness evolved later in eutherians and was selected for because it conferred the benefits of heterozygousity in females (Cooper et al., 1993). Furthermore, there is at least one study that shows parent of origin influencing X-inactivation not only in extra embryonic tissue, but also in the mouse embryo proper, where differences in the ratio of inactivation between the paternal and maternal Xs have been observed in cardiac and skeletal muscle cells (Bittner et al., 1997). The inability to demonstrate paternal inactivation in human extraembryonic tissues was thought to contradict this hypothesis (Zeng and Yankowitz, 2003) and point to independent evolution of paternal inactivation in marsupials and rodents (perhaps to mitigate the maternal-fetal incompatibility). However, the recent discovery of paternal inactivation in the bovine embryo established that it was a more widespread eutherian state, and makes it more likely that paternal inactivation was an ancestral therian character. In addition, X-inactivation in extraembryonic tissues of mouse, as in marsupials, is less stable, incomplete and does not involve DNA methylation.

Thus the hypothesis that the X-inactivation systems of marsupials and eutherians have been achieved by recruitment, in a common ancestor, of a pre-existing chromatin regulatory system to modulate expression of an entire chromosome is a very attractive model. This model suggests imprinted X-inactivation in marsupials represents a carryover of MSCI through fertilization to the embryo that results in inheritance of an already silent X chromosome from the father. This view has recently gained support by the demonstration of transcriptional inactivation of the paternal X during gametogenesis (Hornecker et al., 2007). Results of two recent studies confirmed that MSCI occurs in *Monodelphis domestica,* where the sex chromosomes accumulate marks of meiotic silencing and exclude markers of active transcription (Namekawa et al., 2007; Hornecker et al., 2007). In addition, Hornecker et al. (2007) observed no indication of postmeiotic reactivation of the X chromosome in *M. domestica*. In the absence of any evidence for reactivation in sperm, or in the early embryo, it is possible that the paternal X is inactivated at meiosis and simply remains inactive in the embryo. The partial inactivation of some loci in some tissues would therefore represent a tissue- and locus-specific reactivation of the paternal X. Due to the similarities observed between MSCI in marsupials and eutherian mammals, MSCI could be the ancestral force behind imprinted X-inactivation.

This makes it critical to resolve whether or not the paternal X enters the egg in an inactive state in marsupials and mice, and emphasizes the need for observations of X-inactivation in early marsupial embryos. Without knowing the state of the X in the marsupial embryo, it is not possible to distinguish between the hypotheses that the initial inactivating event was incomplete, or alternatively that the entire X was inactive and particular genes were reactivated in different tissues.

Evolution and refinement of X-inactivation mechanisms might also explain some of the differences in the structure and organization of the X chromosome in eutherians and marsupials. The conserved order of genes on the X chromosome in most eutherian species (except mice) contrasts with the several major rearrangements between the opossum and the wallaby X (Deakin et al., 2008b), and suggests that

there is selection against structural rearrangement of the eutherian (but not the marsupial) X, perhaps because of the need for *Xist* transcripts to spread along the X from the XIC. Lower LINE1 density on the opossum X, in comparison to the eutherian X (Mikkelsen et al., 2007), suggests that LINEs have accumulated only to the extent expected because of the lower recombination rate of the X, and are not involved in the propagation of inactivation along the marsupial X. Yet, recent research suggests that LINEs might have a less critical role in the eutherian X-inactivation process. X-inactivation in a group of South American rodents shows X-inactivation still occurs in spite of extinction of LINE activity (Cantrell et al., 2009).

Perhaps most significant of all the differences between marsupial and eutherian X-inactivation is the absence of *XIST* in marsupials. The significance of *XIST* in X-inactivation must be re-examined, since its absence in marsupials contradicts its seemingly fundamental role in eutherian X chromosome inactivation, both random and imprinted (Marahrens et al., 1997). The evolution of *XIST* might explain why features of eutherian X-inactivation initiation, spreading and maintenance are not shared with marsupials, and could explain why marsupials have a less stable and "leaky" inactivation mechanism compared to the hyper-stable system in human and mouse. However, mice may well be an exception, and X-inactivation may prove to be leaky for many therian mammals, as is apparent for many genes that escape inactivation on the human X. This supports the idea that partial inactivation represents a more basic form of mammalian X-inactivation, which became subject to tighter control during eutherian and particularly rodent evolution. Koina et al. (2009) suggested that *XIST* binding to the Xi is responsible for DNA methylation differences, the accumulation of active histone marks, and the molecular differences that are lacking from the inactive marsupial X. Perhaps the features of X-inactivation observed in eutherians but not marsupials – randomness and hyper-stability – are conferred by *XIST*.

13.6 Conclusion

The gradual loss of genes from the proto-Y chromosome during sex chromosome differentiation is believed to have driven inactivation of homologous X-linked genes. The sex chromosome dosage compensation in the therian ancestor might then have been more like the variable and partial dosage compensation observed in birds and monotremes. The elements of silencing, known to be shared between marsupial and eutherian X-inactivation were established in this therian ancestor, possibly by utilising a pre-existing chromatin regulatory system, such as MSCI. This resulted in a co-ordinated transcription from only one of the sex chromosomes, however, inactivation of the homologous X-linked genes was achieved in a piecemeal fashion. In marsupials no additional epigenetic events evolved to change the inactive state during embryogenesis, whereas in eutherians the benefits of heterozygousity and stability perhaps resulted in evolution of random X-inactivation and further layers of epigenetic silencing.

References

Amar LC, Dandolo L, Hanauer A, et al. (1988) Conservation and reorganization of loci on the mammalian X chromosome: a molecular framework for the identification of homologous subchromosomal regions in man and mouse. Genomics 2:220–230.

Avner P, Heard E (2001) X-chromosome inactivation: counting, choice and initiation. Nat Rev Genet 2:59–67.

Bailey JA, Carrel L, Chakravarti A, Eichler EE (2000) Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proc Natl Acad Sci U S A 97:6634–6639.

Barr ML, Carr DH (1962) Correlations between sex chromatin and sex chromosomes. Acta Cytol 6:34–45.

Bittner RE, Popoff I, Shorny S, Hoger H, Wachtler F (1997) Dystrophin expression in heterozygous mdx/+ mice indicates imprinting of X chromosome inactivation by parent-of-origin-, tissue-, strain- and position-dependent factors. Anat Embryol 195:175–182.

Boyle AL, Ballard SG Ward DC (1990) Differential distribution of long and short interspersed element sequences in the mouse genome: chromosome karyotyping by fluorescence in situ hybridization. Proc Natl Acad Sci U S A 87:7757–7761.

Cantrell MA, Cartens BC, Wihcman HA (2009) X chromosome inactivation and Xist evolution in a rodent lacking LINE-1 activity. PLoS One 4:e6252.

Carrel L, Cottle AA, Goglin KC, Willard HF (1999) A first-generation X-inactivation profile of the human X chromosome. Proc Natl Acad Sci U S A 96:14440–14444.

Carrel L, Park C, Tyekucheva S, Dunn J, Chiaromonte F, Mokova KD (2006) Genomic environment predicts expression patterns on the human inactive X chromosome. PLoS Genet 2:e151.

Carrel L, Willard HF (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434:400–404.

Chadwick BP, Willard HF (2003) Barring gene expression after XIST: maintaining facultative heterochromatin on the inactive X. Semin Cell Dev Biol 14:359–367.

Chao W, Huynh KD, Spencer RJ, Davidow LS, Lee JT (2002) CTCF, a candidate trans-acting factor for X-inactivation choice. Science 295:345–347.

Charlesworth B (1996) The evolution of chromosomal sex determination and dosage compensation. Curr Biol 6:149–162.

Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. Philos Trans R Soc Lond B Biol Sci 355:1563–1572.

Chaumeil J, Le Baccon P, Wutz A, Heard E (2006) A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes Dev 20:2223–2237.

Chu E, Grumbach EH, Morishima A (1960) Karyotypic analysis of a male pseudohermaphrodite with the syndrome of feminizing testes. J Clin Endocrinol Metab 20:1608–1613.

Chureau C, Prissette M, Bourdet A, et al. (2002) Comparative sequence analysis of the X-inactivation center region in mouse, human, and bovine. Genome Res 12:894–908.

Clerc P, Avner P (2003) Multiple elements within the Xic regulate random X inactivation in mice. Semin Cell Dev Biol 14:85–92.

Cooper DW, Johnston PG, Watson JM, Graves JAM (1993) X-inactivation in marsupials and monotremes. Semin Dev Biol 4:117–128.

Cooper DW, VandeBerg JL, Sharman GB, Poole WE (1971) Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. Nat New Biol 230:155–157.

Cooper DW, Woolley PA, Maynes GM, Herman FS, Poole WE (1983) Studies on metatherian sex chromosomes. XII. Sex-linked inheritance and probable paternal X-inactivation of alpha-galactosidase A in Australian marsupials. Aust J Biol Sci 36:511–517.

Davidow LS, Breen M, Duke SE, Samollow PB, McCarrey JR, Lee JT (2007) The search for a marsupial XIC reveals a break with vertebrate synteny. Chromosome Res 15:137–146.

Deakin JE, Hore TA, Koina E, Graves JAM (2008a) The status of dosage compensation in the multiple X chromosomes of the platypus. PLoS Genet 4:e1000140.

Deakin JE, Koina E, Waters PD, et al. (2008b) Physical map of two tammar wallaby chromosomes: a strategy for mapping in non-model mammals. Chromosome Res 16:1159–1175.

Deakin JE, Chaumeil J, Hore TA, Graves JAM (2009) Unravelling the evolutionary origins of X chromosome inactivation in mammals: insights from marsupials and monotremes. Chromosome Res 17:671–685

Dindot SV, Farin PW, Farin CE, et al. (2004) Epigenetic and genomic imprinting analysis in nuclear transfer derived Bos gaurus/Bos taurus hybrid fetuses. Biol Reprod 71:470–478.

Disteche CM (1995) Escape from X inactivation in human and mouse. Trends Genet 11:17–22.

Duret L, Chureau C, Samain S, Weissenbach J, Avner P (2006) The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. Science 312:1653–1655.

Ezaz T Stiglec R, Veyrunes F, Graves JAM (2006) Relationships between vertebrate ZW and XY sex chromosome systems. Curr Biol 16:R736–R743.

Fridolfsson AK, Cheng H, Copeland NG, et al. (1998) Evolution of the avian sex chromosomes from an ancestral pair of autosomes. Proc Natl Acad Sci U S A 95:8147–8152.

Goto T, Wright E, Monk M (1997) Paternal X-chromosome inactivation in human trophoblastic cells. Mol Hum Reprod 3:77–80.

Graves JAM (1967) DNA synthesis in chromosomes of cultured leucocytes from two marsupial species. Exp Cell Res 46:37–57.

Graves JAM (1982) 5-azacytidine-induced re-expression of alleles on the inactive X chromosome in a hybrid mouse cell line. Exp Cell Res 141:99–105.

Graves JAM (1995) The evolution of mammalian sex chromosomes and the origin of sex determining genes. Philos Trans R Soc Lond B Biol Sci 350:305–311.

Graves JAM (2001) From brain determination to testis determination: evolution of the mammalian sex-determining gene. Reprod Fertil Dev 13:665–672.

Graves JAM (2006) Sex chromosome specialization and degeneration in mammals. Cell 124:901–914.

Graves JAM, Gartler SM (1986) Mammalian X chromosome inactivation: testing the hypothesis of transcriptional control. Somat Cell Mol Genet 12:275–280.

Graves JAM, Schmidt MM (1992) Mammalian sex chromosomes: design or accident? Curr Opin Genet Dev 2:890–901.

Graves JAM, Wakefield MJ, Toder R (1998) The origin and evolution of the pseudoautosomal regions of human sex chromosomes. Hum Mol Genet 7:1991–1996.

Grützner F, Rens W, Tsend-Ayush E, El-Morgharbel N, et al. (2004) In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. Nature 432:913–917.

Hansen RS (2003) X inactivation-specific methylation of LINE-1 elements by DNMT3B: implications for the Lyon repeat hypothesis. Hum Mol Genet 12:2559–2567.

Hansen RS, Canfield TK, Fjeld AD, Garlter SM (1996) Role of late replication timing in the silencing of X-linked genes. Hum Mol Genet 5:1345–1353.

Harper MI Fosten M, Monk M (1982) Preferential paternal X inactivation in extraembryonic tissues of early mouse embryos. J Embryol Exp Morphol 67:127–135.

Harrison KB (1989) X-chromosome inactivation in the human cytotrophoblast. Cytogenet Cell Genet 52:37–41.

Heard E (2005) Delving into the diversity of facultative heterochromatin: the epigenetics of the inactive X chromosome. Curr Opin Genet Dev 15:482–489.

Hendrich BD, Brown CJ, Willard HF (1993) Evolutionary conservation of possible functional domains of the human and murine XIST genes. Hum Mol Genet 2:663–672.

Hore TA, Koina E, Wakefield MJ, Graves JA (2007) The region homologous to the X-chromosome inactivation centre has been disrupted in marsupial and monotreme mammals. Chromosome Res 15:147–161.

Hornecker JL, Samollow PB, Robinson ES, VandeBerg JL, McCarrey JR (2007) Meiotic sex chromosome inactivation in the marsupial Monodelphis domestica. Genesis 45:696–708.

Itoh Y, Melamed E, Yang X (2007) Dosage compensation is less effective in birds than in mammals. J Biol 6:2.

Johnston PG, Cattanach BM (1981) Controlling elements in the mouse. IV. Evidence of non-random X-inactivation. Genet Res 37:151–160.

Johnston CM, Lovell FL, Leongamornlert DA, Stranger BE, Dermitzakis ET, Ross MT (2008) Large-scale population study of human cell lines indicates that dosage compensation is virtually complete. PLoS Genet 4:e9.

Johnston PG, Robinson ES (1986) Lack of correlation between Gpd expression and X chromosome late replication in cultured fibroblasts of the kangaroo *Macropus robustus*. Aust J Biol Sci 39:37–45.

Johnston PG, Robinson ES (1987) X chromosome inactivation in female embryos of a marsupial mouse (Antechinus stuartii). Chromosoma 95:419–423.

Johnston PG, Robinson ES Johnston DM (1985) Dictyate oocytes of a kangaroo (Macropus robustus) show paternal inactivation at the X-linked Gpd locus. Aust J Biol Sci 38:79–84.

Kalantry S, Purushothaman S, Bowen RB, Starmer J, Magnuson T (2009) Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation. Nature 460:647–651.

Kaslow DC, Migeon BR (1987) DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. Proc Natl Acad Sci U S A 84:6210–6214.

Kay GF Penny GD, Patel D, Ashworth A, Brockdorff N, Rastan S (1993) Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. Cell 72:171–182.

Koina E, Chaumeil J, Greaves I, Tremethick D, Graves JAM (2009) Specific patterns of histone marks accompany X chromosome inactivation in a marsupial. Chromosome Res 17:115–126.

Koina E, Wakefield MJ, Walcher C, et al. (2005) Isolation, X location and activity of the marsupial homologue of SLC16A2, an XIST-flanking gene in eutherian mammals. Chromosome Res 13:687–698.

Korenberg JR, Rykowski MC (1988) Human genome organization: Alu, lines, and the molecular structure of metaphase chromosome bands. Cell 53:391–400.

Kuroda Y, Arai N, Arita M, Teranishi M, Hori T, Harata M, Mizuno S (2001). Absence of Z-chromosome inactivation for five genes in male chickens. Chromosome Res 9:457–468.

Kuroiwa A, Yokomine T, Sasaki H, et al. (2002) Biallelic expression of Z-linked genes in male chickens. Cytogenet Genome Res 99:310–314.

Lee JT, Lu N (1999) Targeted mutagenesis of Tsix leads to nonrandom X inactivation. Cell 99:47–57.

Lock LF, Melton DW, Caskey CT, Martin GR (1986) Methylation of the mouse hprt gene differs on the active and inactive X chromosomes. Mol Cell Biol 6:914–924.

Lock LF, Takagi N, Martin GR (1987) Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. Cell 48:39–46.

Loebel DA, Johnston PG (1993) Analysis of DNase 1 sensitivity and methylation of active and inactive X chromosomes of kangaroos (*Macropus robustus*) by in situ nick translation. Chromosoma 102:81–87.

Loebel DA, Johnston PG (1996) Methylation analysis of a marsupial X-linked CpG island by bisulfite genomic sequencing. Genome Res 6:114–123.

Lyon MF (1961) Gene action in the X-chromosome of the mouse (*Mus musculus L.*). Nature 190:372–373.

Lyon MF (1998a) X-chromosome inactivation spreads itself: effects in autosomes. Am J Hum Genet 63:17–19.

Lyon MF (1998b) X-chromosome inactivation: a repeat hypothesis. Cytogenet Cell Genet 80:133–137.

Lyon MF Rastan S (1984) Parental source of chromosome imprinting and its relevance for X chromosome inactivation. Differentiation 26:63–67.

Mahahevaiah SK, Royo H, VandeBerg JL, McCarrey JR, MacKay S, Turner JM (2009) Key features of the X inactivation process are conserved between Marsupials and Eutherians. Curr Biol 19:1478–1484.

Marahens Y, Loring J, Jaenisch R (1998) Role of the Xist gene in X chromosome choosing. Cell 92:657–664.

Marahens Y, Panning B, Dausman J, Strauss W, Jaenisch R (1997) Xist-deficient mice are defective in dosage compensation but not spermatogenesis. Genes Dev 11:156–166.

McKay LM, Wrigley JM, Graves JAM (1987) Evolution of mammalian X-chromosome inactivation: sex chromatin in monotremes and marsupials. Aust J Biol Sci 40:397–404.

McLaren A, Monk M (1981) X-chromosome activity in the germ cells of sex-reversed mouse embryos. J Reprod Fertil 63:533–537.

Migeon BR (2003) Is Tsix repression of Xist specific to mouse? Nat Genet 33:337.

Migeon BR, Jan de Beur S, Axelman J (1989) Frequent derepression of G6PD and HPRT on the marsupial inactive X chromosome associated with cell proliferation in vitro. Exp Cell Res 182:597–609.

Migeon BR, Wolf SF, Axelman J, Kaslow DC, Schmidt M (1985) Incomplete X chromosome dosage compensation in chorionic villi of human placenta. Proc Natl Acad Sci U S A 82:3390–3394.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Mohandas TK, Passage MB, Williams JW, Sparkes RS, Yen PH, Shapiro LJ (1989) X-chromosome inactivation in cultured cells from human chorionic villi. Somat Cell Mol Genet 15:131–136.

Mohandas T, Sparkes RS, Shapiro LJ (1981) Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. Science 211:393–396.

Murphy WJ, Davis B, David VA, et al. (2007) A 1.5-Mb-resolution radiation hybrid map of the cat genome and comparative analysis with the canine and human genomes. Genomics 93:299–304.

Namekawa SH, VandeBerg JL, McCarrey JR, Lee JT (2007) Sex chromosome silencing in the marsupial male germ line. Proc Natl Acad Sci U S A 104:9730–9735.

Nanda I, Shan Z, Schartl M, et al. (1999) 300 million years of conserved synteny between chicken Z and human chromosome 9. Nat Genet 21:258–259.

Nesterova TB, Slobodyanyuk SY, Elisaphenko EA, et al. (2001) Characterization of the genomic Xist locus in rodents reveals conservation of overall gene structure and tandem repeats but rapid evolution of unique sequence. Genome Res 11:833–849.

Okamoto I, Arnaud D, Le Baccon P, et al. (2005) Evidence for de novo imprinted X-chromosome inactivation independent of meiotic inactivation in mice. Nature 438:369–373.

Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E (2004) Epigenetic dynamics of imprinted X inactivation during early mouse development. Science 303:644–649.

Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N (1996) Requirement for Xist in X chromosome inactivation. Nature 379:131–137.

Plenge RM, Hendrich BD, Schwartz C, et al. (1997) A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. Nat Genet 17:353–356.

Potrzebowski L, Vinckenbosch N, Marques AC, et al. (2008) Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. PLoS Biol 6:e80

Priest JH, Heady JE, Priest RE (1967) Delayed onset of replication of human X chromosomes. J Cell Biol 35:483–487.

Quilter CR, Blott SC, Mileham AJ, Affara NA, Sargent CA, Griffin DK (2002) A mapping and evolutionary study of porcine sex chromosome genes. Mamm Genome 13:588–594.

Raudsepp T. Lee EJ, Kata SR, et al. (2004) Exceptional conservation of horse-human gene order on X chromosome revealed by high-resolution radiation hybrid mapping. Proc Natl Acad Sci U S A 101:2386–2391.

Reik W, Lewis A (2005) Co-evolution of X-chromosome inactivation and imprinting in mammals. Nat Rev Genet 6:403–410.

Rens W, Grutzner F, O'Brien PC, Fairclogh H, Graves JAM, Ferguson-Smith MA (2004) Resolution and evolution of the duck-billed platypus karyotype with an X1Y1X2Y2X3Y3X4Y4X5Y5 male sex chromosome constitution. Proc Natl Acad Sci U S A 101:16257–16261.

Richardson BJ, Czuppon AB, Sharman GB (1971) Inheritance of glucose-6-phosphate dehydrogenase variation in kangaroos. Nat New Biol 230:154–155.

Robinson ES, Samollow PB, VandeBerg JL, Johnston PG (1994) X-chromosome replication patterns in adult, newborn and prenatal opossums. Reprod Fertil Dev 6:533–540.

Rodriguez Delgado CL, Waters PD, Gilbert C, Robinson TJ, Graves JAM (2009) Physical mapping of the elephant X chromosome: conservation of gene order over 105 million years. Chromosome Res DOI:10.1007/s10577-009-9079-1.

Ross MT, Grafham DV, Coffey AJ, et al. (2005) The DNA sequence of the human X chromosome. Nature 434:325–337.

Samollow PB, Ford AL, VandeBerg JL (1987) X-linked gene expression in the Virginia opossum: differences between the paternally derived Gpd and Pgk-A loci. Genetics 115:185–195.

Samollow PB, Johnston PG, Ford AL, VandeBerg JL (1989) X-linked gene expression in metatherian fibroblasts: evidence from the Gpd and Pgk-A loci of the Virginia opossum and the red-necked wallaby. Biochem Genet 27:313–320.

Samollow PB, Robinson ES, Ford AL, VandeBerg JL (1995) Developmental progression of Gpd expression from the inactive X chromosome of the Virginia opossum. Dev Genet 16:367–378.

Schmidt M, Migeon BR (1990) Asynchronous replication of homologous loci on human active and inactive X chromosomes. Proc Natl Acad Sci U S A 87:3685–3689.

Sharman GB (1971) Late DNA replication in the paternally derived X chromosome of female kangaroos. Nature 230:231–232.

Sinclair AH, Foster JW, Spencer JA (1988) Sequences homologous to ZFY, a candidate human sex-determining gene, are autosomal in marsupials. Nature 336:780–783.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, et al. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837.

Spencer JA, Watson JM, Graves JA (1991) The X chromosome of marsupials shares a highly conserved region with eutherians. Genomics 9:598–604.

Takagi N, Sasaki M (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature 256:640–642.

Taylor JH (1960) Asynchronous duplication of chromosomes in cultured cells of Chinese hamster. J Biophys Biochem Cytol 7:455–464.

Torchia BS, Call LM, Migeon BR (1994) DNA replication analysis of FMR1, XIST, and factor 8C loci by FISH shows nontranscribed X-linked genes replicate late. Am J Hum Genet 55:96–104.

Turner JM (2007) Meiotic sex chromosome inactivation. Development 134:1823–1831.

VandeBerg JL, Thiel JE, Hope RM, Cooper DW (1979) Expression of PGK-A in the Australian brush-tailed possum, Trichosurus vulpecula (Kerr), consistent with paternal X inactivation. Biochem Genet 17:325–332.

VandeBerg JL, Cooper DW, Sharman GB, Poole WE (1977) Studies on metatherian sex chromosomes. IV. X linkage of PGK-A with paternal X inactivation confirmed in erythrocytes of grey kangaroos by pedigree analysis. Aust J Biol Sci 30:115–125.

Veyrunes F, Waters PD, Miethke P, et al. (2008) Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res 18:965–973.

Vigneau S, Augui S, Navarro P, Avner P, Clerc P (2006) An essential role for the DXPas34 tandem repeat and Tsix transcription in the counting process of X chromosome inactivation. Proc Natl Acad Sci U S A 103:7390–7395.

Wake N, Takagi N, Sasaki M (1976) Non-random inactivation of X chromosome in the rat yolk sac. Nature 262:580–581.

Wakefield MJ, Keohane AM, Turner BM, Graves JAM (1997) Histone underacetylation is an ancient component of mammalian X chromosome inactivation. Proc Natl Acad Sci U S A 94:9665–9668.
Wang Z, Willard HF, Mukherjee S, Furey TS (2006) Evidence of influence of genomic DNA sequence on human X chromosome inactivation. PLoS Comput Biol 2:e113.
Waters PD, Wallis MC, Graves JAM (2007) Mammalian sex-Origin and evolution of the Y chromosome and SRY. Semin Cell Dev Biol 18:389–400.
Waterston RH, Lindblad-Toh K, Birney E, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562.
West JD, Frels WI, Chapman VM, Papaioannou VE (1977) Preferential expression of the maternally derived X chromosome in the mouse yolk sac. Cell 12:873–882.
Wrigley JM, Graves JAM (1988) Sex chromosome homology and incomplete, tissue-specific X-inactivation suggest that monotremes represent an intermediate stage of mammalian sex chromosome evolution. J Hered 79:115–118.
Wutz A, Jaenisch R (2000) A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol Cell 5:695–705.
Zeng SM, Yankowitz J (2003) X-inactivation patterns in human embryonic and extra-embryonic tissues. Placenta 24:270–275.

Part VI
Marsupial Reproductive and Developmental Genetics

Jennifer A. Marshall Graves

If anything defines a marsupial in the popular literature, it is their distinctive reproduction, most obviously the development of the young externally in a pouch, rather than internally in the uterus. Indeed, the popular name of mammalian Infracalass Metatheria is taken from "marsupium", meaning pouch. Although not all marsupial species are pouched, the young of every species are born tiny and at a much less advanced stage ("altricial") than are eutherian young, and are suckled with milk for proportionately longer. This profound difference in the developmental program is the result of constraint in the size and location of the uteri, which are prevented from fusing into one large uterus by the position of the ureters (Tyndale-Biscoe, 2005).

This rather minor difference in plumbing has profound consequences for the development and nurture of the young, affecting every aspect of the nutrition and development of marsupial mammals. There is a fully functional yolk sac placenta that regulates carbohydrate and protein supply to the fetus and makes hormones requited for birth (making dubious the claim of uniqueness for "placental" mammals), but it is short lived and less developed, and does not secrete progesterone as does the placenta of eutherians. Marsupial nutrition is largely via the milk rather than the umbilical cord, so it is no surprise that marsupials have an extremely sophisticated and complex lactation system. In this part these two aspects of marsupial development are explored in detail.

As has been extensively documented in classic studies (Tyndale-Biscoe and Renfree, 1987) the earliest stages of marsupial development are different in critical ways from development in placental mammals; in particular the marsupial embryo does not develop an inner cell mass as do eutherian embryos, and the position of stem cells is enigmatic. However, the genes that are critical for early development in eutherians evidently all have important roles in marsupials, and Pask and Renfree, in Chapter 14 describe studies of the expression profiles in the tammar wallaby of genes that are known to be critical for development in eutherians. These genes include the big four; *OCT4, SOX2, NANOG* and *CDX2* that are believed to be the critical cues that define the differentiation of the trophectoderm and inner cell mass in mouse.

Amongst the many special features of marsupial development is the rare ability to switch embryonic development on and off. Although some eutherians also subscribe to this embryonic diapause, it has been best studied in the tammar wallaby, and the signals to resume development are well understood. Not so the signals to enter diapause, and suspend development, up to 11 months, and Pask and Renfree speculate that *LIF1* is critical in this step.

One extremely beneficial consequence of the external development of the marsupial young is that they are accessible to view, and even to experimental manipulation, at much earlier developmental stages than, say, mouse embryos. Growth and development of organs takes place in the pouch and can be investigated directly. For instance, the undifferentiated gonad does not differentiate into testis until a day after birth in kangaroos (O et al., 1988). Marsupials offer, therefore, a unique model for observing and experimenting on mammalian development (Renfree et al., 1995). Pask and Renfree describe how these differences in developmental plan have been exploited to discover fundamental truths about fetal and placental development, germcell development, sex determination, and embryonic growth and development in marsupials. This has immediate relevance to understanding development in all mammals, including humans. Several breakthroughs in mammalian development have been made using marsupials; for instance, the discovery that oestrogen can sex reverse an XY embryo in mammals as well as reptiles, delineation of a new pathway of steroid hormone biosynthesis (Wilson et al., 1999) and the discovery that *SRY*, not *ZFY*, is the sex determining gene in marsupials (Sinclair et al., 1988), and eutherian mammals including humans.

Not surprisingly, marsupial milk is complex and its composition changes during lactation as the young progresses from fetus to newborn to joey. As detailed by Sharp et al., in Chapter 15, the composition changes from protein-rich to carbohydrate and lipid rich as many of these components wax and wane during the year-long lactation cycle in tammar wallabies. These milk components, and the complex genetic control that orchestrates their concentrations, provide a model for control of lactation that is of considerable interest to the dairy industry, as well as for the fundamental knowledge of mammalian nutrition and growth (Brennan et al., 2007). Marsupial lactation is now going genomic (Lefèvre et al., 2007), raising the possibilities of an understanding at the molecular level.

However, marsupial milk is also loaded with growth factors that have very specific phenotypic effects on the pouch young. Cross-fostering experiments show that milk contains growth factors specific to the stage of development. For instance, milk from an advanced stage teat initiates fur growth in a much younger animal. Many of these factors are likely to be homologous to factors that regulate growth and development of eutherian embryos, and are delivered through the umbilical cord within the uterus.

Both papers make the point that marsupial reproduction, though very different from eutherian reproduction, is in no way to be considered a second-rate evolutionary intermediate. The short gestation of an altricial young, and the nurturing of an expendable pouch young through a prolonged lactation period, makes great sense for an animal living in the dry and uncertain conditions that Australia affords.

References

Brennan AJ, Sharp JA, Digby MR, Nicholas KR (2007) The tammar wallaby: a model to examine endocrine and local control of lactation. IUBMB Life 59: 146–150.

Lefèvre CM, Digby MR, Whitley JC, Strahm Y, Nicholas KR (2007) Lactation transcriptomics in the Australian marsupial, *Macropus eugenii*: transcript sequencing and quantification. BMC Genomics 8:417.

Renfree MB, Harry JL, Shaw G (1995) The marsupial male: a role model for sexual development. Philos Trans R Soc Lond B Biol Sci 350:243–251.

Sinclair AH, Foster JW, Spencer JA, et al. (1988) Sequences homologous to ZFY, a candidate human sex-determining gene, are autosomal in marsupials. Nature 336:780–783.

Tyndale-Biscoe H (2005) Life of Marsupials. CSIRO Publishing, Collingwood.

Tyndale-Biscoe H, Renfree MB (1987) Reproductive Physiology of Marsupials. Cambridge University Press, Cambridge.

O WS, Short RV, Renfree MB, Shaw G (1988) Primary genetic control of somatic sexual differentiation in a mammal. Nature 331:716–717.

Wilson JD, George FW, Shaw G, Renfree MB (1999) Virilisation of the male pouch young of the tammar wallaby does not appear to be mediated by plasma androgens. Biol Reprod 61:471–475.

Chapter 14
Molecular Regulation of Marsupial Reproduction and Development

Andrew J. Pask and Marilyn B. Renfree

Abstract The three extant mammalian lineages are primarily defined by their differing strategies of reproduction. Monotremes lay eggs while eutherians generally have a long period of gestation and give birth to well developed young. Marsupials sit between the two strategies and have a relatively short gestation, giving birth to poorly developed young that complete the majority of their early development external to the mother while suckling. This unique reproductive strategy provides an ideal mammalian model system in which to study and manipulate early development. Here we summarise our current understanding of the molecular regulation of marsupial reproduction and development, and highlight how marsupial studies have informed our understanding of the function of genes in mammalian development.

Keywords Early embryonic development · Embryonic diapauses · Fetal and placental development · Gametes and germ cells · Sex determination and sexual differentiation

14.1 Introduction

Three different modes of reproduction characterise the three extant mammalian lineages: the monotremes, the marsupials and the eutherians. Monotremes, like all mammals, lactate, but deliver their young from an egg that is laid when about half of the development of the embryo is complete. Eutherian mammals have extended the period of maternal protection within their mother's bodies through a relatively long gestation, and many have young that are almost independent at birth. Marsupials have adopted an alternative, but equally successful, reproductive strategy to that

A.J. Pask (✉)
Department of Zoology, ARC Centre of Excellence for Kangaroo Genomics, The University of Melbourne, Melbourne, VIC 3010, AustraliaDepartment of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269, USA
e-mail: andrew.pask@uconn.edu

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_14, © Springer Science+Business Media B.V. 2010

of the eutherian mammals (Renfree, 1981), and instead of retaining the fetus for a lengthy period inside the uterus, marsupials have elaborated lactation, regardless of whether they have a pouch or are pouch-less (Renfree, 1983). It is this strategy that makes marsupials such a useful model for research in developmental and reproductive biology.

The reproductive tracts of the three groups of mammals differ in the way that the ureters make their connection to the bladder to store urine (Tyndale-Biscoe and Renfree, 1987; Renfree, 1993; Renfree et al., 1996). In monotremes, the ureters open onto the dorsal wall of the urogenital sinus opposite the urethral opening of the bladder. The urine must cross the urogenital sinus to be stored in the bladder. The urogenital sinus opens into the single opening – the cloaca – that also receives the colon. Thus in monotremes, the eggs, urine and faeces pass out of a single cloacal sphincter. In the therian mammals, what begins as a trivial difference in the migratory route that the urinary ducts take in relation to the genital ducts in the fetus, results in profound differences in the anatomy of the reproductive tract. In marsupials, the ureters pass medially to the genital ducts, so preventing their fusion into a single vagina. This means that marsupials have a tri-partite system of vaginal canals, while in eutherians, since the ureters pass laterally to the genital ducts, the embryonic Mullerian ducts have been able to fuse to form a single vagina (Tyndale-Biscoe and Renfree, 1987; Renfree, 1993; Renfree et al., 1996). There are two external openings: one at the urogenital opening and one at the colon in both groups, but in marsupials these two openings are hidden by a sphincter that looks like a single opening (Renfree, 1992).

Marsupials have many features that make them ideal model species for understanding many basic mammalian developmental processes. Although no male marsupials have a pouch, and many of the smaller marsupial females are pouchless, the pouch or marsupium is the feature that gives the group its name. All deliver a tiny, altricial, relatively undeveloped young (Fig. 14.1a) after a very short period of pregnancy supported by a fully functional placenta (Tyndale-Biscoe and Renfree, 1987). Because the young is so altricial at birth, most organ growth occurs after birth whilst the young are readily accessible within the pouch. This unique developmental strategy lends itself to simple manipulations of cellular and organ differentiation at stages that occur in utero in eutherian species. For example, sex reversal can be induced by at least three different methods: by gonadal transplant, or by culturing in minimal medium or with the addition of Anti-Mullerian Hormone (AMH) (Whitworth et al., 1996) or by treatment with oestradiol (Coveney et al., 2001). Models of fetal growth restriction have shown that early nutrition is essential in establishing longevity. Since marsupials have a long and sophisticated lactation they also provide a model in which to address the role of nutrition during development. Milk composition changes throughout lactation, regulating the development of the young in the pouch (see Sharp et al., Chapter 15). The ability to manipulate growth of young by transferring them to the pouches of foster mothers (cross-fostering) whose milk is at a later or earlier stage of lactation is a unique way of manipulating growth and development that we have exploited (Trott et al., 2003; Menzies

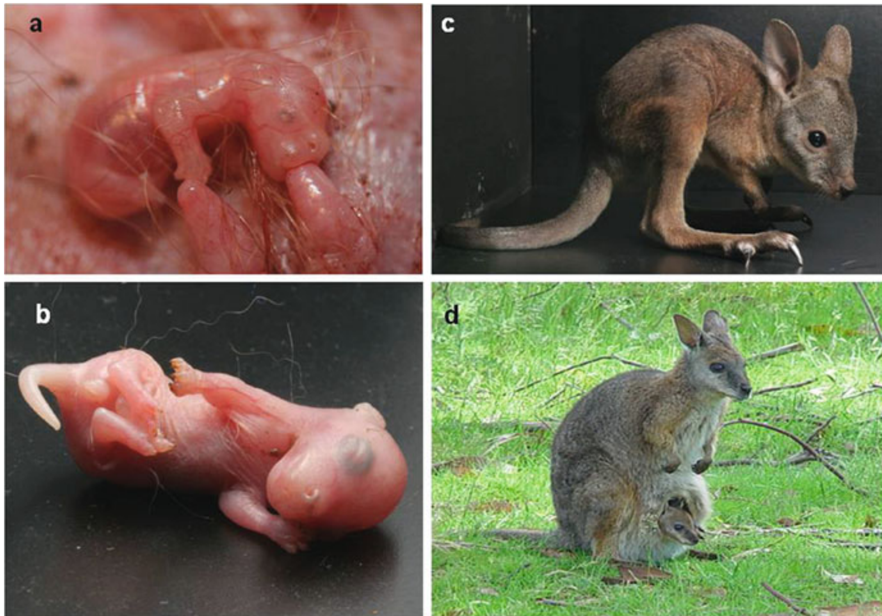


Fig. 14.1 Tammar wallaby neonate (**a**) a neonatal young about 12 h old; (**b**) pouch young aged 25 days (**c**) furred pouch young at weaning (300 days) and (**d**) an adult female tammar carrying a large furry young

et al., 2007, 2008). Here we review the molecular aspects of reproduction and development that have thus far been described in marsupials.

14.2 Early Embryonic Development

The zygote of the marsupial embryo passes rapidly down the oviduct to reach the uterus within 24 h as a single cell embryo. Cleavage patterns differ somewhat between the different marsupial groups, but all zygotes are enclosed in a shell coat, mucin layer and zona pellucida (Tyndale-Biscoe and Renfree, 1987). In addition, the blastomeres are all adherent to the zona pellucia, in contrast to eutherian mammals in which they adhere to each other (Fig. 14.2). The brushtail possum (*Trichosurus vulpecula*) zona pellucida (ZP) is composed of three major glycoproteins, designated ZP1, ZP2, and ZP3 based on their size and homology with eutherian ZP proteins (Mate et al., 2003). The brushtail possum zona pellucida protein ZP2 is 48–55% identical to that of eutherian mammals (Mate and McCartney, 1998) and ZP3 has a 46% amino acid identity with eutherian ZP3 (McCartney and Mate, 1999). The brushtail possum ZP2 and ZP3 are expressed exclusively in the ovary. The shell coat of the brushtail possum is comprised of a novel coat protein CP4

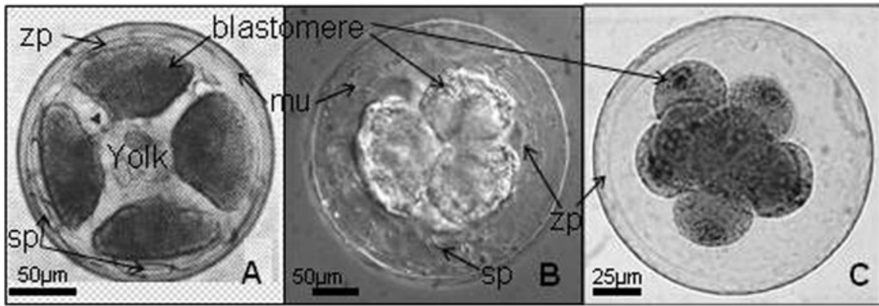


Fig. 14.2 Mammalian cleavage stage embryos. (**a**) dasyurid four cell embryo kindly supplied by (Selwood et al., 1997) (**b**) tammar four cell embryo (Renfree and Lewis, 1996) and (**c**) hamster eight cell embryo. Note that the marsupial zygote is about twice the size of the eutherian one. There is a "yolk" mass in the dunnart egg, but no obvious yolk in the tammar and none in the hamster. mu = mucin layer; sh = shell coat; sp = sperm trapped in the mucin layer; zp = zona pellucida

(Cui and Selwood, 2003), and there is uterine expression of this coat protein gene orthologue in the stripe-faced dunnart, *Sminthopsis macroura* that is under progesterone control (Menkhorst et al., 2009). Development from cleavage stages and up to somitogenesis are all completed while enclosed in these coats.

There are differences in the cleavage pattern between marsupial groups. During cleavage in dasyurid marsupials, the cells of the embryos of these stages are tightly apposed to the zona pellucida without making any contact with each other, and they can be distinguished as large and small cells by the 16-cell stage (Selwood, 1986, 1992, 2000; Selwood et al., 1997). When cells of early cleavage stages are destroyed in the dasyurid brown antechinus (*Antechinus stuartii*), only partial embryos develop, but once the pluriblast and trophoblast have formed (between the 16- and 32-cell stages), damage to the pluriblast epithelium is completely restored after further development in vitro (Selwood, 1986). In contrast, in macropodids, all the cells at cleavage stages are indistinguishable from one another throughout the cleavage stages and they contact each other as well as the zona (Fig. 14.2; Renfree and Lewis, 1996), so it is possible that the blastomeres during early cleavage are totipotent.

There are now molecular markers of potency available to detect expression in both totipotent and pluripotent cell types. The cloning of the gene *Pou5f1* (*OCT4*), which is expressed in totipotent and pluripotent cell lineages, has allowed localisation of these cells in the conceptus by in situ hybridisation (Schöler et al., 1990). *POU5F1* mRNA is expressed in oogonia and in oocytes of growing follicles of the brushtail possum, but not in early meiotic oocytes (Frankenberg et al., 2001). We now have a tammar wallaby specific antibody for *POU5F1* and have detected protein in the germ cells (Frankenberg et al., 2010). Other molecular markers for the trophectoderm (TE) and inner cell mass (ICM) lineages have recently been described in the mouse (Tanaka et al., 2005). In the mouse, it is not until the blastocyst stage that the two lineages are irreversibly determined and it appears that positional factors influence this segregation, with the outside, polar cells giving rise

to the TE cells and the inside, apolar cells forming the ICM (Kunath et al., 2004). In the mouse *Oct4* directly represses the transcription of several trophoblast-specific genes (Liu et al., 1997) but remains active in the ICM and is down-regulated in the TE cells. The HMG box transcription factor, *Sox2*, collaborates with *Oct4* to mediate transcription of target genes such as *Fgf4* (Ambrosetti et al., 1997) in the ICM. From the late morula stage, the homeobox gene *nanog* is strongly expressed in the mouse inner apolar cells, and is also detected in the ICM of the blastocysts and the primordial germ cells at d11.5pc (Chambers et al., 2003; Mitsui et al., 2003). *Nanog* is not detected in cleavage stages, trophectoderm or primitive endoderm. It appears that both *Oct4* and *nanog* are required for ICM determination, and both repress differentiation of the extraembryonic lineages. *Oct4* represses trophoblast, and *Nanog* represses primitive endoderm.

Trophectoderm determination requires *Cdx2*, whose expression pattern is complementary to that of *Nanog* in the early embryo (Kunath et al., 2004). Thus the reciprocal expression of *Oct4* and *Nanog* in the ICM and *Cdx2 in the* TE cells, are believed to be the critical cues that define the differentiation of these tissues. Inactivation of *Cdx2* is preimplantation embryonic lethal in the mouse (Chawengsaksophak et al., 2004). If the morphological distinction between inner cell mass equivalent cells and outer trophblast determines whether or not they express *Nanog* and *Cdx2*, then only the reactivated, and not diapausing, marsupial blastocysts cells will show differential expression of these genes. We now have preliminary evidence that this is so in the tammar wallaby and that the blastocyst in diapause and during early reactivation expresses both *CDX2*, *OCT4*, *NANOG* and *SOX2* in every cell, suggesting that they remain totipotent until relatively later stages of development (Renfree, Freyer, Frankenberg, Shaw and Pask, unpublished results).

14.3 Embryonic Diapause

Embryonic diapause provides a powerful, but as yet poorly exploited, tool to understand the control of early embryonic development. Whilst about 60 eutherian species, including carnivores, rodents and bats, and about 30 marsupial species can have diapause, the control is better understood in the tammar than in most other species (Tyndale-Biscoe and Renfree, 1987; Renfree and Tyndale-Biscoe, 1978). Tammars can completely halt development of their embryos at the unilaminar blastocyst stage when there are only 80–100 cells (Fig. 14.3). During diapause there is no cell division, and no blastocyst growth. Diapause can be maintained for at least 11 months with no loss of embryonic viability, but, remarkably, reactivation requires only a single hormone – progesterone. There are three phases of diapause: the entry into diapause and arrest of cell division; the maintenance of diapause; and reactivation after diapause. All attention to date has been on the latter two phases, but almost nothing is known of the mechanisms by which embryonic development is halted. Characterisation of the signals controlling entry into diapause would provide

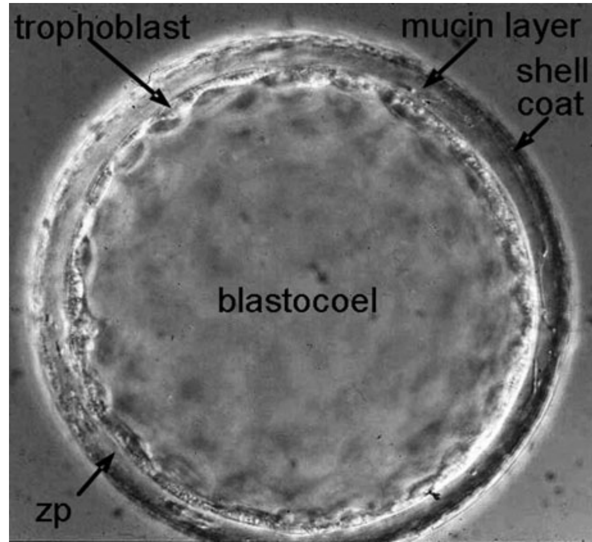


Fig. 14.3 Tammar wallaby blastocyst in diapause. There is a single layer of trophoblast cells and no inner cell mass. The 80–100 cell tammar blastocyst is 0.25 mm in diameter. zp = zona pellucida

an important conceptual advance in our knowledge of pre-implantation embryonic growth and for in vitro culture.

The one most unexpected differences between the marsupial blastocyst and that of eutherians that is that marsupials lack an inner cell mass (ICM) (see Fig. 14.3). This group of cells, on the inside of the eutherian blastocyst, are the source of embryonic stem cells and are responsible for forming the body of the embryo whilst the trophoblast cells (the outer layer) form the placenta. There are now many molecular markers to differentiate expression between totipotent and pluripotent cell types in eutherians (see above).

In the mouse, microarray analysis recently detected a suite of genes in the diapausing blastocyst and after reactivation (Hamatani et al., 2004; Hondo and Stewart, 2005). Only 229 (1%) of the >20,000 genes examined were differentially expressed between blastocysts in diapause (80 genes) and reactivated blastocysts (149 genes) (Hamatani et al., 2004). These 229 genes consisted of major functional categories, including the cell cycle, cell signalling, adhesion molecules and metabolic pathways. Some of these genes presumably control the arrested growth that occurs in diapause. One of these, leukemia inhibitory factor (*LIF*) appears to have multiple roles in regulating blastocyst implantation, diapause and blastocyst viability in mice (Hondo and Stewart, 2005). *LIF* has been cloned in the brushtail possum (that does not have diapause) and *tvLIF* transcripts were detected in most of the adult tissues and in the reproductive tracts of pregnant females (Cui and Selwood, 2000). We have now have evidence that *LIF* is critical for entry into, and reactivation from, diapause as in the mouse (MB Renfree, CM Hearn, H. Gehring, G. Shaw and AJ Pask, unpublished data).

Once development has resumed, the blastocyst expands and an embryonic disc develops. Little molecular biology has been carried out on the somite stage

marsupial embryos, but their planar morphology makes them perfect models for examining gene expression in a mammalian embryo from gastrulation to early fetal development. Marsupial embryos are also being studied because of the differences in the timing of developmental events (heterochrony) (e.g. Smith, 2006). Much novel information will emerge from in situ and immunocytochemistry studies of the developing marsupial embryo.

14.4 Fetal and Placental Development

Much more is known about the molecular control of fetal development, but almost all of this relates to the onset of sexual differentiation (see below). However, there is a growing list of developmental genes now characterised in marsupials, predominantly in the tammar. All marsupials have a chorio-vitelline placenta, and some also have a short lived chorio-allantoic placenta (see Tyndale-Biscoe and Renfree, 1987). The chorio-vitelline (or yolk sac) placenta is fully functional (see Fig. 14.4) and is the site of physiological exchange that supports the pregnancy. It actively regulates glucose, amino acids and protein supply to the fetus throughout pregnancy (Renfree, 1973, 1983; Freyer et al., 2003; Freyer and Renfree, 2009) and also synthesises the key hormones that initiate parturition, including cortisol (Ingram et al., 1999) and prostaglandin F2α (Shaw et al., 1999). However, it does not make much

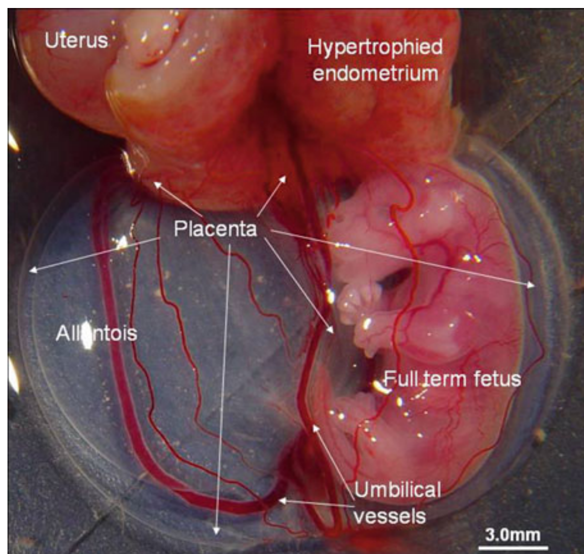


Fig. 14.4 Tammar wallaby full term fetus within its placenta. The fetus is 16 mm crown rump length and 7.2 mm head length. The chorio-vitelline placenta covers the fetus. The allantois is covered by the yolk sac membranes and the prominent vitelline vessels that make contact with the uterine endometrium. The umbilical cord enters at the abdomen

progesterone (Heap et al., 1980). Preliminary results using a bioassay showed that placental material from the tammar has gonadotropic activity (Renfree and Wide, cited in Tyndale-Biscoe and Renfree, 1987), and we have now identified the genes for the gonadotrophin, LHβ, as well as the genes for prolactin and growth hormone in the tammar placenta (BM Menzies, AJ Pask and MB Renfree, unpublished results). Growth hormone receptor is also expressed by the tammar yolk sac placenta (BM Menzies, AJ Pask and MB Renfree, unpublished results) and renin is detected in the yolk sac fluid (Lumbers et al., 1994). In addition, the fetal membranes (the yolk sac and the allantois) contain a cathepsin (Denker and Tyndale-Biscoe, 1986) which could be responsible for activation of renin (Lumbers et al., 1994). A number of imprinted genes have been also characterised in the tammar placenta but not all placentally expressed genes examined so far in the marsupial are imprinted (see Hore et al., Chapter 12). Nevertheless, it is clear that the idea that marsupials are not placental mammals is well and truly out-dated!

14.5 Post Natal Growth

Although the common perception is that marsupials all deliver young that are at the same poorly developed stage, there are several grades of organisation within neonatal marsupials (Fig.14.5) (Hughes and Hall, 1988). The dasyurids are the least developed at birth (Grade 1), and the macropodids the most well developed,

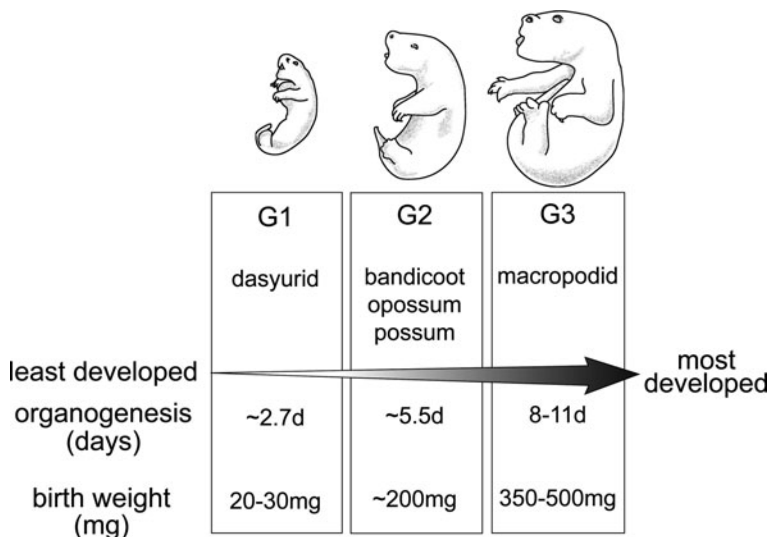


Fig. 14.5 Stages of development at birth. There are three grades of development reached by the time of birth in marsupials, together with increasing periods of organogenesis and birth weights. Grade 1 is characterised by the dasyurids, grade 2 by bandicoots and possums and opossums and grade 3 by the macropodids. Data from Hughes and Hall (1988)

with peramelids intermediate. The young of all groups spend a relatively long time dependent on lactation when most organ systems complete their development and mature (Tyndale-Biscoe and Janssens, 1988) (see Fig. 14.1). Lactation and its genetic control is now well characterised (see Sharp et al., Chapter 15) but less attention has been directed to the growth and maturation of the young.

There is a gradual development of the endocrine growth axis and growth hormone (GH)-regulated growth over the course of an extended lactation period, in contrast to the peri-natal period of development in eutherian mammals, providing a mechanism by which the young can develop nutritional independence from the mother gradually. Growth hormone (GH) release from the pituitary can be stimulated by ghrelin, a hormone that regulates appetite in mammals, and appears in the cells of the stomach during late fetal life. The altricial marsupial young must feed and digest milk at a comparatively immature stage of development, so it is perhaps not surprising that gastric ghrelin cells appear in the developing mesenchyme of the stomach from day 10 post partum (pp) to day 150 pp, but only remain in the hind-stomach after day 150 pp (Menzies et al., 2009). Pituitary ghsr-1alpha expression and plasma concentrations of ghrelin increase significantly up to day 70–120 pp, while GH expression also becomes elevated, declining with GH to reach adult levels by day 180 pp. Thus neonatal marsupials have an early onset of gastric ghrelin expression in concert with a functional stomach at a relatively earlier stage than that of developmentally more mature eutherian young (Menzies et al., 2009; Kwek et al., 2009).

14.6 Gonadotrophic Control of Reproduction and the Developing Pituitary

Gonadotrophin-releasing hormone (GnRH) plays a pivotal role in the endocrine control of both reproduction and embryonic development. The marsupial GnRH receptor (GnRH-R) gene in the tammar is composed of 328 amino acid residues with seven transmembrane domains and is a member of the G protein-coupled receptor family (Cheung and Hearn, 2002). There is a putative protein kinase A phosphorylation site and a putative protein kinase C (PKC) phosphorylation site in the first intracellular loop, and an additional PKC phosphorylation site was located in the third intracellular loop. Comparisons with the eutherian GnRH-Rs show a greater diversity in the N-terminal extracellular domain. Tammar GnRH-R has approximately 80% amino acid sequence homology with eutherian GnRH-Rs and 93% homology with the brushtail possum.

Gonadotropin-releasing hormone receptor (GnRH-R) consists of two splice variants (GnRH-R $\gamma 1$ and GnRH-R $\gamma 2$) Tammar GnRH-R, like eutherian GnRH-Rs, contains three exons and two introns. GnRH-R $\gamma 1$ has a 291 bp deletion from nucleotide positions 232–522 within exon 1. This transcript appears to be distinctive in the wallaby and has not been reported in other species. GnRH-R $\gamma 2$ has a 220 bp deletion from nucleotide positions 523–742, corresponding to exon 2. Both the

wild type receptor and the splice variants are membrane-associated molecules. The different pattern of expression of the wild type receptor and the variants transcripts found in adult and neonatal tissues suggests a specific developmental regulation of the GnRH-R γ2 transcript. The timing of expression suggests that GnRH-R may be involved in the regulation of early development in the testis and ovary of the developing pouch young (Cheung and Hearn, 2003, 2005). Similarly, the tammar FSH-R has 94% amino acid similarity with human FSH-R and is expressed in both the adult testis and ovary suggesting a similar function for this gene in both marsupials and eutherians (Mattiske et al., 2002). There are alternate splices of tFSH-R that produces four transcripts consistent with the splice variants seen in eutherians. This alternate splicing appears to be of functional significance in the mammalian ovary (Mattiske et al., 2002). Thus the gonadotrophic genes that control reproduction in mammals are highly conserved.

In eutherian mammals, the gonadotrophins (LH and FSH) are synthesized and stored in gonadotroph cells under the regulation of multiple mechanisms including GnRH. Very little is known about the regulation of gonadotrophin secretion and storage in pituitary glands of marsupials. In the brushtail possum, LHβ mRNA expression levels are constant over the oestrous cycle, regardless of the presence of a pre-ovulatory LH surge. There is a positive correlation between gene expression of FSHβ and plasma levels of FSH at different stages of the oestrous cycle. LHβ and FSHβ mRNA are both expressed in the pituitary, and there is protein storage of gonadotrophins within the pituitary gland (Crawford et al., 2009). The marsupial cDNAs encoding the follicle stimulating hormone β subunit, α-subunit of LH (also shared by a number of other glycoprotein hormones) and the LH-specific β-subunit are all highly conserved with their orthologues in eutherian species (Belov et al., 1998; Harrison et al., 1998; Fidler et al., 1998).

The developing pituitary of the pouch young is immunopositive for pituitary hormones (Leatherland and Renfree, 1983). In the pars distalis of pouch-young, presumptive somatotrops constitute up to 70% of the gland. There are small numbers (together representing less than 30% of the pars distalis) of immunoreactive mammotrophs, thyrotrophs, gonadotrophs and corticotrophs. The presumptive mammotrophs, gonadotrophs and thyrotrophs increase in number and apparent activity during the first 50 days postpartum. Presumptive corticotroph cells in 25–30 day-old animals were relatively most numerous, and apparently more active than at any other stage of pouch, but it is not known why this should be so. The changes in number and activity of cell types in the pars distalis correlate well with major developmental events such as the onset of adrenal activity, the rapid growth phase in the first 100 days postpartum, and the generally low thyroid activity in pouch-young of less than 50 days of age (Leatherland and Renfree, 1983).

Prolactin is a key regulator of reproduction in macropodid marsupials and critical for the long lactation. Analysis of the molecular evolution of the marsupial prolactin gene has shown that among vertebrates, is most closely related to the prolactin of eutherian mammals (as expected) but since it branched off from eutherian prolactin it has shown reduced rates of evolution compared to those seen in eutherian species (Curlewis et al., 1998).

14.7 Gametes and Germ Cells

Germ cells (the progenitors of the egg and sperm), are the only potentially immortal cells in the body since they are the only ones capable of transmitting DNA to the next generation (Short, 2003). In order to maintain their specialised state, germ cells have unique, and sex specific, developmental pathways. Almost all our information on the origin and fate of the germ cells has been derived from the mouse, where they arise in the embryonic epiblast and then leave the embryo, moving into the extra-embryonic tissues. This may protect them from the DNA methylation by which the other cells of the embryo are genomically imprinted and lineage restricted. Germ cells later return to the embryo and migrate to the developing gonad, controlled by a suite of cytokines, growth factors and cell adhesion molecules. They proliferate during their migration from their extra-embryonic location to the gonadal ridges. Once in the gonads, the germ cells continue to proliferate before initiating meiosis (in the developing ovary) or mitotic arrest (in the developing testis); in most species no further development of germ cells occurs until puberty.

Marsupials have several advantages as models for the study of germ cell biology. Their planar embryonic development makes it easier to visualise germ cells and trace the events involved in their specification. Furthermore, in contrast to the mouse and human, germ cell proliferation and subsequent entry into meiosis occurs post-natally in marsupials, making these processes easy to study. However, there is relatively little information on marsupial germ cells at a genetic level.

Migration of PGCs from their position outside the embryo to the gonads is critical for their survival and function. In the mouse, PGCs migrate from the primitive streak into the endoderm, which forms the gut. They then migrate via the anterior aspect of the gut into the gonads (Anderson et al., 2000). In the tammar, they never enter the gut but instead migrate via the gut mesentery (Ullmann et al., 1997; Jiang et al., 1997), but as yet we know nothing of the signals that direct them. Germ cells have been identified at the 17 somite stage embryo with alkaline phosphatase staining and with POU5F1 (Ullmann et al., 1997; Frankenberg et al., 2010) but earlier identification is currently under investigation (Hickford, Shaw, Pask Frankenberg and Renfree). In the brushtail possum, c-kit gene expression is detected in the germ cells and somatic cells during the first 15 days of life (Eckery et al., 2002a). However, after Day 30 and into adult life, c-kit expression becomes exclusive to germ cells. This pattern of ovarian development is similar in marsupials to eutherians, and suggests that c-kit may play a key role in germ cell development at various stages throughout life.

Although the sex-chromosome composition of a germ cell influences gamete function (McLaren, 1981, 2003), the decision as to whether a germ cell enters meiotic or mitotic arrest depends on the surrounding somatic cell environment of the ovary or testis. Murine germ cells enter the gonads between E10.5 and 11.5, when gonadal differentiation is occurring (Fig. 14.1). Similarly in the tammar germ cells enter the gonads during early gonadal differentiation (Alcorn and Robinson, 1983). In mice, they continue to proliferate until around E13.5, when oogonia stop proliferating and enter prophase of the first meiotic division while spermatogonia undergo

mitotic arrest until early puberty when they enter meiosis. Entry into meiosis is induced by retinoic acid (RA), which activates the expression of *Stra8*, a cytoplasmic protein required for pre-meiotic DNA replication (Koubova et al., 2006). Early entry into meiosis in males is prevented by expression in the testis of *Cyp26b1*, an enzyme in the cytochrome P450 family that degrades RA (Bowles et al., 2006; Koubova et al., 2006; Vernet et al., 2006). We have cloned and characterised tammar *Stra8* and *Cyp26b1* in the tammar, but do not yet have functional studies to confirm that a similar pathway exists in the tammar.

14.7.1 The Ooctye and the Adult Ovary

The marsupial oocyte is larger than that of eutherians, averaging around 0.25 mm compared to the 0.135 mm oocyte of the mouse (see also Fig. 14.2). Marsupial oocytes are surrounded by a zona pellucida, a proteinaceous mucin layer and a shell coat. Each of these three egg coats contains specific proteins. Vap1 is a protein isolated from the vesicle-rich hemisphere of the brushtail possum oocyte, possibly ovary-specific, which occurs during follicle formation and growth and in adult ovaries (Cui et al., 2005). VAP1 is being trialled as an immune target in contraceptive studies of the possum (Cui et al., 2005; Nation et al., 2008). Two other oocyte proteins ZP2 and c-mos have ovary-specific expression in the dunnart (Au et al., 2008): ZP2 is expressed in the cytoplasm of oocytes of primordial, primary and secondary follicles but expression is highest in oocytes of primary follicles. ZP2 and is not expressed in granulosa cells of any follicles (Au et al., 2008).

Both LH and FSH play a central role in controlling ovarian function in mammals. However, little is known about the type of ovarian cells that are responsive to LH and FSH in marsupials. The receptors for LH and FSH in ovaries of brushtail possums are found in granulosa cells of healthy follicles containing at least two complete layers of cells at the time of antrum formation (Eckery et al., 2002b). Cells of the theca interna expressed LH-R mRNA but not FSH-R mRNA. Neither FSH-R nor LH-R mRNA was detected in atretic follicles. Both FSH-R and LH-R mRNAs are found in the luteal cells, but only LH-R mRNA was observed in interstitial cells (Eckery et al., 2002b).

Interstitial tissue is a prominent feature of the marsupial ovary, occupying more than 2/3 of its volume (Tyndale-Biscoe and Renfree, 1987). In the tammar, the cells of the interstitial tissue resemble those of the corpora lutea and have the appearance of a steroid secreting tissue (Renfree and Tyndale-Biscoe, 1973). They do have some steroid synthetic capacity and can metabolise precursors to form progesterone (Renfree et al., 1984). In the brushtail possum, both the luteal cells and the interstitial tissue cells express receptors for prolactin and luteinising hormone (LH), suggesting that prolactin and LH may be important in the regulation of steroidogenesis in these cells (Eckery et al., 2002b). In the tammar, prolactin inhibits luteal growth during diapause (Tyndale-Biscoe and Renfree, 1987).

The oocyte derived growth differentiation factor 9 (GDF9) and bone morphogenetic protein 15 (BMP15; also known as GDF9b) are essential for normal

follicular growth. These two proteins are present in the developing follicles of the brushtail possum, and the mRNAs encoding *GDF9* and *BMP15* are detected in oocytes of newly-formed primordial follicles (i.e. days 78–85) (Eckery et al., 2002c). Expression of both mRNAs is restricted to the oocyte and persists through all stages of folliculogenesis, from non-growing primordial follicles to follicles undergoing preantral or antral development. GDF9 and BMP15 may have a possible role in the maintenance of primordial follicles and during follicular development (Eckery et al., 2002c).

The ovary of the brushtail possum secretes steroids, and the expression of numerous steroidogenic genes have been investigated. Steroidogenic factor-1 (SF-1), steroidogenic acute regulatory protein (StAR), cytochrome p450 side chain cleavage (P450scc), 3β-hydroxysteroid dehydrogenase/Delta5, Delta4 isomerase (3βHSD), cytochrome p45017alphahydroxylase (p45017αOH), and p450 aromatase (p450arom) are not expressed in oocytes at any stage of follicular development (Whale et al., 2003). However, SF-1 is detected in granulosa cells from the type 2 or the primary stage of development and thereafter to the preovulatory stage. In addition, the theca interna of small and medium-size antral, but not preovulatory, follicles and the interstitial glands and corpora lutea expressed SF-1 mRNA. Granulosa cells of preantral and small to medium-size antral follicles do not contain p450scc mRNA. The corpus luteum expresses p450arom, indicating that this tissue also has the potential to secrete estrogens in the brushtail possum (Whale et al., 2003) as it does in the tammar wallaby (Fletcher and Renfree, 1988). Corpora lutea of tammar wallabies contained oestradiol-17 β and oestrone during embryonic diapause and at all stages of pregnancy studied after blastocyst activation (Renfree et al., 1984). Aromatase is undetectable in corpora lutea or in interstitial tissue, but was present in the ovarian tissues (including follicles) which remained after removal of corpora lutea (Renfree et al., 1984). Oestrogen (ERα and ERβ), androgen (AR) and progesterone (PR) receptors are present during ovarian and follicular development in the brushtail possum (Haydon et al., 2008) and of the tammar wallaby (Calatayud et al., 2010; Butler et al., 1998).

Inhibin α in the brushtail possum is expressed throughout all stages of follicular growth and is also detected in cells of the corpus luteum (Vanmontfort et al., 1998). Inhibin α-subunit protein has been highly conserved since the divergence of the marsupial and eutherian mammalian lineages. Inhibin α and β have also been cloned in the tammar, and are more highly expressed in the ovary than in any other tissue (C Borchers, G Shaw, D Robinson and M Renfree, unpublished results).

14.7.2 Spermatozoa and the Adult Testis

Marsupials offer a unique model to examine spermatogenic gene function and evolution. While there are many similarities in the spermatogenesis process between eutherians and marsupials, there are also many differences. These similarities and differences can be exploited to gain a greater understanding of the molecular control of this process. One of the key differences between marsupial and eutherian

spermatozoa is their morphology and the position of the acrosome. American marsupials, with the exception of *Dromiciops gliroides*, have sperm that are paired, joining at their heads at the acrosomal surfaces (Tyndale-Biscoe and Renfree, 1987). There are eight stages of spermatogenesis and 14 stages of spermiogenesis recognized in the tammar and the spermatogenic cycle takes 16 days (Setchell, 1977; Harris and Rodger, 2005). Spermatogonia undergo a series of mitoses to produce the primary spermatocytes. Sperm maturation follows a similar pattern to that seen in eutherians (Harris and Rodger, 2005; Tyndale-Biscoe and Renfree, 1987). However, in marsupials the process of spermiogenesis, when the spermatids become spermatozoa, is complicated and dynamic and differs in some respects from that of eutherians (Lin et al., 1997). Sperm of most marsupials have a dorso-ventrally flattened nuclear surface (Temple-Smith, 1994) and there is an acrosome on the dorsal nuclear surface (Mate and Rodger, 1991; Sistina et al., 1993). During epididymal transit and capacitation, there is a distinct reorientation of the sperm head which rotates on its axis from a T-shape relative to the tail (Temple-Smith, 1994; Bedford and Breed, 1994; Lin and Rodger, 1999). The marsupial sperm head becomes compact but there does not appear to be any special stabilization of chromatin in the nucleus by disulphide bonding or in the sperm head membranes (Cummins, 1980; Temple-Smith and Bedford, 1976). In vitro capacitation cannot be achieved unless the sperm are co-cultured with oviductal cells (Sidhu et al., 1999). cAMP induces tyrosine phosphorylation of sperm proteins and the "T"-shaped orientation is associated with capacitation in marsupials (Sidhu et al., 2004). These differences in structural organization suggest that there may also be differences in the genes controlling spermatogenesis and sperm function and motility. GAPDs was identified as a conserved mammalian energy source for sperm motility in marsupials (Ricci and Breed, 2005), suggesting that marsupials employ similar molecular regulators as eutherians during spermatogenesis. For example, at fertilisation, spermatozoa bind to the zona pellucida, and in eutherian mammals, β1,4-galactosyltransferase-I (GalTase-I) is one of the key molecules on the sperm surface that is likely to be involved. This protein is also present in possum spermatozoa, suggesting a similar role as a gamete receptor molecule on the sperm surface of the brushtail possum (Braundmeier et al., 2008).

In further characterising the molecular controls of the spermatogenesis process in marsupials, research has focused so far on the role of X-linked genes in the process. The mammalian X-chromosome has long been suspected to contain a disproportionate number of genes associated with sex, compared to the rest of the genome (Ohno, 1967). Recent analyses suggests that the eutherian X contains, in particular, a high number of genes involved in early spermatogenesis, while it is deficient in those that control late spermatogenesis (Khil et al., 2004). Since marsupials have a smaller X-chromosome then eutherians it is possible to examine whether genes that are spermatogenic and X-linked in eutherians but autosomal in marsupials evolved their spermatogenic role after their recruitment to the X-chromosome, or if genes that already had a function in spermatogenesis were selectively moved to the X-chromosome in eutherian evolution. We can also determine if genes that

are conserved on both the marsupial and eutherian X have maintained a role in spermatogenesis through therian evolution.

AKAP4 is an X-linked member of *AKAP* gene family (Turner et al., 1998) that encodes the most abundant protein so far identified on the sperm fibrous sheath, and is present in all mammals so far examined (Carrera et al., 1994, 1996; Colledge and Scott, 1999; Eddy et al., 2003; Fulcher et al., 1995; Turner et al., 1998). AKAP4 anchors cAMP-dependent protein kinase A to the fibrous sheath, where it is thought to be required for the regulation of spermatozoal motility (Brown et al., 2003; Edwards and Scott, 2000). *Akap4* mRNA is expressed only in the post-meiotic phase of spermatogenesis and AKAP4 protein is restricted to the spermatogenic cells from the round spermatid stage (Miki et al., 2002). *Akap4* deletions result in shorter sperm flagellum with an incomplete fibrous sheath causing infertility due to loss of motility, suggesting AKAP4 is a scaffold protein involved in regulating flagellum function (Miki et al., 2002). In the tammar wallaby, *AKAP4* gene structure, sequence and predicted protein was highly conserved with that of eutherian orthologues. As expected from its location on the long arm of the eutherian X, it is located on the marsupial X-chromosome. There is no *AKAP4* expression detected in the developing young but in the adult, *AKAP4* expression is abundant and limited to the testis. *AKAP4* mRNA is detected in the cytoplasm of round and elongated spermatids while AKAP4 protein is found on the principal piece of the flagellum in the sperm tail. This is consistent with its expression in other mammals. Thus *AKAP4* appears to have had a conserved role in spermatogenesis for at least the last 166 million years of mammalian evolution (Hu et al., 2009), so is presumed to have had this role on the original therian X-chromosome.

14.8 Sex Determination and Sexual Differentiation

Disorders of gonadal development are among the most common birth defects in humans (1 in 4,000 births). These disorders frequently result in infertility and can have profound psychological and medical consequences upon the individual, family, and society. Some intersexual conditions are the result of inappropriate exposure to hormones during fetal life, while others are due to spontaneous or inherited gene mutations. Dysregulation of the gonadal differentiation pathway may play a role in gonadal cancer since ovarian tissue can form in testicular tumours, and testicular cords can form in ovarian tumours. Defining the genes and their interactions during gonad organogenesis will inform us about the causes and consequences of normal and abnormal sexual development, infertility, and gonadal malignancies.

The gonad is also an ideal model tissue in which to study the process of organogenesis, since two completely different organs arise from a bipotential primordium. Any errors in development are readily identifiable because of the resulting intersexual phenotypes. The process that results in the sexually dimorphic development of the gonad is one of the most fundamental, yet surprisingly poorly understood,

areas of developmental biology. The majority of studies to date on mammalian sex determination and differentiation have been carried out in mice. While mice are greatly amenable to genetic manipulation their embryonic development is complex and extremely rapid. The gonad forms in mice around E10.5 and differentiates into a testis just 24 h later in XY males. The rapid process of development makes it extremely difficult to dissect the progression and relative expression of genes during this process. In the tammar wallaby the gonad forms around day 22 of gestation and does not differentiate into a testis until around 6 days later and an ovary around 10 days later. Furthermore, both testicular and ovarian differentiation occur after birth, making them accessible for experimental manipulations both in vivo and in vitro. Studies of the expression and function of sex determining genes in the tammar have revealed some exciting findings that expanded our understanding of their function in the sex determination pathway of all mammals.

14.8.1 Localisation and Characterisation of the Testis Determining Factor SRY

The *SRY* (Sex determining region of the Y) gene is the testis determining gene in both humans (Sinclair et al., 1990) and mice (Koopman et al., 1991) and is located on the Y chromosome of most eutherian mammals (Sinclair et al., 1990), and marsupials (Foster et al., 1992). Although it is poorly conserved with human *SRY* (around 70% in the HMG box and almost none in the flanking arms), and deletion or mutation analysis is not possible, expression of tammar *SRY* is consistent with a conserved role as the testis determining factor in marsupials. In the mouse *SRY* reaches peak levels of expression shortly before testis formation at around E11–11.5 (Koopman et al., 1990) and acts as a switch to initiate Sertoli cell differentiation and the expression of *Sox9* and *Amh* (both critical for male development). In the tammar wallaby *SRY* was broadly expressed in the developing embryo (Harry et al., 1995), unlike eutherians where it is restricted mainly to the testis. However, these early studies examined *SRY* expression only by PCR, so it is possible that there is a more restricted expression pattern. In the developing tammar testis, *SRY* reaches a peak in expression just before testis cord formation (as in mice and humans) but *SOX9* and *AMH* have a sexually dimorphic expression up to 3 days prior to the *SRY* peak (AJ Pask and MB Renfree, unpublished data). This suggests that in the extended developmental timing of the tammar gonad, *SRY* expression begins to induce a sexually dimorphic expression before cord formation and the peak in its own expression (AJ Pask and MB Renfree, unpublished data).

14.8.2 Pouch and Scrotum

Unlike eutherians, not all aspects of marsupial male development depend on formation of a testis as a result of *SRY* action and the production of androgens. There are at least four structures that depend on a gene or genes on the X-chromosome, namely

the processus vaginalis, that lines the inguinal canal, the gubernaculum, the ligament that directs the testis into the scrotum, the scrotum itself, and the pouch within which the mammary primordial develop. On the day of birth, these four structures are sexually dimorphic, despite the absence of a differentiated testis, suggesting that their differentiation is directly genetically controlled and independent of testicular hormones (O et al., 1988). Taking this further back into the fetal stages, we found that the scrotal bulges in males and mammary primordial in females are present on the abdominal wall of the fetus by d22 of the 26.5 day pregnancy in the tammar, a time when the gonadal ridge is barely formed. Examination of the intersexual phenotypes of animals with abnormal sex chromosome constitutions led to the conclusion that a single X chromosome directs the appearance of a scrotum, while two X-chromosomes are correlated with the development of a pouch (Renfree and Short, 1988). XXY marsupials have abdominal testes and a penis, but no scrotum, and have a pouch, while XO marsupials have ovaries, but no pouch and they develop a scrotum (Renfree and Short, 1988). This hormonal independence of sexual dimorphism is confirmed by treatment of young with exogenous steroids that had effects on the genital tract but did not induce or prevent scrotum, mammary and pouch formation (Shaw et al., 1988). The mechanism of this X-dosage is not clear, and an alternative that Cooper (1993) suggested is that it could relate to the parent of origin of the X chromosome. The most puzzling of all are the intersexes those marsupials that are bilateral gynandromorphs, with half a pouch on one side of the body and half a scrotum on the other (Sharman et al., 1970, 1990). These examples of sexual dimorphisms that are not under the influence of androgens overturned the Jost (1947, 1970) paradigm that a testis, once formed, controlled all sexual dimorphisms. Subsequently, the description of other previously unrecognised sexual dimorphisms that occur before gonad formation emerged. In eutherian mammals, it was noted that Y bearing blastocysts are larger than XX counterparts in both rat and cattle, furthermore there is sexually dimorphic timing of neural tube development in the rat (see Shaw et al., 1990; Burgoyne et al., 1995). Later, a bilateral gynandromorphic bird was also described with male characteristics on one side and female on the other defining plumage and song development in the brain, again suggesting that is also under genetic, and not hormonal, control (Arnold, 2009).

14.8.3 X-Linked Genes Involved in Sexual Differentiation

As mentioned above, a disproportionate number of genes involved in sex and reproduction reside on the eutherian X-chromosome. Many of these have been now characterised in marsupials (Fig. 14.6) and all have provided new insights into the evolution of sex-linked genes.

ATRX: The X-linked *alpha-thalassemia and mental retardation gene* (*ATRX*) is essential for normal testicular development in humans. Marsupial have two orthologues of this gene, one on the X and the other on the Y. Marsupial *ATRX* is widely expressed as in eutherians, but unlike its eutherian orthologue, it is not expressed in

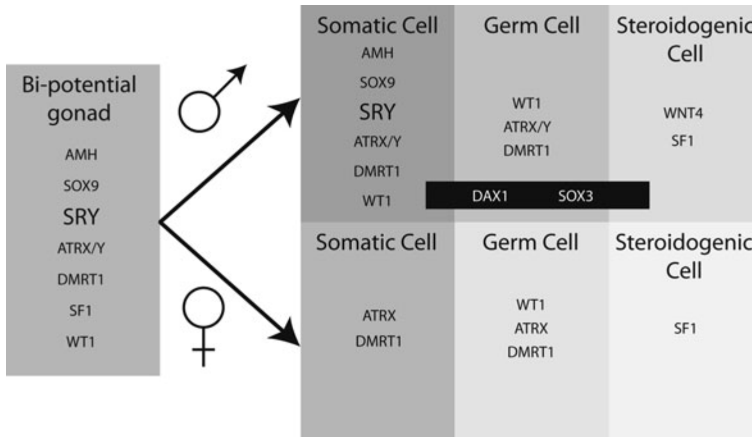


Fig. 14.6 Summary of expression of sexual differentiation genes in the marsupial gonad. The genes with known expression patterns are shown, along with the component of the gonad in which they are expressed. Somatic cells represent the granulosa cells of the ovary or Sertoli cells of the testis; the steroidogenic cells represent the theca cells of the ovary or the Leydig cells of the testis. Genes shown in the *black box* are known to be expressed in the adult testis or ovary (by RT-PCR), but their cellular localisation in unknown. SRY is only present in XY gonads and is not present in the bipotential XX gonad

the developing testis. The testis-specific Y-orthologue *ATRY* is unique to marsupials. It is exclusively expressed in the developing and adult testis and at a low level in some other components of the male urogenital system (Pask et al., 2000a). This suggests the *ATRY* may have been retained in the marsupial lineage due to its specialisation in male reproduction, and the loss of testicular expression from the *ATRX* orthologue. In eutherians this specialisation of the Y-linked allele may not have occurred, *ATRX* continued to fulfil its role in the testis and *ATRY* was degraded and lost. These two genes are of exceptional interest because they allow us to calculate rates of evolution occurring in X and Y homologues that acquire specialised functions.

SOX3: Almost all genes on the mammalian Y chromosome represent degraded orthologues of loci on the X-chromosome. *SOX3* (*SRY-like*, *HMG box containing gene 3*) has been the subject of much interest in gonadal differentiation since the demonstration that it is the X-borne homologue of *SRY* (Foster and Graves, 1994). *SOX3* is expressed in the developing mouse and human testis (Weiss et al., 2003) and so it was possible that it may interact with *SRY* to determine testicular development. However, in knockout mice it has no effect on sex determination, but rather plays a role later in sexual differentiation and particularly gonadogenesis (Weiss et al., 2003). In marsupials, the sequence of *SOX3* is highly conserved and X-linked, but it is not expressed in the developing gonads (Pask et al., 2000b). This is consistent with the absence of *SOX3* function in early sex determination and suggests that this gene and its Y-linked orthologue (from which *SRY* evolved) on the original therian sex chromosomes, was not expressed in the ancestral developing testis. Therefore

ATRY, and not *SRY*, is the earliest known Y-linked gene with conserved expression in the developing gonad, and could represent an ancestral sex determining gene that was later supplanted by *SRY*, before the marsupial/eutherian divergence.

14.8.4 Autosomal Genes Involved in Sexual Differentiation

The majority of genes critical for gonad development and its subsequent differentiation are autosomal in mammals. Again, marsupial sexual differentiation is typically mammalian, and the genes so far described for eutherian sexual differentiation are essentially the same as in marsupials (Fig. 14.6).

SF1: Steroidogenic factor 1 is critical for the formation of the early gonadal primordia, and later for gonadal function in both mice and humans (reviewed in Parker, 1998). The amino acid sequence of SF1 is highly conserved among vertebrate species, and marsupials are no exception. Tammar *SF1* is expressed by both ovaries and testes on the day of birth, just prior to the onset of testicular differentiation, and persists until at least 8 days after birth in both sexes. *SF1* mRNA is predominant in the Sertoli cells of the developing testis and the pre-granulosa cells of the developing ovary, as in mice and humans (Whitworth et al., 2001). In the adult ovary *SF1* is detected in the steroidogenic corpus luteum, and in the granulosa cells and theca of small to medium-sized antral follicles, but is absent from large antral follicles. These findings are consistent with a conserved role for SF1 in both early gonad formation in marsupials and gonadal development and steroidogenesis.

WT1: Another gene critical for early gonad formation in mammals is Wilms tumor 1 (*WT1*). WT1 is expressed in the early developing gonad in mouse and is believed to be important for initiating SRY transcription in the developing male gonad (Matsuzawa-Watanabe et al., 2003). Wt1 also plays an essential role in mouse germ cell survival (Natoli et al., 2004). Like SF1, WT1 is exceptionally conserved in marsupials with its eutherian orthologue at both the nucleotide and amino acid level. However, tammar WT1 is predominantly expressed as a novel isoform in which a portion of exon 1 is removed, partially deleting the RNA recognition motif (RRM), believed to be important for WT1 function (Pask et al., 2007). Expression of tammar WT1 is similar to mouse and human, and protein is detected in the Sertoli and granulosa cells of the developing testis and ovary, respectively. In contrast to eutherians, WT1 is also localised in the germ cells of both males and females at all stages of gonadal development, and appears to be dynamically regulated during spermatogenesis and oogenesis. Despite partial absence of the RRM, WT1 is still localised to RNA rich regions of the tammar oocyte including speckled bodies within the nucleus, in the nucleolus and the perinucleolar compartment (Pask et al., 2007). The perinucleolar compartment is important for RNA processing and metabolism within the cell. The location of WT1 in this region in the tammar suggests that the RRM is not required for WT1 co-localisation with RNA.

AMH: Anti-Müllerian hormone (AMH), also known as Müllerian inhibiting substance (MIS) is critical for the regression of the female urogenital ducts in the developing male. Müllerian duct regression occurs in utero in eutherian mammals,

but occurs entirely postnatally in marsupials. The tammar AMH gene is highly conserved both within the protein coding domain and 5′ promoter sequence. Like eutherian AMH, the tammar orthologue contains binding sites for SF1, SOX9, and GATA factors but also contains a putative SRY-binding site (Pask et al., 2004). AMH expression in the developing testis begins shortly after Sertoli cell differentiation and seminiferous cord formation at around 2 days post partum (Hutson et al., 1988; Pask et al., 2004). Müllerian duct regression is first apparent in male pouch young of the tammar wallaby 6–7 days after birth and, as in eutherian mammals, is characterised by a condensation of the periductal mesenchyme into a whorl around the ductal epithelial cells (Whitworth et al., 1997) and it is not fully regressed until around day 13 pp in males (Renfree et al., 1996). *AMH* is localized in the cytoplasm of the Sertoli cells in the developing testis, as in mice and humans, and is lost by adulthood. There is no detectable *AMH* expression in the developing ovary at the times examined (days 0–8 pp), but in adults it is produced by the granulosa cells of primary and secondary follicles (Pask et al., 2004). This expression pattern is consistent with a conserved role for AMH during mammalian evolution and is intimately linked to upstream sex determination mechanisms. A similar pattern of expression occurs in the brushtailed possum (Juengel et al., 2002). Half of the neonatal pouch young examined have *AMH* mRNA in the neonatal ovary, but it rapidly diminishes and is not seen again until after day 78 pp when follicles first form in the possum ovary. *AMH* mRNA is limited to the granulosa cells once follicles have formed. In the possum testis, there is strong expression of *AMH* in Sertoli cells around the time of sexual differentiation of the gonad, but expression decreases to very low levels in adult possums (Juengel et al., 2002).

SOX9: *SRY-like HMG box containing gene 9* is one of the most critical genes in testicular development in all vertebrates. SRY may do little more than to ensure SOX9 is upregulated in the developing XY gonad. Once activated, SOX9 triggers the rest of the testicular development pathway. However, SRY could not be the sole regulator, since SOX9 is also upregulated in the developing male gonad of chickens, turtles, alligators, and rainbow trout, all species that lack SRY (Oreal et al., 1998; Western et al., 1999; Spotila et al., 1998; Takamatsu et al., 1997). In the mouse SOX9 is upregulated by synergistic binding of Sry and Sf1 to a testis specific enhancer element TESCO. Once Sry expression decreases, *Sox9* can bind to its own promoter and maintain its expression (Sekido and Lovell-Badge, 2008).

The expression of both SF1 and SRY in the tammar is consistent with a similar synergistic role in upregulating *SOX9* (AJ Pask and MB Renfree, unpublished data) and TESCO is conserved in marsupials (Bagheri-Fam et al., 2009). However, in the tammar *SOX9* mRNA levels are sexually dimorphic 4 days before the peak in *SRY* expression (AJ Pask and MB Renfree, unpublished data). This could reflect the extended duration of SRY expression in the tammar compared to the 1 day of expression seen in mouse (Morais de Silva et al., 1996); as a result *SOX9* upregulation is switched on much earlier in tammar males. Alternatively, it could reflect a gradual increase to reach threshold levels to effect its action. In the tammar gonad, *SOX9* expression can be detected in pre-Sertoli cells and then in the mature Sertoli cells of the testicular cords throughout development, as in the mouse and human (Pask et al., 2002; Hanley et al., 2000; de Santa Barbara et al., 2000).

Tammar SOX9 is also detected in the developing Wolffian duct epithelium in the male mesonephros (Pask et al., 2002). This has not been previously described in other mammals and suggests that SOX9 has additional roles in the differentiation of the marsupial reproductive system.

DMRT1: *Doublesex* and *mab3* related transcript (DMRT1) is one of the most conserved genes in sexual differentiation, showing homology to both *doublesex* and *mab3* from *Drosophila melanogaster* and *Caenorhabditis elegans* respectively (Raymond et al., 1998), and is now known to be the long sought sex determining gene in birds (Smith et al., 2009). In humans *DMRT1* maps to 9p24.3, within a region whose deletion is know to cause human sex reversal. In the tammar, *DMRT1* is highly conserved at the genetic level and as in eutherian mammals, DMRT1 protein is detected in the germ cells and the Sertoli cells of the developing testis (Pask et al., 2003; Raymond et al., 1999a, 1999b). In the differentiated testis, DMRT1 was also detected in the Leydig cells, peri-tubular myoid cells and the acrosome of the sperm heads. In tammar females, DMRT1 occurs in the fetal and adult ovary in the pre-granulosa, granulosa and germ cells. This report of DMRT1 expression in the mature ovary prompted further investigations in mice where it was also present in the germ and granulosa cells (Pask et al., 2003). Marsupial investigations have therefore helped to uncover a wider role for this gene in mammals in both testicular and ovarian function.

WNT4: Wingless-type MMTV integration site family, member 4 (*WNT4*) is a key regulator of early embryogenesis and in the development of the gonadal in both humans and mice (Vainio et al., 1999). *Wnt4* knockout mice have masculinised gonads (Vainio et al., 1999) and endothelial and steroidogenic cells migrate into the developing ovary (Jeays-Ward et al., 2003). Over-expression of Wnt4 in the developing testis interferes with testicular vascular development and Sertoli cell differentiation, demonstrating that Wnt4 has specific and distinct roles in both male and female gonadal development (Jordan et al., 2003; Jeays-Ward et al., 2003).

The expression of WNT4 in the tammar gonads supports a dual role for this gene in testicular and ovary development. In the XX gonad, levels rise during early development, reaching a peak during the time the ovary is undergoing differentiation and then steadily declines. In males the level of *WNT4* decreases and is basal during the time the testis is differentiating and then steadily rises (Yu et al., 2006). Prior to the tammar wallaby study, WNT4 had not been immuno-localised in the developing gonads of any species. In the tammar, WNT4 protein localised to the Leydig cells of the testis (the site of testosterone production) consistent with a role in steroidogenesis. This suggests WNT4 acts to directly block male steroidogenesis in the developing ovary. Comparative analyses using tammar WNT4 was also used to define the core promoter region required for its transcription. Marsupials are particularly useful for promoter analysis in mammals and provide a powerful comparative model to mice and humans when looking for conserved and potentially functional elements. In addition the function of promoters can be examined in primary tammar cell lines and even in human cell lines using standard reporter assays (Yu et al., 2009). Using this approach, a highly conserved 89 bp minimal promoter region in both tammar and human *WNT4* was defined. Conserved transcription factor binding

sites in the proximal promoter were also determined, including sites for SP1, MyoD, NFκB and AP2. Furthermore, there is a highly conserved CpG island within the human, mouse and marsupial promoters, indicating that DNA methylation may play an important role in *WNT4* transcriptional regulation, explaining the absence of a traditional TATA box from this gene (Yu et al., 2009). Here, comparative tammar studies were able to provide new information on the transcriptional regulation of this essential mammalian developmental gene.

NR0B1: Nuclear receptor subfamily 0, group B, member 1, also known as DAX1; *Dosage sensitive sex reversal*, *adrenal hypoplasia congenita gene on the X-chromosome 1* gene maps to a region of the human X-chromosome that, when duplicated, causes male to female sex reversal (Zanaria et al., 1994). Nr0b1 appears to be crucial for testis differentiation by regulating the development of peritubular myoid cells and the formation of intact testis cords (Meeks et al., 2003). It was suggested to represent a vestige of an ancestral sex determining mechanism that was controlled by X-chromosome dosage and not the Y-chromosome or SRY. However, *NR0B1* resides in a region of the X-chromosome added recently during eutherian evolution. Cloning and mapping in the tammar confirmed that *NR0B1* resides on chromosome 5 alongside many other genes that are on the eutherian X-chromosome (Pask et al., 1997). This suggests that *NR0B1* was autosomal in the mammalian ancestor, and therefore cannot represent an ancestral sex-determining gene. Furthermore, testicular development can clearly occur normally in marsupials in the presence of two active copies of the gene, suggesting *NR0B1* only developed the ability to block normal testis development after its recruitment to the X-chromosome.

14.8.5 The Effect of Oestrogen on Sex Determination

In the lower vertebrates gonadal sex is primarily determined by the presence of absence of oestrogen. Aromatase is produced in the early female (but not the male) gonad causing the production of oestrogen, which drives the formation of the ovary. In the absence of aromatase (and therefore the absence of oestrogen production) a testis develops. The importance of oestrogen in non-mammlaian vertebrate sexual development was further demonstrated by experimental manipulations. Oestrogen administration to male embryos, or a block of oestrogen production in female embryos, results in sex reversal in almost all cases examined (Ramsey and Crews, 2009). In contrast, oestrogen is not needed for early sexual differentiation in mammals. Instead, sex is determined by the presence or absence of the *SRY* gene. In the presence of *SRY* a testis will develop, while ovarian development occurs in its absence. Marsupials are a useful tool to examine how or why mammals made a switch from primarily hormonal control of sex determination.

Mice missing one or both of the oestrogen receptors or the oestrogen synthesizing enzyme aromatase show normal early gonadal development (Couse and Korach, 1999). However, in adulthoood the gonads later fail in both male and female mice. The most severely affected is the aromatase knockout mouse. Shortly after birth,

germ cells are lost in the aromatase deficient mouse, and the stromal tissue is remodelled to form cord-like structures that show increased *SOX9* and *AMH* expression, indicative of Sertoli cell development (Britt et al., 2004). Exogenous oestrogen administered to these animals transdifferentiated the tissue back to an ovarian architecture, and suppressed *SOX9* levels (Britt et al., 2004). Thus, while oestrogen deprivation does not affect early sex determination in mice, oestrogen still plays an essential role in maintaining ovarian somatic cell fate.

Marsupials provide a unique developmental system in which to examine the effect of exogenous oestrogen during early development. As sexual differentiation in marsupials occurs after birth when the young are readily accessible in the mother's pouch, administration of carefully measured doses can be performed directly on the developing young (Renfree et al., 1995). Daily oral administration of oestradiol-17β to male pouch young can induce ovarian differentiation from an XY gonad and entry of XY germ cells into meiosis (Coveney et al., 2001; Renfree et al., 2001). A similar response is also seen in opossum young painted with estradiol dipropionate on the abdomen (Burns, 1955, 1961). This provides a novel model to examine the effects of early oestrogen administration of the molecular biology of the developing gonads. It is now clear that oestrogen can prevent male specific upregulation of *SRY*, and as a consequence *AMH* also fails to upregulate. Interestingly, *SOX9* mRNA levels are not affected by oestrogen but the protein localisation is critically affected. In the presence of oestrogen, *SOX9* is restricted to the cytoplasm of the somatic cells of the gonad and therefore could not initiate the male developmental program. In the absence of oestrogen, SOX9 rapidly becomes nuclear and initiates testis development (AJ Pask, N Calatayud, W Wood, G Shaw and MB Renfree, unpublished results).

Thus these sex reversal studies have revealed many similarities with mice and humans, but also some important differences that can help us better define the precise function and role of genes in mammalian sexual differentiation pathways.

14.8.6 Androgens and Virilization

A great deal is now known about the endocrine control of sexual differentiation in the postnatal stages of marsupial development (Wilson et al., 2003). Again, the tammar wallaby is perfectly suited for studies of virilisation, since the entire process occurs postnatally, unlike most eutherian species that show clear virilisation at birth. Each of the key structures and processes – the formation of the Wolffian ducts, the prostate, the phallus and testicular descent – all occur at different post-natal ages, making it possible to examine the hormonal and genetic control of each of these organs separately (Renfree et al., 1995).

The tammar androgen receptor protein is similar to that of other mammals. It is expressed in the fetal gonad and brain as early as Day 19 of the 26.5-day gestation, 7 days earlier than the first rise in testicular testosterone (Butler et al., 1998). Immunoreactive AR is detected in all the developing male tissues except the scrotum, and in the female urogenital sinus tissues. Similarly, AR is present in

the vas deferens, gubernaculum, testis, inguinal and scrotal areas, in the urogenital sinus and in the genital tubercle of the grey short tailed opossum (Sonea et al., 1997), although Russell et al. (2003) report that androgen receptor immunoreactivity can be detected in the scrotal anlagen of male opossum pups as early as 1 day following birth.

However, despite the importance of AR, it is not the initiating signal for virilization of the marsupial male (Butler et al., 1998) and the signal does not appear to be testosterone. This enigma was solved by a detailed study of androgen control of virilisation in the tammar which uncovered a new, alternate pathway of androgen biosynthesis that forms dihydrotestosterone (DHT), and in which the potent androgen androstanediol (adiol) is the key circulating androgen in the developing young instead of testosterone (Wilson et al., 2003). Androstanediol is back-converted to DHT in the target tissues, and then binds to the androgen receptor (Shaw et al., 2000). This is the first new pathway of steroid hormone biosynthesis described in the last 50 years. This alternate pathway also functions in the mouse (Mahendroo et al., 2004; Wilson et al., 2005) and explains some rare cases of virilization seen in girls with P450 oxidoreductase deficiency and congenital adrenal hyperplasia (Arlt et al., 2004; Homma et al., 2006). Investigations are now underway to characterize the expression profiles of androgen biosynthesis enzyme production during these critical stages of virilisation during development of the tammar pouch young.

14.9 Conclusions

Marsupials form a unique class of mammals that are especially valuable for studies of mammalian developmental biology. Their unique reproductive strategy provides access to young at developmental stages that occur entirely in utero in eutherian mammals. By exploiting this unique situation, we and others have provided new information about the processes of sex determination and sexual differentiation in all mammals. Such investigations have helped identify the sex-determining gene *SRY*, uncover novel roles for genes in mammalian sexual development, determine the ancestry of key developmental genes and shed further light on pathways of differentiation in humans. The evolutionary position of marsupials makes comparisons of gene and promoter structure exceptionally powerful and they will continue to play a critical role in defining key regulatory elements for mammalian gene function.

References

Alcorn GT, Robinson ES (1983) Germ cell development in female pouch young of the tammar wallaby (*Macropus eugenii*). J Reprod Fertil 67:319–325.

Ambrosetti DC, Basilico C, Dailey L (1997) Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. Mol Cell Biol 17:6321–6329.

Anderson R, Copeland TK, Scholer H, Heasman J, Wylie C (2000) The onset of germ cell migration in the mouse embryo. Mech Dev 91:61–68.

Arlt W, Walker EA, Draper N, et al. (2004) Congenital adrenal hyperplasia caused by mutant P450 oxidoreductase and human androgen synthesis: analytical study. Lancet 363: 2128–2135.

Arnold AP (2009) The organizational-activational hypothesis as the foundation for a unified theory of sexual differentiation of all mammalian tissues. Horm Behav 55:570–578.

Au PC, Whitley J, Vaux D, Selwood L, Familari M (2008) Identification of novel and known ovary-specific genes including ZP2, in a marsupial, the stripe-faced dunnart. Mol Reprod Dev 75:318–325.

Bagheri-Fam S, Sinclair AH, Koopman P, Harley VR (2009) Conserved regulatory modules in the Sox9 testis-specific enhancer predict roles for SOX TCF/LEF Forkhead DMRT and GATA proteins in vertebrate sex determination. Int J Biochem Cell Biol 42:472–477.

Bedford JM, Breed WG (1994) Regulated storage and subsequent transformation of spermatozoa in the fallopian tubes of an Australian marsupial Sminthopsis crassicaudata. Biol Reprod 50:845–854.

Belov K, Harrison GA, Cooper DW (1998) Cloning of the red kangaroo (Macropus rufus) follicle stimulating hormone beta subunit. Reprod Fertil Dev 10:289–291.

Bowles J, Knight D, Smith C, et al. (2006) Retinoid signaling determines germ cell fate in mice. Science 312:596–600.

Braundmeier AG, Breed WG, Miller DJ (2008) Spermatozoa from a marsupial the brushtail possum contain beta14-galactosyltransferase. Reprod Fertil Dev 20:402–407.

Britt KL, Stanton PG, Misso M, Simpson ER, Findlay JK (2004) The effects of estrogen on the expression of genes underlying the differentiation of somatic cells in the murine gonad. Endocrinology 145:3950–3960.

Brown PR, Miki K, Harper DB, Eddy EM (2003) A-kinase anchoring protein 4 binding proteins in the fibrous sheath of the sperm flagellum. Biology Reprod 68:2241–2248.

Burgoyne PS, Thornhill AR, Boudrean SK, Darling SM, Bishop CE, Evans EP (1995) The genetic basis of XX–XY differences present before gonadal sex differentiation in the mouse Philos. Trans R Soc Lond B Biol Sci 350:253–260.

Burns RK (1955) Experimental reversal of sex in the gonads of the opossum Didelphis virginiana. PNAS 41:669–676.

Burns RK (1961) Role of hormones in the differentiation of sex. In: Young WC (eds) Sex and Internal Secretions, Vol 1. Williams and Wilkins, Baltimore, pp. 76–158.

Butler CM, Harry JL, Deakin JE, Cooper DW, Renfree MB (1998) Developmental expression of the androgen receptor during virilization of the urogenital system of a marsupial. Biol Reprod 59:725–732.

Calatayud NE, Pask AJ, Shaw G, Renfree MB (2010) Oestrogen receptors ERα and ERb in the tammar wallaby. Reproduction 139:599–611.

Carrera A, Gerton GL, Moss SB (1994) The major fibrous sheath polypeptide of mouse sperm: structural and functional similarities to the A-kinase anchoring proteins. Dev Biol 165: 272–284.

Carrera A, Moos J, Ning XP, et al. (1996) Regulation of protein tyrosine phosphorylation in human sperm by a calcium/calmodulin-dependent mechanism: identification of A kinase anchor proteins as major substrates for tyrosine phosphorylation. Dev Biol 180:284–296.

Chambers I, Colby D, Robertson M, et al. (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113:643–655.

Chawengsaksophak K, de Graaff W, Rossant J, Deschamps J, Beck F (2004) Cdx2 is essential for axial elongation in mouse development. Proc Natl Acad Sci USA 101:7641–7645.

Cheung TC, Hearn JP (2002) Molecular cloning and tissue expression of the gonadotrophin-releasing hormone receptor in the tammar wallaby (*Macropus eugenii*). Reprod Fertil Dev 14:157–164.

Cheung TC, Hearn JP (2003) Developmental expression and subcellular localization of wallaby gonadotropin-releasing hormone receptor and its splice variants. Gen Comp Endocrinol 133:88–99.

Cheung TC, Hearn JP (2005) Dimerizations of the wallaby gonadotropin-releasing hormone receptor and its splice variants. Gen Comp Endocrinol 144:280–288.

Colledge M, Scott JD (1999) AKAPs: from structure to function. Trends Cell Biol 9:216–221.

Cooper DW (1993) The evolution of sex determination sex chromosome dimorphism and X-inactivation in Therian mammals: a comparison of metatherians (marsupials) and eutherians ('placentals'). In: Reed KC Graves JAM (eds) Sex Chromosomes and Sex Determining Genes. Harwood Academic Publishers, GmbH Chur, Switzerland.

Couse JF, Korach KS (1999) Estrogen receptor null mice: what have we learned and where will they lead us? Endocr Rev 20:358–417.

Coveney D, Shaw G, Renfree MB (2001) Estrogen-induced gonadal sex reversal in the tammar wallaby. Biol Reprod 65:613–621.

Crawford JL, Heath DA, Haydon LJ, Thomson BP, Eckery DC (2009) Gene expression and secretion of LH and FSH in relation to gene expression of GnRH receptors in the brushtail possum (*Trichosurus vulpecula*) demonstrates highly conserved mechanisms. Reproduction 137:129–140.

Cui S, Selwood L (2000) cDNA cloning characterization expression and recombinant protein production of leukemia inhibitory factor (LIF) from the marsupial the brushtail possum (*Trichosurus vulpecula*). Gene 243:167–178.

Cui S, Selwood L (2003) Cloning and expression of a novel cDNA encoding shell coat protein cp4 from the brushtail possum (*Trichosurus vulpecula*). Mol Reprod Dev 65:141–147.

Cui S, Nikolovski S, Nanayakkara K, Selwood L (2005) VAP1 with cystatin C motif an oocyte protein encoded by a novel ovarian-specific gene during oogenesis in the common brushtail possum (Trichosurus vulpecula). Mol Reprod Dev 71:19–28.

Cummins JM (1980) Decondensation of sperm nuclei of Australian marsupials: effects of air drying and of calcium and magnesium. Gamete Res 3:351–367.

Curlewis JD, Saunders MC, Kuang J, Harrison GA, Cooper DW (1998) Cloning and sequence analysis of a pituitary prolactin cDNA from the brushtail possum (Trichosurus vulpecula). Gen Comp Endocrinol 111:61–67.

de Santa Barbara P, Moniot B, Poulat F, Berta P (2000) Expression and subcellular localisation of SF1 SOX9 WT1 and AMH proteins during early human testicular development. Dev Dyn 217:293–298.

Denker HW, Tyndale-Biscoe CH (1986) Embryo implantation and proteinase activities in a marsupial (*Macropus eugenii*). Histochemical patterns of proteinases in various gestational stages. Cell Tissue Res 246:279–291.

Eckery DC, Lawrence SB, Juengel JL, Greenwood P, McNatty KP, Fidler AE (2002a) Gene expression of the tyrosine kinase receptor c-kit during ovarian development in the brushtail possum (*Trichosurus vulpecula*). Biol Reprod 66:346–353.

Eckery DC, Lun S, Thomson BP, Chie WN, Moore LG, Juengel JL (2002b) Ovarian expression of messenger RNA encoding the receptors for luteinizing hormone and follicle-stimulating hormone in a marsupial the brushtail possum (Trichosurus vulpecula). Biol Reprod 66: 1310–1317.

Eckery DC, Whale LJ, Lawrence SB, Wylde KA, McNatty KP, Juengel JL (2002c) Expression of mRNA encoding growth differentiation factor 9 and bone morphogenetic protein 15 during follicular formation and growth in a marsupial the brushtail possum (Trichosurus vulpecula). Mol Cell Endocrinol 192:115–126.

Eddy EM, Toshimori K, O'Brien DA (2003) Fibrous sheath of mammalian spermatozoa. Microsc Res Tech 61:103–115.

Edwards AS, Scott JD (2000) A-kinase anchoring proteins: protein kinase A and beyond. Curr Opin Cell Biol 12:217–221.

Fidler AE, Lawrence SB, Vanmontfort DM, Tisdall DJ, McNatty KP (1998) The Australian brushtail possum (*Trichosurus vulpecula*) gonadotrophin alpha-subunit: analysis of cDNA sequence and pattern of expression. J Mol Endocrinol 20:345–353.

Fletcher TP, Renfree MB (1988) Effects of corpus luteum removal on progesterone estradiol-17 beta and LH in early pregnancy of the tammar wallaby Macropus eugenii. J Reprod Fertil 83:185–191.

Foster JW, Brennan FE, Hampikian GK, Goodfellow PN, Sinclair AH, Lovell-Badge R, Selwood L, Renfree MB, Cooper DW, Graves JA (1992) Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials. Nature 359:531–533.

Foster JW, Graves JAM (1994) An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene. Proc Natl Acad Sci USA 91:1927–1931.

Frankenberg S, Tisdall D, Selwood L (2001) Identification of a homologue of POU5F1 (OCT3/4) in a marsupial the brushtail possum. Mol Reprod Dev 58:255–261.

Frankenberg S, Pask AJ, Renfree MB (2010) The evolution of class V POU domain transcription factors in vertebrates and their characterisation in a marsupial Developmental Biology 337:162–170.

Freyer C, Renfree MB (2009) The mammalian yolk sac placenta. J Exp Zool B Mol Dev Evol 312:545–554.

Freyer C, Zeller U, Renfree MB (2003) The marsupial placenta: a phylogenetic analysis. J Exp Zool A Comp Exp Biol 299:59–77.

Fulcher KD, Mori C, Welch JE, O'Brien DA, Klapper DG, Eddy EM (1995) Characterization of Fsc1 cDNA for a mouse sperm fibrous sheath component. Biol Reprod 52:41–49.

Hamatani T, Daikoku T, Wang H, et al. (2004) Global gene expression analysis identifies molecular pathways distinguishing blastocyst dormancy and activation. Proc Natl Acad Sci USA 101:10326–10331.

Hanley NA, Hagan DM, Clement-Jones M, et al. (2000) SRY SOX9 and DAX1 expression patterns during human sex determination and gonadal development. Mech Dev 91: 403–407.

Harris MS, Rodger JC (2005) Characterisation of an epitope shared by an acrosomal acrosin-like protein and the surface of tammar wallaby (*Macropus eugenii*) spermatozoa. J Exp Zool A Comp Exp Biol 303:713–721.

Harrison GA, Deane EM, Cooper DW (1998) cDNA cloning of luteinizing hormone subunits from brushtail possum and red kangaroo. Mamm Genome 9:638–642.

Harry JL, Koopman P, Brennan FE, Graves JAM, Renfree MB (1995) Widespread expression of the testis determining gene SRY in a marsupial. Nat Genet 11:347–349.

Haydon LJ, Juengel JL, Thomson BP, Eckery DC (2008) Expression of mRNAs encoding oestrogen receptor (ER) alpha and ERbeta, androgen receptor and progesterone receptor during gonadal and follicular development in the marsupial brushtail possum (*Trichosurus vulpecula*). Fertil Dev 20:335–349.

Heap RB, Renfree MB, Burton RD (1980) Steroid metabolism in the yolk sac placenta and endometrium of the tammar wallaby, *Macropus eugenii*. J Endocrinol 87:339–349.

Homma K, Hasegawa T, Nagai T, et al. (2006) Urine steroid hormone profile analysis in cytochrome P450 oxidoreductase deficiency: implication for the backdoor pathway to dihydrotestosterone. J Clin Endocrinol Metab 91:2643–2649.

Hondo E, Stewart CL (2005) Profiling gene expression in growth-arrested mouse embryos in diapause. Genome Biol 6:202.

Hu Y, Yu H, Pask AJ, O'Brien DA, Shaw G, Renfree MB (2009) A-kinase anchoring protein 4 has a conserved role in mammalian spermatogenesis. Reproduction 137:645–653.

Hughes RL, Hall LS (1988) Structural adaptations of the newborn marsupial. In: Tyndale-Biscoe CH Janssens PA (eds) The Developing Marsupial. Springer, Berlin, pp. 8–27.

Hutson JM, Shaw G, O WS, Short RV, Renfree MB (1988) The ontogeny of Mullerian inhibiting substance production and testicular differentiation migration and descent in the pouch young of a marsupial. Development 104:549–556.

Ingram JN, Shaw G, Renfree MB (1999) Cortisol in fetal fluids and the fetal adrenal at parturition in the tammar wallaby (*Macropus eugenii*). Biol Reprod 60:651–655.

Jeays-Ward K, Hoyle C, Brennan J, et al. (2003) Endothelial and steroidogenic cell migration are regulated by WNT4 in the developing mammalian gonad. Development 130:3663–3670.

Jiang FX, Clark J, Renfree MB (1997) Ultrastructural characteristics of primordial germ cells and their amoeboid movement to the gonadal ridges in the tammar wallaby. Anat Embryol 195:473–481.

Jordan BK, Shen JH, Olaso R, Ingraham HA, Vilain E (2003) Wnt4 over-expression disrupts normal testicular vasculature and inhibits testosterone synthesis by repressing steroidogenic factor 1/ beta-catenin synergy. Proc Natl Acad Sci USA 100:10866–10871.

Jost A (1947) Recherches sur la differenciation sexuelle de l'embryon de lapin. J Arch Anat Microsc Morphol Exp 36:117–121.

Jost A (1970) Hormonal factors in the sex differentiation of the mammalian foetus. Philos Trans R Soc Lond B Biol Sci 259:119–130.

Juengel JL, Whale LJ, Wylde KA, Greenwood P, McNatty KP, Eckery DC (2002) Expression of anti-Müllerian hormone mRNA during gonadal and follicular development in the brushtail possum (Trichosurus vulpecula). Reprod Fertil Dev 14:345–353.

Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD (2004) The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. Nat Genet 36:642–646.

Koopman P, Münsterberg A, Capel B, Vivian N, Lovell-Badge R (1990) Expression of a candidate sex-determining gene during mouse testis differentiation. Nature 348:450–452.

Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R (1991) Male development of chromosomal female mice transgenic for SRY. Nature 351:117–121.

Koubova J, Menke DB, Zhou Q, Capel B, Griswold MD, Page DC (2006) Retinoic acid regulates sex-specific timing of meiotic initiation in mice. Proc Natl Acad Sci USA 103:2474–2479.

Kunath T, Strumpf D, Rossant J (2004) Early trophoblast determination and stem cell maintenance in the mouse–a review. Placenta 25:S32–S38.

Kwek JHL, De Iongh R, Digby MR, Renfree MB, Nicholas KR, Familari M (2009) Cross-fostering of the tammar wallaby (*Macropus eugenii*) pouch young accelerates fore-stomach maturation. MOD 126:449–463.

Leatherland JF, Renfree MB (1983) Structure of the pars distalis in pouch-young tammar wallabies (*Macropus eugenii*). Cell Tissue Res 230:587–603.

Lin M, Harman A, Rodger JC (1997) Spermiogenesis and spermiation in a marsupial the tammar wallaby (*Macropus eugenii*). J Anat 190:377–395.

Lin M, Rodger JC (1999) Acrosome formation during sperm transit through the epididymis in two marsupials the tammar wallaby (*Macropus eugenii*) and the brushtail possum (*Trichosurus vulpecu*la). J Anat 194:223–232.

Liu L, Leaman D, Villalta M, Roberts RM (1997) Silencing of the gene for the alpha-subunit of human chorionic gonadotropin by the embryonic transcription factor Oct-3/4. Mol Endocrinol 11:1651–1658.

Lumbers ER, Menzies RI, Gibson KJ, Tyndale-Biscoe CH (1994) Levels of active and inactive reninlike enzymes in plasma and reproductive tract of the pregnant wallaby. Am J Physiol 266:R1353–R1358.

Mahendroo M, Wilson JD, Richardson JA, Auchus RJ (2004) Steroid 5alpha-reductase 1 promotes 5alpha-androstane-3alpha17beta-diol synthesis in immature mouse testes by two pathways. Mol Cell Endocrinol 222:113–120.

Mate KE, McCartney CA (1998) Sequence and analysis of zona pellucida 2 cDNA (ZP2) from a marsupial the brushtail possum *Trichosurus vulpecula*. Mol Reprod Dev 51:322–329.

Mate KE, Rodger JC (1991) Stability of the acrosome of the brush-tailed possum (*Trichosurus vulpecula*) and tammar wallaby (*Macropus eugenii*) in vitro and after exposure to conditions and agents known to cause capacitation or acrosome reaction of eutherian spermatozoa. J Reprod Fertil 91:41–48.

Mate KE, Buist JM, Duckworth JA (2003) Expression in Escherichia coli and immunological characterization of three zona pellucida proteins (ZP1 ZP2 and ZP3) from a marsupial the brushtail possum (*Trichosurus vulpecula*). Mol Reprod Dev 64:136–143.

Matsuzawa-Watanabe Y, Inoue J, Semba K (2003) Transcriptional activity of testis-determining factor SRY is modulated by the Wilms' tumor 1 gene product WT1. Oncogene 22:7900–7904.

Mattiske D, Pask AJ, Shaw JM, Shaw G (2002) Structure and expression of the follicle-stimulating hormone receptor gene in a marsupial Macropus eugenii. Mol Reprod Dev 63:24–31.

McCartney CA, Mate KE (1999) Cloning and characterisation of a zona pellucida 3 cDNA from a marsupial the brushtail possum Trichosurus vulpecular. Zygote 7:1–9.
McLaren A (1981) The fate of germ cells in the testis of fetal Sex-reversed mice. J Reprod Fertil 61:461–467.
McLaren A (2003) Primordial germ cells in mice. Dev Biol 262:1–15.
Meeks JJ, Crawford SE, Russell TA, Morohashi K, Weiss J, Jameson JL (2003) Dax1 regulates testis cord organization during gonadal differentiation. Development 130:1029–1036.
Menkhorst E, Selwood L, Cui S (2009) Uterine expression of cp4 gene homolog in the Stripe-faced Dunnart Sminthopsis macroura: relationship with conceptus development and progesterone profile. Mol Reprod Dev 76:863–872.
Menzies BR, Shaw G, Fletcher TP, Renfree MB (2007) Perturbed growth and development in marsupial young after reciprocal cross-fostering between species. Reprod Fertil Dev 19: 976–983.
Menzies BR, Shaw G, Fletcher TP, Pask AJ, Renfree MB (2008) Exon 3 of the growth hormone receptor (GH-R) is specific to eutherian mammals. Mol Cell Endocrinol 296:64–68.
Menzies BR, Shaw G, Fletcher TP, Renfree MB (2009) Early onset of ghrelin production in a marsupial. Mol Cell Endocrinol 299:266–273.
Miki K, Willis WD, Brown PR, Goulding EH, Fulcher KD, Eddy EM (2002) Targeted disruption of the Akap4 gene causes defects in sperm flagellum and motility. Dev Biol 248:331–342.
Mitsui K, Tokuzawa Y, Itoh H, et al. (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 113:631–642.
Morais da Silva S, Hacker A, Harley V, et al. (1996) SOX9 expression during gonadal development implies a conserved role for the gene in Sertoli cell differentiation in mammals and birds. Nat Genet 14:62–68.
Nation A, Cui S, Selwood L (2008) Vesicle-associated protein 1: a novel ovarian immuno-contraceptive target in the common brushtail possum *Trichosurus vulpecula*. Reproduction 136:657–665.
Natoli TA, Alberta JA, Bortvin A, et al. (2004) Wt1 functions in the development of germ cells in addition to somatic cell lineages of the testis. Dev Biol 268:429–440.
O WS, Short RV, Renfree MB, Shaw G (1988) Primary genetic control of somatic sexual differentiation in a mammal. Nature 332:716–717.
Ohno S (1967) Sex Chromosomes and Sex Linked Genes. Springer, Berlin.
Oreal E, Pieau C, Mattei M, Picard J, Carre-eusebe D, Magre S (1998) MIS in chicken embryonic gonads precedes testicular SOX9 expression. Dev Dyn 21:522–532.
Parker KL (1998) The roles of steroidogenic factor 1 in endocrine development and function. Mol Cell Endocrinol 140:59–63.
Pask A, Renfree MB, Graves JAM (2000a) The human sex reversing ATRX gene has a homologue on the marsupial Y chromosome ATRY: implications for the evolution of mammalian sex determination. Proc Natl Acad Sci USA 97:13198–13202.
Pask A, Harry JL, Renfree MB, Graves JAM (2000b) Genesis 27:145–152.
Pask A, Toder R, Wilcox SA, Camerino G, Graves JAM (1997) The candidate sex-reversing DAX1 gene is autosomal in marsupials: implications for the evolution of sex determination in mammals. Genomics 41:422–426.
Pask AJ, Harry JL, Graves JAM, et al. (2002) SOX9 has both conserved and novel roles in marsupial sexual differentiation. Genesis 33:131–139.
Pask AJ, Behringer RR, Renfree MB (2003) Expression of DMRT1 in the mammalian ovary and testis – from marsupials to mice. Cytogenet Genome Res 101:229–236.
Pask AJ, Whitworth DJ, Mao C-A, et al. (2004) Marsupial anti-Müllerian hormone gene structure regulatory elements and expression. Biol Reprod 70:160–167.
Pask AJ, Paplinska JZ, Shaw G, Graves JAM, Renfree MB (2007) Marsupial WT1 has a novel isoform and is expressed in both somatic and germ cells in the developing ovary and testis. Sex Dev 1:169–180.

Ramsey M, Crews D (2009) Steroid signaling and temperature-dependent sex determination-Reviewing the evidence for early action of estrogen during ovarian determination in turtles. Semin Cell Dev Biol 20:283–292.

Raymond CS, Shamu CE, Shen MM, et al. (1998) Evidence for evolutionary conservation of sex-determining genes. Nature 391:691–695.

Raymond CS, Parker ED, Kettlewell JR (1999a) A region of human chromosome 9p required for testis development contains two genes related to known sexual regulators. Hum molec Genet 8:989–996.

Raymond CS, Kettlewell JR, Hirsch B, Bardwell VJ, Zarkower D (1999b) Expression of Dmrt1 in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. Dev Biol 215:208–220.

Renfree MB (1973) The composition of fetal fluids of the marsupial *Macropus eugenii*. Dev Biol 33:62–79.

Renfree MB (1981) Marsupials: alternative mammals. Nature 293:100–101.

Renfree MB (1983) Marsupial reproduction: the choice between placentation and lactation. In: Finn CA (ed) Oxford Reviews of Reproductive Biology, Vol 5. Oxford University Press, New York, pp. 1–29.

Renfree MB (1992) The role of genes and hormones in marsupial sexual differentiation. J Zool Lond 226:165–173.

Renfree MB (1993) Ontogeny genetic control and phylogeny of female reproduction in monotreme and therian mammals. In: Szalay FS Novacek JJ McKenna MC (eds) Mammal Phylogeny. Springer, New York, pp. 4–20.

Renfree MB, Lewis AM (1996) Cleavage in vivo and in vitro in the marsupial *Macropus eugenii*. Reprod Fertil Dev 8:725–742.

Renfree MB, Short RV (1988) Sex determination in marsupials: evidence for a marsupial-eutherian dichotomy. Phil Trans R Soc Lond B 322:41–54.

Renfree MB, Tyndale-Biscoe CH (1973) Intrauterine development after diapause in the marsupial Macropus eugenii. Devl Biol 32:28–40.

Renfree MB, Tyndale-Biscoe CH (1978) Manipulation of marsupial embryos and pouch young. In: Daniel JC (ed) Methods of Mammalian Reproduction, Academic Press, pp. 307–331.

Renfree MB, Flint AP, Green SW, Heap RB (1984) Ovarian steroid metabolism and oestrogens in the corpus luteum of the tammar wallaby. J Endocrinol 101:231–240.

Renfree MB, Harry JL, Shaw G (1995) The marsupial male: a role model for sexual development. Phil Trans R Soc Lond B 350:243–251.

Renfree MB, O WS, Short RV, Shaw G (1996) Sexual differentiation of the urogenital system of the fetal and neonatal tammar wallaby *Macropus eugenii*. Anat Embryol 194: 111–134.

Renfree MB, Coveney D, Shaw G (2001) The influence of estrogen on the developing male marsupial. Reprod Fertil Dev 13:231–240.

Ricci M, Breed WG (2005) Morphogenesis of the fibrous sheath in the marsupial spermatozoon. J Anat 207:155–164.

Russell AJ, Gilmore DP, Mackay S, Ullmann SL, Baker PJ, Payne AP (2003) The role of androgens in development of the scrotum of the grey short-tailed Brazilian opossum (*Monodelphis domestica*). Anat Embryol (Berl) 206:381–389.

Schöler HR, Dressler GR, Balling R, Rohdewohld R, Gruss P (1990) Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. EMBO J 9:2185–2195.

Sekido R, Lovell-Badge R (2008) Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. Nature 453:930–9344.

Selwood L (1986) Cleavage in vitro following destruction of some blastomeres in the marsupial *Antechinus stuartii* (Macleay). J Embryol Exp Morphol 92:71–84.

Selwood L (1992) Mechanisms underlying the development of pattern in marsupial embryos. Curr Top Dev Biol 27:175–233.

Selwood L (2000) Mechanisms for pattern formation leading to axis formation and lineage allocation in mammals: a marsupial perspective. Reproduction 121:677–683.

Selwood L, Robinson ES, Pedersen RA, Vandeberg JL (1997) Development in vitro of Marsupials: a comparative review of species and a timetable of cleavage and early blastocyst stages of development in Monodelphis domestica. Int J Dev Biol 41:397–410.

Setchell BBP (1977) Reproduction in male marsupials. In: Stonehouse B Gilmorepp D (eds) The Biology of Marsupials. Macmillan, London, pp. 411–457.

Sharman GB, Robinson ES, Walton SM, Berger RJ (1970) Sex chromosomes and reproductive anatomy of some intersexual marsupials. J Reprod Fertil 21:57–68.

Sharman GB, Hughes RL, Cooper DW (1990) The chromosomal basis of sex differentiation in marsupials. Aust J Zool 37:443–450.

Shaw G, Renfree MB, Short RV, O WS (1988) Experimental manipulation of sexual differentiation in wallaby pouch young treated with exogenous steroids. Development 104:689–701.

Shaw G, Renfree MB, Short RV (1990) Primary genetic control of sexual differentiation in marsupials. Aust J Zool 37:443–450.

Shaw G, Gehring HM, Bell EC (1999) Production of prostaglandin f2alpha and its metabolite by endometrium and yolk sac placenta in late gestation in the tammar wallaby Macropus Eugenii. Biol Reprod 60:611–614.

Shaw G, Renfree MB, Leihy MW, Shackleton CHL, Roitman E, Wilson JD (2000) Prostate formation in a marsupial is mediated by the testicular androgen 5α-androstane-3α, 17β-diol. Proc Natl Acad Sci USA 97:12256–12259.

Short RV (2003) The magic and mystery of the oocyte: ex ovo omnia. In: Trounson AO Gosden RG (eds) The Biology and Pathology of the Oocyte. Cambridge University Press, New York.

Sidhu KS, Mate KE, Molinia FC, Rodger JC (1999) Induction of thumbtack sperm during coculture with oviduct epithelial cell monolayers in a marsupial the brushtail possum (*Trichosurus vulpecula*). Biol Reprod 61:1356–1361.

Sidhu KS, Mate KE, Gunasekera T, et al. (2004) A flow cytometric assay for global estimation of tyrosine phosphorylation associated with capacitation of spermatozoa from two marsupial species the tammar wallaby (*Macropus eugenii*) and the brushtail possum (*Trichosurus vulpecula*). Reproduction 127:95–103.

Sinclair AH, Berta P, Palmer MS, et al. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA binding motif. Nature 346:240–244.

Sistina Y, Lin M, Mate KE, Robinson ES, Rodger JC (1993) The unique stability of the marsupial sperm acrosomal membranes examined by unprotected freeze-thawing and treatment with the detergent Triton X-100. Reprod Fertil Dev 5:1–14.

Smith CA, Roeszler KN, Ohnesorg T, et al. (2009) The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature 461:267–271.

Smith KK (2006) Craniofacial development in marsupial mammals: developmental origins of evolutionary change. Dev Dyn 235:1181–1193.

Sonea IM, Iqbal J, Prins GS, Jacobson CD (1997) Ontogeny of androgen receptor-like immunoreactivity in the reproductive tract of male Monodelphis domestica. Biol Reprod 56: 852–860.

Spotila L, Spotila JR, Hall SE (1998) Sequence and expression analysis of WT-1 and SOX9 in the red-eared slider turtle Trachemys scripta. J Exp Zool 281:417–427.

Takamatsu N, Kanda H, Ito M, Yamashita A, Yamashita S, Shiba T (1997) Rainbow trout SOX9: cDNA cloning gene structure and expression. Gene 202:167–170.

Tanaka SS, Yamaguchi YL, Tsoi B, Lickert H, Tam PP (2005) IFITM/Mil/fragilis family proteins IFITM1 and IFITM3 play distinct roles in mouse primordial germ cell homing and repulsion. Dev Cell 9:745–756.

Temple-Smith PD (1994) Comparative structure and function of marsupial spermatozoa. Reprod Fertil Dev 6:421–435.

Temple-Smith PD, Bedford JM (1976) The features of sperm maturation in the epididymis of a marsupial the brushtailed possum *Trichosurus vulpecula*. Am J Anat 147:471–499.

Trott JF, Simpson KJ, Moyle RL, et al. (2003) Maternal regulation of milk composition milk production and pouch young development during lactation in the tammar wallaby (*Macropus eugenii*). Biol Reprod 68:929–936.

Turner RM, Johnson LR, Haig-Ladewig L, Gerton GL, Moss SB (1998) An X-linked gene encodes a major human sperm fibrous sheath protein hAKAP82 Genomic organization protein kinase A-RII binding and distribution of the precursor in the sperm tail. J Biol Chem 273: 32135–32141.

Tyndale-Biscoe CH, Janssens PA (eds) (1988) The Developing Marsupial. Models for Biomedical Research. Springer, Berlin.

Tyndale-Biscoe H, Renfree M (1987) Reproductive Physiology of Marsupials. Cambridge University Press, Cambridge, 476 pp.

Ullmann SL, Shaw G, Alcorn GT, Renfree MB (1997) Migration of primordial germ cells to the developing gonadal ridges in the tammar wallaby *Macropus eugenii*. J Reprod Fertil 110: 135–143.

Vainio S, Heikkila M, Kispert A, Chin N, McMahon AP (1999) Female development in mammals is regulated by Wnt-4 signalling. Nature 397:405–409.

Vanmontfort D, Fidler AE, Heath DA, et al. (1998) cDNA sequence analysis gene expression and protein localisation of the inhibin alpha-subunit of Australian brushtail possum (*Trichosurus vulpecula*). J Mol Endocrinol 21:141–152.

Vernet N, Dennefeld C, Rochette-Egly C, et al. (2006) Retinoic acid metabolism and signaling pathways in the adult and developing mouse testis. Endocrinology 147:96–110.

Weiss J, Meeks JJ, Hurley L, Raverot G, Frassetto A, Jameson JL (2003) Sox3 is required for gonadal function but not sex determination in males and females. Mol Cell Biol 23:8084–8091.

Western PS, Harry JL, Graves JAM, Sinclair AH (1999) Temperature dependent sex determination: upregulation of SOX9 expression after commitment to male development. Dev Dyn 214: 171–177.

Whale LJ, Eckery DC, Juengel JL (2003) Determination of steroidogenic potential of ovarian cells of the brushtail possum (*Trichosurus vulpecula*). Biol Reprod 69:947–958.

Whitworth DJ, Shaw G, Renfree MB (1996) Gonadal sex reversal of the developing marsupial ovary in vivo and in vitro. Development 122:4057–4063.

Whitworth DJ, Shaw G, Renfree MB (1997) Müllerian duct regression in a marsupial the tammar wallaby. Anat Embryol 196:39–46.

Wilson JD, Leihy MW, Shaw G, Renfree MB (2003) Unsolved problems in male physiology: studies in a marsupial. Mol Cell Endocrinol 211:33–36.

Whitworth DJ, Pask AJ, Shaw G, Marshall Graves JA, Behringer RR, Renfree MB (2001) Characterization of steroidogenic factor 1 during sexual differentiation in a marsupial. Gene 277:209–219.

Wilson JD, Shaw G, Renfree MB, Auchus RJ, Leihy MW, Eckery DC (2005) Ontogeny and pathway of formation of 5alpha-androstane-3alpha, 17beta-diol in the testes of the immature brushtail possum *Trichosurus vulpecula*. Reprod Fertil Dev 17:603–609.

Yu H, Pask AJ, Shaw G, Renfree MB (2006) Differential expression of WNT4 in testicular and ovarian development in a marsupial. BMC Dev Biol 6:44.

Yu H, Pask AJ, Shaw G, Renfree MB (2009) Comparative analysis of the mammalian WNT4 promoter. BMC Genomics 10:416.

Zanaria E, Muscatelli F, Bardoni B, et al. (1994) An unusual member of the nuclear hormone receptor super family responsible for X-linked adrenal hypoplasia congenita. Nature 372: 635–641.

Chapter 15
Marsupial Milk – Identifying Signals for Regulating Mammary Function and Development of the Young

Julie A. Sharp, Christophe Lefevre, Joly Kwek, Denijal Topcic, Laurine Buscara, Cate Pooley, Robb de Iongh, Mary Familari, and Kevin R. Nicholas

Abstract The role of milk in providing nutrition for the young is well established. However, it is becoming apparent that milk has a more comprehensive role in programming and regulating growth and development of the suckled young, and an autocrine impact on the mammary gland so that it functions appropriately during the lactation cycle. This central role of milk is best studied in animal models, such as marsupials that have evolved a different lactation strategy to eutherians and allow researchers to more easily identify regulatory mechanisms that are not as readily apparent in eutherian species. For example, the tammar wallaby (*Macropus eugenii*) has evolved with a unique reproductive strategy of a short gestation, birth of an altricial young and a relatively long lactation during which the mother progressively changes the composition of the major, and many of the minor components of milk. Thus, in contrast to eutherians, there is a far greater investment in development of the young during lactation and it is likely that many of the signals that regulate development of eutherian embryos in utero are delivered by the milk. This requires the co-ordinated development and function of the mammary gland. Inappropriate timing of these signalling events in mammals may result in either limited or abnormal development of the young, and potentially a higher incidence of mature onset disease. The tammar is emerging as an attractive model to better understand the role of milk factors in these processes.

Keywords Milk · Mammary · Stomach · Development · Neonate

K.R. Nicholas (✉)
Institute of Technology Research and Innovation, Deakin University, Geelong, VIC 3214, Australia
e-mail: kevin.nicholas@deakin.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_15, © Springer Science+Business Media B.V. 2010

15.1 Introduction

Milk is no longer simply considered a source of nutrients. Milk has a more comprehensive role in regulating growth and development of the suckled young, and autocrine control of the mammary gland so that it functions appropriately during the lactation cycle. Identification of regulatory factors in milk is more readily achieved in animal models that have evolved extended lactation strategies compared to eutherians (Brennan et al., 2007; Sharp et al., 2009).

Marsupials have evolved a unique reproductive strategy consisting of a short gestation, birth of an altricial young and a relatively long lactation, during which the mother progressively changes the composition of all the major, and many of the minor components of milk (Brennan et al., 2007; Sharp et al., 2009). These changes in milk composition are required for development of the suckled young (Trott et al., 2003). Thus there is a far greater investment in development of the young during lactation. In contrast, early development of the eutherian young is largely programmed and regulated in utero. Inappropriate signalling results in either limited or abnormal development, and potentially a higher incidence of mature onset disease. Therefore, the timing of delivery of these signals is crucial to setting biological clocks for tissue development.

The newborn tammar is similar to a late stage eutherian foetus, and must rapidly develop gut and lung function to successfully thrive. The neonatal stomach is not mature, and considerable development occurs in the pouch following birth of the young (Kwek et al., 2009b). Many of the necessary signals that regulate development of eutherian embryos in utero appear to be mediated by milk, and require co-ordinated development and function of the mammary gland (Brennan et al., 2007). Indeed, many studies have reported the capacity of the macropod marsupial to practise concurrent asynchronous lactation whereby the mother produces two very different milks from adjacent mammary glands (Nicholas, 1988a). These data suggest that, in addition to the well described endocrine regulation of milk synthesis and secretion (Nicholas et al., 1995), the tammar utilises a local, intrinsic, mechanism to regulate milk composition and rate of milk production. Moreover, this local regulation in the mammary gland determines the rate of pouch young (PY) growth and development, irrespective of the age of the PY (Trott et al., 2003; Waite et al., 2005; Kwek et al., 2009b).

15.2 The Lactation Cycle in the Tammar Wallaby (*Macropus eugenii*)

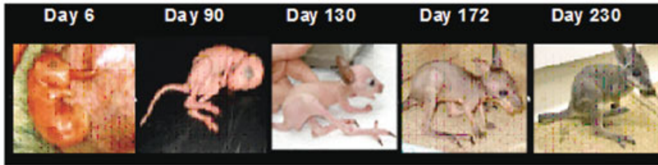
The lactation cycle is common to all mammals, although marsupials have evolved a reproductive strategy that is very different from that of most eutherians (Tyndale-Biscoe and Janssens, 1988). Eutherians have a long gestation relative to their lactation period, and the composition of the milk does not change substantially. In contrast, reproduction in marsupials such as the tammar wallaby (*Macropus*

eugenii) is characterized by a short gestation followed by a long lactation, and all the major milk constituents change progressively during lactation (Tyndale-Biscoe and Janssens, 1988). The altricial tammar neonate remains attached to the teat for the first 100 days of lactation, and may be considered as a foetus that is maintained in the pouch rather than the uterus (Fig. 15.1a). The conversion of milk to body mass in marsupials is similar to the conversion of nutrients to body mass observed in the eutherian foetus (Tyndale-Biscoe and Janssens, 1988). Therefore, examination of marsupial milk provides a unique opportunity to identify specific molecules that play a primary role in regulating the development of the young.

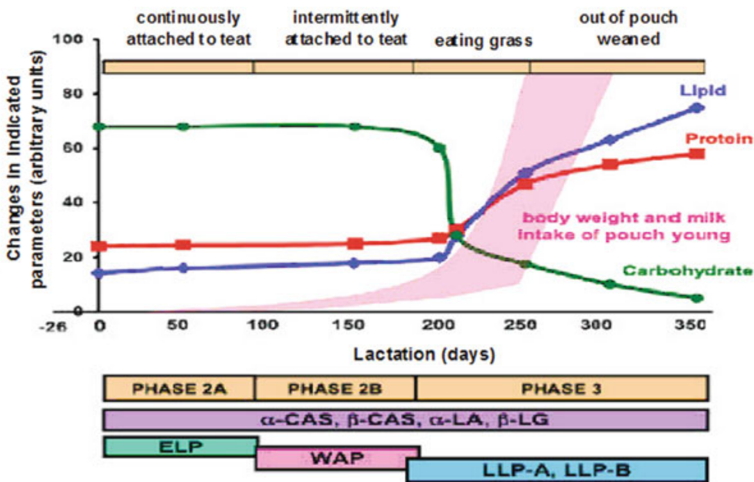
Lactation in the tammar has been divided into phases (Fig. 15.1b) defined by the composition of milk and apparent sucking pattern of the young (Nicholas et al., 1997b). Phase 1 is comprised of a 26.5-day pregnancy followed by parturition, and the subsequent 200 days of Phase 2 is characterized by lactogenesis and the secretion of small volumes of dilute milk high in complex carbohydrate and low in fat and protein. The PY remains attached to the teat for approximately the first 100 days (Phase 2A), after which it relinquishes permanent attachment to the teat and presumably sucks less frequently while remaining in the pouch (Phase 2B). The onset of Phase 3 of lactation (200–330 days) is characterized by temporary exit from the pouch by the young, a large increase in milk production and a change in the composition of milk to include elevated levels of protein and lipid and low levels of carbohydrate. There is a progressive increase in protein concentration (Nicholas, 1988a) and protein production during lactation, and the composition of the proteins changes considerably (Nicholas, 1988a, b; Nicholas et al., 1997a; Simpson et al., 2000; Trott et al., 2002) (Fig. 15.1b). As mentioned previously, macropodids such as the tammar can practice concurrent asynchronous lactation whereby the mother provides a concentrated milk high in protein and fat for an older animal which is out of the pouch, and a dilute milk low in fat and protein but high in complex carbohydrate from an adjacent mammary gland for a newborn in the pouch (Nicholas, 1988a, 1988b). This phenomenon suggests the mammary gland is controlled locally, and it is consistent with the concept that specific milk factors contribute to this process (Trott et al., 2003; Sharp et al., 2008; Kwek et al., 2009b).

To provide more detailed information on changes in milk protein during lactation in the tammar we recently developed a database of 15,000 sequenced tammar mammary ESTs (Lefèvre et al., 2007) and a custom tammar mammary EST microarray printed with 10,000 sequenced cDNAs representing >7,000 genes expressed across the lactation cycle (Sharp et al., 2009). The microarray was used to transcript profile the mammary gland at all the major stages of pregnancy and lactation (Sharp et al., 2009) (Fig. 15.1c). In addition the ESTs sequenced from mammary cDNA libraries prepared from tammars in pregnancy and lactation were analysed to provide a greater understanding of the mammary transcriptome during the lactation cycle (Lefèvre et al., 2007). The transcriptome data and the microarray analysis confirmed earlier published results using northern blot analysis examining the timing and level of expression of the major milk protein genes (Sharp et al., 2009) at each phase of the lactation cycle (Fig. 15.1d). There were two temporally-different patterns of milk protein gene expression during the lactation cycle; one group

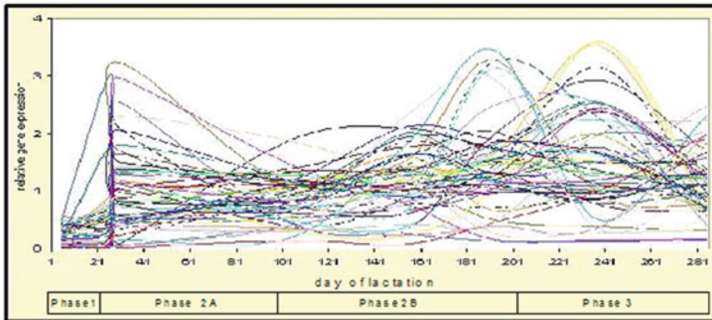
A



B



C



D

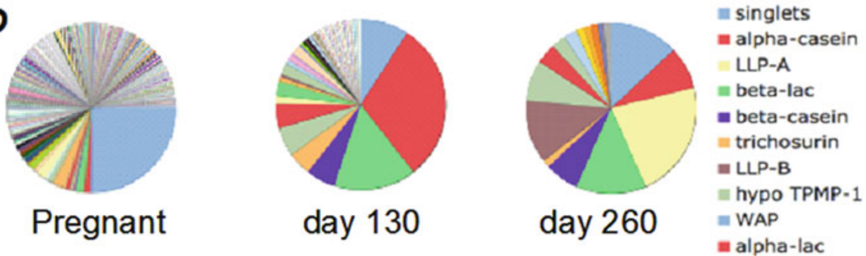


Fig. 15.1 Lactation in the tammar wallaby. (**a**) Developmental stages of the tammar wallaby pouch young during lactation. (**b**) Profile of tammar milk composition across the phases of the lactation cycle showing changes in total carbohydrate, protein and lipid secretion. The pattern of secretion of the major milk proteins are shown; α-casein (α-CAS), β-casein (β-CAS), α-lactalbumin (α-LA), β-lactoglobulin (β-LA), Early Lactation Protein (ELP), Whey Acid Protein (WAP), Late Lactation

of genes was induced to high levels around parturition and expressed throughout lactation and a second group of genes expressed highly only during specific phases of lactation (Sharp et al., 2009).

15.3 Regulation of the Lactation Cycle

The expression of milk protein genes is regulated concurrently by systemic endocrine factors, paracrine factors such as the extracellular matrix and by autocrine factors secreted in the milk (Sharp et al., 2009). More recently it has become evident that microRNA (miRNA) may have a role in regulating development and function of the mammary gland.

Micro RNA (miRNA) have important roles in proliferation, differentiation and apoptosis of various cell types. The main known function of miRNA is to silence expression of specific genes, either by degrading the targeted mRNA or preventing translation of messenger RNA (mRNA). Currently, it is thought that these miRNA represent 2–3% of the genome but potentially regulate approximately 20% of the genome (Silveri et al., 2006).

Preliminary studies, reported by Wang and Li (2007) show specific expression patterns for miRNA at each stage of mammary development in rodents. Furthermore, analysis with the miChip, showed that 21 miRNAs were down regulated and 17 up regulated in the gestation/lactation stages compared to virgin/involution stages, suggesting that miRNA have important roles in the mammary gland during onset, maintenance and cessation of lactation.

More recently, Sdassi et al. (2009) extended the repertoire of mouse mammary miRNA using a cloning approach that led to the identification and characterisation of 33 new miRNA. Although no miRNA were found to be mammary gland specific, six were found to be mouse specific (Sdassi et al., 2009). A recent study by Tanaka et al. (2009) showed that over expression of miR-101 was found to inhibit the cyclo-oxygenase-2 (cox-2) gene which encodes a protein mediating cellular proliferation, and also resulted in the suppression of β-casein gene (CSN2) expression. Thus, these data suggest that development defaults, such as incorrect mammary gland proliferation, may lead to altered expression of milk protein genes (Tanaka et al., 2009). These data support a significant role for miRNA in regulating the

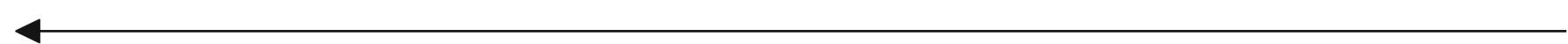


Fig. 15.1 (continued) Protein-A (LLP-A) and Late Lacation Preotin-B (LLP-B). (**c**) A total of 75 gene transcript profiles are shown which were predicted to encode for secreted proteins expressed in tammar mammary gland during lactation. This analysis was performed using microarray of tammar mammary tissue collected from the different phases of lactation. (**d**) Transcriptome analysis of three representative stages within the tammar lactation cycle. During pregnancy the gland produces a wide variety of transcripts, while the mammary gland at day 130 lactation (Phase 2B) and day 260 lactation (Phase 3) show an abundance of major transcripts relating to the various stages of milk secretion

timing of expression of milk protein genes in the mammary gland of mice, and indicate a likely role in the mammary gland of the marsupial, which undergoes a more complex pattern of mammary gland development and function.

15.4 Identification of Milk Bioactives in Tammar Milk Using a Functional Genomics Platform

The most commonly used approach to identify bioactive proteins secreted in milk is a proteomics-based analysis that fractionates milk using a variety of separation technologies, and examines fractions for bioactivity (Sharp et al., 2009). We have used an alternative approach that utilises a genomics platform to identify differentially expressed genes that encode for secreted proteins in mammary glands.

The tammar EST database and the microarray data were used to identify cDNAs coding for proteins that contained a signal peptide and that were differentially expressed during different phases of lactation (Lefèvre et al., 2007). This approach identified 75 novel genes that were expressed at specific stages of the lactation cycle. To determine the functional significance of these genes, individual milk proteins encoded by each cDNA were synthesized in vitro and subsequently shown to have activity for immune modulation, inflammatory responses, growth and differentiation effects in cell-based assays (Sharp et al., 2009). The candidate proteins were incubated with a range of cell lines to measure the activation of specific endogenous cellular proteins linked to the indicated cellular responses.

In addition, some of these synthesized proteins showed a capacity to alter development of tammar mammary epithelial cells in a mammosphere culture model by increasing their size and rate of formation. The mammosphere is a three-dimensional structure that closely resembles the mammary acini, the functional component of the mammary gland. This raised the possibility that these proteins may contribute to growth and function of the mammary gland during lactation. These data suggest that milk contains a variety of bioactives with a number of stimulatory activities and demonstrates how milk has the potential to play an integral role in development and function of the mammary gland as well as a variety of other tissues.

15.5 Changes in Milk Composition Regulates Growth of the Tammar Pouch Young

A study by Trott et al. (2003) confirmed that the sucking pattern of pouch young does not control the pattern of milk secretion in the tammar wallaby. Groups of 60 day old pouch young were fostered at 2-weekly intervals onto one group of host mothers so that a constant sucking stimulus on the mammary gland was maintained for 56 days and the lactational stage progressed ahead of the age of the young. The timing of changes on the concentration of protein and carbohydrate in milk

from fostered and control groups was unaffected by altering the sucking pattern. However, the rate of growth and development of the fostered PY was significantly increased relative to the control PY. It was concluded from these studies that at this stage of lactation the tammar wallaby regulates both milk composition and the rate of milk production, which subsequently determine the rates of PY growth and development. Therefore, the tammar provides a unique model to examine the correlation of milk composition with the defined developmental stages of the suckled young (Sharp et al., 2009) as well as new opportunities to describe function of specific milk components.

These studies established a role for milk composition in regulating growth of the young but did not evaluate effects of milk composition on development of specific tissues or the frequency of mature onset disease. It is now well established that additional signals are required for setting biological clocks that regulate development into adulthood. Growth of the foetus and size of the young at birth are determined by the supply of oxygen and nutrients, but it is the latter which is the dominant influence on fetal growth rate, birth weight and phenotype. In human populations, variations in maternal size, nutritional state, and placental function during late gestation can substantially account for the observed variation in birth weight. The in utero and postnatal environments are critical periods in programming growth of the foetus. Both low body weight or length at birth are associated with development of glucose intolerance, type 2 diabetes, obesity and hypertension in adults (Phillips, 2001). Recent studies also suggest that catch-up growth in infancy and accelerated growth later, particularly in those of low birth weight, independently increases the risk of these diseases in adults (Forsen et al., 2000; Eriksson et al., 2003). Therefore, the roles of the uterine and lactation environments in programming of adult diseases are clearly important. A question that remains to be addressed is whether the tammar mammary gland provides signals in the milk to the tammar neonate that are similar to those delivered to the foetus of a eutherian.

15.6 A Role for Milk in Regulation of Stomach Maturation

The stomach is lined with a self-renewing epithelial monolayer that is organised into pit/gland units, and each pit/gland contains a heterogeneous population of cells that originate from the stem cells (Karam, 1995, 1999) whose cytodifferentiation is influenced by factors from the external milieu. Recently Waite and colleagues (2005) used immunohistological techniques to examine development of the stomach in the tammar pouch young. Dramatic changes in tissue morphology take place around day 170 post-partum (Waite et al., 2005). In what will become the adult hindstomach region, parietal cells increase in number, gastric glands enlarge and adopt the adult-like phenotype of very long, thin glands (Waite et al., 2005) and peptic activity becomes elevated (Davis, 1981). In contrast, the forestomach region undergoes a transition from an immature gastric glandular phenotype in which there is a progressive loss of parietal cells, and the cardia glandular phenotype becomes apparent in

the region that will become the adult forestomach. By day 230 post-partum, parietal cells are undetectable in the forestomach mucosa (Waite et al., 2005; Kwek et al., 2009a). This phenotypic change in the forestomach is accompanied by functional changes, such as an increase in pH to neutrality (Janssens and Ternouth, 1987), a decline in peptic activity (Davis, 1981) and the gastric glandular cell type gene markers, prochymosin and pepsinogen are down-regulated (Kwek et al., 2009a). As these changes in forestomach morphology during tammar wallaby development are correlated with significant changes in milk composition, it raises the possibility that these processes may be regulated by specific components in milk.

To address this hypothesis Waite and colleagues (2005) transferred pouch young at 67 days of age to host mothers that were at day 100 of lactation. After the 30-day cross-fostering period, there was no obvious morphological changes in either stomach regions of non-fostered control or the fostered pouch young. These data suggest that while factors in later stage milk accelerated gross development as shown previously (Trott et al., 2003), there was no effect on stomach maturation. However, in the study by Waite and colleagues (2005), the period selected for cross-fostering across pouch young development was between day 60 and 100 post-partum, a stage during which stomach mucosal morphology normally remains uniform for at least another 70 days.

To examine whether milk from later stages of lactation can regulate changes in morphology in different regions of the stomach, a recent study targeted a later period of pouch young development when there are dramatic changes in stomach morphology (Kwek et al., 2009b). In this study, pouch young at day 120 of age were cross-fostered to host mothers at day 170 of lactation (Fig. 15.2a–d). In the fostered PY fore-stomach the parietal cell population was significantly reduced (Fig. 15.2e–g), and expression of gastric glandular phenotype marker genes (*ATP4A*, *GKN2*, *GHRL* and *NDRG2*) was down-regulated, suggesting down-regulation of gastric phenotype in fostered PY fore-stomach. The expression of cardia glandular phenotype genes (*MUC4*, *KRT20*, *CSTB*, *ITLN2* and *LPLUNC1*) was not changed in

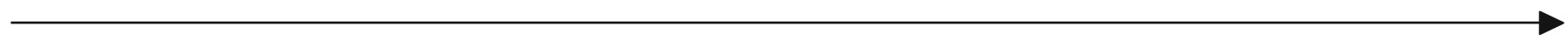


Fig. 15.2 Milk regulates gut maturation in the tammar wallaby. Photograph of pouch young at 120 days receiving Phase 2B milk **(a** and **b)**. **(b)** Photograph of pouch young **(a)** at 180 days receiving Phase 2A milk for an additional 60 days. **(c)** Photograph of pouch young **(b)** at 180 days after transfer to mother producing Phase 3 milk for 60 days (crossfostered). A dramatic difference in growth and development can be seen in the cross-fostered pouch young suggesting Phase 3 milk is developmentally regulating the growth of this animal. (**e**–**g** and **n**) Immunochemical studies show parietal cells can be detected by immune-reactivity against the α-subunit of H+K+ATPase (*dark staining cells*) in the forestomach of **(e)** tammar pouch young at 180 days of age receiving Phase 2B milk but are significantly reduced in tammar pouch young at 180 days of age receiving Phase 3 milk (cross-fostered) **(f** and **n)**. **(g)** Control animal shows absence of staining of parietal cells in the forestomach of 230 day old pouch young receiving Phase 3 milk. (**h**–**p**) Analysis of apoptotic cells (*arrowed*) and proliferating cells (*arrowed*) in the tammar forestomach were measured by (**h**–**j**) TUNNEL assay and (**k**–**m**) reactivity to anti-phoshpo-histone H3, respectively and nuclei were visualized by use of Hoechst dye (*Blue*). (**h** and **k**) show staining of forestomach tissue from pouch young at 180 days receiving Phase 2B milk, (**i** and **l**)

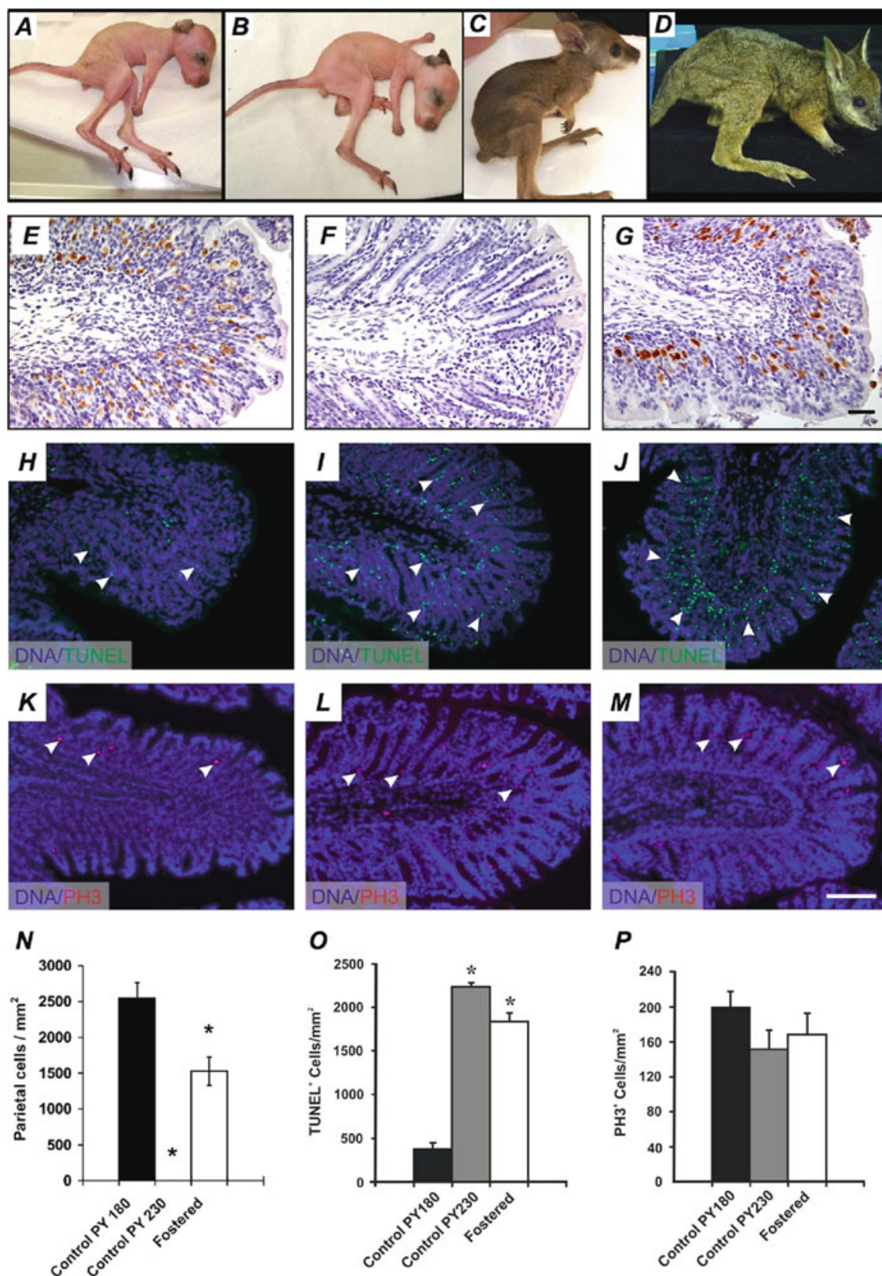


Fig. 15.2 (continued) show staining of forestomach tissue from pouch young at 230 days receiving Phase 3 milk and (**j** and **m**) show staining of forestomach tissue from pouch young at 180 days receiving Phase 3 milk (cross-fostered). Scale bar represnts 50 μm. These experiments show a significant increase in the number of apoptotic cells in the cross-fostered animal (**o**) while the number of proliferating cells remain the same (**p**). $^*P$<0.001 when compared to control PY180

fostered PY. In fostered PY fore-stomach, there was also an increased apoptosis, but no change in cell proliferation (Kwek et al., 2009b) (Fig. 15.2 h–p). These data suggest that fore-stomach maturation proceeds via two temporally distinct processes: down-regulation of gastric glandular phenotype and initiation of cardia glandular phenotype. In fostered PY, these two processes appear uncoupled, as gastric glandular phenotype was down-regulated but cardia glandular phenotype was not initiated. It remains to be determined what factors regulate these two apparently distinct processes. Candidate contributory factors include molecular factors in late stage lactation milk, or even contributions from herbage consumed by the pouch young during later lactation stage and the attendant colonisation of the stomach by micro-organisms.

In addition to the loss of parietal cells, there may also be other gastric glandular type cells undergoing apoptosis in the fostered PY forestomach, as many of the cells reactive for active caspase-3 or TUNEL did not co-localise with $H^{+}K^{+}$ATPase (*Atp4b*). Moreover, qPCR analysis showed down-regulated expression of other gastric glandular specialised cell marker genes such as *GHRL* and *GKN2*. Several studies indicate that normal parietal cell development is critical for homeostasis of other gastric epithelial cell types in the mucosa (Li et al., 1995, 1996; Franic et al., 2001; Jain et al., 2008). Mice deficient in the β-subunit of *Atp4b* gene lacked functional parietal cells and had a depleted number of zymogenic cells and increased number of immature gastric glandular cells undergoing apoptosis (Franic et al., 2001). Similarly, mice deficient for Huntington interacting protein 1 related (an F-actin binding protein expressed on canalicular membranes of parietal cell) also had abnormal parietal cell morphology, reduced numbers of parietal and chief cells and increased apoptosis (Jain et al., 2008). As much of the cell death occurs in the isthmus region of the mucosal gland of the tammar wallaby forestomach, it is possible that undifferentiated progenitors, as well as differentiated cells are affected and thus may contribute to the altered differentiation profile of the forestomach epithelium. However, the identity of other gastric glandular cell types lost from the forestomach of control PY 230 and fostered PY remains to be determined.

15.7 The Role of Gastric Microflora in Regulating Stomach Development

The mammalian gastrointestinal tract harbours a complex microbiota consisting of around 500–1,000 distinct microbial species (Backhed et al., 2005). Various animal models have provided convincing evidence that the inter-relationships between the host gastrointestinal tract and gut bacteria are important in regulating immune system development (Kelly et al., 2007) and lipid metabolism of the host (Backhed et al., 2004). Most importantly, there are now many studies that show that the establishment of gut microbiota plays a critical role in cytodifferentiation of the developing gut (Uribe et al., 1994; Bry et al., 1996; Bates et al., 2006).

At the end of the 60-day transfer period in the experiments described by Kwek et al. (2009b), the fostered PY were still consuming milk with high levels of carbohydrates and some parietal cells were still present in the fostered PY forestomach, suggesting the presence of acidic contents in the forestomach. In addition, expression of genes involved in bacterial recognition (e.g. *ITLN2*) and gut immunity (e.g. *CSTB* and *LPLUNC1*) were extremely low in forestomach of the fostered PY. It is possible that specific carbohydrates prevent establishment of specific microbial populations in the forestomach of the fostered PY (Kunz et al., 2000; Kunz and Rudloff, 2008) as some milk oligosaccharides can reduce microbial colonisation by acting as receptor analogues to inhibit adhesion of pathogens or microorganisms on the gut mucosal epithelial surface (Holmgren et al., 1983; Andersson et al., 1986; Kunz et al., 2000; Kunz and Rudloff, 2008). Complex carbohydrate and oligosaccharides found in human milk demonstrate prebiotic roles in the maintenance of gut microbiota (Boehm et al., 2004; Ward et al., 2006), and strongly suggests that there are specific factors within the breast milk that can promote the growth of beneficial bacteria and suppress the colonisation of harmful bacteria in the gastrointestinal tract of the infant (Newburg, 1996, 1997; Kunz and Rudloff, 2008).

In studies using the germ-free zebrafish, differentiation of the larval gut epithelium was arrested (Bates et al., 2006). Similarly, in studies with the germ-free rats, a lack of microbiota resulted in greatly reduced numbers of goblet cells (Ishikawa et al., 1986). Another study by Bry and colleagues (1996), using germ-free mice model, clearly demonstrated that development in the intestine was affected in the absence of microflora. Importantly, re-introduction of microflora to these germ-free animals restored the normal gut phenotype and function comparable to the conventionally-reared animals (Bry et al., 1996; Bates et al., 2006). This suggests that microbiota can regulate cytodifferentiation of the gut by directly influencing cell fate decisions in the gut epithelium, and that milk carbohydrate may play a role in this process.

15.8 A Role for Milk in the Control of Mammary Function in the Tammar

There is increasing evidence to suggest that milk plays an important role in regulating mammary epithelial function and survival, and this is particularly evident during involution (Brennan et al., 2007). For example, apoptosis was induced preferentially in the sealed teats of lactating mice (Li et al., 1997; Marti et al., 1997), whereas the litter suckled successfully on the remaining teats, indicating that cell death is stimulated by an intra-mammary mechanism sensitive to milk accumulation (Quarrie et al., 1996). A protein known as the feedback inhibitor of lactation (FIL), suggested as a candidate, is also secreted in the milk of the tammar (Hendry et al., 1998; Marti et al., 1999) and other species. It acts specifically through interaction with the apical surface of the mammary epithelial cell to reduce milk secretion

(Wilde et al., 1995; Hendry et al., 1998). An alternative hypothesis has proposed that engorgement of the mammary gland signals a down regulation of secretion and onset of apoptosis by a mechanism of physico-chemical signalling (Boudreau et al., 1995; Clark and Brugge, 1995).

Recent studies using the tammar mammary explant culture model (Nicholas and Tyndale-Biscoe, 1985) to examine the process of involution has confirmed the likely role of milk, and particularly of putative autocrine factors, for directly controlling mammary function during involution (Brennan et al., 2007). Mammary explants, small (1–2 mg) pieces of mammary tissue prepared from pregnant tammars, were cultured on siliconised lens paper in Medium 199 with lactogenic hormones (insulin, cortisol and prolactin) to induce milk protein gene expression; subsequent removal of all hormones for 10 days down-regulated expression of the milk protein genes and mimicked involution. Surprisingly, the explants retained the same level of lactogenic response during a subsequent challenge with lactogenic hormones. The maintenance of epithelial cell viability and hormone responsiveness in explants cultured in the absence of hormones is consistent with the idea that accumulation of local factors in the milk is the primary stimulus for apoptosis of mammary epithelial cells in the tammar wallaby mammary gland. This model will permit analysis of how hormone and milk regulation of involution are uncoupled, and will provide evidence for the extraordinary capacity for survival and maintenance of hormone responsiveness by tammar mammary epithelial cells cultured in a chemically defined medium in the absence of exogenous hormones and growth factors. Other studies in our laboratory, using the fur seal as a model to examine the uncoupling of involution from cessation of suckling by the young, have identified the milk protein α-lactalbumin as a candidate for initiating phase one involution and apoptosis (Sharp et al., 2008).

The functions of milk proteins such as β-lactoglobulin (BLG) and Whey Acidic Protein (WAP) in mammary development have remained elusive. WAP is a major whey protein and is secreted throughout lactation in many eutherian species including rat, mouse, rabbit, camel and pig (Simpson et al., 2000; Simpson and Nicholas, 2002). WAP is also secreted in the milk of all marsupials studied to date including the tammar (Simpson et al., 2000), red kangaroo (Nicholas et al., 2001), brush-tailed possum (Demmer et al., 2001), and the fat-tailed dunnart (De Leo et al., 2006). In addition, WAP mRNA has been detected in the lactating mammary gland of a

Fig. 15.3 Analysis of Whey Acidic Protein function in the mammary gland. (**a**) Recombinant tammar WAP (tWAP) constructs were transfected into HEK-293 cells. The constructs comprised full length tWAP, 4-DSC domain I and II, and 4DSC domain III cloned into the pCMV expression vector. Evidence of production of recombinants proteins was measured by collection of conditioned media and Western blot analysis (*right side panel*) showing the correct recombinant protein sizes (tWAP = 22.13 kDa, DI–II = 15.3 kDa and DIII = 8.9 kDa). (**b**) Proliferation of primary cultures of wallaby cells (Wall-MEC) and HC11 cells grown in conditioned medium containing tWAP or DIII was significantly increased at 72 h compared to cells in conditioned media containing DI–II or media alone. (**c**) The rate of DNA synthesis, as measued by [3H]-thymidine incorporation, was significantly higher in both Wall-MEC and HC11 cells grown

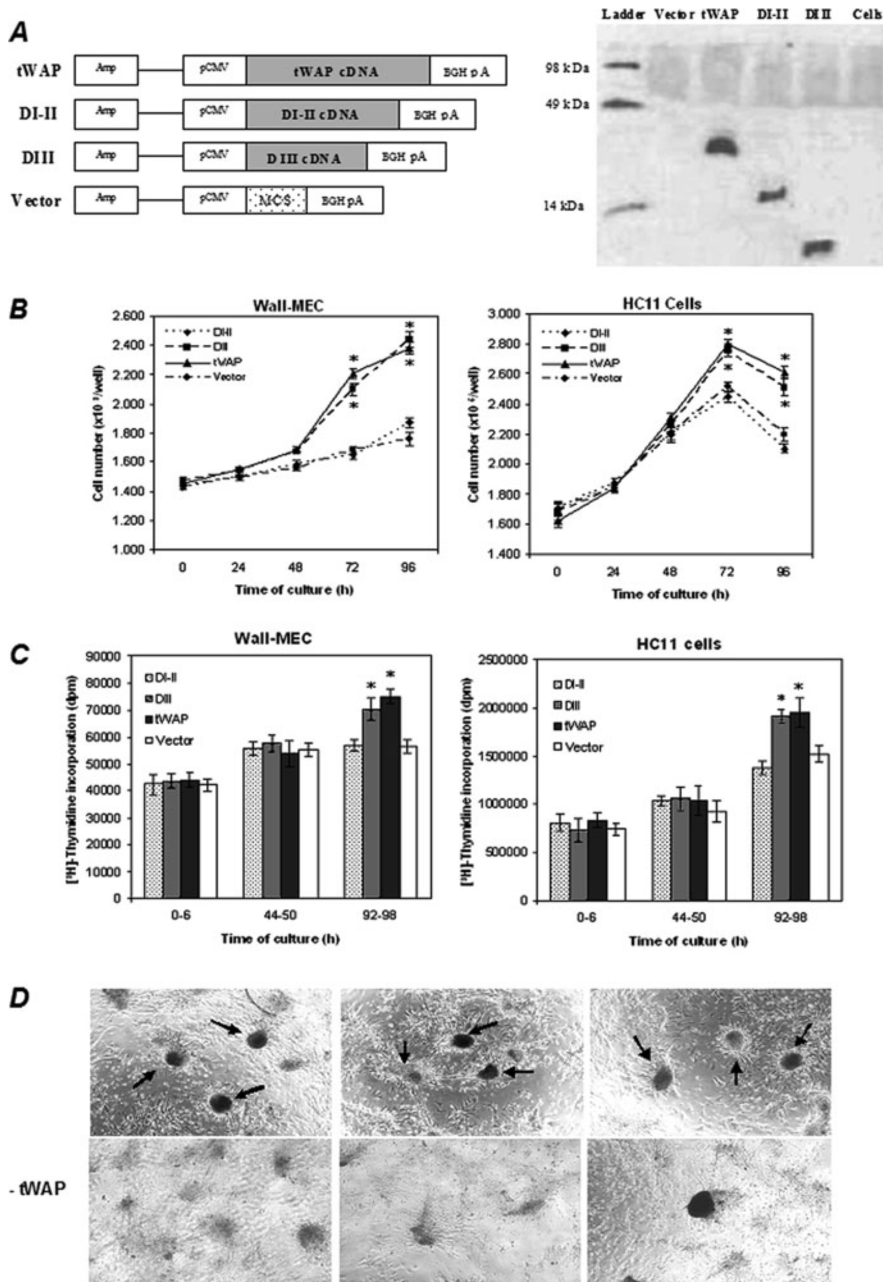


Fig. 15.3 (continued) in the presence of either tWAP or DIII compared to growth in the presence of DI–DII or media alone after 98 h. (**d**) The ability to stimulate mammosphere formation within a culture of primary mammary cells was used as a measure for promotion of mammary gland devlopment. The number of mammospheres (*arrowed*) was significantly higher when Wall-MEC cells were grown in medium containing tWAP for 5 days compared to cells grown in media alone

stripe faced dunnart and short-tailed opossum (Topcic and Nicholas, unpublished). However, in these species the secretion of the protein is specific to a phase of lactation and therefore the tammar provides a naturally occurring, transient gene knockout model to assess the role of this protein in regulating either specific stages of development of the young, the mammary gland or potentially both. The amino acid sequence of WAP from two monotremes, platypus and echidna, has been reported (Sharp et al., 2007; Simpson et al., 2000) but it is unknown whether the pattern of secretion in these species is similar to the marsupial.

Alignment of the amino acid sequence of WAPs from marsupial, monotreme and eutherian species shows limited sequence identity (Sharp et al., 2007; Simpson et al., 2000). However, these proteins are recognized by the presence of the WAP motif and a four disulphide core (4-DSC) domain (Ranganathan et al., 1999). WAP secreted in the milk of eutherians has two 4-DSC domains (Simpson and Nicholas, 2002), whereas WAP in all marsupials and monotremes studied to date has three domains (Simpson et al., 2000; Demmer et al., 2001; De Leo et al., 2006; Sharp et al., 2007). The five exons of the marsupial and monotreme WAP genes contrast with the eutherian WAP genes, which have only four exons (De Leo et al., 2006; Sharp et al., 2007). The significance of the loss of the third domain in eutherians remains unclear, particularly as a biological function of milk WAP remains to be established.

Recent studies have shown that mouse WAP added to culture medium of mouse HC-11 cells (a mammary epithelial cell line) is antiproliferative, and may act by an autocrine or paracrine mechanism (Nukumi et al., 2004). This is consistent with reports showing that over-expression of WAP in transgenic mice inhibits development of the mammary gland and secretion of milk (Burdon et al., 1991). In contrast, studies in our laboratory using full length tammar WAP and the individual 4-DSC tammar WAP domains (Fig. 15.3a) have shown that tammar WAP added to culture media stimulates proliferation of tammar mammary epithelial cells and increases expression of the cell cycle genes cyclin D1 and CDK-4 (Topcic et al., 2009) (Fig. 15.3b). The studies also show that this stimulation is entirely due to the presence of the third tammar WAP domain (DIII) which is absent from eutherian WAP genes. Earlier studies have shown that DNA synthesis in the tammar mammary gland is higher in Phase 2 than Phase 3 (Nicholas, 1988b), which is consistent with a potential role of tammar WAP in development of the mammary gland. This is further supported by studies showing that tammar WAP stimulates DNA synthesis in primary cultures of tammar mammary epithelial cells (Topcic et al., 2009) (Fig. 15.3c) and promotes in vitro growth of 3D alveoli-like structures (or mammospheres), which mimic in vivo mammary gland development and function (Fig. 15.3d). Interestingly, studies using a WAP gene knock-out mouse have shown mammary development is normal. However, the pups had decreased development at the later stages of lactation (Triplett et al., 2005). As tammar WAP is a major milk protein secreted during the middle third of lactation when the pouch young's diet comprises only milk, it could be speculated that this protein plays a specific role in development of both the mammary gland and the suckled young at this time. The milk WAP gene in human, cow, ewe and goat has been sequenced and mutations

within these genes suggest the gene is a non-functional psuedogene (Hajjoubi et al., 2006). It is conceivable that a putative function for WAP in mammary development is only relevant in marsupials and monotremes. The activity has most likely been lost in human, cow, ewe and goat due to loss of evolutionary pressure on this protein that relates to the changes in reproductive strategy of eutherians. Therefore, the marsupial may be a more appropriate model to explore the potential of other major milk proteins such as β-lactoglobulin, which, like WAP, is not expressed in all eutherians.

15.9 Conclusion

Early development of the eutherian young is programmed and regulated in utero. Developmental signals may be provided to the young, either continually or at specific times during development, and be provided by the amniotic fluid, placenta, colostrum and milk. Changes in the timing of these signals may result in abnormal development and a higher frequency of mature onset disease. The marsupial gives birth to an altricial young and much of the early development is regulated by milk. It is now apparent that new roles for milk are emerging, and future studies using the marsupial model will allow researchers to more fully understand the central role of milk to deliver time-dependent signals for both growth and development of the young, protect the young and mammary gland from infection and regulate the development and function of the mammary gland. A better understanding of the temporal delivery of these signals may provide new opportunities for treatment and prevention of disease.

References

Andersson B, Porras O, Hanson LA, Lagergard T, Svanborg-Eden C (1986) Inhibition of attachment of *Streptococcus pneumoniae* and *Haemophilus influenzae* by human milk and receptor oligosaccharides. J Infect Dis 153:232–237.

Backhed F, Ding H, Wang T, et al. (2004) The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci USA 101:15718–15723.

Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI (2005) Host-bacterial mutualism in the human intestine. Science 307:1915–1920.

Bates JM, Mittge E, Kuhlman J, Baden KN, Cheesman SE, Guillemin K (2006) Distinct signals from the microbiota promote different aspects of zebrafish gut differentiation. Dev Biol 297:374–386.

Boehm G, Jelinek J, Stahl B, et al. (2004) Prebiotics in infant formulas. J Clin Gastroenterol 38:S76–S79.

Boudreau N, Sympson CJ, Werb Z, Bissell MJ (1995) Suppression of ICE and apoptosis in mammary epithelial cells by extracellular matrix. Science 267:891–893.

Brennan AJ, Sharp JA, Lefevre C, et al. (2007) The tammar wallaby and fur seal: models to examine local control of lactation. J Dairy Sci 90:E66–E75.

Bry L, Falk PG, Midtvedt T, Gordon JI (1996) A model of host-microbial interactions in an open mammalian ecosystem. Science 273:1380–1383.

Burdon T, Sankaran L, Wall RJ, Spencer M, Hennighausen L (1991) Expression of a whey acidic protein transgene during mammary development. Evidence for different mechanisms of regulation during pregnancy and lactation. J Biol Chem 266:6909–6914.

Clark EA, Brugge JS (1995) Integrins and signal transduction pathways: the road taken. Science 268:233–239.

Davis PH (1981) The Post-natal Development of Digestive Proteinases in the Rat (*Rattus rattus*) and the Tammar Wallaby (*Macropus eugenii*). University of Sydney, Australia.

De Leo AA, Lefevre C, Topcic D, et al. (2006) Characterization of two whey protein genes in the Australian dasyurid marsupial, the stripe-faced dunnart (*Sminthopsis macroura*). Cytogenet Genome Res 115:62–69.

Demmer J, Stasiuk SJ, Grigor MR, Simpson KJ, Nicholas KR (2001) Differential expression of the whey acidic protein gene during lactation in the brushtail possum (*Trichosurus vulpecula*). Biochim Biophys Acta 1522:187–194.

Eriksson JG, Forsen T, Tuomilehto J, Osmond C, Barker DJ (2003) Early adiposity rebound in childhood and risk of type 2 diabetes in adult life. Diabetologia 46:190–194.

Forsen T, Eriksson J, Tuomilehto J, Reunanen A, Osmond C, Barker D (2000) The fetal and childhood growth of persons who develop type 2 diabetes. Ann Intern Med 133:176–182.

Franic TV, Judd LM, Robinson D, et al. (2001) Regulation of gastric epithelial cell development revealed in H(+)/K(+)-ATPase beta-subunit- and gastrin-deficient mice. Am J Physiol Gastrointest Liver Physiol 281:G1502–G1511.

Hajjoubi S, Rival-Gervier S, Hayes H, et al. (2006) Ruminants genome no longer contains Whey Acidic Protein gene but only a pseudogene. Gene 370:104–112.

Hendry KA, Simpson KJ, Nicholas KR, Wilde CJ (1998) Autocrine inhibition of milk secretion in the lactating tammar wallaby (*Macropus eugenii*). J Mol Endocrinol 21:169–177.

Holmgren J, Svennerholm AM, Lindblad M (1983) Receptor-like glycocompounds in human milk that inhibit classical and El Tor Vibrio cholerae cell adherence (hemagglutination). Infect Immun 39:147–154.

Ishikawa K, Satoh Y, Tanaka H, Ono K (1986) Influence of conventionalization on small-intestinal mucosa of germ-free Wistar rats: quantitative light microscopic observations. Acta Anat 127:296–302.

Jain RN, Al-Menhali AA, Keeley TM, et al. (2008) Hip1r is expressed in gastric parietal cells and is required for tubulovesicle formation and cell survival in mice. J Clin Invest 118: 2459–2470.

Janssens PA, Ternouth JH (1987) The transition from milk to forage diets. In: Hacker JB, Ternouth JH (eds) The Nutrition of Herbivores. Academic Press, London and New York.

Karam SM (1995) New insights into the stem cells and the precursors of the gastric epithelium. Nutrition 11:607–613.

Karam SM (1999) Lineage commitment and maturation of epithelial cells in the gut. Front Biosci 4:D286–D298.

Kelly D, King T, Aminov R (2007) Importance of microbial colonization of the gut in early life to the development of immunity. Mutat Res 622:58–69.

Kunz C, Rudloff S (2008) Potential anti-inflammatory and anti-infectious effects of human milk oligosaccharides. Adv Exp Med Biol 606:455–465.

Kunz C, Rudloff S, Baier W, Klein N, Strobel S (2000) Oligosaccharides in human milk: structural, functional, and metabolic aspects. Annu Rev Nutr 20:699–722.

Kwek J, De Iongh R, Nicholas K, Familari M (2009a) Molecular insights into evolution of the vertebrate gut: focus on stomach and parietal cells in the marsupial, *Macropus eugenii*. J Exp Zool B Mol Dev Evol 312:613–624.

Kwek JH, Iongh RD, Digby MR, Renfree MB, Nicholas KR, Familari M (2009b) Cross-fostering of the tammar wallaby (*Macropus eugenii*) pouch young accelerates fore-stomach maturation. Mech Dev 126:449–463.

Lefèvre CM, Digby MR, Whitley JC, Strahm Y, Nicholas KR (2007) Lactation transcriptomics in the Australian marsupial, *Macropus eugenii*: transcript sequencing and quantification. BMC Genomics 8:417–425.

Li M, Liu X, Robinson G, et al. (1997) Mammary-derived signals activate programmed cell death during the first stage of mammary gland involution. Proc Natl Acad Sci USA 94:3425–3430.

Li Q, Karam SM, Gordon JI (1995) Simian virus 40 T antigen-induced amplification of pre-parietal cells in transgenic mice. Effects on other gastric epithelial cell lineages and evidence for a p53-independent apoptotic mechanism that operates in a committed progenitor. J Biol Chem 270:15777–15788.

Li Q, Karam SM, Gordon JI (1996) Diphtheria toxin-mediated ablation of parietal cells in the stomach of transgenic mice. J Biol Chem 271:3671–3676.

Marti A, Feng Z, Altermatt HJ, Jaggi R (1997) Milk accumulation triggers apoptosis of mammary epithelial cells. Eur J Cell Biol 73:158–165.

Marti A, Lazar H, Ritter P, Jaggi R (1999) Transcription factor activities and gene expression during mouse mammary gland involution. J Mammary Gland Biol 4:145–152.

Newburg DS (1996) Oligosaccharides and glycoconjugates in human milk: their role in host defense. J Mammary Gland Biol 1:271–283.

Newburg DS (1997) Do the binding properties of oligosaccharides in milk protect human infants from gastrointestinal bacteria? J Nutr 127:980S–984S.

Nicholas K, Simpson K, Wilson M, Trott J, Shaw D (1997a) The tammar wallaby: a model to study putative autocrine-induced changes in milk composition. J Mammary Gland Biol 2:299–310.

Nicholas KR (1988a) Asynchronous dual lactation in a marsupial, the tammar wallaby (*Macropus eugenii*). Biochem Biophys Res Commun 154:529–536.

Nicholas KR (1988b) Control of milk protein synthesis in the tammar wallaby:a model system to study prolactin-dependent development. In: Tyndale-Biscoe CH, Janssens PA (eds) The Developing Marsupial: Models for Biomedical Research. Springer, Heidelberg.

Nicholas KR, Tyndale-Biscoe CH (1985) Prolactin-dependent accumulation of alpha-lactalbumin in mammary gland explants from the pregnant tammar wallaby (*Macropus eugenii*). J Endocrinol 106:337–342.

Nicholas KR, Wilde C, Bird K, Hendry K, Tregenza K, Warner B (1995) Asynchronous concurrent secretion of milk proteins in the tammar wallaby (*Macropus eugenii*). In: Wilde C, Knight C, Peaker M (eds) Intercellular Signalling in the Mammary Gland. Plenum Press, London.

Nicholas KR, Simpson K, Wilson M, Trott J, Shaw D (1997b) The tammar wallaby: a model to study putative autocrine-induced changes in milk composition. J Mammary Gland Biol 2: 299–310.

Nicholas KR, Fisher JA, Muths E, et al. (2001) Secretion of whey acidic protein and cystatin is down regulated at mid-lactation in the red kangaroo (*Macropus rufus*). Comp Biochem Physiol A Mol Integr Physiol 129:851–858.

Nukumi N, Ikeda K, Osawa M, Iwamori T, Naito K, Tojo H (2004) Regulatory function of whey acidic protein in the proliferation of mouse mammary epithelial cells in vivo and in vitro. Dev Biol 274:31–44.

Phillips DIW (2001) Non-insulin-dependent diabetes and obesity. In: Barker DJ (ed) Fetal Origins of Cardiovascular and Lung Disease. Marcel Dekker, New York.

Quarrie LH, Addey CVP, Wilde CJ (1996) Programmed cell death during mammary tissue involution induced by weaning, litter removal and milk stasis. J Cell Physiol 168:559–569.

Ranganathan S, Simpson KJ, Shaw DC, Nicholas KR (1999) The whey acidic protein family: a new signature motif and three-dimensional structure by comparative modeling. J Mol Graph Model 17:106–113, 134–136.

Sdassi N, Silveri L, Laubier J, et al. (2009) Identification and characterization of new miRNAs cloned from normal mouse mammary gland. BMC Genomics 10:149.

Silveri L, Tilly G, Vilotte JL, Le Provost F (2006) MicroRNA involvement in mammary gland development and breast cancer. Reprod Nutr Dev 5:549.

Sharp JA, Lefevre C, Nicholas KR (2007) Molecular evolution of monotreme and marsupial whey acidic protein genes. Evol Dev 9:378–392.

Sharp JA, Lefevre C, Nicholas KR (2008) Lack of functional alpha-lactalbumin prevents involution in Cape fur seals and identifies the protein is an apoptotic milk factor in mammary gland involution. BMC Biol 6:48.

Sharp JA, Digby M, Lefevre C, et al. (2009) The comparative genomics of tammar wallaby and fur seal lactation; models to examine function of milk proteins. In: Thompson A, Boland M, Singh H (eds) Milk Proteins: From Expression to Food. Academic Press, New York.

Simpson KJ, Ranganathan S, Fisher JA, Janssens PA, Shaw DC, Nicholas KR (2000) The gene for a novel member of the whey acidic protein family encodes three four-disulfide core domains and is asynchronously expressed during lactation. J Biol Chem 275:23074–23081.

Simpson KJ, Nicholas KR (2002) The comparative biology of whey proteins. J Mammary Gland Biol 7:313–326.

Tanaka T, Haneda S, Imakawa K, Sakai S, Nagaoka K (2009) A microRNA, miR-101a, controls mammary gland development by regulating cyclooxygenase-2 expression. Differentiation 77:181–187.

Topcic D, Auguste A, De Leo AA, Lefevre C, Digby MR, Nicholas KR (2009) Characterization of the tammar wallaby (*Macropus eugenii*) whey acidic protein gene: new insights into the function of the protein. Evol Dev 11:363–375.

Triplett AA, Sakamoto K, Matulka LA, Shen L, Smith GH, Wagner KU (2005) Expression of the whey acidic protein (Wap) is necessary for adequate nourishment of the offspring but not functional differentiation of mammary epithelial cells. Genesis 43:1–11.

Trott JF, Wilson MJ, Hovey RC, Shaw DC, Nicholas KR (2002) Expression of novel lipocalin-like milk protein gene is developmentally-regulated during lactation in the tammar wallaby, *Macropus eugenii*. Gene 283:287–297.

Trott JF, Simpson KJ, Moyle RL, et al. (2003) Maternal regulation of milk composition, milk production, and pouch young development during lactation in the tammar wallaby (*Macropus eugenii*). Biol Reprod 68:929–936.

Tyndale-Biscoe CH, Janssens PA (1988) The Developing Marsupial: Models for Biomedical Research. Springer, Heidelberg.

Uribe A, Alam M, Johansson O, Midtvedt T, Theodorsson E (1994) Microflora modulates endocrine cells in the gastrointestinal mucosa of the rat. Gastroenterology 107:1259–1269.

Waite R, Giraud A, Old J, et al. (2005) Cross-fostering in *Macropus eugenii* leads to increased weight but not accelerated gastrointestinal maturation. J Exp Zool Part A 303:331–344.

Wang C, Li Q (2007) Identification of differentially expressed microRNAs during the development of Chinese murine mammary gland. J Genet Genomics 34:966–973.

Ward RE, Ninonuevo M, Mills DA, Lebrilla CB, German JB (2006) In vitro fermentation of breast milk oligosaccharides by Bifidobacterium infantis and Lactobacillus gasseri. Appl Environ Microb 72:4497–4499.

Wilde CJ, Addey CVP, Boddy-Finch L, Peaker M (1995) Autocrine control of milk secretion: from concept to application. In: Intercellular Signalling in the Mammary Gland. Plenum Press, New York.

Part VII
Marsupial Immune Genes

Janine E. Deakin

Marsupial young are born without an immune system, and yet survive in an environment filled with potentially harmful microorganisms. Determining how these immature young survive in the absence of an immune system has been the subject of a number of studies over the past several decades. Likewise, comparisons of the adult immune response of marsupials with that of eutherian mammals have been carried out in attempt to gain insight into the similarities and differences between immune responses in these two mammalian lineages. The amazing survival of the "altricial" young would suggest that marsupials are amply able to immunologically protect themselves. However, early studies in this field, hampered by a lack of marsupial-specific reagents and sequence information for key immune genes, concluded that the immune response of marsupials was inferior to that of their eutherian counterparts (Jurd, 1994; Wilkinson et al., 1994; Stone et al., 1998). With the release of the grey short-tailed opossum and tammar wallaby genome sequences, major advances have been made in the field, overturning this idea of inferiority.

The first genomics study of immune genes in marsupials was carried out on the Major Histocompatibility Complex (MHC), a region of the genome essential to the immune response. In the first chapter of this part by Cheng et al. we learn of the complexity and organisation of this region in the opossum and wallaby, and how sequence information from these two model organisms has aided studies into genetic diversity at MHC loci. Such diversity studies are essential when species conservation is a concern due to the threat of disease, with the Tasmanian devil being an obvious example due to its recent addition to the endangered species list as a result of devil facial tumour disease (DFTD). These studies are equally important for determining the efficacy of immunocontraception as an approach to controlling pest populations of marsupials, such as the brushtail possum in New Zealand.

Immunoglobulins (Igs), antigen receptors central to the adaptive immune response, have been studied for over 30 years in marsupials, with initial studies isolating some of immunoglobulin proteins from several different species (reviewed in Miller and Belov, 2000). Over the past decade, further advances into the characterisation of marsupial Igs were made by sequencing the genes encoding Igs, and examining the genomic context of these loci as reviewed here by Baker and colleagues (Chapter 17). Marsupial-specific antisera to the different types of Igs have now been developed from these sequence data (Rawson et al., 2002) and

have been used to characterise the immune response of tammar wallabies and brushtail possums (Deakin et al., 2005). Undoubtedly, these antisera will continue to be a tremendous resource for immunological studies in marsupials and further characterisation of the immune response.

Early studies of the marsupial immune response hypothesised that T cell mediated responses were inferior to those of eutherian mammals due mainly to poor mixed lymphocyte responses and sluggish antibody isotype switching. Despite these observations, one of the most interesting findings from genomic studies into marsupial immune genes comes from an investigation of the receptors found on the surface of T cells. In addition to four T cell receptors (TCRs) traditionally found in eutherian mammals and other jawed vertebrates, marsupials have a novel TCR, as discussed by Baker et al. (Chapter 17).

Perhaps one of the greatest advances in this field is highlighted in Chapter 18, where Morris et al. describe the recent identification of many of the divergent immune genes, such as the cytokines and antimicrobials. These genes were virtually impossible to isolate by standard laboratory methods due to low levels of sequence similarity across taxa. With the release of genomic sequence data for both the opossum and wallaby, and the use of sophisticated bioinformatics, the identification of these genes has been made possible with remarkable results.

Overall, genomic studies of marsupial immune genes have shown that the marsupial immune gene repertoire is just as complex as that of eutherians, dispelling the insinuation of marsupial inferiority (Belov et al., 2007). This perceived "primitive" nature of the immune response in marsupials could be attributed to the use of reagents raised against components of the eutherian immune response, which gene sequencing has shown, are often too divergent to be detected by such reagents. Thanks to the marsupial genome projects, many more marsupial-specific reagents can now be generated.

Two key areas of continued research are highlighted by the chapters in this part. Firstly, the protection of the altricial young will continue to be a primary focus as it is likely that research in this area could have implications for humans, such as identifying potential antimicrobial therapies. Secondly, the fate of the Tasmanian devil population, which is under threat of extinction due to DFTD, may rely not only on studies investigating the diversity of MHC genes but determining how the tumour evades detection by the devil's immune system. Marsupial genomics studies have greatly advanced our understanding of the genes involved in the immune response in marsupials. Now we are able to move from simply generating a catalogue of marsupial immune genes to determining the function and physiology of the immune system in marsupials.

References

Belov K, Sanderson CE, Deakin JE, et al. (2007) Characterization of the opossum immune genome provides insights into the evolution of the mammalian immune system. Genome Res 17: 982–991.

Deakin JE, Belov K, Curach NC, Green P, Cooper DW (2005) High levels of variability in immune response using antigens from two reproductive proteins in brushtail possums. Wildl Res 32:1–6.
Jurd RD (1994) "Not proper mammals": Immunity in monotremes and marsupials. Comp Immunol Microbiol Infect Dis 7:41–52.
Miller RD, Belov K (2000) Immunoglobulin genetics of marsupials. Developmental and comparative immunology 24:485–490.
Rawson RR, Belov K, Gidley-Baird AA, Cooper DW (2002) Characterisation of antisera to recombinant IgA of the common brushtail possum (*Trichosurus vulpecula*). Vet Immunol Immunopathol 88:89–95.
Stone WH, Brunn DA, Foster EB, et al. (1998) Absence of a significant mixed lymphocyte reaction in a marsupial (*Monodelphis domestica*). Lab Anim Sci 48:184–189.
Wilkinson R, Kotlarski I, Barton M (1994) Further characterisation of the immune response of the koala. Vet Immunol Immunopathol 40:325–339.

Chapter 16
The Marsupial Major Histocompatibility Complex

Yuanyuan Cheng, Hannah V. Siddle, and Katherine Belov

Abstract The major histocompatibility complex (MHC) contains genes that play essential roles in the innate and adaptive immune responses, and has been one of the most studied genomic regions among vertebrates. The marsupial MHC has been intensively investigated in several species, with the genomic organization having been characterized in the gray short-tailed opossum and tammar wallaby, and gene sequences and genetic diversity examined in a number of representative marsupials. These studies have significantly contributed to our understanding of the marsupial immune system, evolution of MHC, and wildlife health and conservation. This review will provide a comprehensive summary of current knowledge on the marsupial MHC.

Keywords Marsupials · MHC · Immunogenetics · Wildlife health

16.1 Introduction

The major histocompatibility complex (MHC) is a gene-dense region found in all jawed vertebrate genomes which contains a large number of genes involved in the immune response. This cluster of genes was first identified for its critical role in determining tissue compatibility in transplantation, for which it received its name. Based on the structure and function of their encoded proteins, the MHC genes are grouped into three classes (class I, class II and class III), and class I and class II genes are further classified by function as classical and nonclassical (Parham and Strominger, 1982). Classical class I and class II molecules are involved in antigen presentation and have high levels of genetic polymorphism within their peptide binding regions in most species. Classical class I (class Ia) genes are ubiquitously expressed and encode antigens that function in the recognition and destruction of foreign, virus-infected or malignant cells by cytotoxic T cells (reviewed in Cresswell

K. Belov (✉)
Faculty of Veterinary Science, The University of Sydney, Sydney, NSW 2006, Australia
e-mail: kbelov@usyd.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_16, © Springer Science+Business Media B.V. 2010

et al., 2005). Therefore, they are also believed to be the major antigens involved in self/nonself discrimination and graft rejection. Nonclassical class I (class Ib) genes encode class I-like molecules, but have varied functions. Compared to class Ia genes, they also have lower expression levels, tissue specific expression and lower levels of polymorphism (Stroynowski and Lindahl, 1994). MHC class II molecules are responsible for presenting antigens derived from intravesicular and extracellular pathogens to CD4 helper T cells, which produce chemical signals triggering the production of antigen-specific antibodies and an inflammatory response that kills the pathogens (reviewed in Pieters, 2000). The class III region consists of a heterogeneous mixture of genes and is classified primarily basing on the location of these genes, which is between class I and class II regions in eutherian mammals, rather than functional commonality (Klein, 1986).

Early studies of MHC-related immune responses in marsupials, which comprise studies of skin allograft and mixed lymphocyte culture responses, date back to three to four decades ago (LaPlante et al., 1969; Yadav et al., 1974; Rowlands, 1976; Wilson et al., 1976). Marsupial MHC genes were not clearly defined until 1987, when Stone and co-workers identified for the first time a marsupial class I locus in the gray short-tailed opossum (*Mododelphis domestica*) (Stone et al., 1987). Since then, more studies on the marsupial MHC have been undertaken, driven by the establishment of new marsupial model species and advances in experimental technology. The strong research interest in the marsupial MHC has arisen not only from its essential role in the immune response and disease resistance, but also because of its important position phylogenetically in the field of comparative genomics. This review will present and summarize our current understanding of the marsupial MHC.

16.2 Marsupial MHC Class I Genes

An MHC class I molecule is a heterodimer of a membrane-spanning α chain and a β_2-microglobulin chain. The α chain is encoded by the MHC class I gene, comprised of eight exons translating to three extracellular domains, transmembrane and cytoplasmic domains. The extracellular α1 and α2 domains comprise the peptide binding domains. The β_2-microglobulin chain is not encoded in the MHC region. Under selection pressure from changes in pathogens and environment, MHC genes evolve rapidly in a "birth and death" manner (Nei et al., 1997). A new copy of an MHC gene is "born" from gene duplication, and then subsequently accumulates mutations, which may provide the gene with a change in function or cause it to become a pseudogene ("death" of the gene). As a consequence of this mode of gene evolution, there are multiple copies of class I genes (including classical and nonclassical genes and pseudogenes) in the genome of most vertebrate species, which has also been observed in marsupials. In addition to their rapid evolution, MHC class I genes undergo species-specific expansions and orthologous relationships between class I genes of different species are difficult to detect.

MHC class I genes and sequences that have been reported in marsupial species are summarized in Table 16.1. Up to now, cDNA transcripts of MHC class I genes

Table 16.1 Marsupial MHC class I genes and sequences (Gene names shown in italic)

Species	*Locus*/Sequence	Description	Accession number	References
Gray short-tailed opossum	*Modo-UA1*	Classical	CH465496	Miska and Miller (1999)
	Modo-UG	Nonclassical	(opossum scaffold_42)	Gouin et al. (2006b)
	Modo-UE, *UK*, *UI*, *UJ* and *UM*	Likely nonclassical	AF522352	Baker et al. (2009)
		Not expressed		Belov et al. (2006)
	Modo-UA2, *UH*, *UF* and *UL*	Un-linked to MHC, low expression		Miska et al. (2004)
	Modo-UB and *UC*			
Red-necked wallaby	Maru-UA*01, 02 and UB*01	Complete cDNA sequences	L04950–2	Mayer et al. (1993)
Tammar wallaby	*Maeu-UK*, *UM*, *UE* and *UO*	In MHC region, likely nonclassical		Siddle et al. (2009)
	Maeu-UP and *UL*	In MHC region, not expressed		
	Maeu-UA, *UB* and *UC*	Un-linked to MHC, likely classical	DQ304109–14	Siddle et al. (2006)
	Maeu-UD, *UF*, *UH* and *UN*	Un-linked to MHC		
	Maeu-UI and *UJ*	Un-linked to MHC, not expressed		
	Maeu1–6	Complete cDNA		
	Maeu-2*01–04 and Maeu-3*01–02	Complete cDNA	EF602311–6	Daly et al. (2007)
Brushtail possum	Trvu-UB*01	Complete cDNA, likely classical	AF359509	Lam et al. (2001b)
	Trvu01–20	Partial cDNA, α1–α3 domain	EU570810–29	Holland et al. (2008b)
Koala	Phci-U	Partial cDNA, α1 and α2 domain	U33807–15	Houlden et al. (1996)
Tasmanian devil	SahaI*01–13	Complete cDNA	EF591089–101	Siddle et al. (2007b)
	3297T1–6	Class I sequences from tumour	EF591108–12, 20	Siddle et al. (2007a)
	Saha*1–26	α1 domain	EU094443–68	

have been identified and sequenced in the following marsupial species: red-necked wallaby (*Macropus rufogriseus*) (Mayer et al., 1993), koala (*Phascolarctos cinereus*) (Houlden et al., 1996), gray short-tailed opossum (*Monodelphis domestica*) (Miska and Miller, 1999), brushtail possum (*Trichosurus vulpecula*) (Lam et al., 2001b; Holland et al., 2008b), tammar wallaby (*Macropus eugenii*) (Siddle et al., 2006; Daly et al., 2007) and Tasmanian devil (*Sarcophilus harrisii*) (Siddle et al., 2007a, b). It has been estimated that there are at least two, three and seven class I loci in the red-necked wallaby (Mayer et al., 1993), koala (Houlden et al., 1996) and Tasmanian devil (Siddle et al., 2007b), respectively. Due to the extensive level of polymorphism in the sequences and the high degree of similarity between genes, it is usually impossible to determine whether two cDNA sequences represent different genes or alleles from a single locus. Fortunately, the availability of the genome sequence of gray short-tailed opossum and tammar wallaby has allowed investigation into the MHC genes at the genomic level in these two species and provided better understanding of marsupial class I genes.

In the opossum, there are 13 MHC class I loci: *Modo-UA1*, *UA2*, *UB*, *UC*, *UE*, *UF*, *UG*, *UH*, *UI*, *UJ*, *UK*, *UL* and *UM* (Belov et al., 2006). These genes can be divided into three groups by expression level. (1) Of all opossum class I genes, *Modo-UA1* is the only confirmed classical locus, which is ubiquitously expressed and highly polymorphic. Transcripts of *Modo-UA1* have been detected in all tissues tested and account for all class Ia cDNAs reported earlier (Miska and Miller, 1999). (2) Six genes, *Modo-UG*, *UE*, *UK*, *UI*, *UJ* and *UM*, are proposed to be expressed class Ib loci. *Modo-UG* is expressed as three alternatively spliced mRNA variants, encoding a short cytoplasmic tail that lacks the phosphorylation sites conserved in classical class I molecules, and showing no polymorphism in the peptide binding regions (Gouin et al., 2006b). *Modo-UE*, *UK*, *UI*, *UJ* and *UM* all have tissue-specific expression and show limited levels of polymorphism in peptide binding grooves; *Modo-UJ* and *UM* are expressed as alternative splice variants (Baker et al., 2009). Another two genes, *Modo-UB* and *UC*, are suggested to be products of recent gene duplication events (Miska et al., 2004). They are highly similar to *Modo-UA1* in sequence, but have much lower expression levels and are unlinked to the MHC. Without data showing the polymorphism levels, whether they are classical or nonclassical genes remains to be determined. (3) The remaining four class I loci (*Modo-UA2*, *UH*, *UF* and *UL*) have not been found expressed in any tissue. Lacking a predicted open reading frame (ORF), *UA2* and *UH* are predicted to be pseudogenes. In addition, orthologues of several human class Ib genes have been identified in the opossum, including a *MIC* gene and a *CD1* pseudogene (Baker and Miller, 2007).

The number of class I loci in the tammar wallaby is similar to that observed in the opossum. The wallaby has 15 class I loci, 11 of which are transcribed (Siddle et al., 2009). Investigations on the expression pattern, polymorphism level, phylogenetic relationship and promoter elements have provided independent lines of evidence supporting that three (*Maeu-UA*, *UB* and *UC*) of the 15 class I loci are likely classical class I genes. Four loci, *Maeu-UK*, *UM*, *UE* and *UO*, have characteristics of class Ib genes. First, *Maeu-UK* was found expressed in all tissues examined except

skin, whereas *Maeu-UM*, *UE* and *UO* showed tissue specific expression patterns. Second, all four genes have extremely low levels of polymorphism. Finally, the promoter elements of these genes show little sequence similarity to the wallaby or opossum class Ia genes. The remaining four transcribed loci, *Maeu-UD*, *UF*, *UH* and *UN*, are expressed at low levels and whether they are class Ia or class Ib genes remains to be determined. *Maeu-UL*, *UP*, *UI* and *UJ* show no evidence of expression.

The phylogenetic relationship between marsupial class I genes is shown in Fig. 16.1. As mentioned previously, MHC class I genes evolve rapidly, which generates species-specific expansions of genes. As a result, class I genes of marsupial, avian, monotreme and eutherian species form separate clades upon phylogenetic analysis. The marsupial class I genes form three distinct clades. Clade 1 consists of class I sequences from Australian marsupials, including the three class Ia genes (*Maeu-UA*, *UB* and *UC*) of tammar wallaby as well as class I transcripts from red-necked wallaby (MaruUB*01), brushtail possum (TrvuUB*01) and Tasmanian devil (Saha-I*01 and Saha-I*12) that are presumed to be classical. Class Ia genes of the opossum, representing the American marsupial species, fall into a sister clade (clade 2) to clade 1, which is not surprising as these species belong to different orders and are separated by ∼64 million years of evolution in different environments (Nilsson et al., 2003). Clade 3 is of particular interest as it contains all the wallaby class Ib genes (*Maeu-UK*, *UM*, *UE* and *UO*), three opossum class Ib genes (*Modo-UK*, *UM* and *UE*) and two class I sequences isolated from brushtail possum (*Trvu19* and *Trvu20*). *Trvu19* and *Trvu20* are highly divergent from the possum classical *TrvuUB*01* (Holland et al., 2008b). The close phylogenetic relationships between *Trvu19* and *Trvu20* and class Ib genes of the opossum and the wallaby suggest that these two sequences may represent class Ib genes of the brushtail possum. It is interesting to notice that, within clade 3, the wallaby *Maeu-UM* and *Maeu-UK* show clear orthologous relationships with the opossum *Modo-UM* and *Modo-UK*. Such orthology between class Ib genes of distantly related species is highly unusual (Hughes and Nei, 1989). Furthermore, the existence of a *UK* orthologue in the brushtail possum (Trvu19) suggests that all Australian marsupials may share this class Ib locus. This long-term conservation of *UM* and *UK* loci throughout marsupial species may be an indication of a certain critical, marsupial-specific function of these genes, such as providing immunological protection to the marsupial pouch young, which are born immuno-naïve (Siddle et al., 2009).

16.3 Marsupial MHC Class II Genes

The MHC class II molecule consists of an α chain and a β chain, which are encoded by an α chain gene (*A*) and a β chain gene (*B*) located in the MHC. MHC class II genes also evolve in the manner of gene "birth and death", but at a slower rate compared to class I genes (Hughes and Nei, 1990; Nei et al., 1997). For this reason, unlike class I genes, which usually show no clear orthology between species, class II genes from different species group together into several gene families. Five class II gene families have been identified in eutherian mammals, designated *DM*, *DN*, *DP*,

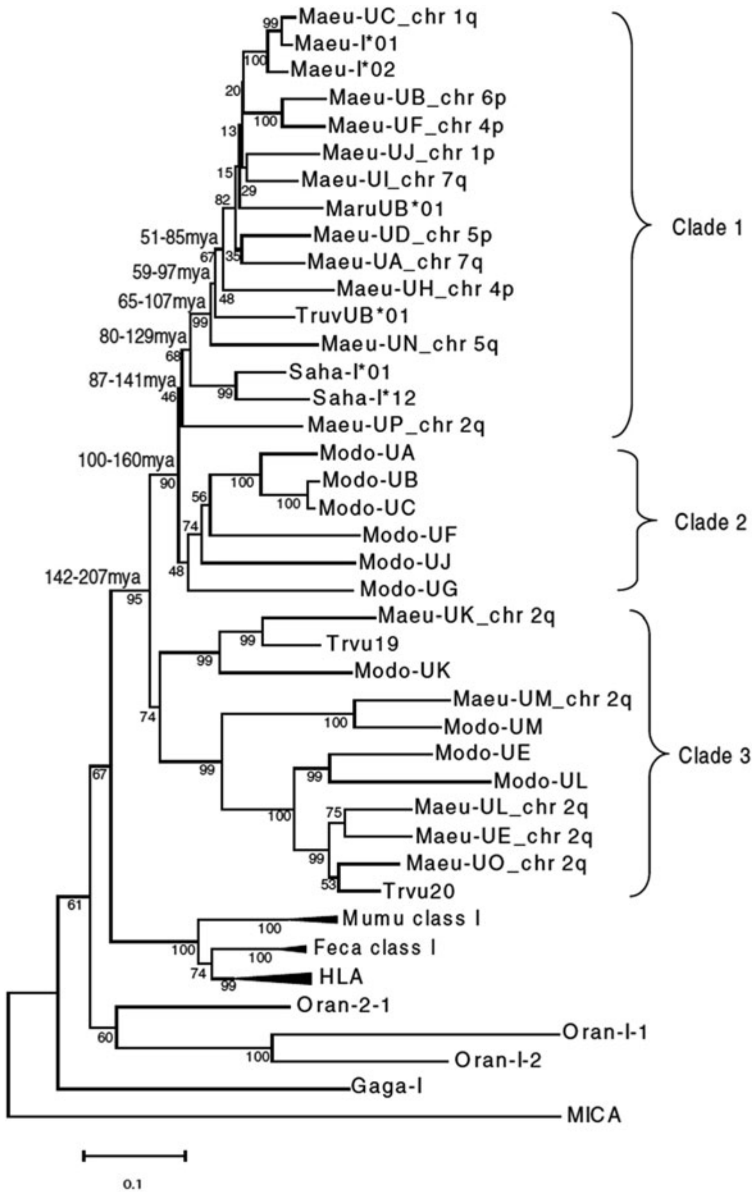


Fig. 16.1 Phylogenetic analysis of the relationship between marsupial class I genes (Siddle et al., 2009). Estimated divergence times are shown adjacent to the relevant node. Maeu, Maru, Trvu, Saha, Modo, Mumu, Feca, HLA, Oran and Gaga represent sequences from the tammar wallaby, *red*-necked wallaby, brushtail possum, Tasmanian devil, *gray* short-tailed opossum, mouse, domestic cat, human, platypus and chicken, respectively

DQ and *DR*, and the copy number of genes of each family can vary considerably between species. In marsupials, there has been a debate upon the nomenclature of MHC class II gene families. The nomenclature of marsupial *DAB* and *DBB* genes was introduced in the first article to report marsupial class II β sequences, where Schneider and co-workers (1991) suggested that marsupial class II β genes represent new gene families that are not orthologous to eutherian β gene families. This suggestion was challenged by several authors, who argued that the *DAB* genes should belong to the eutherian *DRB* family (Stone et al., 1999; Lam et al., 2001a). However, an intensive phylogenetic analysis using full-length marsupial class II sequences from the opossum, red-necked wallaby and brushtail possum, supports the distinction of *DAB* and *DBB* from eutherian class II gene families (Belov et al., 2004). Since then, the *DA* and *DB* nomenclature has been reinstated and adopted in marsupial MHC class II studies.

Immunohistological studies have demonstrated the expression of MHC class II molecules in several marsupial species, including the Brazilian white-belly opossum (*Didelphis albiventris*) (Coutinho et al., 1993, 1994), koala (Canfield et al., 1996), northern brown bandicoot (*Isoodon macrourus*) (Cisternas and Armati, 2000) and the endangered red-tailed phascogale (*Phascogale calura*) (Old et al., 2006). At the sequence level, MHC class II genes or cDNA transcripts have been isolated and characterized in a number of Australian and American marsupial species (Table 16.2). Until now, four class II gene families have been defined in marsupial species – *DA*, *DB*, *DC* and *DM*. Marsupial *DM* genes, which are orthologous to the eutherian nonclassical class II *DM*, have been reported in gray short-tailed opossum (O'hUigin et al., 1998; Gouin et al., 2006a) and tammar wallaby (Siddle et al., 2009). *DC* is a third marsupial-specific class II gene family that has only been identified in the opossum, while the classical *DA* and *DB* genes have been characterized in most marsupial species studied. Similar to eutherian class II genes, the copy number of *DA* and *DB* genes varies greatly between marsupial species as a result of species-specific gene expansion. For instance, there is one pair of *DA* genes (*DAA* and *DAB*) in the opossum (Belov et al., 2006), whereas there are up to 11 *DAB* loci in the tammar wallaby (Personal Communication), one *DAA* and at least three *DAB* in the brushtail possum (Holland et al., 2008c), at least five *DAB* genes in the Tasmanian devil (Personal Communication), and at least four and two *DAB* in the Brazilian gracile mouse opossum (*Gracilinanus microtarsus*) and gray slender mouse opossum (*Marmosops incanus*) respectively (Meyer-Lucht et al., 2008). Similarly, the number of *DB* loci also varies between species, with four in the opossum (Belov et al., 2006), at least four β chain genes in the tammar wallaby (Cheng et al., 2009a), and at least two α chain and two β chain genes in the brushtail possum (Holland et al., 2008c).

16.4 Marsupial Class III Genes

The MHC class III region contains a number of genes involved in the inflammatory response and includes members of the tumour necrosis family. Forty-eight class III genes have been identified in the opossum MHC (Belov et al., 2006). In

Table 16.2 Marsupial MHC class II genes and sequences (gene names shown in italic)

Species	*Locus*/Sequence	Description	Accession number	References
Gray short-tailed opossum	*DMA and DMB*	Nonclassical	CH465496	O'hUigin et al. (1998)
	DAA, *DBA1*, *DBB1*, *DBA2*, *DBB2*, *DCA* and *DCB*	Presumably classical	(opossum scaffold_42)	Belov et al. (2006)
	DAB	Originally named as *DRB*	AF010497 (*DAB*)	Stone et al. (1999)
Red-necked wallaby	MaruDAB1	*DAB* sequence, complete cDNA	M81624	Schneider et al. (1991)
	MaruDBB	*DBB* sequence, complete cDNA	M81625	
	MaruDRA	α chain complete cDNA	U18109	Slade and Mayer (1995)
	MaruDNA	α chain complete cDNA	U18110	
Tammar wallaby	*DNA*-like gene sequences	α1 and α2 domain	L12121–4	Slade et al. (1994)
	MaeuDBB*01–04	*DBB* sequences	AY438038–41	Browning et al. (2004)
	MaeuDBB*05–46	*DBB* sequences, β1 and β2 domain	FJ041154–95	Cheng et al. (2009a)
Brushtail possum	TrvuDAB2, TrvuDAB3	*DAB* sequences, complete cDNA, originally named as *DRB*	AF312029–30	Lam et al. (2001a)
	TrvuDBB	*DBB* sequence, complete cDNA	AY271265	Belov et al. (2004
	TrvuDAA*01	*DAA* sequences, α1 domain	EU500871–5	Holland et al. (2008c)
	TrvuDAB*0101–0116	*DAB* sequences, β1 domain	EU500876–94	
	TrvuDBA*01–03	*DBA* sequences, α1 domain	EU500895–906	
	TrvuDBB*0102–0105	*DBB* sequences, β1 domain	EU500907–11	
	TrvuDAB*0117–0127	*DAB* sequences, β1 domain	EU849585–95	Holland et al. (2008a)
Tasmanian devil	SahaDAB*1–6	*DAB* sequences, complete cDNA	EF591102–7	Siddle et al. (2007b)
	3297DABT1–4	*DAB* sequences from tumour	EF591116, 3, 4, 1	Siddle et al. (2007a)
Brazilian gracile mouse opossum	Grmi-DAB*01a, 01b, 02–46	*DAB* sequences, β1 domain	EU350150–96	Meyer-Lucht et al. (2008)
Gray slender mouse opossum	Main-DAB*01–08	*DAB* sequences, β1 domain	EU350142–9	Meyer-Lucht et al. (2008)

the tammar wallaby, a 165 kb BAC (bacterial artificial chromosome) clone containing the tammar class III region has been sequenced and 14 class III genes (*MCCD1*, *BAT1*, *ATP6V1G2*, *NFKBIL1*, *LTA*, *TNF*, *LTB*, *NCR3*, *AIF1*, *BAT2*, *BAT3*, *APOM* and *C6orf47*) have been characterized (Cross et al., 2005; Deakin et al., 2006). The organization and sequence of these marsupial class III genes are highly conserved compared to their eutherian orthologues, with some genes remaining clustered for over 450 million years, predating the divergence of mammals from fish (Deakin et al., 2006). Details of the marsupial cytokine genes that are located in the MHC class III region are reviewed by Morris et al. (Chapter 18).

16.5 Genetic Organization of Marsupial MHC

The MHC is a mosaic of stretches of conserved framework genes, serving as landmarks in comparative genomic analyses, and stretches of nonconserved class I and class II genes, which evolve rapidly through species-specific duplications and diversification (reviewed in Kumanovics et al., 2003; Kelley et al., 2005). In most vertebrate species studied, the overall organization of the MHC is well conserved, with class I, class II and class III genes clustered together. It has been suggested that clustering of MHC genes is functionally significant (Trowsdale and Powis, 1992). However, the exception of bony fishes, whose class I, class II and class III regions are not linked (Fig. 16.2), demonstrates that linkage is not obligatory for gene function (Sato et al., 2000). Comparisons between eutherian and non-mammalian MHCs reveal a major difference in gene organization. The eutherian MHC is organized in the order of class I-III-II regions, whereas in the non-mammalian MHC, class I and class II regions are adjacent (reviewed in Kelley et al., 2005).

The opossum MHC has been physically mapped to the centromeric region of opossum Chromosome 2q (Belov et al., 2006). It is similar to the eutherian MHC in both size (3.95 Mb) and gene complexity (114 genes). However, there is a major difference between the opossum and eutherian MHC – within the eutherian MHC, class I and class II regions are separated by the class III region and class I genes are interspersed among a set of framework genes; in the opossum MHC, class I genes are located in the class II region instead of the framework region, giving rise to a class I/II region, which also contains the antigen processing genes (Fig. 16.2). Such organization resembles that seen in non-mammals. In contrast to class I and class II loci, the gene organization in class III and framework regions is highly conserved compared to that of eutherian mammals. Based on these findings, a MHC model of a therian ancestor of marsupials and eutherians has been constructed, in which class I and class II loci are located together (Fig. 16.2) (Belov et al., 2006). It was also deduced that a relocation of class I genes from the class II region to the framework region took place after the divergence of marsupials and eutherians, but prior to the divergence of primates and rodents.

The tammar wallaby MHC is also located on Chromosome 2q (Deakin et al., 2007; Siddle et al., 2009). The genetic organization of the tammar wallaby MHC shows some similarity with that of the opossum MHC as it has a region that

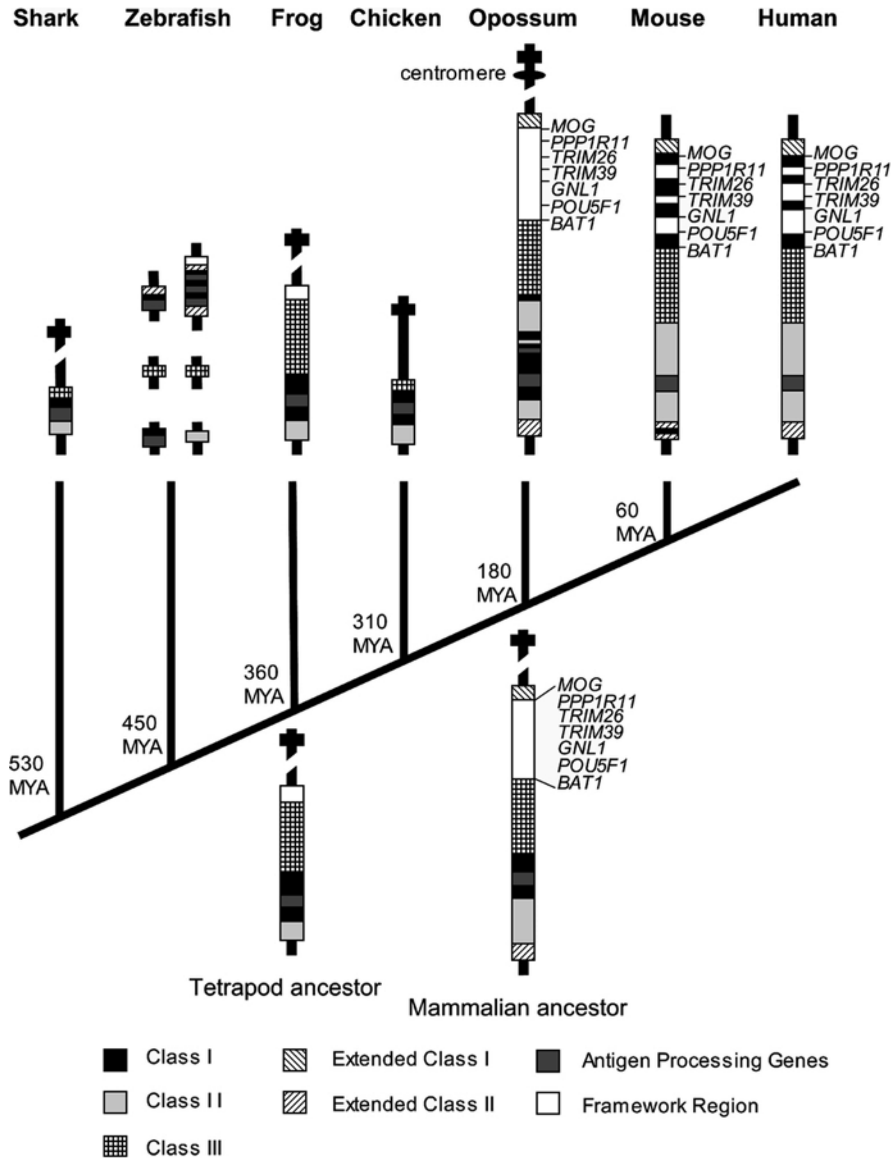


Fig. 16.2 A model of the evolution of the mammalian MHC (modified from Belov et al., 2006). The size and gene complexity of opossum MHC is similar to that of eutherian mammals, but the organization resembles that seen in non-mammals by a classI/II region

contains interspersed class I, class II and antigen processing genes, indicating that this organization is characteristic of the marsupial MHC. However, there is a unique feature of the tammar wallaby MHC – instead of clustering in the MHC region, 15 class I genes are scattered throughout the genome in ten different chromosomal

locations, across six different chromosomes. Of the 15 tammar class I loci, six are located within the MHC region, while the other nine are un-linked to the MHC. In addition, it has been found that the six MHC-linked class I loci are either non-classical or not expressed, while all un-linked class I genes share high levels of sequence identity and include the three putatively classical class I (*Maeu-UA*, *UB* and *UC*). It has been proposed that the Kangaroo Endogenous RetroViral elements (KERVs) flanking the MHC un-linked class I genes may have been involved in the duplication and movement of these class I loci. Further studies on the MHC of other marsupials are needed to determine whether class I loci are similarly distributed in other species. Since the relocation events have been roughly estimated to have taken place between 107 and 15 million years ago by molecular dating, we predict that class I loci that are un-linked to the MHC also exist in many other marsupial species. In fact, in the opossum, there are two class I genes (*Modo-UB* and *UC*) that have been mapped to the telomere of Chromosome 2p, distant from the MHC at Chromosome 2q centromere (Belov et al., 2006). These two genes are highly similar to the opossum classical class I gene and have been found expressed. These MHC un-linked classical class I genes of marsupials have again raised questions regarding the functional significance of clustering of MHC genes.

16.6 The Level of MHC Diversity in Marsupials

MHC gene diversity, referring to the genetic variation and polymorphism of MHC molecules, is a key factor in the successful defence of a host immune system against pathogen infection. In an individual, such genetic diversity is generated by the expression of multiple functional MHC genes and/or polymorphisms in these expressed genes, particularly in the peptide binding region (PBR). A substantial variety of MHC molecules enable a population to recognize an extensive range of pathogenic antigens (reviewed in Potts and Slev, 1995).

High levels of class I diversity have been observed in most marsupial species studied. Based on skin graft experiments, Stone and co-workers (1997) suggested that the opossum has a high degree of class I polymorphism and rejects allogeneic skin grafts in a manner similar to, though more slowly than, the eutherian mammals. Consistently, high variability at class I loci has been reported for the brushtail possum (Holland et al., 2008b) and tammar wallaby (Siddle et al., 2009), supported by sequencing data.

It has been proposed that marsupials have low MHC class II variability (McKenzie and Cooper, 1994). This suspicion primarily came from the results of mixed lymphocyte reaction experiments in several marsupial species. Mixed lymphocyte reactions are a cellular assay that is used to detect differences in MHC alleles between individuals. No or extremely low mixed lymphocyte response was observed in the opossum (Wilson et al., 1976; Infante et al., 1991; Stone et al., 1997), koala (Wilkinson et al., 1992) and wallaby (Cooper and McKenzie, 1997), suggesting limited diversity of class II molecules in these species. It was also

suspected that the lack of class II diversity in marsupials might be related to their unique reproductive characteristics (Infante et al., 1991; Cooper and McKenzie, 1997).

However, recent evidence challenges these hypotheses. RFLP evidence of abundant genetic variation at class II loci has been reported in the opossum (Stone et al., 1999). More convincingly, recent sequencing analyses in several marsupials have demonstrated that the class II diversity is actually high in the examined species, including the brushtail possum *DAB* (Holland et al., 2008a), Brazilian gracile mouse opossum *DAB* (Meyer-Lucht et al., 2008) and tammar wallaby *DBB* (Cheng et al., 2009a) and *DAB* (Personal communication). A comparative analysis between 44 tammar wallaby *DBB* sequences, five brushtail possum *DBB* sequences and 650 human *DRB* (*DRB1*, *DRB3*, *DRB4* and *DRB5*) alleles at polymorphic PBR amino acid residues revealed that the level of genetic variation at class II loci of the wallaby is comparable to that of humans (Table 16.3). Although the possum *DBB* appears to be less diverse, data have shown that this species has highly polymorphic *DAB* loci (Holland et al., 2008a). As discussed earlier, different marsupial species have expanded different class II gene families. Furthermore, balancing selection, a form of natural selection favouring polymorphism, has been observed in these marsupial class II genes through the investigation of the non-synonymous/synonymous substitution patterns in the PBR of these sequences. Such selection pressure is imposed on the MHC genes by the ever-changing pathogenic environment and helps generate and maintain MHC diversity during evolution.

Table 16.3 Comparison of the level of class II genetic variation between marsupial *DBB* and the human *DRB* (Cheng et al., 2009a)

	Number of amino acid substitutions		
Amino acid position	Tammar wallaby *DBB*	Brushtail possum *DBB*	Human *DRB*
13	3	2	7
26	4	2	3
28	4	2	4
30	5	1	6
37	5	2	6
38	3	2	3
47	3	2	2
57	2	1	4
67	4	2	3
70	2	3	3
71	4	3	4
74	4	3	5
78	2	2	2
85	3	3	2
86	5	4	2
89	2	1	1
90	2	1	1

In addition to sequencing analysis, microsatellite loci that are located in the MHC region have also been applied in the investigation of MHC diversity in marsupials. Due to their linkage to MHC genes and ease of use, MHC-linked microsatellites have been increasingly employed as proxies for inferring the level of polymorphism at MHC loci (Aguilar et al., 2004; Doxiadis et al., 2007; Santucci et al., 2007; de Groot et al., 2008; Cheng et al., 2009a). Up to now, ten class I-linked and two class II *DBB*-linked microsatellites have been isolated and characterized in the tammar wallaby, some of which can also be used cross-species (Cheng et al., 2009a, b). Consistent with sequencing results obtained from their linked MHC genes, genotyping data at these microsatellites showed high levels of genetic variation at MHC loci in the tammar wallaby.

16.6.1 MHC Diversity and Wildlife Health

Because of the significant role of MHC genes in immune response and disease resistance, the level of MHC diversity can be an important indicator of the immunological fitness of wild populations (Edwards and Potts, 1996). Populations with low MHC polymorphism will have limited capability to respond to changes in the environment, such as new diseases (O'Brien and Evermann, 1988). The example of Tasmanian devil and devil facial tumour disease (DFTD) has vividly demonstrated that a high level of genetic variation at MHC loci is crucial for wildlife population health.

Triggered by the newly emerged and fast spreading DFTD (see Bender, Chapter 23), a great deal of effort has been put into characterizing the immune system of devils in recent years. DFTD cells are transferred between individuals as allografts without inducing immune rejection (Pearse and Swift, 2006). Investigations on devil MHC genes, especially the class I genes, which play a central role in the self/nonself recognition and graft rejection, have provided valuable insights into understanding this unusual disease. Using mixed lymphocyte culture, single strand conformation polymorphism (SSCP) and gene sequencing approaches, it has been shown that Tasmanian devil populations have low levels of genetic variation at MHC class I loci (Siddle et al., 2007a). As the host shares similar class I molecules with the tumour cells, the host immune system fails to recognize the DFTD cells as foreign tissue. It is believed that a lack of sufficient MHC diversity is the most likely explanation for the spread of DFTD in devil populations. A recent large-scale MHC diversity study of devil populations has revealed variation in MHC class I copy number in Tasmanian devils and raises hopes that some animals have MHC types that are sufficiently different from those of DFTD cells to allow immunological recognition of tumour cells (Siddle et al., 2010).

MHC genetic diversity studies have also been carried out in the koala, with diversity studies in class II gene families in different koala populations being evaluated in our laboratory (Jobbins et al., unpublished data). The study of MHC class II genes in the koala is particularly important, because infectious diseases, such as chlamydial conjunctivitis and lower urogenital tract disease, are highly prevalent and have been

major causes of morbidity and mortality in koala populations (Brown et al., 1987; Hemsley and Canfield, 1997).

Reduced MHC diversity in wild populations can be caused by genetic bottlenecks, such as in the Tasmanian devil, which has undergone population crashes in recent history and has a low overall genetic diversity level (Jones et al., 2004), and through genetic drift, especially in populations that are confined to islands (Frankham, 1997; Miller and Lambert, 2004). There are many populations of marsupial species located on offshore islands in Australia. Investigations into MHC diversity in these populations will facilitate the evaluation of disease susceptibility and genetic fitness of the population. Such studies have been carried out on tammar wallabies on Kangaroo Island in South Australia. Results have suggested that this wallaby population is highly genetically diverse at both MHC class I and class II loci and capable of responding to new disease threats (Cheng et al., 2009a; Siddle et al., 2009).

16.6.2 MHC Diversity and Immunocontraception

Immunocontraception is a method of fertility control that occurs through the generation of an autoimmune response to a self-antigen resulting in infertility. Immunocontraceptive approaches are currently being developed to control the number of brushtail possums in New Zealand (Cowan, 2000). The brushtail possum was introduced from Australia and is now considered as a major invasive pest in New Zealand. Possums are not only causing extensive damage to native plants and animals, but are also the main wildlife host for bovine tuberculosis and are spreading the infection to farm animals (Coleman and Caley, 2000). Genetic variation can lead to the variability of responses to immunocontraceptive vaccines between individuals (Cooper and Larsen, 2006). Therefore, together with the development of immunocontraceptive vaccines, a series of in-depth studies have been carried out to investigate the genetic variation of possum MHC genes (Holland et al., 2008a–c). High levels of MHC diversity were observed at possum class I and class II loci and this variation needs to be taken into account during the design of immunocontraceptive vaccines. To any vaccine, some individuals in the population are likely to be resistant. These individuals will be able to breed, rapidly leading to selection for non-responders (Cooper and Larsen, 2006).

16.7 Future Perspectives

So far, the MHC has only been well characterized for two marsupial species, the gray short-tailed opossum and tammar wallaby. The difference in gene organization between these species highlights the importance of comparative studies. Further investigations of the MHC from other marsupial species will improve our understanding of the marsupial MHC and the evolutionary history of MHC gene families. Currently, two Tasmanian devil BAC libraries are under construction and

will be used to characterize the devil MHC. MHC diversity studies will continue to expand our understanding of wildlife health and conservation, especially for marsupial species or populations constrained to small islands or threatened by disease. Future studies will need to focus on characterization of MHC gene function, with Tasmanian devil specific anti-MHC antibodies currently being developed (Siddle et al., unpublished data).

Acknowledgments This work was funded by the Australian Research Council. We thank our fabulous collaborators Stephan Beck (University College London), Janine Deakin (Australian National University), Rob Miller (University of New Mexico) and Michelle Baker (University of New Mexico) for their contributions to this work.

References

Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK (2004) High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. Proc Natl Acad Sci USA 101:3490–3494.

Baker ML, Miller RD (2007) Evolution of mammalian CD1: marsupial CD1 is not orthologous to the eutherian isoforms and is a pseudogene in the opossum *Monodelphis domestica*. Immunology 121:113–121.

Baker ML, Melman SD, Huntley J, Miller RD (2009) Evolution of the opossum major histocompatibility complex: evidence for diverse alternative splice patterns and low polymorphism among class I genes. Immunology 128:e418–e431.

Belov K, Lam MKP, Colgan DJ (2004) Marsupial MHC class II beta genes are not orthologous to the eutherian beta gene families. J Hered 95:338–345.

Belov K, Deakin JE, Papenfuss AT, et al. (2006) Reconstructing an ancestral mammalian immune supercomplex from a marsupial Major Histocompatibility Complex. PLoS Biology 4:e46.

Brown AS, Girjes AA, Lavin MF, Timms P, Woolcock JB (1987) Chlamydial disease in koalas. Aust Vet J 64:346–350.

Browning TL, Belov K, Miller RD, Eldridge MDB (2004) Molecular cloning and characterization of the polymorphic MHC class II DBB from the tammar wallaby (*Macropus eugenii*). Immunogenetics 55:791–795.

Canfield PJ, Hemsley S, Connolly J (1996) Histological and immunohistological study of the developing and involuting superficial cervical thymus in the koala (*Phascolarctos cinereus*). J Anat 189:159–169.

Cheng Y, Siddle HV, Beck S, Eldridge MDB, Belov K (2009a) High levels of genetic variation at MHC class II DBB loci in the tammar wallaby (*Macropus eugenii*). Immunogenetics 61:111–118.

Cheng Y, Wakefield M, Siddle HV, et al. (2009b) Isolation and characterization of 10 MHC Class I-associated microsatellite loci in tammar wallaby (*Macropus eugenii*). Mol Ecol Resour 9:346–349.

Cisternas PA, Armati PJ (2000) Immune system cell markers in the northern brown bandicoot, Isoodon macrourus. Dev Comp Immunol 24:771–782.

Coleman J, Caley P (2000) Possums as a reservoir of bovine TB. In: Montague TL (ed) The Brushtail Possum: Biology, Impact and Management of an Introduced Marsupial. Manaaki Whenua, Lincoln.

Cooper DW, Larsen E (2006) Immunocontraception of mammalian wildlife: ecological and immunogenic issues. Reproduction 132:821–828.

Cooper DW, McKenzie LM (1997) Genetics of tammar wallabies. In: Saunders N, Hinds L (eds) Marsupial Biology: Recent Research, New Perspectives. UNSW Press, Sydney.

Coutinho HB, King G, Sewell HF, et al. (1993) Immunocytochemical study of Peyer's patches follicular-associated epithelium in the marsupial, Didelphis albiventris. Dev Comp Immunol 17:537–548.

Coutinho HB, Nogueira JC, King G, et al. (1994) Immunocytochemical study of the ontogeny of Peyer's patches in the Brazilian marsupial Didelphis albiventris. J Anat 185:347–354.

Cowan E (2000) Biological control of possums. In: Montague TL (ed) The Brushtail Possum: Biology, Impact and Management of an Introduced Marsupial. Manaaki Whenua, Lincoln.

Cresswell P, Ackerman AL, Giodini A, Peaper DR, Wearsch PA (2005) Mechanisms of MHC class I-restricted antigen processing and cross-presentation. Immunol Rev 207:145–157.

Cross JGR, Harrison GA, Coggill P, et al. (2005) Analysis of the genomic region containing the tammar wallaby (*Macropus eugenii*) orthologues of MHC class III genes. Cytogenet Genome Res 111:110–117.

Daly K, Church WB, Nicholas K, Williamson P (2007) Comparative modeling of marsupial MHC class I molecules identifies structural polymorphisms affecting functional motifs. J Exp Zool 307A:611–624.

de Groot N, Doxiadis GGM, de Vos-Rouweler AJM, de Groot NG, Verschoor EJ, Bontrop RE (2008) Comparative genetics of a highly divergent DRB microsatellite in different macaque species. Immunogenetics 60:737–748.

Deakin JE, Papenfuss AT, Belov K, et al. (2006) Evolution and comparative analysis of the MHC Class III inflammatory region. BMC Genomics 7:281.

Deakin JE, Siddle HV, Cross JGR, Belov K, Graves JAM (2007) Class I genes have split from the MHC in the tammar wallaby. Cytogenet Genome Res 116:205–211.

Doxiadis GGM, de Groot N, Claas FHJ, Doxiadis IIN, van Rood JJ, Bontrop RE (2007) A highly divergent microsatellite facilitating fast and accurate DRB haplotyping in humans and rhesus macaques. Proc Natl Acad Sci USA 104:8907–8912.

Edwards SV, Potts WK (1996) Polymorphism of genes in the major histocompatibility complex: implications for conservation genetics of vertebrates. In: Smith TB, Wayne RK (eds) Molecular Genetic Approaches in Conservation. Oxford University Press, Oxford.

Frankham R (1997) Do island populations have less genetic variation than mainland populations? Heredity 78:311–327.

Gouin N, Deakin JE, Miska KB, et al. (2006a) Linkage mapping and physical localization of the major histocompatibility complex region of the marsupial Monodelphis domestica. Cytogenet Genome Res 112:277–285.

Gouin N, Wright AM, Miska KB, et al. (2006b) Modo-UG, a marsupial nonclassical MHC class I locus. Immunogenetics 58:396–406.

Hemsley S, Canfield PJ (1997) Histopathological and immunohistochemical investigation of naturally occurring chlamydial conjunctivitis and urogenital inflammation in koalas (Phascolarctos cinereus). J Comp Pathol 116:273–290.

Holland OJ, Cowan PE, Gleeson DM, Chamley LW (2008a) High variability in the MHC class II DA beta chain of the brushtail possum (*Trichosurus vulpecula*). Immunogenetics 60: 775–781.

Holland OJ, Cowan PE, Gleeson DM, Chamley LW (2008b) Identification of novel major histocompatibility complex class I sequences in a marsupial, the brushtail possum (*Trichosurus vulpecula*). Immunogenetics 60:609–619.

Holland OJ, Cowan PE, Gleeson DM, Chamley LW (2008c) Novel alleles in classical major histocompatibility complex class II loci of the brushtail possum (*Trichosurus vulpecula*). Immunogenetics 60:449–460.

Houlden B, Greville W, Sherwin W (1996) Evolution of MHC class I loci in marsupials: characterization of sequences from Koala (*Phascolarctos cinerus*). Mol Biol Evol 13:1119–1127.

Hughes AL, Nei M (1989) Evolution of the major histocompatibility complex: independent origin of nonclassical class I genes in different groups of mammals. Mol Biol Evol 6:559–579.

Hughes AL, Nei M (1990) Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. Mol Biol Evol 7:491–514.

Infante AJ, Samples NK, Croix DA, Redding TS, VandeBerg JL, Stone WH (1991) Cellular immune response of a marsupial, Monodelphis domestica. Dev Comp Immunol 15:189–199.
Jones ME, Paetkau D, Geffen E, Moritz C (2004) Genetic diversity and population structure of Tasmanian devils, the largest marsupial carnivore. Mol Ecol 13:2197–2209.
Kelley J, Walter L, Trowsdale J (2005) Comparative genomics of major histocompatibility complexs. Immunogenetics 56:683–695.
Klein J (1986) Natural History of the Major Histocompatibility Complex. Wiley, New York.
Kumanovics A, Takada T, Lindahl KF (2003) Genomic organization of the mammalian MHC. Annu Rev Immunol 21:629–657.
Lam MKP, Belov K, Harrison GA, Cooper DW (2001a) Cloning of the MHC class II DRB cDNA from the brushtail possum (*Trichosurus vulpecula*). Immunol Letters 76:31–36.
Lam MKP, Belov K, Harrison GA, Cooper DW (2001b) An Mhc class I gene in the Australian brushtail possum (*Trichosurus vulpecula*). Immunogenetics 53:430–433.
LaPlante ES, Burrell R, Watre AL, Taylor DL, Zimmermann B (1969) Skin allograft studies in the pouch of the young opossum. Transplantation 7:67–72.
Mayer WE, Williams NS, O'hUigin C, Vincek V, Zaleska-Rutczynska Z, Klein J (1993) Class I major histocompatibility complex genes of the red-necked wallaby, *Macropus rufogriseus*. Mol Phylogenet Evol 2:23–30.
McKenzie LM, Cooper DW (1994) Low MHC class II variability in a marsupial. Reprod Fert Dev 6:721–726.
Meyer-Lucht Y, Otten C, Puttker T, Sommer S (2008) Selection, diversity and evolutionary patterns of the MHC class II DAB in free-ranging Neotropical marsupials. BMC Genetics 9:39.
Miller HC, Lambert DM (2004) Genetic drift outweighs balancing selection in shaping post-bottleneck major histocompatibility complex variation in New Zealand robins (Petroicidae). Mol Ecol 13:3709–3721.
Miska KB, Miller RD (1999) Marsupial Mhc class I: classical sequences from the opossum, *Monodelphis domestica*. Immunogenetics 50:89–93.
Miska KB, Wright AM, Lundgren R, et al. (2004) Analysis of a marsupial MHC region containing two recently duplicated class I loci. Mamm Genome 15:851–864.
Nei M, Gu X, Sitnikova T (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci USA 94:7799–7806.
Nilsson MA, Gullberg A, Spotorno AE, Arnason U, Janke A (2003) Radiation of extant marsupials after the K/T boundary: evidence from complete mitochondrial genomes. J Mol Evol 57:S3–S12.
O'Brien S, Evermann J (1988) Interactive influence of infectious disease and genetic diversity in natural populations. Trends Ecol Evol 3:254–259.
O'hUigin C, Sultmann H, Tichy H, Murray BW (1998) Isolation of Mhc class II DMA and DMB cDNA sequences in a marsupial: the gray short-tailed opossum (Monodelphis domestica). J Mol Evol 47:578–585.
Old JM, Carman RL, Fry G, Deane EM (2006) The immune tissues of the endangered red-tailed phascogale (Phascogale calura). J Anat 208:381–387.
Parham P, Strominger JL (eds) (1982) Histocompatibility Antigens Structure and Function. Receptors and Recognition Series B, Vol 14. Chapman and Hall, New York.
Pearse AM, Swift K (2006) Allograft theory: transmission of devil facial-tumour disease. Nature 439:549–549.
Pieters J (2000) MHC class II-restricted antigen processing and presentation. Adv Immunol 75:159–208.
Potts WK, Slev PR (1995) Pathogen-based models favoring MHC genetic diversity. Immunol Rev 143:181–197.
Rowlands DT (1976) Immunity in prototherian and metatherian mammals. In: Marchalonis JJ (ed) Comparative Immunology. Blackwell Scientific, London, pp. 367–389.
Santucci F, Ibrahim KM, Bruzzone A, Hewit GM (2007) Selection on MHC-linked microsatellite loci in sheep populations. Heredity 99:340–348.

Sato A, Figueroa F, Murray BW, et al. (2000) Nonlinkage of major histocompatibility complex classI and classII loci in bony fishes. Immunogenetics 51:108–116.

Schneider S, Vincek V, Tichy H, Figueroa F, Klein J (1991) MHC class II genes of a marsupial, the red-necked wallaby (*Macropus rufogriseus*): identification of new gene families. Mol Biol Evol 8:753–766.

Siddle HV, Deakin JE, Baker ML, Miller RD, Belov K (2006) Isolation of major histocompatibility complex Class I genes from the tammar wallaby (*Macropus eugenii*). Immunogenetics 58:487–493.

Siddle HV, Kreiss A, Eldridge MDB, et al. (2007a) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA 104:16221–16226.

Siddle HV, Marzec J, Cheng Y, Jones M, Belov K (2010) MHC gene copy number variation in Tasmanian devils: implications for the spread of a contagious cancer. Proc Biol Sci. Published online March 10, 2010 doi:10.1098/rspb.2009.2362.

Siddle HV, Sanderson C, Belov K (2007b) Characterization of major histocompatibility complex class I and class II genes from the Tasmanian devil (*Sarcophilus harrisii*). Immunogenetics 59:753–760.

Siddle HV, Deakin JE, Coggill P, et al. (2009) MHC-linked and un-linked class I genes in the wallaby. BMC Genomics 10:310.

Slade RW, Mayer WE (1995) The expressed class II alpha-chain genes of the marsupial major histocompatibility complex belong to eutherian mammal gene families. Mol Biol Evol 12:441–450.

Slade RW, Hale PT, Francis DI, Graves JAM, Sturm RA (1994) The marsupial MHC: the tammar wallaby, *Macropus eugenii*, contains an expressed DNA-like gene on chromosome 1. J Mol Evol 38:496–505.

Stone WH, Samples NK, Kraig E, VandeBerg JL (1987) Definition of class I antigens of the MHC of a marsupial (*Monodelphis domestica*). Arch Zootec 36:219–235.

Stone WH, Manis GS, Hoffman ES, Saphire DG, Hubbard GB, VandeBerg JL (1997) Fate of allogeneic skin transplantations in a marsupial (*Monodelphis domestica*). Lab Anim Sci 47:283–287.

Stone WH, Bruun DA, Fuqua C, et al. (1999) Identification and sequence analysis of an Mhc class II B gene in a marsupial (*Monodelphis domestica*). Immunogenetics 49:461–463.

Stroynowski I, Lindahl KF (1994) Antigen presentation by non-classical class I molecules. Curr Opin Immunol 6:38–44.

Trowsdale J, Powis SH (1992) The Mhc: relationship between linkage and function. Curr Opin Genet Dev 2:492–497.

Wilkinson R, Kotlarski I, Barton M (1992) Koala lymphoid cells: analysis of antigen-specific responses. Vet Immunol Immunop 33:237–247.

Wilson DB, Fox DH, Rowlands DT (1976) Proliferative reactivity of opossum peripheral blood leukocytes to allogeneic cells, mitogens and specific antigens. Transplantation 21:164–167.

Yadav M, Waring H, Stanley NF (1974) Effect of thymectomy on skin allograft survival in a macropod marsupial *Setonix brachyurus*. Transplantation 17:30–36.

Chapter 17
Marsupial Immunoglobulin and T Cell Receptor Genomics

Michelle L. Baker, Xinxin Wang, and Robert D. Miller

Abstract The development of marsupial genomic resources has provided a wealth of information for the study of the development and function of the adaptive immune system in these remarkable mammals. Here is reviewed the organization, content, and evolutionary history of the genes encoding the antigen specific receptors of the adaptive immune system, the immunoglobulins and T cell receptors. Included is one of the more unexpected outcomes of marsupial genomics, the discovery of a previously unknown TCR chain, not found in eutherian mammals. In addition, how the genomic complexity of the Ig and TCR genes influences their expression and its relationship to what is known about marsupial immune responses is also presented. Although this review focuses primarily on work done for *Monodelphis domestica*, research on other marsupial species is also included.

Keywords Genomics · Immunity · Immunoglobulin · Marsupial · T cell receptor

17.1 Introduction

The evolution of the gnathostomata, or jawed vertebrates, was accompanied by a number of innovations in the adaptive immune system, the hallmarks of which are antigen specific responses resulting not only in the control and elimination of pathogens but also the establishment of immune memory. Unique to the jawed vertebrates are two antigen receptor systems that mediate specific recognition of antigens: the immunoglobulins and the T cell receptors. Immunoglobulins (Ig) are synthesized by B lymphocytes and are expressed both as cell surface molecules (the B cell receptor [BCR]) where they participate in antigen recognition and cell activation, and also as secreted forms (antibodies) where they directly mediate pathogen

R.D. Miller (✉)
Department of Biology, Center for Evolutionary and Theoretical Immunology, The University of New Mexico, Albuquerque, NM 87131, USA
e-mail: rdmiller@unm.edu

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_17, © Springer Science+Business Media B.V. 2010

clearance. T cell receptors (TCR) on the other hand exist solely as recognition and activating receptors on the surface of T lymphocytes. The mode of antigen recognition also differs between Ig and TCR. Ig generally bind intact, native antigen structures, whereas TCR bind antigens that have been processed into typically small peptide fragments that are presented on the cell surface bound by molecules of the Major Histocompatibility Complex (MHC).

The protein domains of both Ig and TCR that directly contact antigens are called the variable (V) domains. A unique feature of these domains is that they are encoded by exons that exist as incomplete or fragmented gene segments in the germ-line genome (Tonegawa, 1983). The gene segments, named V, D (diversity), and J (joining) genes, are assembled into a complete exon through somatic DNA rearrangements in developing lymphocytes, the so-called V(D)J recombination. This process is dependent on the recombination activating gene (RAG) recombinase system, which is an endonuclease complex that recognizes specific nucleotide sequences, called the recombination signal sequences (RSS), that flank each gene segment (reviewed in Sekiguchi et al., 2004). The possible combinations of gene segments, along with variation in nucleotides at the junctions between recombined segments, results in the typical mammalian Ig and TCR loci being able to encode on the order of 10^{15}–10^{17} different receptor specificities. It is this vast potential receptor repertoire that is the basis for having clonally diverse B and T cells, providing a broad recognition coverage of rapidly evolving pathogens.

The V, D, and J gene segments, along with the exon encoding the constant (C) regions of the Ig and TCR, are generally found organized in two possible ways: translocon or cluster. The translocon organization is the most common and the V, D, and J gene segments exist in distinct arrays as $(V)_n - (D)_n - (J)_n - C$. This organization is found in all conventional TCR genes analyzed so far, and the Ig genes of bony fish and tetrapods (Flajnik, 2002). Based on recent analysis of the genomic organization TCRγ chain in a cartilaginous fish it appears that the translocon is the ancestral form for TCR (Chen et al., 2009). For clarification, the term "conventional" TCR is meant to denote the four TCR chains that are known to be present in all jawed-vertebrates, α, β, γ and δ, and to distinguish them from the atypical TCRμ, found only so far in marsupials and monotremes (Parra et al., 2007, 2008). Genes with the cluster organization have the gene segments organized in cassettes of $(V - [D] - J - C)_n$ with multiple, sometimes hundreds of such cassettes or clusters being found throughout the genome. This organization has mostly been found in the Ig genes of cartilaginous fishes, and the number of D segments per cluster can vary (Flajnik, 2002).

Although clearly sharing a common origin, Ig and TCR have followed distinct evolutionary paths, resulting in contrasting degrees of conservation and plasticity between the two receptor systems. For example, two isotypes of TCR have been identified in all jawed vertebrates: the αβTCR, composed of a heterodimer of an alpha and beta chain, and the γδTCR composed of gamma and delta chains. The conservation of these two isotypes across the whole of jawed vertebrates may be a reflection of their use solely as receptors for antigen recognition. Only recently have TCR isotypes other than the conventional α, β, γ, and δ chains been discovered.

These include the TCRμ chain found in marsupials and monotremes, but apparently lacking in eutherian mammals, as well as the NAR-TCR chain in cartilaginous fishes (Baker et al., 2005a; Criscitiello et al., 2006; Parra et al., 2007, 2008).

In contrast to the relatively conserved nature of the TCR, the Ig have evolved a wider array of isotypes, some of which appear to be specific to particular vertebrate lineages. The typical Ig is composed of four proteins, two identical heavy chains and two identical light chains, with the isotype of the secreted antibody being determined by its heavy chain. Differences in the C regions of the heavy chains impart different functional roles on the isotype of the secreted antibody. Some isotypes, such as IgM, have been conserved throughout gnathostome evolution, whereas others such as IgG and IgE are lineage specific (mammalian in this case). IgD, on the other hand, is an apparently ancient Ig isotype but appears to have a spotty distribution across vertebrate lineages (Ohta and Flajnik, 2006).

In addition to there being a greater diversity of effector functions associated with Ig heavy chain C regions, there is also greater variation in the genetic contributions to diversity of binding specificities in antibodies among species. Notably there is an association between genomic organization and complexity of the Ig genes and patterns of B cell ontogeny (reviewed in Butler, 1997). In rabbits and chickens for example, V(D)J recombination alone is not sufficient to produce a diverse primary antibody repertoire and in both these species developing B cells migrate to a secondary site (the appendix in rabbits and the Bursa of Fabricius in chickens) where the rearranged V(D)J segments are further mutated before the B cells enter the primary pool. In contrast mice and humans are able to generate diverse antibody specificities solely through V(D)J recombination due to a diverse and complex array of gene segments (reviewed in Butler, 1997). Therefore, even lineages as closely related as lagomorphs (rabbits) and rodents (mice) can be distinctly different in how their primary antibody repertoires are constructed and these are a direct reflection of the genomic content and organization of their Ig genes.

The altricial nature of the newborn marsupial makes them ideal models to study early development in the immune system and the evolution of maternal immunity. Some cross-reactive antibodies directed against evolutionarily conserved peptide sequences present in human T and B cells and a limited number of marsupial specific reagents have been used successfully for immunohistochemistry on marsupial tissues and cells from adults (Jones et al., 1993; Wilkinson et al., 1995; Hemsley et al., 1995; Baker et al., 1999) and developing marsupials (Coutinho et al., 1993, 1995; Canfield et al., 1996; Baker et al., 1999; Cisternas and Armati, 2000). Overall, these studies have demonstrated a pattern of T and B cell distribution in lymphoid tissues similar to that of eutherian mammals. However, overall such research has been limited by a scarcity of marsupial specific reagents.

The completion of the first marsupial genome project along with extensive analysis of expressed Ig and TCR chains has provided a near complete picture of the genomic contributions to Ig and TCR diversity and expression in the gray short-tailed opossum, *Monodelphis domestica* (Mikkelsen et al., 2007; Parra et al., 2008, 2009; Wang et al., 2009). The level of detail available helps put this species on par with more developed eutherian model organisms such as mice and humans.

In addition, the Ig and TCR genes of several other marsupial species have been analyzed at least at the level of cDNA clones providing the opportunity to make some generalizations regarding the evolution and function of these receptors in metatherians.

17.2 Marsupial B Cell Development, Antibody Responses, and Ig Genetics

Some of the earliest experiments performed on marsupial immune systems were measuring antibody responses (Kalmutz, 1962). This is not surprising given that the methods for detection and quantification of serum Ig were developed by the early twentieth century. Early studies of antibody responses in marsupials were consistent with differences in both the kinetics and magnitude of primary and secondary antibody responses compared with eutherian mammals. Prolonged primary antibody responses have been reported in *M. domestica*, *Setonix brachyurus* and *Trichosurus vulpecula*, with elevated IgG responses lasting 37, 26 and 15 weeks respectively (Yadav, 1971; Croix et al., 1989; Deakin et al., 2005). Secondary responses have been described as being delayed and lower in magnitude compared with those of eutherian mammals. In addition, the presence of a substantial amount of IgM in the secondary response supports poor isotype switching in marsupials (Croix et al., 1989; Stanley et al., 1972; Wilkinson et al., 1992a; Shearer et al., 1995). More recently, studies using variations on the traditional immunization protocols used for eutherian mammals have reported more comparable antibody responses to those of eutherian mammals (Kitchener et al., 2002; Deakin et al., 2005; McLelland et al., 2005; Asquith et al., 2006).

17.2.1 The Immunoglobulin Heavy Chain Locus

From the analysis of cDNA sequences it was determined that marsupials have a typically mammalian repertoire of Ig isotypes. Four Ig heavy chain (IgH) isotypes (M, G, E, and A) have been described in marsupials (Miller et al., 1998; Aveskogh and Hellman, 1998; Aveskogh et al., 1999; Belov et al., 1999a, b, c; Daly et al., 2007). The only isotype that appears to be missing is IgD, which is found in the other two mammalian lineages, the eutherians and the monotremes, and some non-mammalian lineages (Ohta and Flajnik, 2006; Wang et al., 2009).

Prior to the completion of the *M. domestica* genome, the locations of the genes encoding the IgH chains (*IGH*) were localized to chromosome (chr) 1 using in situ hybridization on metaphase chromosomes (Deakin et al., 2006a). The availability of the opossum whole genome sequence has provided a more detailed view of the genomic landscape surrounding the *IGH* locus and its overall organization. The loci on the centromeric side of *IGH* share a high degree of conserved synteny between opossum chr 1 and human and mouse chr 14 and 12, respectively (Fig. 17.1). Genes

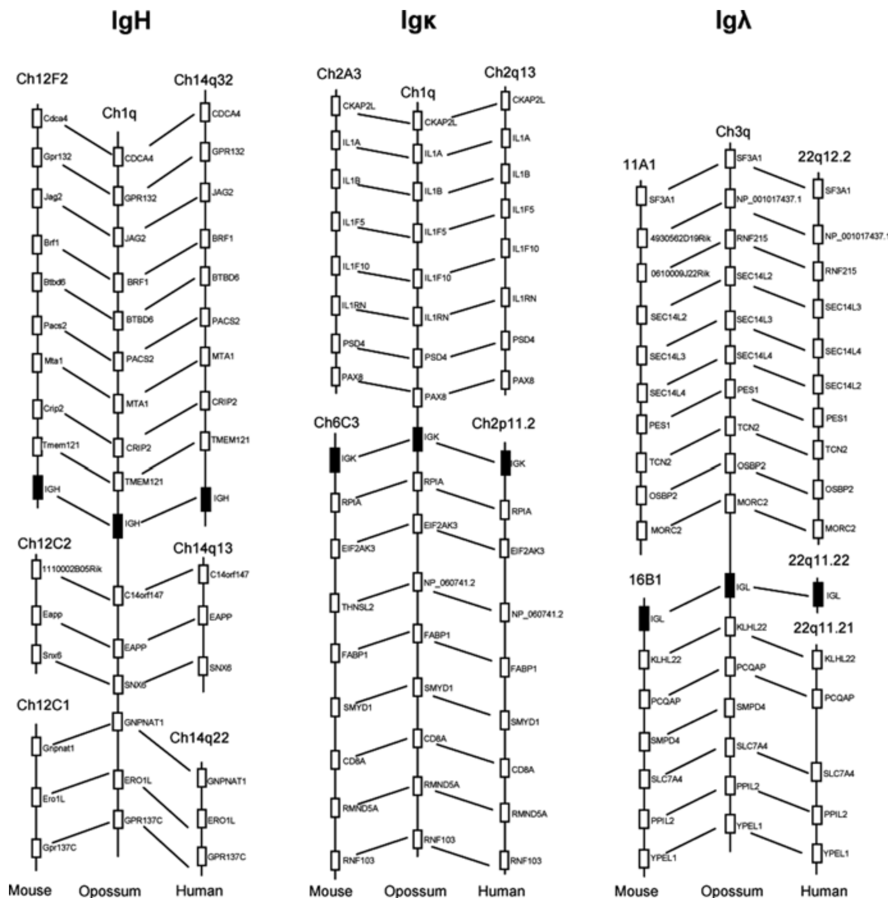


Fig. 17.1 Chromosome maps of the regions containing the *IGH* (*right*), *IGK* (*center*), and *IGL* (*left*) loci, comparing the opossum (*M. domestica*) with the regions in human and mouse bearing conserved synteny. The maps are oriented with the opossum chromosome oriented as centromeric at the *top* and telomeric at the *bottom*. Gene lists and locations used are based on the current Ensembl annotations for opossum, human, and mouse genome assemblies (www.ensembl.org)

on the telomeric side of opossum *IGH* are also found on human chr 14 and mouse chr 12. This is not contiguous, conserved synteny since, in contrast to opossum, human and mouse *IGH* is the most telomeric locus on their respective chromosomes. The overall architecture of the opossum *IGH* locus is fairly typical for mammalian orthologues with a translocon organization (Fig. 17.2). This pattern is only broken by the presence of two D segments interspersed amongst the V segments and a duplication of a region containing three J segments and partial, pseudogene copy of the IgM constant regions (Fig. 17.2, Wang et al., 2009). This latter duplication contains an insertion of a long interspersed (LINE) element. Insertion of retro-elements, LINE and endogenous retroviral (ERV) elements, has clearly contributed to the evolution of the *IGH* locus in the opossum, particularly in the region encoding

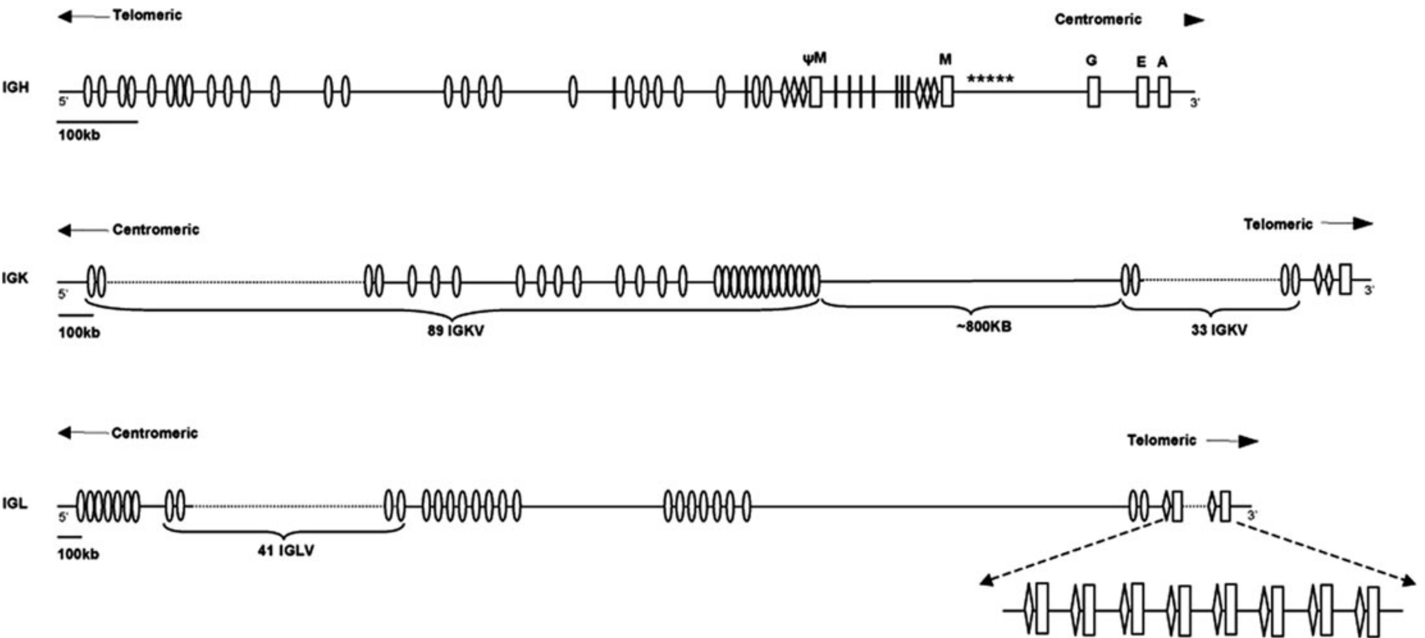


Fig. 17.2 Gene maps of the *IGH, IGK,* and *IGL* loci. *Vertical ovals* indicate V gene segments, *vertical lines* indicate D segments, *vertical diamonds* indicate J segments, and *vertical rectangles* indicate C regions. For the *IGH* locus, C region genes are distinguished based on isotype (*M*, *G*, *E*, and *A*) by letters above the rectangle. ψ indicates pseudo-gene. *Asterisks* indicate region rich in L1 and ERV elements immediately 3′ of the functional IgM constant region exons. *Brackets below* the IGK locus indicate the two large clusters of V gene segments separated by a region devoid of these coding sequences. An expanded view of the eight J–C clusters in the *IGL* locus are shown below the main line

the constant genes (Wang et al., 2009). They are not only associated with both an unusual duplication that resulted in a pseudogene copy of IgM but also may have contributed to the loss of IgD in *M. domestica* (see below). The exons encoding the constant regions are also in the conserved order of 5′-M > G > E > A-3′ (Fig. 17.2).

From the analysis of Ig heavy chain transcripts in the opossum and other marsupial species there appeared to be only a single IgG. This is in contrast to predictions made from analysis of serum Ig in *M. domestica* and other marsupials (Bell et al., 1974; Bell, 1977; Shearer et al., 1995). Analysis of the whole genome assembly revealed only a single set of exons encoding conserved IgG constant regions. The genomic organization and content are consistent with what was predicted from the cDNA analysis of there being three extra-cellular constant region domains and a single hinge region encoding exon in marsupial IgG (Aveskogh and Hellman, 1998; Belov et al., 1999a). The discrepancy between the serum protein data and the genomic data may have a number of possible explanations. First Ig heavy chains are often glycosylated and it is possible that there are distinct patterns of glycosylation that made it appear that there was more than one IgG subclass based on molecular weight. Alternatively, the methods used to purify the serum IgG may have led to false conclusions. The majority of these experiments used Staphylococcal Protein A (SpA) based purification that is expected to preferentially enrich for IgG through binding to its C domain. However, marsupial IgM has also been shown to bind SpA, so it is possible that this purification method inadvertently enriches for multiple heavy chain isotypes. SpA has also been shown to act as a super-antigen for human Ig containing clan III V domains (Silverman and Goodyear, 2002). Since all marsupial VH are also clan III it is possible that SpA is binding the V domains of some marsupial Ig rather than the C domains (Baker et al., 2005b; Wang et al., 2009).

The exons encoding IgD are typically immediately downstream of those encoding IgM and IgD and IgM can be co-expressed in some species through alternative mRNA splicing. A thorough search of the *M. domestica* genome sequence failed to identify any coding sequences with homology to IgD (Wang et al., 2009). Furthermore, the region of the opossum genome expected to contain IgD contains two large repetitive insertions, each over 15 kb in length composed of a mixture of LINE and ERV elements (Wang et al., 2009). Whether a similar insertion is present in other marsupials isn't known. How old these insertions are and if they contributed to the loss of IgD in the opossum lineage are also not known. Nonetheless, it is pretty clear that the absence of IgD in the opossum is due to a gene loss given the ancient nature of this isotype and its presence in the other mammalian lineages (Ohta and Flajnik, 2006; Wang et al., 2009).

The diversity of expressed IgH V domains (VH) has been published for two marsupial species, *M. domestica* and the brushtail possum *T. vulpecula* (Miller et al., 1998; Aveskogh et al., 1999; Baker et al., 2005b). In addition the VH repertoire of the tammar wallaby *Macropus eugenii* has also been extensively sampled at the cDNA level (Fig. 17.3). Several general conclusions could be drawn from the analysis of the diversity of marsupial Ig cDNA clones. The first is all expressed marsupial VH belong to a single evolutionary clan, called clan III. This is an ancient and

Fig. 17.3 Phylogenetic analysis of VH gene sequences from marsupials and other representative mammalian and non-mammalian species. The Neighbor Joining method of Saitou and Nei contained within the MEGA package was used for this analysis (Kumar et al., 2004). A VH sequence from the horn shark, *Heterodontus francisci* (Hefr) is used as the outgroup. On the *right*, the three major clans of VH genes are indicated as I, II and III. The single large marsupial specific VH clade

ubiquitous clan and is the most conserved of all the VH clans so it is not surprising that they are present in marsupials (Tutter and Riblet, 1989). What is surprising is that all marsupials species studied so far have only clan III VH genes and they are all highly similar, forming a "marsupial-only" monophyletic clade (Baker et al., 2005b). This is not due to the analyses being performed on species that are closely related. The divergence times for *M. domestica, T. vulpecula,* and *M. eugenii* are on par with that of eutherian lineages such as humans and mice that have more diverse VH repertoires (Baker et al., 2005b; Das et al., 2008).

One of the more unusual features in the opossum *IGH* locus is the presence of a partially germ-line joined V gene, designated VH3.1 (Wang et al., 2009). This gene segment is longer than a typical V gene by 14 codons and appears to be a V segment already fused to a D segment in the germ-line DNA. The sequence and genomic organization of this VH gene segment has been confirmed by direct sequencing of *M. domestica* genomic DNA to ensure that it is not an assembly artifact (see GenBank accession no. EU592040). Typical Ig and TCR V gene segments are composed of two exons separated by a short intron. Exon 1 encodes most of the leader (L) sequence and exon 2 encodes most of the extracellular V domain and ends at the RSS. The partially germ-line joined VH3.1 gene segment retains both the intron and RSS and the latter contains a 12 base pair spacer typical of DH segments, rather than the 23 bp spacer of a VH segment RSS. VH3.1, therefore, appears to be the result of RAG-mediated recombination (Wang et al., 2009). Partial or fully germ-line joined gene segments have been found in the Ig loci of cartilaginous and bony fishes and birds, however this is the first such Ig gene segment to be described in a mammal (Kokubu et al., 1988; Reynaud et al., 1989; Ventura-Holman and Lobb, 2002). Their origin is thought to be due to ectopic activation of V(D)J recombination in germ cells (Lee et al., 2000). Full or partially germ-line joined V genes have been shown to contribute to the antibody repertoires in non-mammalian species. In sharks they are used in the early neonatal antibody repertoire (Rumfelt et al., 2001). In chicken they contribute by being a source of sequence variation by modifying the primary antibody repertoire through gene conversion mechanisms (Reynaud et al., 1989). Interestingly the structure of VH3.1 resembles that of V pseudogenes found in the chicken *IGH* locus, which are also partially germline joined V to D gene segments (Reynaud et al., 1989). VH3.1 appears to be completely functional, with an open reading frame (ORF) and canonical RSS (Wang et al., 2009). Whether it contributes to antibody diversity in the opossum is not known, however, and transcripts containing VH3.1 in adult opossum splenic RNA have not been found

Fig. 17.3 (continued) is also indicated and has high bootstrap value. The two large clades containing tammar wallaby and opossum VH1 are also *bracketed*. Other species included are: the Northern brown bandicoot *Isoodon macrourus* (Isma), brushtail possum *Trichosurus vulpecula* (Trvu), *M. domestica* (Modo) VH that are not VH1, Virginia opossum *Didelphis virginiana* (Divi), mouse *Mus musculus* (Mumu), human *Homo sapien* (Hosa), cow *Bos taurus* (Bota), and sheep *Ovis aries* (Ovar)

(unpublished observation). So far there also does not appear to evidence of gene conversion influencing the opossum repertoire either.

17.2.2 The Ig Light Chain Loci

Marsupials, like all mammals investigated so far, have two Ig light chain isotypes: kappa (Igκ) and lambda (Igλ). Light chain gene sequences have been reported for *M. domestica, T. vulpecula,* and *M. eugenii* (Lucero et al., 1998; Miller et al., 1999: Belov et al., 2001, 2002a; Baker et al., 2005b; Daly et al., 2007). The *IGK* and *IGL* loci have been localized to the long arms of chromosomes 1 and 3, respectively, in *M. domestica* (Deakin et al., 2006a). The region of opossum chr 3 containing *IGL* has conserved synteny with human chr 22 (Fig. 17.1). This conservation is not as well maintained in mice, however, where chr 16 has conserved synteny with human chr 22 and opossum chr 3 on the telomeric side of *IGL*, and 11 on the centromeric side. The *IGK* locus in opossum is also at a break point in synteny with eutherian mammals. This region of opossum chr 1 is conserved with regions on both short and long arms of human chr 2 and with mouse chr 2 and 6 (Fig. 17.1).

The overall genomic organization of the opossum *IGK* and *IGL* loci are typical of mammalian light chain genes and match what was predicted from analysis of cDNA clones (Miller et al., 1999; Lucero et al., 1998; Wang et al., 2009). This includes a single C region and only two J gene segments in *IGK*, and eight J and C pairs in *IGL* (Fig. 17.2). The V gene segments in *IGK* are organized in two large clusters, with inverted reading orientation, separated by an approximately 800 kb region. Inverted V genes are not a problem since V(D)J recombination can occur through inversion rather than deletion of intervening DNA (Sekiguchi et al., 2004). This organization is similar to that seen in human *IGK* where it is clear that a large inverted duplication occurred during the evolution of this locus (Zachau, 2004). This does not appear to be what has occurred in *M. domestica*, however. Phylogenetic analysis of the V gene segments in each cluster reveal that they belong to distantly related subgroups (Miller et al., 1999; Wang et al., 2009).

In contrast to the limited diversity of marsupial VH genes, the light chain V region genes, Vκ and Vλ, are highly diverse and the different marsupial Vκ and Vλ clades intersperse amongst that of other mammalian lineages suggesting the retention of a more ancient diversity than in VH (Miller et al., 1999; Lucero et al., 1998; Baker et al., 2005b; Wang et al., 2009). This has resulted in speculation that Ig light chains contribute more to the overall antibody diversity in marsupials than do heavy chains (Baker et al., 2005b).

17.2.3 Marsupial B Cell Ontogeny and Immuno-Competence

The availability of fully annotated *IGH, IGK* and *IGL* has facilitated the analysis of both B cell ontogeny and diversity. Transcripts containing IgM heavy chains with

completely recombined V genes can be detected within the first 24 h of postnatal life in *M. domestica* (unpublished observation). This is significantly earlier than what has been reported previously for the brushtail possum *T. vulpecula*, where the earliest time-point investigated was day 10 (Belov et al., 2002b). The diversity of the IgH chains at these early time-points is limited and the V-D-J junctions are devoid of additional nucleotides added by the enzyme terminal deoxynucleotidyl transferase, which adds N nucleotides without the use of a template. N nucleotides are found in the adult repertoire (Miller et al., 1998; unpublished observations). Whether this limited diversity is due to limited numbers of B cells in the newborn, or represents a programmed bias, as has been seen in eutherian fetal antibodies, is not known. The absence of N nucleotides in early opossum IgH chains is similar to what has been seen for mouse neonatal antibodies and may be an analogous ontogenic pattern (Feeney, 1992; Benedict et al., 2000).

In contrast to IgH, light chain gene transcription is not detectable until day 8 of postnatal life in *M. domestica* (unpublished observation). This suggests that, although B cell ontogeny is initiated peri-natally, mature functional B cells are not present until the second week. These results support that mature B cells also appear later in development than mature conventional T cells in the opossum (Parra et al., 2009). These results are also consistent with early studies of antibody responses in a related marsupial species, the Virginia opossum *Didelphis virginiana*, which found antibody responses to a T dependent antigen not being detectable until the newborn opossums were at least two weeks of age (Kalmutz, 1962; Rowlands et al., 1964). Therefore, the newborn marsupial is fully dependent on maternal antibodies for protection during the first week, which, in *M. domestica*, are entirely transferred via the milk (Samples et al., 1986).

17.3 T Cell Receptor Genomics

Critical to the function of T cells are their TCR, which establish not only the antigen specificity of individual clones but also their phenotype and function. T cells that utilize αβ TCR differentiate into the well-known helper, cytotoxic (killer), and regulatory (suppressor) phenotypes. This differentiation is determined by selection for combinations of binding affinity for self-peptides and self-MHC molecules in the thymus. The random combinations of V, D and J segments encoded at the TCR loci establish these binding affinities. Hence, understanding the structure and organization of the TCR genes in any given species is critical to understanding T cell function and fate.

Marsupial T cell mediated responses have attracted attention, in part, because of a general sense that they are deficient relative to that of eutherian mammals. For example, rejection of primary skin grafts occurs with similar kinetics in marsupials and eutherians, however, marsupial second set rejection is relatively slower (Infante et al., 1991; Stone et al., 1997). Delayed type hypersensitivity (DTH) responses are also slower and weaker in magnitude than those of eutherians (Turner et al., 1972; Taylor and Burrell, 1968). In vitro mitogenic responses in all marsupials

tested to date are similar to eutherian mammals, however, mixed lymphocyte reactions (MLR) are negligible or absent (Fox et al., 1976; Wilkinson et al., 1992a, b; Stone et al., 1997, 1998; Baker et al., 1998; Baker and Gemmell, 1999; Baker et al., 1999; Turner et al., 1972; Woods et al., 2007). There is no clear explanation for the weak MLR or slow DTH responses in marsupials. Low MHC polymorphism could account for the absence of MLR observed in Tasmanian devils, which have undergone a considerable population bottleneck (Woods et al., 2007). However, in at least one other marsupial, *M. domestica*, polymorphism at classical MHC genes is comparable to human HLA, so absence of MHC diversity in this species is not a likely explanation for poor MLR (Gouin et al., 2006). In addition, antibody isotype switch in secondary immune responses in marsupials are not as robust as they are in eutherian mammals (Stanley et al., 1972; Croix et al., 1989; Wilkinson et al., 1992a; Stone et al., 1996). Since isotype switching is also T cell dependent, a number of investigators have questioned the sufficiency of T cells in humoral immune responses in marsupials as well. The availability of marsupial genomic resources has provided the means to investigate T cell development and diversity to gain some insight into marsupial T cell biology.

T cell maturation in all gnathostomes is dependent on a functional thymus, the gland-like lymphoid organ typically located in the upper thoracic region. It is in the thymus where T cells commit to a lineage, αβ or γδ, and where they undergo V(D)J recombination and selection to establish their antigen specificity and MHC restriction. A unique, and so far not fully understood, aspect of thymus biology in marsupials is the presence of an extra set of thymuses in the cervical vertebrae region in some family groups. The diprotodont species, including Burramydae, Phalangeridae, Tarsipedidae and Macropodidae, have both cervical and thoracic thymuses with the exception of the koala that possesses only a cervical thymus. The polyprotodont species, including Caenolestidae, Dasyuridae, Thylacinidae, Notoryctidae and Peramelidae, possess only a thoracic thymus, and Didelphidae marsupials have paired thoracic thymuses (Yadav, 1973; Deane and Cooper, 1988; Hubbard et al., 1991; Haynes, 2001).

Although the thymus is present in newborn marsupials it is generally undifferentiated and undergoes much of its development during postnatal life. In marsupials with both cervical and thoracic thymuses, both appear to be histologically identical and to have the same function, but the development of the thoracic thymus generally lags behind that of the cervical thymus (Kathiresan, 1969; Stanley et al., 1972). T lymphocytes differentiate in the thymus before populating other sites in the body. Peripheral lymphoid organs such as the spleen, lymph nodes and gut associated lymphoid tissue typically develop into functional lymphoid tissue only after the thymus has reached maturity. T cell responses are correspondingly low at birth, and neonates are dependent on maternal antibody for protection until their immune systems are functional (Samples et al., 1986; Hubbard et al., 1991).

Immunohistochemistry using cross-reactive antibodies on developing lymphoid tissue from marsupials has provided evidence for the appearance of T cells soon after birth. $CD3^+$ T cells have been identified in the thoracic thymus as early as

day 2 postpartum in *T. vulpecula* and *D. albiventris* (Coutinho et al., 1995; Baker et al., 1999). In contrast, the thymus of *M. eugenii* does not have CD3 positive T cells earlier than day 12 postpartum (Old and Deane, 2003). Subsequent appearance of $CD3^+$ T cells appear in the spleen by day 21 and 25 postpartum in *M. eugenii* and *T. vulpecula* respectively (Baker et al., 1999; Old and Deane, 2003). These immunohistochemistry results are somewhat consistent with recent analysis of T cell ontogeny in the opossum *M. domestica* based on expression of mature, rearranged TCR transcripts described below (Parra et al., 2009).

As with the Ig genes, the annotation of the TCR loci was greatly assisted by the published analysis of marsupial cDNAs encoding TCR chains, analysis which has already revealed considerable homology between marsupial TCRα, β, γ and δ chains to those of eutherian mammals. TCR genes encoding the α and β chains of the αβ TCR have been characterized in *M. domestica*, *T. vulpecula* and *M. eugenii*, and for the γδ TCR, the δ chain has been characterized in *M. eugenii*, *M. domestica* and *Isoodon macrourus* and the γ chain has been identified among expressed sequence tags (EST) from *I. macrourus* (Zuccolotto et al., 2000; Baker et al., 2001; Harrison et al., 2003; Baker et al., 2005a, 2007). The diversity of expressed V domains in cDNAs has been described only for the TCRδ chain from *M. eugenii* and the TCRα and β chains of developing *M. domestica*. At least three different subfamilies of *M. eugenii* Vδ genes were expressed in cDNA from mammary gland associated lymphoid tissue, consistent with a high level of V segment diversity in *M. eugenii* (Harrison et al., 2003). In *M. domestica*, a highly diverse Vα and Vβ repertoire was observed, with preferential use of V segments during development as discussed below (Parra et al., 2009).

Physical mapping of the conventional TCR loci in *M. domestica* has localized TCRα/δ to chromosome 1p, TCRβ to chr 8q, and TCRγ to chr 6q by in situ hybridization of metaphase chromosomes (Deakin et al., 2006b). The availability of the whole genome sequence of *M. domestica* has provided a more detailed view of the genomic organization of TCR loci in this marsupial, demonstrating a high degree of synteny with TCR loci from other mammals and a translocon arrangement (Figs. 17.4 and 17.5, Parra et al., 2008).

17.3.1 TCR α/δ Locus

Consistent with physical mapping of the TCR loci, the TCRδ genes in the genome assembly are clustered within the TCRα locus, resembling the organization found in other amniotes (Fig. 17.5, Chien et al., 1987; Kubota et al., 1999; Deakin et al., 2006b). The TCRα/δ locus is the largest and most complex of the conventional TCR loci. The V segments within the TCRα/δ locus are either utilized exclusively by TCRα or by both TCRα and TCRδ chains (Parra et al., 2008, 2009).

The number and complexity of D and J segments in the opossum's TCR loci also appears to be comparable to eutherian mammals and non-mammalian vertebrates (Fig. 17.5, Partula et al., 1996; Parra et al., 2008). For example, there are 53 TCRα

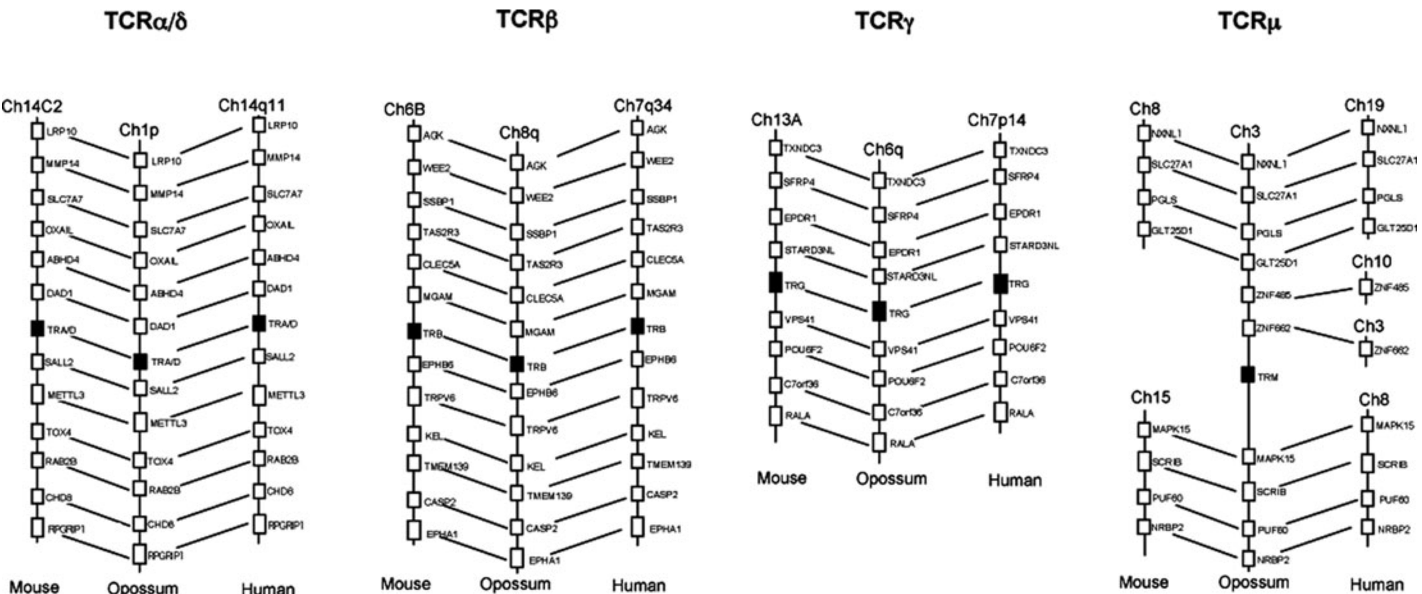


Fig. 17.4 Chromosome maps of the regions containing, *left* to *right*, the TRA/D, TRB, TRG, and TRM loci, comparing the opossum (*M. domestica*) with the regions in human and mouse bearing conserved synteny. The nomenclature is by the convention where the locus encoding the TCRα/δ chains is TRA/D, TCRβ is TRB, and TCRγ is TRG, TCRμ is TRM. The maps are oriented with the opossum chromosome oriented as centromeric at the *top* and telomeric at the *bottom*. Gene lists and locations used are based on the current Ensembl annotations for opossum, human, and mouse genome assemblies (www.ensembl.org)

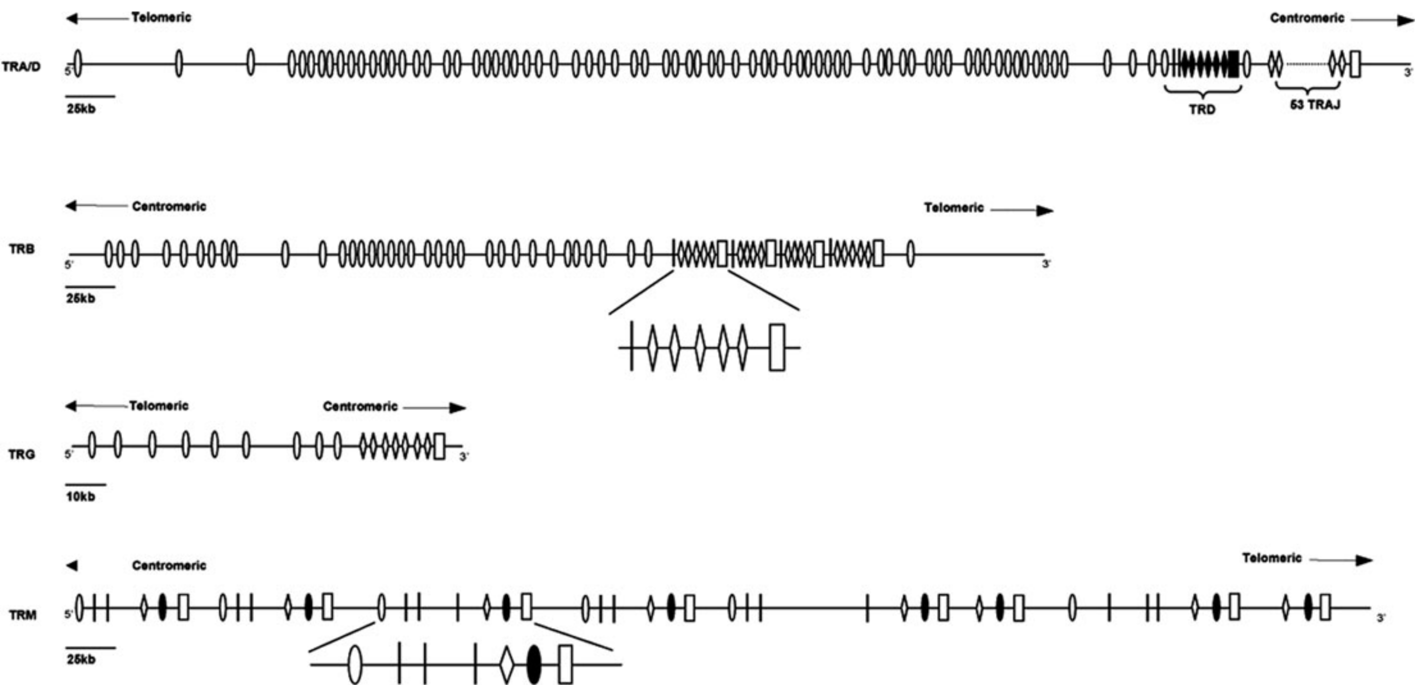


Fig. 17.5 Gene maps of the loci encoding the TCRα, β, γ, δ, and μ chains. *Vertical ovals* indicate V gene segments, *vertical lines* indicate D segments, *vertical diamonds* indicate J segments, and *vertical rectangles* indicate C regions. The nomenclature is by the convention where the locus encoding the TCRα/δ chains is TRA/D, TCRβ is TRB, TCRγ is TRG, and TCRμ is TRM. In the case of the TRA/D locus the Jα and Cα genes are indicated with open symbols and the Jδ and Cδ as filled symbols. In the TRM locus the filled oval indicates the germ-line joined gene segment Vμj. *Below* the TRB locus is an expanded view of a single TCRβ D, J, and C cluster. *Below* the TRM locus is an expanded view of a single TCRμ cluster

J segments located between the TCRδ C and TCRα C all of which appear to be functional. This high number of Jα segments is thought to confer the ability of developing T cells to undergo secondary V to J rearrangements, increasing the possibility of successfully generating a functional TCRα chain (Marolleau et al., 1988).

17.3.2 TCR β Locus

In addition to the high level of conserved synteny surrounding the TCRβ locus, the structure of the locus is also highly conserved (Figs. 17.4 and 17.5). As in humans and mice, the opossum TCRβ D, J and C segments are arranged in tandem cassettes each containing a single D, four or five J, and a single C. There are four such cassettes in the *M. domestica* TCRβ locus, all of which appear to be functional (Fig. 17.5; Parra et al., 2008). Noteworthy is the presence of a single V gene segment on the 3′ side of the terminal C region exon in the opossum (Fig. 17.5). This Vβ gene segment is in the opposite transcriptional orientation to the D, J, and C segments and can and is used by V(D)J recombination creating an inversion rather than a deletion (Parra et al., 2008). This gene segment is conserved in eutherians as well, consistent with this arrangement being ancient in therian mammals (Parra et al., 2008).

17.3.3 TCR γ Locus

The locus encoding TCRγ chains is the smallest of the conventional TCR loci in *M. domestica*, with the fewest number of V gene segments and only a single C region gene. It is also highly generic in its organization, a model for the definition of the translocon organization (Fig. 17.5).

17.3.4 Genomic Organization Influences Early Diversity of αβT Cells

Once the content and organization of the conventional TCR has been determined, both the germ-line and somatic contributions to T cell diversity can be determined during ontogeny. Such studies have been conducted using *M. domestica* as the model species. mRNA transcribed from functionally recombined TCRα and TCRβ genes are both detectable within the first postnatal day in newborn opossums, indicating that they are born with T cells that have at least reached the point of development when they can undergo selection for binding MHC molecules and self-peptides (Parra et al., 2009). This development is taking place prior to histological evidence of a clear thymus in this species (Hubbard et al., 1991). TCRδ transcripts were also detectable within the first postnatal 24 h. In contrast to the other three conventional TCR chains, TCRγ is the last to be expressed during postnatal development in *M. domestica*, not detected until the second postnatal week (Parra et al., 2009). This was a surprising finding since it suggests that αβT cells appear in ontogeny prior to

γδT cells, which is the reverse of eutherian mammals where γδT cells develop first (Allison and Havran, 1991; Parra et al., 2009). What biological significance there may be to this difference is not known.

Developing αβT cells rearrange their TCRβ chain genes prior to TCRα and analysis of the V gene segments being used in these early rearrangements indicates that their genomic position in the TCRβ locus does not influence their use. However, there is a preferential use of TCRβ gene segments that share micro-homology between the V, D and J (Parra et al., 2009). This micro-homology is evident in the rearrangements as nucleotides at the VD and/or DJ junctions that could be contributed by either germ-line gene segment. This is not unlike what has been seen in fetal Ig and TCR repertoires in some eutherian species (Feeney, 1992; Zhang et al., 1995). In contrast, micro-homology does not appear to bias the rearranged gene segments in TCRα chains. Rather, genomic position of V segments appears to have an influence in early TCRα rearrangements, with there being a bias for the most 3′ or C proximal V genes. By the second week of life both TCRα and TCRβ chains are random and diverse, and similar to adult (Parra et al., 2009).

In summary, having spatial information as well as content has revealed differences in the way V(D)J rearrangements are being biased in early developing marsupial T cells, in this case even between different TCR loci within the same cell. The bias in gene segment use early in development clearly limits the diversity of T cells, but how this influences function during the first week of postnatal life remains to be determined. The presence of completely rearranged TCR loci does not necessarily indicate that there are functionally mature T cells present. Early studies of early immuno-competence in marsupials have found T dependent responses, both humoral and cell mediated, lacking during the first week of postnatal life (Kalmutz, 1962; La Via et al., 1963; Rowlands et al., 1964; Ashman et al., 1975). This was thought to be due to lack of functional T cells at the earlier time-points, but may also be explained by limited diversity in the pool of functional T cells available in the neonate (Parra et al., 2009).

17.3.5 A Novel TCR in Marsupials

One of the more remarkable discoveries made while examining marsupial genomes, at least from an immunological perspective, was presence of a fifth TCR chain, TCRμ (Baker et al., 2005a; Parra et al., 2007). TCRμ has all the canonical amino acid residues necessary to form a heterodimer with another chain, however, what partner it pairs with remains unknown. TCRμ was originally described as a divergent TCRδ chain due to the homology of its C region (Baker et al., 2005a). However, the genes encoding this chain are a separate locus from the conventional TCR loci (Parra et al., 2007). In addition to sequence homology to other TCR, it is clear that TCRμ is a *bona fide* TCR in that it undergoes V(D)J recombination in developing thymocytes (Parra et al., 2007). Based on limited phylogenetic surveys TCRμ also appears to be common to all marsupials but is absent from all eutherian species checked so far (Parra et al., 2007, 2008). Not surprisingly then, the region that flanks

the opossum's TCRμ locus on chromosome 3 shares no conserved synteny with the genomes of eutherian mammals (Fig. 17.4; Parra et al., 2008). This lack of conservation between the chromosomal regions in the opossum that contains TCRμ genes with that of non-marsupial species has confounded uncovering the evolutionary history of this locus. This is particularly unfortunate given the apparent hybrid nature of this unusual TCR chain; TCRμ appears, based on sequence similarity, to be a hybrid between a TCR and an Ig (Parra et al., 2007). The constant region of TCRμ is most closely related to TCRδ, while the variable domain is homologous to the variable domains of immunoglobulin heavy chains. It is therefore likely that TCRμ may have been the result of a recombination between ancient TCRδ and Ig loci (Parra et al., 2008). Any evidence of such a recombination, however, appears to be lost in that the regions flanking the TCRα/δ locus in amniotes is highly conserved, showing no sign of a break in synteny (Fig. 17.4, Parra et al., 2008).

In addition to its apparent hybrid features, *M. domestica* TCRμ is also the first, and so far only, TCR to have a clear cluster style organization, rather than translocon (Fig. 17.5). The number of clusters may vary between species with the opossum having eight and the Northern brown bandicoot, *I. macrourus*, likely having only two for example (Baker et al., 2005a; Parra et al., 2008). In the opossum six of the eight clusters are complete with each containing two classes of V segments, as well as D, J and C genes. The two classes of V segments are a single non-rearranged V gene segment (Vμ) that can be somatically recombined with the D and J segments, and pre-joined V segment (Vμj) that is a complete exon that appears to be derived from V, D, and J gene segments already recombined in the germ-line DNA. The two partial clusters lack Vμ and Dμ gene segments. Phylogenetic analysis of the gene segments in each cluster reveals that the clusters are due to whole cluster duplications. Unlike the Ig genes of cartilaginous fishes, which are arranged in clusters that are scattered throughout the genome, the opossum TCRμ clusters are in a tandem array (Parra et al., 2008).

TCRμ is also unusual for a TCR in that it is expressed in at least two mRNA isoforms, designated TCRμ1.0 and TCRμ2.0. Both isoforms are transcribed in the thymus and spleen but TCRμ2.0 appears to predominate in peripheral lymphoid tissue (Parra et al., 2007). The chain predicted to be encoded by the TCRμ2.0 isoform contains two V domains; one that is somatically recombined and another that is Vμj. TCRμ1.0 contains only the pre-joined Vμj domain and is predicted to be structurally more similar to conventional TCR chains (Parra et al., 2007). The TCRμ2.0 isoform is generated when the unrearranged Vμ gene segment is recombined to the D and J segments and this exon is spliced to the Vμj exon during mRNA processing (Parra et al., 2007). The TCRμ1.0 isoform is generated by initiated transcription upstream of the Vμj exon and does not require V(D)J recombination. Given the paucity of TCRμ1.0 transcripts in peripheral lymphoid tissues it is not clear if this is a functionally translated form. Furthermore, only two of the eight opossum TCRμ clusters are capable of producing the TCRμ1.0 isoform due to in frame stops in the leader

sequence of Vμj in the other six clusters. All six complete clusters are capable of generating the TCRμ2.0 isoform, and one of the partial clusters is capable of producing the TCRμ2.0 isoform by using Vμ gene segments from upstream clusters (Parra et al., 2007). This latter observation may explain the evolutionary advantage to maintaining the clusters as a tandem array even though they can function independently. It is possible that inter-cluster V(D)J recombination provides the opportunity to generate additional diversity that would not be possible if the clusters were not closely physically linked.

Vμj is the first, and so far only germ-line joined V segment to be described in a TCR. It and the VH3.1 gene segment in the *M. domestica IGH* locus are the only germ-line joined V genes to be described in mammals for either TCR or Ig, although VH3.1 is a partially joined gene. In contrast to VH3.1, which clearly appears to be due to the action of RAG based recombination on the DNA in a germ cell, Vμj appears to be the result of retro-transposition, involving an RNA intermediate (Parra et al., 2007; Wang et al., 2009). How or where this took place is not clear, although it seems likely to also have been in a germ cell for it to be in the germ-line. The recent discovery of a TCRμ homologue in the duckbill platypus genome, lacking the germ-line joined Vμj gene segment will hopefully give some insight into its origins (Parra et al., 2008, unpublished observations).

The discovery of TCRμ in the marsupial genome illustrates not only the dynamic nature of the evolution of the antigen receptor loci but also the value of investigating the genomes of more, distantly related species. Not long ago, the TCR appeared to be fairly static with all jawed vertebrates having a similar, if not identical, complement of receptor chains. This view is more likely due to the fact that much of comparative immunology was proceeded by discoveries being made first in humans and mice followed by identifying homologues in other species. It is harder to ask the question "what does a non-model species have that is lacking in humans and mice?" Comparative genomics makes asking such questions feasible. What does TCRμ do? Now that's a tough question to answer.

17.4 Conclusion

Comparative genomics provides not only the fodder for understanding the evolution of organisms and their genomes, but also the resources for physiological studies of non-model species that do not always have as sufficient a tool-kit as would be desired. The genomic analysis of at least one marsupial species, *M. domestica*, has put this species on par with humans, mice and a very short list of other gnathostomate species in regards to being able to study their adaptive immune responses. In 1977 Ashman lamented the scarcity of research on marsupial immune systems and asked if this field would have a brighter future (Ashman, 1977). The past decade has seen amazing growth and discoveries in the studies of marsupial immunity. Given the emerging problems such as the Tasmanian devil facial tumor disease, it is hopefully not too late (McCallum and Jones, 2006).

Acknowledgments The authors acknowledge support from the U.S. National Science Foundation.

References

Allison JP, Havran WL (1991) The immunobiology of T cells with invariant gamma delta antigen receptors. Ann Rev Immunol 9:679–705.

Ashman RB (1977) Marsupial immunology: a brighter future? Dev Comp Immunol 1:283–284.

Ashman R, Keast D, Stanley NF, Waring H (1975) The immunological responses of marsupials. Am Nat 15:155–166.

Asquith KL, Kitchener AL, Kay DJ (2006) Immunisation of the male tammar wallaby (*Macropus eugenii*) with spermatozoa elicits epididymal antigen-specific antibody secretion and compromised fertilisation rate. J Reprod Immunol 69:127–147.

Aveskogh M, Hellman L (1998) Evidence for an early appearance of modern post-switch isotypes in mammalian evolution; cloning of IgE, IgG and IgA from the marsupial Monodelphis domestica. Eur J Immunol 28:2738–2750.

Aveskogh M, Pilström L, Hellman L (1999) Cloning and structural analysis of IgM (μ chain) and the heavy chain V region repertoire in the marsupial *Monodelphis domestica*. Dev Comp Immunol 23:597–606.

Baker ML, Gemmell RT (1999) Physiological changes in brushtail possums relocated from Armidale into captivity in Brisbane, Australia. J Exper Zool 284:42–49.

Baker ML, Gemmell E, Gemmell RT (1998) Physiological changes in brushtail possums, *Trichosurus vulpecula*, transferred from the wild to captivity. J Exper Zool 280:203–212.

Baker ML, Gemmell E, Gemmell RT (1999) Ontogeny of the immune system of the brushtail possum, *Trichosurus vulpecula*. Anat Record 256:354–365.

Baker ML, Rosenberg GH, Zuccolotto P, Harrison G, Deane EM, Miller RD (2001) Further characterization of T cell receptor chains of marsupials. Dev Comp Immunol 25:495–507.

Baker ML, Osterman AK, Brumburgh S (2005a) Divergent T cell receptor delta chains from marsupials. Immunogenetics 57:665–673.

Baker ML, Belov K, Miller RD (2005b) Unusually similar patterns of antibody V segment diversity in distantly related marsupials. J Immunol 174:5665–5671.

Baker ML, Indiviglio S, Nyberg AM, et al. (2007) Analysis of a set of Australian northern brown bandicoot expressed sequence tags with comparison to the genome sequence of the South American grey short tailed opossum. BMC Genomics 8:50.

Bell RG (1977) Marsupial immunglobulins: the distribution and evolution of macropod IgG2, IgM and light chain antigenic markers within the sub-class metatheria. Immunol 33:917–924.

Bell RG, Lynch NR, Turner KJ (1974) Immunoglobulins of the marsupial *Setonix brachyurus* (quokka). Immunol 27:1103–1115.

Belov K, Harrison GA, Miller RD, Cooper DW (1999a) Isolation and sequence of a cDNA coding for the heavy chain constant region of IgG from the Australian brushtail possum, *Trichosurus vulpecula*. Mol Immunol 36:535–541.

Belov K, Harrison GA, Miller RD, Cooper DW (1999b) Molecular cloning of the brushtail possum (*Trichosurus vulpecula*) immunglobulin E heavy chain constant region. Mol Immunol 36: 1255–1261.

Belov K, Harrison GA, Rosenberg GH, Miller RD, Cooper DW (1999c) Isolation and comparison of the IgM heavy chain constant region from Australian (*Trichosurus vulpecula*) and American (Monodelphis domestica) marsupials. Dev Comp Immunol 23:649–656.

Belov K, Harrison GA, Miller RD, Cooper DW (2001) Characterisation of the kappa light chain of the brushtail possum (*Trichosurus vulpecula*). Vet Immunol Immunopathol 78:317–324.

Belov K, Harrison GA, Miller RD, Cooper DW (2002a) Molecular cloning of four lambda light chain cDNAs from the Australian brushtail possum (*Trichosurus vulpecula*). Eur J Immunogenet 29:95–99.

Belov K, Nguyen M-AT, Zenger KR, Cooper DW (2002b) Ontogeny of immunoglobulin expression in the brushtail possum (*Trichosurus vulpecula*). Dev Comp Immunol 26:599–602.

Benedict CL, Gillian S, Thai T, Kearney JF (2000) Terminal deoxynucleotidyl transferase and repertoire development. Immunol Rev 175:150–157.

Butler JE (1997) Immunoglobulin gene organization and the mechanism of repertoire development. Scand J Immunol 45:455–462.

Canfield P, Hemsley S, Connolly J (1996) Histological and immunohistological study of the developing and involuting superficial cervical thymus in the koala (*Phascolarctos cinereus*). J Anat 189:159–169.

Chen H, Kshirsagar S, Jensen I, et al. (2009) Characterization of arrangement and expression of the T cell receptor γ locus in the sandbar shark. Proc Natl Acad Sci U S A 106:8591–8596.

Chien YH, Iwashima M, Kaplan KB, Elliot JF, Davis MM (1987) A new T-cell receptor gene located within the alpha locus and expressed early in T-cell differentiation. Nature 327: 677–682.

Cisternas PA, Armati PJ (2000) Immune system cell markers in the northern brown bandicoot, *Isoodon macrourus*. Dev Comp Immunol 24:771–782.

Coutinho HB, King G, Sewell HF, et al. (1993) Immunocytochemical study of Peyer's patches follicular-associated epithelium in the marsupial, *Didelphis albiventris*. Dev Comp Immunol 17:537–548.

Coutinho HB, Sewell HF, Tighe P, et al. (1995) Immunocytochemical study of the ontogeny of the marsupial *Didelphis albiventris* immune system. J Anat 187:37–46.

Criscitiello MF, Saltis M, Flajnik MF (2006) An evolutionarily mobile antigen receptor variable region gene: doubly rearranging NAR-TcR genes in sharks. Proc Natl Acad Sci U S A 103:5036–5041.

Croix DA, Samples NK, VandeBerg JL, Stone WH (1989) Immune response of a marsupial (*Monodelphis domestica*) to sheep red blood cells. Dev Comp Immunol 13:73–78.

Daly KA, Digby M, Lefèvre C, et al. (2007) Analysis of the expression of immunoglobulins throughout lactation suggests two periods of immune transfer in the tammar wallaby (*Macropus eugenii*). Vet Immunol Immunopathol 120:187–200.

Das S, Nozawa M, Klein J, Nei M (2008) Evolutionary dynamics of the immunoglobulin heavy chain variable region genes in vertebrates. Immunogenetics 60:47–55.

Deakin JE, Belov K, Curach NC, Green P, Cooper DW (2005) High levels of variability in immune response using antigens from two reproductive proteins in brushtail possums. Wild Res 32:1–6.

Deakin JE, Olp JJ, Graves JAM, Miller RD (2006a) Physical mapping of immunoglobulin loci IGH@, IGK@, and IGL@ in the opossum (*Monodelphis domestica*). Cytogenet Genome Res 114:94.

Deakin JE, Parra ZE, Graves JAM, Miller RD (2006b) Physical Mapping of T cell receptor loci (TRA@, TRB@, TRD@ and TRG@) in the opossum (*Monodelphis domestica*). Cytogenet Genome Res 112:342 K.

Deane EM, Cooper DW (1988) Immunological development in pouch young marsupials. In: Tyndale-Biscoe CH, Janssens PA (eds) The Developing Marsupial. Springer, Berlin.

Feeney AJ (1992) Predominance of VH-D-JH junctions occurring at sites of short sequence homology results in limited junctional diversity in neonatal antibodies. J Immunol 149:222–229.

Fox DH, Rowlands DT Jr, Wilson DB (1976) Proliferative reactivity of opossum peripheral blood leukocytes allogeneic cells, mitogens, and specific antigens. Transplantation 21:164–168.

Flajnik MF (2002) Comparative analyses of immunoglobulin genes: surprises and portents. Nat Rev Immunol 2:688–698.

Gouin N, Wright AM, Miska KB, et al. (2006) Modo-UG, a marsupial non-classical MHC class I locus. Immunogenetics 58:396–406.

Harrison GA, Taylor CL, Miller RD, Deane EM (2003) Primary Structure and Variation of the T-Cell Receptor δ Chain from a Marsupial, *Macropus eugenii*. Immunol Lett 88:117–125.

Haynes JI (2001) The marsupial and monotreme thymus, revisited. J Zool Lond 253:167–173.

Hemsley SW, Canfield PJ, Husband AJ (1995) Immunohistological staining of lymphoid tissue in four Australian marsupial species using species cross-reactive antibodies. Immunol Cell Biol 73:321–325.

Hubbard GB, Saphire DG, Hackleman SM, Silva MV, VandBerg JL, Stone WH (1991) Ontogeny of the thymus gland of a marsupial (*Monodelphis domestica*). Lab Anim Sci 41:227–232.

Infante AJ, Samples NK, Croix DA, Redding TS, VandeBerg JL, Stone WH (1991) Cellular immune response of a marsupial, *Monodelphis domestica*. Dev Comp Immunol 15:189–199.

Jones M, Cordell JL, Beyers A, Tse AGD, Mason DY (1993) Detection of T and B cells in many animal species using cross-reactive antipeptide antibodies. J Immunol 150:5429–5435.

Kalmutz SE (1962) Antibody production in the opossum embryo. Nature 193:851–853.

Kathiresan S (1969) Morphological studies on the thymus. Part I. Thymus of the marsupial *Trichosurus vulpecula*. Indian J Med Res 57:939–945.

Kitchener AL, Edds LM, Molinia FC, Kay DJ (2002) Porcine zonae pellucidae immunization of tammar wallabies (*Macropus eugenii*): fertility and immune responses. Reprod Fertil Dev 14:215–223.

Kokubu F, Litman R, Shamblott MJ, Hinds K, Litman GW (1988) Diverse organization of immunoglobulin VH gene loci in a primitive vertebrate. EMBO J 7:3413–3422.

Kubota T, Wang JY, Gobel TWF, Hockett RD, Cooper MD, Chen CH (1999) Characterization of an avian (*Gallus gallus domesticus*) TCR α/δ locus. J Immunol 163:3858–3866.

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163.

La Via MF, Rowlands DT, Block M (1963) Antibody formation in embryos. Science 140: 1219–1220.

Lee SS, Fitch D, Flajnik MF, Hsu E (2000) Rearrangement of immunoglobulin genes in shark germ cells. J Exp Med 191:1637–1648.

Lucero JE, Rosenberg GH, Miller RD (1998) Marsupial light chains: complexity and conservation of Igl in the opossum Monodelphis domestica. J Immunol 161:6724–6732.

Marolleau JP, Fondell JD, Malissen M, et al. (1988) The joining of germ-line Vα to Jα genes replaces the preexisting Vα-Jα complex in a T cell receptor alpha and beta positive T cell line. Cell 55:291–300.

McCallum H, Jones M (2006) To lose both would look like carelessness: tasmanian devil facial tumour disease. PLoS Biology 4:e342.

McLelland DJ, Kirkland PD, Rose KA, Dixon RJ, Smith N (2005) Serologic responses of Barbary sheep (*Ammotragus lervia*), Indian antelope (*Antilope cervicapra*), wallaroos (*Macropus robustus*), and chimpanzees (*Pan troglodytes*) to an inactivated encephalomyocarditis virus vaccine. J Zoo Wildl Med 36:69–73.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals lineage-specific innovation in coding sequences. Nature 447:167–178.

Miller RD, Grabe H, Rosenberg GH (1998) The V_H repertoire of a marsupial: *Monodelphis domestica*. J Immunol 160:259–265.

Miller RD, Bergemann ER, Rosenberg GH (1999) Marsupial light chains: IGK with four V families in the opossum *Monodelphis domestica*. Immunogenetics 50:329–335.

Ohta Y, Flajnik M (2006) IgD, like IgM, is a primordial immunoglobulin class perpetuated in most jawed vertebrates. Proc Natl Acad Sci U S A 103:10723–10728.

Old JM, Deane EM (2003) The detection of mature T- and B-cells during development of the lymphoid tissues of the tammar wallaby (*Macropus eugenii*). J Anat 203:123–131.

Parra ZE, Baker ML, Schwarz R, Deakin JE, Lindblad-Toh K, Miller RD (2007) Discovery of a new T cell receptor in marsupials. Proc Natl Acad Sci U S A 104:9776–9781.

Parra ZE, Baker ML, Hathaway J, et al. (2008) Comparative genomic analysis and evolution of the T cell receptor loci in the opossum *Monodelphis domestica*. BMC Genomics 9:111.

Parra ZE, Baker ML, Lopez AM, Trujillo J, Volpe JM, Miller RD (2009) TCRm recombination and transcription relative to the conventional TCR during postnatal development in opossums. J Immunol 182:154–163.

Partula S, de Guerra A, Fellah JS, Charlemagne J (1996) Structure and diversity of the TCR alpha-chain in a teleost fish. J Immunol 157:207–212.

Reynaud CA, Dahan A, Anquez V, Weill JC (1989) Somatic hyperconversion diversifies the single Vh gene of the chicken with a high incidence in the D region. Cell 59:171–183.

Rowlands DT, Lavia MF, Block MH (1964) The blood forming tissues and blood of the new-born opossum (*Didelphys virginiana*). II. Ontogenesis of antibody formation to flagella of Salmonella typhi. J Immunol 93:157–164.

Rumfelt LL, Avila D, Diaz M, Bartl S, McKinney EC, Flajnik MF (2001) A shark antibody heavy chain encoded by a nonsomatically rearranged VDJ is preferentially expressed in early development and is convergent with mammalian IgG. Proc Natl Acad Sci U S A 98:1775–1780.

Samples NK, VandeBerg JL, Stone WH (1986) Passively acquired immunity in the newborn of a marsupial (*Monodelphis domestica*). Am J Reprod Immunol Microb 11:94–97.

Sekiguchi J, Alt FW, Oettinger M (2004) The mechanisms of V(D)J recombination. In: Honjo T, Alt FW, Neuberger M (eds) Molecular Biology of B Cells. Academic Press, London.

Shearer MH, Robinson ES, VandeBerg JL, Kennedy RC (1995) Humoral immune response in a marsupial *Monodelphis domestica*: anti-isotypic and anti-idiotypic responses detected by species-specific monoclonal anti-immunoglobulin reagents. Dev Comp Immunol 19: 237–246.

Silverman GJ, Goodyear CS (2002) A model B-cell superantigen and the immunobiology of B lymphocytes. Clin Immunol 102:117–134.

Stanley NF, Yadav M, Waring H, Eadie M (1972) The effect of thymectomy on response to various antigens of a marsupial *Setonix brachyrus* (Quokka). Aust J Exp Biol Med Sci 50:689–702.

Stone WH, Manis GH, Hoffman ES, Saphire DG, Hubbard GB, VandeBerg JL (1997) Fate of allogeneic skin transplantations in a marsupial (*Monodelphis domestica*). Lab Anim Sci 47:283–287.

Stone WH, Bruun DA, Manis GS (1996) The immunobiology of the marsupial Monodelphis domestica. In: Stolen JS, Fletcher TC Bayne CJ, et al. (eds) Modulators of Immune Responses, The Evolutionary Trail, Breckenridge Series 2, Vol 11. SOS Publications, Fairhaven.

Stone WH, Bruun DA, Foster EB, et al. (1998) Absence of a significant mixed lymphocyte reaction in a marsupial (*Monodelphis domestica*). Lab Anim Sci 48:184–189.

Taylor DL, Burrell R (1968) The immunological responses of the North American opossum (*Didelphis virginiana*). J Immunol 101:1207–1216.

Tonegawa S (1983) Somatic generation of antibody diversity. Nature 302:575–581.

Turner KJ, Alpers MP, Wight AM (1972) Delayed hypersensitivity in the marsupial *Setonix brachyurus* (Quokka). J Immunol 108:1675–1680.

Tutter A, Riblet R (1989) Conservation of an immunoglobulin variable-region gene family indicates a specific, noncoding function. Proc Natl Acad Sci U S A 86:7460–7464.

Ventura-Holman T, Lobb CJ (2002) Structural organization of the immunoglobulin heavy chain locus in the channel catfish: the IgH locus represents a composite of two gene clusters. Mol Immunol 38:557–564.

Wang X, Olp JJ, Miller RD (2009) On the genomics of immunoglobulins in the gray, short-tailed opossum *Monodelphis domestica*. Immunogenetics 61:581–596.

Wilkinson R, Kotlarski I, Barton M (1992a) Koala lymphoid cells: analysis of antigen-specific responses. Vet Immunol Immunopathol 33:237–247.

Wilkinson R, Kotlarski I, Barton M, Phillips P (1992b) Isolation of koala lymphoid cells and their in vitro responses to mitogens. Vet Immunol Immunopathol 31:21–33.

Wilkinson R, Barton M, Kotlarski I (1995) Identification of koala lymphocytes using an anti-human CD3 antibody. Dev Comp Immunol 19:537–545.

Woods GM, Kreiss A, Belov K, Siddle HV, Obendorf DL, Muller HK (2007) The immune response of the Tasmanian devil (*Sarcophilus harrisii*) and devil facial tumour disease. EcoHealth 4: 338–345.

Yadav M (1971) The transmissions of antibodies across the gut of pouch-young marsupials. Immunology 21:839–851.

Yadav M (1973) The presence of the cervical and thoracic thymus lobes in marsupials. Aust J Zool 21:285–301.

Zachau HG (2004) Immunoglobulin kappa genes of human and mouse. In: Honjo T, Alt FW, Neuberger M (eds) Molecular Biology of B Cells. Academic Press, London.

Zhang Y, Cado D, Asarnow DM, Komori T, Alt FW, Raulet DH, Allison JP (1995) The role of short homology repeats and TdT in generation of the invariant γδ antigen receptor repertoire in the fetal thymus. Immunity 3:439–447.

Zuccolotto PD, Harrison GA, Deane EM (2000) Cloning of marsupial T cell receptor α and β constant region cDNAs. Immunol Cell Biol 78:103–109.

Chapter 18
Use of Genomic Information to Gain Insights into Immune Function in Marsupials: A Review of Divergent Immune Genes

Katrina Morris, Emily S.W. Wong, and Katherine Belov

Abstract Marsupial immune genes, including cytokines, Natural Killer cell receptors and antimicrobial peptides, have been difficult to isolate due to the wide divergence of these genes from their eutherian counterparts. Early studies using conventional laboratory techniques only isolated the most conserved genes. With the sequencing of opossum and tammar wallaby genomes, many of these divergent genes have now been identified. Knowledge of these gene sequences will enable us to study immune function in marsupials. We can develop immunological reagents for studying the immune system of healthy marsupials and look at the immune response of diseased marsupials. We can also investigate marsupial specific immune functions such as the immune protection of pouch young. Additionally, antimicrobial therapies for both humans and marsupials may be designed around some of these immune peptides.

Keywords Marsupial · Opossum · Cytokine · Antimicrobial peptide · Natural killer cell

18.1 Introduction

For many decades our understanding of the marsupial immune system has lagged well behind that of the eutherian immune system. This is due, in part, to the difficulty of isolating divergent immune genes in marsupials. Genes for immune peptides including cytokines, Natural Killer cell receptors and antimicrobial peptides are some of the most rapidly evolving genes within the vertebrate genome. Conventional laboratory techniques have only identified the most conserved immune genes. However, in the last few years, with the sequencing of the South American opossum

K. Belov (✉)
Faculty of Veterinary Science, The University of Sydney, Sydney, NSW 2006, Australia
e-mail: kbelov@usyd.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_18, © Springer Science+Business Media B.V. 2010

(*Monodelphis domestica*) (Mikkelsen et al., 2007) and tammar wallaby (*Macropus eugenii*) genomes (Baylor College of Medicine, 2008), there has been a huge growth in the number of divergent immune genes that have been identified in silico. Here we review recently identified divergent marsupial immune genes (listed in Table 18.1) and discuss their importance for functional immunology studies including development of immunological reagents, understanding marsupial specific immune functions and increased understanding of the evolution of the immune system in vertebrates. This chapter will focus on two aspects of marsupial immunology, the cytokines and the antimicrobial peptides. Natural Killer cell receptors and T cell surface peptides will also be discussed briefly. The use of sequence information of these molecules in understanding immune function and disease in marsupials will be explored.

Table 18.1 Cytokines currently sequenced in marsupials

Cytokines	Species	Method of isolation	References
TNF-α	Wallaby	RT-PCR and RACE	Harrison et al. (1999)
TNF-α	Possum	RT-PCR and RACE	Wedlock et al. (1996)
LT-α/LT-β	Wallaby	Genomic walking and RACE	Harrison and Deane (1999, 2000)
IL1β	Possum	RT-PCR and RACE	Wedlock et al. (1999b)
IL2	Opossum	In silico gene prediction	Wong et al. (2006)
IL4	Opossum	In silico gene prediction	Wong et al. (2006)
IL5	Wallaby	Isolation from lambda clone library	Hawken et al. (1999)
IL5	Opossum	In silico gene prediction	Wong et al. (2006)
IL6	Opossum	In silico gene prediction	Wong et al. (2006)
IL10	Possum	RT-PCR and RACE	Wedlock et al. (1998)
IL10	Opossum	In silico gene prediction	Wong et al. (2006)
IL12A	Opossum	In silico gene prediction	Wong et al. (2006)
IL13	Opossum	In silico gene prediction	Wong et al. (2006)
IL16	Opossum	In silico gene prediction	Wong et al. (2006)
IL19	Opossum	In silico gene prediction	Wong et al. (2006)
IL20	Opossum	In silico gene prediction	Wong et al. (2006)
IL21	Opossum	In silico gene prediction	Wong et al. (2006)
IL22	Opossum	In silico gene prediction	Wong et al. (2006)
IL24	Opossum	In silico gene prediction	Wong et al. (2006)
IL26	Opossum	In silico gene prediction	Wong et al. (2006)
IFN-α	Wallaby	PCR and Southern blotting	Harrison et al. (2003)
IFN-α	Opossum	In silico gene prediction	Wong et al. (2006)
IFN-β	Wallaby	PCR and Southern blotting	Harrison et al. (2003)
IFN-β	Opossum	In silico gene prediction	Wong et al. (2006)
IFN-γ	Opossum	In silico gene prediction	Wong et al. (2006)
IFN-κ	Opossum	In silico gene prediction	Wong et al. (2006)
Chemokines	Opossum	HMMer search	Belov et al. (2007)

Opossum (*Monodelphis domestica*), possum (*Trichosurus vulpecula*), wallaby (*Macropus eugenii*)

18.2 Cytokines

The immune system of mammals is highly complex with many different immune cells being coordinated together to enable the prompt and effective removal of anything recognised as non-self. This system involves a vast array of cells types each with unique functions and a similarly vast array of chemical signals required to coordinate the function and interaction of these cells. The adaptive immune system in mammals can be broadly divided into two branches; the cell mediated arm and the humoural arm. Cell mediated immune defense is primarily directed to intracellular pathogens, especially viruses. When viral peptides are presented on the surface of cells by Major Histocompatibility Complex (MHC) class I molecules they are recognised by cytotoxic T cells. These cytotoxic T cells have CD8 peptides on their surface to enable binding of antigen and T cell receptor. The cytotoxic T cells can then mediate destruction of infected cells. The humoural immune system is primarily directed to the removal of extracellular pathogens. Peptides from phagocytised pathogens will be presented using MHC class II molecules, which are recognised by T helper cells expressing CD4 peptides on their surface. This will lead to activation of B cells and the production of antibodies to destroy the extracellular pathogens (reviewed in Abbas and Janeway, 2000).

T helper cells are often subdivided into T helper 1 (Th1) and T helper 2 (Th2) cells. Th1 cells are required for activation of cytotoxic T cells and promote macrophage activation and so are important for the destruction of intracellular pathogens. Th2 cells stimulate B cell proliferation and immunoglobulin secretion and, as such, are important for humoural immunity. These two subsets of T helper cells are distinguished by the cytokines that they express (reviewed in Abbas et al., 1996).

Cytokines are small glycoproteins that are involved in the regulation of all aspects of the immune system. The cytokine family is a large heterogeneous group with proteins performing a diverse array of functions (reviewed in Dinarello, 2007). These proteins are unified in that they all mediate intra-cellular communication between immune cells, allowing for the highly sophisticated immune response seen in higher vertebrates. Cytokine signaling is involved in immune responses including phagocytosis, degranulation, complement mediated destruction, apoptosis and cytolysis. Cytokine molecules belong to several families based on their structure and function. Interleukins (ILs) are a highly diverse group of cytokines which are grouped together as they regulate interactions between lymphocytes and other leukocytes. Interferons (IFNs) are a group of antiviral cytokines able to interfere with viral RNA and protein synthesis. Tumour Necrosis Factors (TNFs) are produced by macrophages and T cells and are important in inflammation. Chemokines are cytokines that are able to induce directed chemotaxis in nearby cells and serve as leukocyte activators. Other cytokines including GM-CSF and TGF-β serve as growth factors and regulate leukocyte production.

Cytokines are expressed by T helper cells in order to coordinate immune response (Fig. 18.1). Th1 cells are characterised by their production of IL-2, IFN-γ and TNF-β while Th2 cells express IL-4, IL-5, IL-6, IL-10 and IL-13. Through

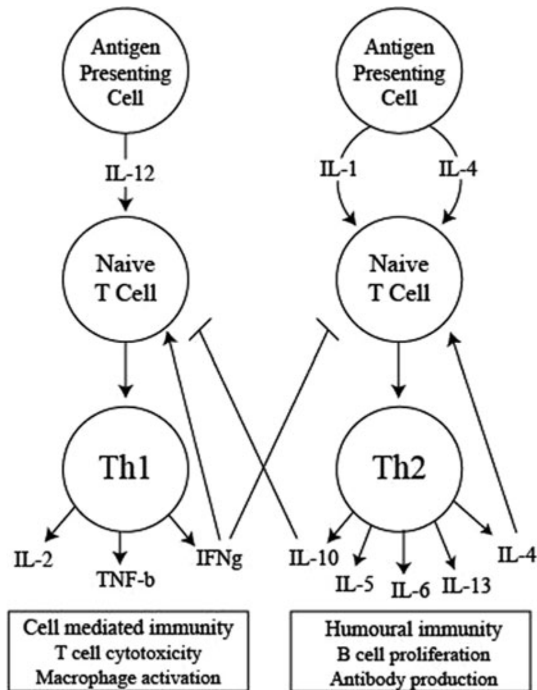


Fig. 18.1 Interleukins produced by antigen presenting cells induce naïve T cells to differentiate into either Th1 or Th2 cells. These T helper subsets produce different sets of cytokines which promote cell mediated immunity (Th1) or humoural immunity (Th2). Th1 and Th2 cells cross-inhibit each other through the production of IFN-γ and IL-10 respectively

the production of cytokines, these subsets of helper cells will cross-inhibit each other. IFN-γ produced by Th1 suppresses Th2 cells, while IL-10 and IL-10 produced by Th2 cells suppresses the functioning of Th1 cells (reviewed in Abbas et al., 1996).

Identifying which class of T helper cells is dominating is important for understanding whether cell mediated or humoural immunity dominates an immune response in a diseased animal. Being able to examine the cytokines produced during disease enables us to identify whether Th1 or Th2 cells dominate. However, investigation of marsupial cytokines, and their use in understanding disease pathology in marsupials, has been hampered due to the difficulty in identifying these genes. Isolating cytokine sequences has been problematic as a result of wide divergence of these molecules from their eutherian counterparts. These genes are rapidly evolving as they are under strong selection because they must allow for coordination of the immune response to tackle a vast array of pathogens and are also targets for immunoevasion by pathogens. Isolation of divergent cytokines has been difficult across vertebrate species. In chickens it took 7 years to isolate the *IL-2* sequence, which was eventually achieved through a large-scale EST approach and not through heterologous cloning techniques (Beck, 1998).

18.2.1 Marsupial Tumour Necrosis Factors

The first marsupial cytokines were identified using conventional molecular experimental strategies such as using degenerate primers for RT-PCR or heterologous probes for screening genetic libraries. These methods were only able to identify the most conserved cytokine sequences. For instance, some of the earliest isolated marsupial cytokine sequences were those belonging to the Tumour Necrosis Factor (TNF) family. This family includes the genes *TNF*, Lymphotoxin α (*LTA*) and Lymphototoxin β (*LTB*). These proteins have a wide range of roles in inflammation and immune regulation. The sequence for *TNF* has been isolated in both the tammar wallaby (Harrison et al., 1999) and brushtail possum (*Trichosurus vulpecula*) (Wedlock et al., 1996). The possum *TNF* cDNA was isolated by RT-PCR using primers designed from conserved regions of eutherian cDNAs. The complete cDNA sequence was isolated from lymphocyte RNA using 5′ and 3′ RACE. The tammar wallaby cDNA sequence was isolated using a similar method. As expected, the two marsupial sequences are more closely related to one another than to eutherians, with 70% sequence identity at the amino acid level. However, the wallaby shows greater sequence identity than the possum to eutherians at the amino acid level (65–70% and 56–68%, respectively). Moreover, the wallaby amino acid sequence is only slightly more similar to the possum (70%) than to eutherians (65–70%). By comparing the rates of substitution at synonymous and non-synonymous sites, Harrison et al. (1999) indicated that the possum gene appears to have been evolving more rapidly at non-synonymous sites than the wallaby gene. The reason for this observation is not clear but it has been suggested that this could be caused by positive selection in response to particular disease outbreaks (Harrison et al., 1999).

Structurally, TNF is very similar in eutherians and marsupials with important structural features being conserved (Harrison and Wedlock, 2000). This includes the conservation of the His^{15} and Tyr^{56} residues which are important for receptor binding (Cha et al., 1998). Recombinant TNF has been produced in possums and its effect on possum physiology investigated (Wedlock et al., 1999a). Recombinant TNF administered to possums caused a profound increase in the number of circulating leukocytes and increased the possums' body temperature. In eutherians, administration of TNF has also been found to increase blood leukocyte counts (Rothel et al., 1998) and cause fever (Dinarello, 1999). TNF is also being investigated for its potential use as an adjuvant in biocontrol vaccines in possums. In eutherians it has been shown that cytokines are capable of increasing the antibody response when used as an adjuvant in vaccines. A study by Wedlock et al. (1999a) demonstrated that recombinant TNF successfully enhanced antibody response when co-administered with the model antigen keyhole limpet hemocyanin (KLH).

LT-α is a cytokine which is secreted as a homotrimer of the LT-α peptide. LT-α can also form a complex on the surface of activated lymphocytes with the membrane bound LT-β peptide. Both *LTA* and *LTB* have been cloned and sequenced from the tammar wallaby (Harrison and Deane, 1999, 2000). This was achieved using a method that utilised the proximity of *LTA* and *LTB* to *TNF* in the genome.

In eutherians these genes are tightly clustered together. Using sequence information from *TNF*, a PCR based "genomic walking" approach was used to obtain partial sequences for *LTA* and *LTB*. These partial sequences were used to design primers which were then used in RACE PCR to derive the entire cDNA sequences. The LT-α and LT-β peptides share 70 and 61% sequence identity to eutherian mammals respectively. Structurally, LT-α and LT-β appear to be very similar to their eutherian counterparts. In eutherian mammals *TNF*, *LTA* and *LTB* lie in close proximity to one another within the class III inflammatory region of the MHC. These genes are located in the same position in both the opossum and tammar wallaby (Deakin et al., 2006), and therefore it is likely that these genes are located within the MHC class III in all marsupials.

18.2.2 Marsupial Interleukins

Interleukin 1β (IL-1β) was the first interleukin to be studied in marsupials. IL-1β is a strong pro-inflammatory cytokine and is also involved in cell differentiation, proliferation and apoptosis. IL-1β activity was first shown in the opossum when the IL-1β peptide was extracted from lippopolysaccharide (LPS) stimulated macrophages (Brozek and Ley, 1991). Subsequently, the cDNA encoding IL-1β in the brushtail possum was sequenced (Wedlock et al., 1999b). Primers based on conserved regions from eutherian sequences were used to amplify cDNA prepared from brushtail possum LPS-stimulated alveolar macrophages. RACE was employed in order to obtain the complete cDNA sequence. IL-1β shared 46% amino acid identity with human IL-1β. Experiments on opossum IL-1β (Brozek and Ley, 1991) and recombinant possum IL-1β (Wedlock et al., 1999b), demonstrate that there is very little serological or functional cross reactivity between the marsupial and eutherian IL-1β. The IL-1β peptide is another possible candidate for use as an adjuvant in biocontrol vaccines in possums. Recombinant possum IL-1β caused an increase in serum antibody levels to KLH when used as an adjuvant (Harrison and Wedlock, 2000).

IL-10 is a cytokine which has an immunosuppressive effect on macrophages and T cell antigen specific proliferation but also has an immunostimulatory effect on B cells. *IL10* was sequenced in the brushtail possum using a similar method to IL-1β (Wedlock et al., 1998). IL-10 is one of the most conserved interleukins and the possum protein shared 53–60% homology with eutherian sequences. An unusual long perfect palindromic sequence 42 base pairs long was found in the 3′ UTR of the possum sequence. Although the function of this is unknown it has been suggested that it may be involved in regulation of IL-10 expression (Harrison and Wedlock, 2000). IL-5, a cytokine which stimulates B cell growth and increases immunoglobulin secretion, was sequenced in the tammar wallaby (Hawken et al., 1999). Primers designed on conserved sequences between mice and humans were used to amplify part of the gene from genomic DNA of the striped faced dunnart (*Sminthopsis macroura)*. The PCR product was used as a probe to identify the tammar wallaby *IL5* sequence from a lambda clone library. FISH was then used to localise this gene to chromosome 1. The IL-5 protein shares 48–63% identity with eutherians. Many

important features in eutherian IL-5 sequences were conserved in the possum IL-5 including two cysteine residues thought to be essential for biological activity as well as key residues involved in the interaction between the protein and its receptor (Hawken et al., 1999). These features suggest that IL-5 has a conserved immune function across mammals.

18.2.3 Interferons

In mammals, interferons are divided into two types. Type I interferons belong to one of several classes including IFN-α, IFN-β, IFN-δ, IFN-ω, IFN-τ and IFN-κ. The number of genes within these classes varies from species to species. In eutherians, IFN-α and IFN-β genes produce distinct clusters in phylogenies, with -δ, -ω and -τ classes clustering with the -α class (Roberts et al., 1998). These -δ, -ω and -τ classes are only found in some species of eutherians and there has been disagreement between authors to whether the -δ and -ω classes are ancient classes or whether these classes have recently diverged from the -α class. IFN-κ sequences are only found in some eutherian species and cluster separately from the –α and –β sequences (Wong et al., 2006). Type II interferons are represented by a single class, IFN-γ, which contains only a single gene. A PCR based survey of interferons in the tammar wallaby revealed a large number of interferon genes rivaling that of eutherian species (Harrison et al., 2003). This survey indicated the presence of ten to twelve IFN-α genes and four IFN-β genes. The authors suggest that the diversification of interferons in marsupials and eutherians may be related to the evolution of viviparity, as echidnas and other vertebrates lack diversification of interferons. Southern blotting and phylogenetic analyses indicated that tammar wallaby interferons showed orthology with both IFN-α and IFN-β sequences from eutherians. This indicates that the divergence of the IFN-α and IFN-β classes occurred prior to the divergence of marsupials and eutherians. However none of the tammar interferons showed homology to the -δ, -ω or -τ subtypes.

18.2.4 Identifying Marsupial Cytokines In Silico

The above studies, using PCR based approaches, have only identified the most conserved cytokines. Many of the more divergent cytokines, including the majority of the interleukins and IFN-γ, have not been successfully sequenced using these laboratory methods due to the lack of sequence similarity with eutherians. Attempts to identify divergent interleukins such as 2, 4, 6 and 12 and IFN-γ using molecular approaches have failed (Harrison and Wedlock, 2000). However, with the recent sequencing of the opossum and tammar wallaby genomes, in silico methods have been utilised in order to identify the sequences of these divergent cytokines. Automated annotation techniques have been used to identify many interleukins in the genomes of the opossum and tammar wallaby. The previously unidentified interleukins 1α, 8, 15, 16, 17α, 19, 21, 22, 25, and 27 have all been identified in the

opossum genome sequence by the automated annotation techniques of *Ensembl*. Interleukins 15, 16, 17α, 19, 21, 25, and 27 have also been identified in the tammar wallaby genome by *Ensembl,* though the full sequence is only available for interleukins 17α and 25 due to gaps in the tammar genome sequence.

However, these automated annotation techniques have still been unable to identify the most divergent cytokines as these techniques are not sensitive enough. Instead, a manual, expert curated approach has been used, utilizing a synteny based strategy to identify the most divergent cytokines in the opossum genome (Wong et al., 2006). This approach involved identifying the genomic region containing the gene of interest by locating conserved flanking segments. After identifying this region, gene prediction was performed. Finally, the predicted genes were compared to known eutherian sequences to identify genes that are similar to the gene of interest. This method was used to successfully identify many highly divergent cytokines in the opossum including many previously unsequenced in marsupials including IL-2, IL-4, IL-13, IFN-γ and IFN-κ.

Thirteen interleukins were identified in this study (Wong et al., 2006), including four that had not been previously sequenced in marsupials. These interleukins had amino acid identity to human sequences ranging from 36.3% (IL-6) to 59.7% (IL-20). Many of the interleukins identified including IL-6, IL-2 and IL-13 had amino acid identities with human which were less than 45%, highlighting the need of new approaches such as these in silico methods to identify divergent cytokines in marsupials. The interleukin sequences in opossums share many features with eutherians that are crucial for the biological activity of these molecules. This includes the conserved leucine and aspartic acid amino acids within helix A of IL-2 which is critical for ligand binding in humans (Eckenberg et al., 2000) and the disulfide bonds which join helix B to the CD loop in IL-4 (Kruse et al., 1991).

This study identified seven IFN–α genes and a single IFN–β gene in the opossum genome (Wong et al., 2006). An IFN–κ gene, previously unidentified in marsupials, was identified in the opossum genome. The highly divergent type II interferon, IFN-γ, was also identified in the opossum genome with amino acid identity of 47% compared to eutherian sequence. This demonstrates that marsupials retain both type I and type II interferon genes. The presence of –α, –β and –κ subclasses indicates that the most divergent interferon subclasses are common to both eutherians and marsupials while the lack of –τ, –ω and –δ sequences in both opossum and tammar indicate that these are likely to have arisen recently within particular eutherian lineages.

Chemokine sequences within the opossum genome have also been identified bioinformatically (Belov et al., 2007). Hidden Markov model (HMM) searches were performed to identify the chemokine sequences (see Papenfuss et al., Chapter 6 for a detailed explanation of this approach). Chemokines are small peptides which are involved in cell activation, migration and differentiation. Thirty-one chemokine sequences were identified in the opossum genome, compared to 47 and 24 in humans and chickens respectively. The opossum sequences have the same conserved cysteine motif that is common to chemokines in all vertebrates (Belov et al., 2007). Chemokines are traditionally divided into four subgroups based on the position of

cysteine residues; these are the C, CC, CXC and CX_3C subgroups. The largest subgroup of chemokines in eutherians and birds is the CC group. In opossums there are 22 members of this subfamily. Within this subfamily are five genes which phylogenetic analyses suggest have arisen by a marsupial or opossum specific lineage gene expansion (Belov et al., 2007). The role of these cytokines is unknown, but they may have an important marsupial specific role. The CXC family in opossums contains seven genes. Some genes within this family in opossum appear to be orthologous to chemokine genes in humans. Many of the chemokines in this family in eutherians have a three residue motif that is important for neutrophil attraction (Laing and Secombes, 2004). Three opossum cytokines also contained this feature. A single member of the C and the CX_3C families were identified in the opossum genome.

Chemokine genes have also been identified within the tammar wallaby genome by *Ensembl*. 18 CC family chemokines and 3 CXC chemokines have been annotated in the wallaby genome. There are likely to be more chemokines within the tammar wallaby genome including CX_3C and C family chemokines, though a more complete wallaby genome and more sensitive searches may be required to identify these.

Another important set of genes for studying immune response in marsupials are the genes encoding the CD4 and CD8 peptides. These peptides are found on the surface of T helper cells and cytotoxic T cells respectively. Generation of marsupial specific reagents to these peptides in order to distinguish the classes of T cells is important to gain a better understanding of the response of marsupial immune systems to disease. The ratio of CD4/CD8 cells is commonly used to understand how the immune system responds to a pathogenic challenge and is also used in disease diagnostics. Identification of the sequence for *CD4, CD8α* and *CD8β* was first achieved by Wong et al. (2006), by performing TBLASTN searches on the opossum genome. However, this method was unable to predict the terminal exon of these genes. In two subsequent studies, the full opossum and tammar cDNAs for CD4 (Duncan et al., 2007) and CD8α/CD8β (Duncan et al., 2009) were isolated. These studies identified the *CD4* and *CD8* sequences from the opossum genome sequence and then used these sequences to isolate full length opossum and tammar cDNAs using RT-PCR and 5′ and 3′ RACE. Both the *CD4* and *CD8* sequences showed conserved synteny and structural similarity to human and mouse sequences. In both the tammar and opossum *CD4* and *CD8*, the sequences were shown to be truncated at the terminal end of the cytoplasmic tail. The functional significance of this is as yet unknown.

Sequencing of marsupial cytokines and CD4/CD8 peptides will provide a springboard for future studies in immunology and disease in marsupials. Studies into the immune response to marsupial diseases have been hampered due to the lack of marsupial specific immune reagents. Previous studies have attempted to used antibodies developed in eutherians to measure cytokines or CD4/CD8 (Higgins et al., 2004; Old and Deane, 2002), but many of these do not show cross-reactivity with marsupial species and they can be unreliable. The sequencing of *CD4* and *CD8* cDNAs will allow us to distinguish cytotoxic and helper T cells. This will be important for establishing baseline information on normal lymphoid tissue in healthy animals and

will allow us to characterise the pathogenesis of marsupial diseases. Key cytokines from all cytokine families have now be identified in marsupials including the members of the TNF family, many interleukins, both type I and type II interferons and chemokines from all of the four chemokine families. Cytokines that are associated with a Th1 response including IL-2, IFN-γ and TNF and those associated with Th2 response including IL-4, IL-5, IL-6, IL-10 and IL-13 have now been sequenced in marsupial species. This will allow us to examine Th1/Th2 cytokine profiles in diseased animals to understand the immune response of these animals.

18.2.5 Uses for Marsupial Cytokine Sequence Data

The development of cytokine sequence information in marsupials can be used for many purposes. The opossum in an important model for cancer immunology since opossums can be induced to accept melanoma cells (Wang and Vandeberg, 2005). With the expanding knowledge of cytokines, researchers can more easily investigate tumour suppression in the opossum. Another use of this sequence information will be in investigating the purpose of the two thymuses which is seen in many marsupials, especially within the diprotodont lineage. Also, researchers can continue to investigate the potential of marsupial cytokines to be used as adjuvants in biocontrol vaccines in overpopulated species such as the brushtail possum and koala. However the most important use of cytokine sequence information is likely to be in studying infectious diseases in marsupials. Two examples of this are in the devil facial tumour disease in Tasmanian devils and Chlamydia in koalas.

The devil facial tumour disease (DFTD) is caused by a transmissible tumour which is spread from one devil to another by biting (see Bender, Chapter 23). This disease has caused the decline of the population by about 70% in the last decade. This form of disease, where cancer cells themselves are the transmissible agent, is very rare as under normal circumstances these cells would be recognised as foreign and cleared by the immune system of the infected individual. There is only one other known naturally occurring transmissible cancer. This is the canine transmissible venereal tumour which is found in domestic and wild canines. One strategy that this tumour uses in order to avoid the immune response is by secreting immunosuppressive cytokines (Hsiao et al., 2004). Secretion of immunosuppressive cytokines is common in cancers and includes cytokines such as IL-10 (Garcia-Hernandez et al., 2002), IL-6 (Wei et al., 2003), IL-4 (Todaro et al., 2007) and TGF-β (Gold, 1999). In Tasmanian devils, there are no documented cases in which devils have been able to overcome the DFTD tumour and no signs of lymphocyte infiltration of the tumour (Loh et al., 2006a, b). One reason for this lack of immune response may be that DFTD cells are producing immunosuppressive cytokines. With sequence information now available for cytokines in marsupials we can investigate whether DFTD cells are producing any immunosuppressive cytokines. Through designing primers based on conserved regions of opossum or tammar cDNAs we can create probes to isolate Tasmanian devil cytokine cDNAs from cDNA libraries. We can then use the devil cDNA sequences in order to perform RT-PCR to determine the level of

expression of these cytokines in DFTD cells. Marsupial cytokine sequences can also be used to develop marsupial specific antibodies against cytokines. By determining whether DFTD cells produce immunosuppressive cytokines we can better understand the pathology of this disease and this will help us in designing treatments or vaccines.

Another use of marsupial cytokine sequence data will be to investigate the pathology of Chlamydia in koalas. Chlamydial infection in koalas is considered to be widespread with both *Chlamydophila pecorum* and *Chlamydophila pneumoniae* infecting koalas. Studies in human Chlamydia have shown that a Th1 dominated host response is important for elimination of disease (Lampe et al., 1998) while Th2 dominated host response results in the persistence of Chlamydia bodies and results in chronic hypersensitivity and fibrosis (Beatty et al., 1995; Darville et al., 1997). Investigation into the balance of Th1/Th2 response during infection of Chlamydia in koalas will be important to understand host and environmental factors which affect susceptibility of koalas to Chlamydial disease. In the past, bovine antibodies have been tested for the efficacy to label IL-4 and IFN-γ, in order to study the Th1/Th2 response of koalas to Chlamydia (Higgins et al., 2004). While the IFN-γ antibody showed some success in labeling marsupial IFN-γ, the IL-4 antibody did not show cross-reactivity to marsupial IL-4. Now that these cytokines, as well as other cytokines that form part of the Th1 and Th2 profiles, have been sequenced in marsupials we can develop marsupial specific antibodies. These will enable us to effectively assess the dominance of Th1 or Th2 response in koalas infected with Chlamydia. This will give us a better understanding of the koala's immune response to the disease and may be important for treating Chlamydia in koalas.

18.3 Natural Killer Cell Receptors

Natural Killer (NK) cells are cytotoxic lymphocytes that play an important role in the innate immune system. These cells play a major part in the destruction of tumour cells and virally infected cells. NK cells' cytotoxicity is controlled by activating and inhibitory signals. MHC class I molecules expressed in normal healthy cells inhibit NK cytotoxicity, while cells not expressing MHC class I such as some virally infected cells or tumour cells will be destroyed. Additionally, the MHC Class I associated (MIC) proteins will activate NK cytotoxicity. MIC proteins are not expressed in healthy cells but are upregulated in tumour cells and virally infected cells (reviewed in Cerwenka and Lanier, 2001). The detection of activating or inhibiting targets in NK cells is mediated by two major classes of receptors. These are the immunoglobulin superfamily (IgSF) which in eutherians is found in the leukocyte receptor complex (LRC) and the C-type lectin superfamily (CLSF) which in eutherians is found in the Natural Killer complex (NKC) (Kelley and Trowsdale, 2005).

NK receptor genes were first described in marsupials in 2005 by Hao and colleagues. This study identified eight putative NKC genes in the opossum genome. Five of these genes showed orthology to eutherian NKC genes. Following on from

this, a study by Belov and colleagues (2007) constructed a gene map of the NKC and identified genes using HMMer and GenomeScan. The location of the LRC was identified using BLAST and then receptor genes within this region were identified using HMMer. In opossums the NKC is located on chromosome 8 and encodes nine C-lectin type receptor genes. Based on orthology of these genes to eutherian NKC genes, at least seven of these C-lectin type receptor genes predated divergence of the eutherian and marsupial lineages. This indicates that the NKC evolved prior to the separation of these lineages. However, expansion in the NKC genes appears to have occurred only within the eutherian lineage and particularly in the rodents. While the opossum NKC only contains nine receptor genes the human NKC contains 29 and rats contain 75 (Hao et al., 2006).

The opossum LRC is located on chromosome 4 and contains five separate clusters of Ig domains (Belov et al., 2007). The LRC contains 154 Ig-like domains while a further 22 Ig-like domains are located on the unordered chromosome. When examined phylogentically, Ig domains from eutherians and avians cluster into two branches (Nikoladis et al., 2005) with mammalian MI and avian CI domains clustering together and mammalian MII and avian CII domains clustering separately. Opossum Ig domains fall into three distinct clades in this phylogenetic tree, designated clade I, Ia and II. Clade I clusters with eutherian MI and avian CI sequences while clade II clusters with MII and CII sequences. Clade Ia forms a sister group to the MI and CI clades and contains only opossum sequences. This clade represents a large expansion of Ig domains within opossums. This diversification is likely to be the result of different pathogenic pressures, particularly on the newborn marsupial which relies on innate immune mechanisms for protection from pathogens.

18.4 Antimicrobial Peptides

Another important group of immunological proteins are the antimicrobial peptides (AMPs). These molecules have a broad spectrum antibiotic effect and as such are an important mediator of the innate immune response. Antimicrobial peptides have a conserved structure usually containing both a hydrophobic and a hydrophilic side enabling them to be soluble in water but also penetrate lipid membranes. Once inside the target microbe, AMPs kill the target through diverse mechanisms. The two main groups of antimicrobial peptides are the defensins and the cathelicidins.

Defensins are cationic peptides with three conserved cysteine sulfide bonds. These peptides are found in epithelial cells, neutrophils and Paneth cells. Defensins function by disrupting the membrane of a microbial target cells causing an efflux of ions and nutrients leading to death of the cell (Ganz, 2003). In mammals defensins are classified into either α, β, or θ families based on their structure. β defensins are the most widely distributed defensins, while α defensins are mostly secreted by neutrophils and Paneth cells (Ganz et al., 1985; Jones and Bevins, 1992). θ defensins have only been discovered in a few primate species (Tang et al. 1999). It has been demonstrated that all mammalian defensins evolved from a common β-defensin like

ancestor with β defensins having been found in a wide range of vertebrates including fish and birds (Xiao et al., 2004).

18.4.1 Defensins

Defensins are one of the most rapidly evolving families in the genome making them difficult to isolate in marsupials. Identification of marsupial defensins has only occurred since the sequencing of the opossum genome. Although defensins are rapidly evolving, they have a highly conserved structure based on six cysteine residues (Selsted, 2004). This conserved feature has been utilised by researchers to identify defensins within the opossum genome. α defensins were previously thought only to have evolved in the branch of eutherian mammals containing rodents and primates, as α defensins could not be identified in the cow, dog or chicken (Bals et al., 1998; Patil et al., 2004). However, two studies (Belov et al., 2007 Lynn and Bradley, 2007) both utilising hidden Markov models to identify defensins, demonstrated that there is a single α gene present in the opossum genome. This gene is expressed in the small intestine epithelium, a tissue in which α defensins are expressed in eutherians, suggesting that this defensin has a similar function (Lynn and Bradley, 2007). The opossum α defensin is located in a region syntenic to the human chromosomal region containing α defensins (Belov et al., 2007; Lynn and Bradley, 2007). There are two types of α defensins found in eutherian mammals; myeloid which are primarily expressed in bone marrow and enteric which are primarily expressed in Paneth cells (Patil et al., 2004). Phylogenetic analysis places the sole opossum α defensin at the base of the eutherian myeloid and enteric α defensin lineages (Belov et al., 2007). This suggests that the myeloid and enteric lineages may have evolved within eutherians from the ancestral α defensin seen in opossums.

In the opossum genome, 32 β-defensin genes have been identified (Belov et al., 2007). This is similar to the number of β-defensins in humans (39) (Patil et al., 2005). The opossum β-defensins are arranged in three chromosomal clusters which are orthologous to β-defensin clusters in eutherian mammals. Through the construction of a phylogenetic tree and synteny map it was demonstrated that some opossum β-defensins are orthologous to eutherian β-defensins, while other opossum β-defensins have been produced by lineage specific expansions (Belov et al., 2007). A large opossum specific expansion is seen within one of the chromosomal clusters and it is possible that these genes may have a marsupial-specific immunological role. β-defensins are encoded by two exons in eutherian mammals, but are encoded by three and four exons in birds and fish respectively. It has been suggested that the fusion of defensin exons occurred as an adaptive event in mammals, allowing for faster mobilisation of these defensins to better cope with invading microbes (Xiao et al., 2004). The β-defensins identified in the opossum genome were all encoded by two exons, indicating that this fusion occurred prior to the divergence of eutherians and marsupials. This also suggests a similar function of β-defensins in marsupials and eutherians. The presence of both α and β defensins in the opossum indicates that the differentiation between α and β defensins occurred much earlier

than once suspected, prior to the divergence of the eutherian and marsupial lineages. θ defensins have only been identified in some primate species and were not found in the opossum genome. These θ defensins are likely to have arisen recently within the primate lineage. The presence of defensins in other marsupial lineages has yet to been investigated.

18.4.2 Cathelicidins

Cathelicidins, like defensins, have a cationic amphipathic structure which allows them to disrupt the membrane of pathogens. Peptides in this family are characterised by a highly conserved cathelin-like prosequence containing four cysteine residues. Cathelicidins are predominately expressed in neutrophils and epithelial cells and are expressed in response to inflammation, cell injury and microbial infection. As well as their antimicrobial effects cathelicidins have an immunomodulatory role (Braff et al. 2005). Cathelicidins can induce changes in cytokine expression, bind neutralise endotoxins and induce chemotaxis of T cells and neutrophils (Bals and Wilson, 2003).

The first marsupial cathelicidin sequences were identified in the opossum genome using HMMer searches (Belov et al., 2007 and Papenfuss et al., Chapter 6). Twelve cathelicidin genes were located in the opossum genome, in a single cluster on chromosome 6. This region shows synteny with the region containing cathelicidins in mammals and birds. Like all other cathelicidin genes in vertebrates, these genes are composed of four exons. The opossum cathelicidin genes do not show orthology to eutherian genes and are clustered separately in a phylogeny. There are more genes in the cathelicidin family in the opossum then in any eutherian mammal studied to date, with many species including humans and mice having only a single cathelicidin gene. While most eutherian cathelicidin sequences cluster closely within species, indicating that they have arisen from recent gene expansions, opossum sequences form at least three distinct clusters in a phylogeny with long branches separating them. The amino acid identity of the mature peptides ranges from 1 to 94%. The level of heterogeneity of these opossum cathelcidins is much higher than in other species, including bird and fish species. The expansion in this gene family suggests that these genes may have an important marsupial-specific role in innate immunity. The high level of heterogeneity may suggest that these genes provide protection against a diverse array of microbial pathogens, which has led to the maintenance of diverse cathelicidin alleles.

Cathelicidin genes have also been identified in the tammar wallaby by using molecular based approaches. A study in 2008 (Daly et al., 2008) identified seven cathelicidin peptides from a tammar wallaby EST mammary gland cDNA library. The average percentage identities of the cathelin-like domain of these peptides was 86%. The identity to eutherians was 46% in wallaby, which was higher than the 36% identity seen in the opossum sequences (Belov et al., 2007). The expression of these cathelicidins was increased in adult wallaby leukocytes when exposed to lippopolysaccharides or lipoteichoic acid (Daly et al., 2008). This suggests that the

wallaby cathelicidins, like eutherian cathelicidins, have an important role in defense and recognition of both gram positive and gram negative bacteria. The wallaby peptides also have many conserved features with eutherians including conserved cysteine residues, a hydrophobic C-terminus and cationic charge. Conservation of these features also suggests that the role of cathelicidins is similar in both lineages.

Cathelicidins have been of particular interest to marsupial researchers due to their potential role in protecting marsupial neonates. In eutherian mammals, including humans and mice, cathelicidins are found at much higher levels in the skin of neonates than in the skin of healthy adults (Dorshner et al., 2003). Human milk (Lipeke et al., 2001) and tracheal aspirates of newborn infants (Schaller-Bals et al., 2002) also contain antimicrobial peptides. Marsupials have a very short gestation and their young undergo the majority of their development in the non-sterile pouch. They are born without a functional immune system and underdeveloped lymphoid tissue. Despite this lack of adult immune function, the marsupial neonates are protected against pathogens, and researchers have observed that surgical incisions in neonatal young do not become infected when returned to the pouch (Renfree and Tyndale-Biscoe, 1978). Cathelicidins have been suggested to play an important role in the protection of marsupial neonates in the pouch. In the study of wallaby cathelicidins by Daly and colleagues (2008) it was found that one cathelicidin, named *Maeucath1*, was expressed in the primary and secondary lymphoid tissue of wallaby neonates. Furthermore, it was found that the expression of this cathelicidin was increased in the liver, spleen, bone marrow, skin, gut and lungs in the first 25 days post-partum. By 90–120 days, the approximate time of immune maturation of tammar neonates, expression had stabilized or decreased in these tissues. A subsequent study isolated a new cathelicidin using RACE-PCR (Carmen et al., 2009). This new cathelicidin was found to be expressed in the spleen and GIT of newborns and was observed in most of the neonates' tissues by day 7 post-partum. These studies demonstrate that cathelicidins are being synthesised by neonates. This finding, along with the observation of expansion and diversification of cathelicidin genes in both opossums and tammar wallabies suggest that cathelicidins play a key role in protecting the young prior to development of an adult-like humoural immune system. The expression of defensins in neonatal tissue has not yet been investigated but these peptides may also be important in protecting pouch young.

18.4.3 Potential Uses for Marsupial Antimicrobial Peptides

Cathelicidins and defensins are also of interest to researchers due to their potential use as therapeutic agents. Antimicrobial peptides have a broad range antimicrobial activity but do not appear to induce antibiotic resistance. A number of naturally occurring antimicrobial peptides and their derivatives have been developed as novel antibiotic therapies. Pexiganan, an analog of the naturally occurring magainin extracted from the skin of *Xenopus laevius* frogs, has been developed for use in treating diabetic foot infection (Gottler and Ramamoorthy, 2009). Other AMPs extracted from animals have been investigated for their use in treating buco-dental

infection, ocular infections (Andreu and Rivas, 1998) and lower respiratory tract infections (Cazzola et al., 2003). Additionally, some AMPs have also been found to have anti-cancer effects (Hoskin and Ramamoorthy, 2008). Pharmaceutical companies are currently developing AMP based therapies for treatment of diseases such as meningitis, acne, sepsis and HIV (Koczulla and Bals, 2003). As marsupials appear to have a wide diversity of AMPs, with many of these peptides being very unique from those found from other vertebrates, marsupials AMPs may be a new target for developing antimicrobial therapies. With continued investigation and experimentation into marsupial antimicrobial peptides new antimicrobial therapies may be discovered for treating both marsupials and humans.

18.5 Conclusion

Research into the immunology of marsupials has provided insight into the evolution of the mammalian immune system. All the main peptides within the cytokine families are conserved in marsupials and eutherians including members of the TNF family, Th1 and Th2 associated interleukins, both Type I and Type II interferons and a diverse array of chemokines. Major divergences within these families are common to both eutherians and marsupials. Studies on marsupial cytokines have shown that the divergence of IFN-κ, IFN-α and IFN-β occurred prior to the separation of the eutherian and marsupial lineages (Harrison et al., 2003; Wong et al., 2006). Additionally both Type I and Type II interferons are preserved in both lineages, as are the four different families of chemokines. These studies have also suggested that some sub-classes within these families are likely to be eutherian specific, such as the IFN-δ and IFN-ω classes. Studies on antimicrobial peptides have shown that the development of α and β defensins occurred prior to the divergence of marsupials and eutherians (Belov et al., 2007).

Investigations into divergent immune molecules have demonstrated that marsupials possess a very similar immune architecture to eutherian mammals. The AMP complement of marsupials is comparable to that of eutherians in terms of the number of peptides within the defensin and cathelicidin families and diversity of these peptides. Important cytokines across all cytokine families are conserved between eutherians and marsupials. Structural and functional analyses on these cytokines and AMPs suggest that they play an equivalent role in marsupials and eutherian mammals. While these studies have demonstrated the similarity between marsupial and eutherian immunology, the sequencing of these molecules will allow for future studies to further investigate difference between the immune system in marsupials and eutherians. One of the most important distinguishing features between marsupials and eutherians are that marsupials are born at very premature stage and undergo most of their development, including their immune development in the non-sterile pouch. Cathelicidins are produced by neonatal young as soon as day 1 postpartum and so these peptides are likely to play an important role in defending the young from pathogens prior to immune development. However, these molecules are likely to be only one of many factors protecting the young during this stage.

The development of sequence information for the divergent immune genes discussed in this chapter will have a huge range of potential uses in understanding marsupial immunology and treating marsupial diseases. Development of reagents against CD4 and CD8 will enable researchers to distinguish cytotoxic T cells and helper T cells, while the sequencing of many cytokines will enable the differentiation of Th1 and Th2 dominated immune responses. This will enable researchers to better understand the immune response of marsupials to disease and will lead to better treatment of diseased animals and prevention of disease. Some of these divergent immune genes may lead to treatment not only of marsupials but also of humans. Research into both defensins and cathelicidins may yield new antimicrobial or anticancer therapies for both humans and animals. Through the increase in the genomic information available for divergent immune sequence our understanding of marsupial immunity and our ability to treat disease in marsupials will improve in the future.

References

Abbas AK, Janeway CA Jr (2000) Immunology: improving on nature in the twenty-first century. Cell 100:129–138.

Abbas AK, Murphy KM, Sher A (1996) Functional diversity of helper T lymphocytes. Nature 383:787–793.

Andreu D, Rivas L (1998) Animal antimicrobial peptides: An overview. Biopolymers 47:415–433.

Bals R, Wilson JM (2003) Cathelicidins – a family of multifunctional antimicrobial peptides. Cell Mol Life Sci 60:711–720.

Bals R, Goldman MJ, Wilson JM (1998) Mouse beta-defensin 1 is a salt-sensitive antimicrobial peptide present in epithelia of the lung and urogenital tract. Infect Immun 66:1225–1232.

Baylor College of Medicine (2008) Retrieved from http://www.hgsc.bcm.tmc.edu/project-species-m-Wallaby.hgsc?pageLocation=Wallaby 20/11/09

Beatty WL, Morrison RP, Byrne GI (1995) Reactivation of persistent *Chlamydia trachomatis* infection in cell culture. Infect Immun 63:199–205.

Beck G (1998) Macrokines: invertebrate cytokine-like molecules? Front Biosci 3:559–569.

Belov K, Sanderson CE, Deakin JE, et al. (2007) Characterization of the opossum immune genome provides insights into the evolution of the mammalian immune system. Genome Res 17: 982–991.

Braff MH, Hawkins MA, Di Nardo A, et al. (2005) Structure–function relationships among human cathelicidin peptides: dissociation of antimicrobial properties from host immunostimulatory activities. J Immunol 174:4271–7278.

Brozek CM, Ley RD (1991) Production of interleukin-1 in a South American opossum (*Monodelphis domestica*). Dev Comp Immunol 15:401–412.

Carman RL, Old JM, Baker M, Jacques NA, Deane EM (2009) Identification and expression of a novel marsupial cathelicidin from the tammar wallaby (*Macropus eugenii*). Vet Immunol Immunopathol 127:269–275.

Cazzola M, Sanduzzi A, Matera MG (2003) Novelties in the field of antimicrobial compounds for the treatment of lower respiratory tract infections. Pulm Pharmacol Ther 16:131–145.

Cha SS, Kim JS, Cho HS, et al. (1998) High resolution crystal structure of a human tumor necrosis factor-alpha mutant with low systemic toxicity. J Biol Chem 273:2153–2160.

Cerwenka A, Lewis LL (2001) Natural killer cells, viruses and cancer. Nat Rev Immunol 1:41–49.

Daly KA, Digby MR, Lefévre C, Nicholas KR, Deane EM, Williamson P (2008) Identification, characterization and expression of cathelicidin in the pouch young of tammar wallaby (*Macropus eugenii*). Comp Biochem Physiol 149:524–533.

Darville T, Andrews CW Jr, Laffoon KK, Shymasani W, Kishen LR, Rank RG (1997) Mouse strain-dependent variation in the course and outcome of chlamydial genital tract infection is associated with differences in host response. Infect Immun 65:3065–3073.

Deakin JE, Papenfuss AT, Belov K, et al. (2006) Evolution and comparative analysis of the MHC Class III inflammatory region. BMC Genomics 7:281.

Dinarello CA (1999) Cytokines as endogenous pyrogens. J Infect Dis 179:S294–S304.

Dinarello CA (2007) Historical insights into cytokines. Eur J Immunol 37:S.

Dorshner RA, Lin KH, Murakami M, Gallo RL (2003) Neonatal skin in mice and humans expresses increased levels of antimicrobial peptides: innate immunity during development of the adaptive response. Pediatr Res 53:566–572.

Duncan LG, Nair SV, Deane NM (2007) Molecular characterisation and expression of CD4 in two distantly related marsupials: the gray short-tailed opossum (*Monodelphis domestica*) and tammar wallaby (*Macropus eugenii*). Mol Immunol 44:3641–3652.

Duncan LG, Nair SV, Deane EM (2009) The marsupial CD8 gene locus: molecular cloning and expression analysis of the alpha and beta sequences in the gray short-tailed opossum (*Monodelphis domestica*) and the tammar wallaby (*Macropus eugenii*). Vet Immunol Immunopathol 129:14–27.

Eckenberg R, Rose T, Moreau JL, et al. (2000) The first alpha helix of interleukin (IL)-2 folds as a homotetramer, acts as an agonist of the IL-2 receptor beta chain, and induces lymphokine-activated killer cells. J Exp Med 191:529–540.

Ganz T, Selsted ME, Szlarek D, et al. (1985) Defensins – natural peptide antibiotics of human neutrophils. J Clin Invest 76:1427–1435.

Ganz T (2003) Defensins: antimicrobial peptides of innate immunity. Nat Rev Immunol 3: 710–720.

Garcia-Hernandez ML, Hernandez-Pando R, Gariglio P, Berumen J (2002) Interleukin-10 promotes B16-melanoma growth by inhibition of macrophage functions and induction of tumour and vascular cell proliferation. Immunology 105:231–243.

Gold LI (1999) The role for transforming growth factor-beta (TGF-beta) in human cancer. Crit Rev Oncogen 10:303–360.

Gottler LM, Ramamoorthy A (2009) Structure, membrane orientation, mechanism, and function of pexiganan – A highly potent antimicrobial peptide designed from magainin. Biochimica et Biophysica Acta 1788:1680–1686.

Hao L, Klein J, Nei M (2006) Heterogeneous but conserved natural killer receptor gene complexes in four major orders of mammals. *Proc* Natl Acad Sci U S A 103:3192–3197.

Harrison GA, Deane EM (1999) cDNA sequence of the lymphotoxin beta chain from a marsupial: *Macropus eugenii* (tammar wallaby). J Interferon Cytokine Res 19:1099–1102.

Harrison GA, Deane EM (2000) cDNA Cloning of lymphotoxin alpha (LT-α) from a marsupial, *Macropus eugenii*. Mitochondrial DNA 10:399–403.

Harrison GA, Wedlock DN (2000) Marsupial cytokines; structure, function and evolution. Dev Comp Immunol 24:473–484.

Harrison GA, Broughton MJ, Young LJ, Cooper DW, Deane EM (1999) Conservation of 3′ untranslated region elements in tammar wallaby (*Macropus eugenii*) TNF-αmRNA. Immunogenetics 49:464–467.

Harrison GA, Young LJ, Watson CM, Miska KB, Miller RD, Deane EM (2003) A survey of type I interferons from a marsupial and monotreme: implications for the evolution of the type I interferon gene family in mammals. Cytokine 21:105–119.

Hawken RJ, Maccarone P, Toder R, Graves JAM, Maddox JF (1999) Isolation and characterization of marsupial IL5 genes. Immunogenetics 49:942–948.

Higgins DP, Hemsley S, Canfield PJ (2004) Assessment of anti-bovine IL4 and IFN gamma antibodies to label IL4 and IFN gamma in lymphocytes of the koala and brushtail possum. Vet Immunol Immunopathol 101:153–160.

Hoskin DW, Ramamoorthy A (2008) Studies on anticancer activities of antimicrobial peptides. BBA-Biomembranes 1778: 357–375.

Hsiao YW, Liao KW, Hung SW, Chu RM (2004) Tumor-infiltrating lymphocyte secretion of IL-6 antagonizes tumor-derived TGF-beta 1 and restores the lymphokine-activated killing activity. J Immunol 172:1508–1514.

Jones DE, Bevins CL (1992) Paneth cells of the human small intestine express an antimicrobial peptide gene. J Biol Chem 267:23216–23225.

Kelley J, Trowsdale J (2005) Features of MHC and NKC gene clusters. Transpl Immunol 14: 129–134.

Koczulla AR, Bals R (2003) Antimicrobial peptides – current status and therapeutic potential. Drugs 63:389–406.

Kruse N, Lehrnbecher T, Sebald W (1991) Site-directed mutagenesis reveals the importance of disulfide bridges and aromatic residues for structure and proliferative activity of human interleukin-4. FEBS Lett 286:58–60.

Laing KJ, Secombes CJ (2004) Chemokines. *Dev* Comp Immunol 28:443–460.

Lampe MF, Wilson CB, Bevan MJ, Starnbach MN (1998) Gamma interferon production by cytotoxic T lymphocytes is required for resolution of *Chlamydia trachomatis* infection. Infect Immun 66:5457–5461.

Liepke C, Zucht H-D, Forssmann W-G, Standerker L (2001) Purification of novel peptide antibiotics from human milk. J Chromatogr 752:369–377.

Loh R, Bergfeld J, Hayes D, et al. (2006a) The pathology of devil facial tumor disease (DFTD) in Tasmanian devils (*Sarcophilus harrisii*). Vet Pathol 43:890–895.

Loh R, Hayes D, Mahjoor A, O'Hara A, Pyecroft S, Raidal S (2006b) The immunohistochemical characterization of devil facial tumor disease (DFTD) in the Tasmanian Devil (*Sarcophilus harrisii*). Vet Pathol 43:896–903.

Lynn DJ, Bradley DG (2007) Discovery of a-defensins in basal mammals. Dev Comp Immunol 31:963–967.

Mikkelsen TS, Wakefield MJ, Aken B, et al. (2007) Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447:167–177.

Nikolaidis N, Klein J, Nei M (2005) Origin and evolution of the Ig-like domains present in mammalian leukocyte receptors: insights from chicken, frog, and fish homologues. Immunogenetics 57:151–157.

Old JM, Deane EM (2002) Immunohistochemistry of the lymphoid tissues of the tammar wallaby, *Macropus eugenii*. J Anat 201:257–266.

Patil A, Hughes AL, Zhang G (2004) Rapid evolution and diversification of mammalian alpha-defensins as revealed by comparative analysis of rodent and primate genes. Physiol Genomics 20:1–11.

Patil A, Cai Y, Sang Y, Blecha F, Zhang G (2005) Cross-species analysis mammalian beta-defensin gene family: presence of syntenic gene clusters and preferential expression in the male reproductive tract. Physiol Genomics 23:5–17.

Renfree MB, Tyndale-Biscoe CH (1978) Manipulation of marsupial embryos and pouch young. In: Daniel JC (ed) Methods in Mammalian Reproduction. Academic Press, New York.

Roberts RM, Liu L, Guo Q, Leaman D, Bixby J (1998) The evolution of the type I interferons. J Interferon Cytokine Res 18:805–816.

Rothel JS, Seow H-F, Lightowlers MW, et al. (1998) The use of recombinant ovine IL-1βand TNF-α as natural adjuvants and their physiological effects in vivo. Immunol Cell Biol 76:167–172.

Schaller-Bals S, Schulze A, Bals R (2002) Increased levels of antimicrobial peptides in tracheal aspirates of newborn infants during infection. Am J Respir Crit Care Med 165:992–995.

Selsted ME (2004) Theta-defensins: cyclic antimicrobial peptides produced by binaligation of truncated alpha-defensins. Curr Protein Pept Sci 5:365–371.

Tang YQ, Yuan J, Osapay G, et al. (1999) A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated alpha-defensins. Science 286:498–502.

Todaro M, Alea MP, Di Stefano AB (2007) Colon cancer stem cells dictate tumor growth and resist cell death by production of interleukin-4. Cell Stem Cell 1:389–402.

Wang Z, Vandeberg JL (2005) Immunotolerance in the laboratory opossum (*Monodelphis domestica*) to xenografted mouse melanoma. Contemp Top Lab Anim Sci 44:39–42.

Wedlock DN, Aldwell FE, Buddle BM (1996) Molecular cloning and characterization of tumour necrosis factor alpha (TNF-α) from the Australian common brushtail possum, *Trichosurus vulpecula*. Immunol Cell Biol 74:151–158.

Wedlock DN, Aldwell FE, Buddle BM (1998) Nucleotide sequence of a marsupial interleukin-10 cDNA from the Australian brushtail possum (*Trichosurus vulpecula*). DNA Seq 9:239–244.

Wedlock DN, Goh LP, McCarthy AR, Midwinter RG, Parlane NA, Buddle BM (1999a) Physiological effects and adjuvanticity of recombinant brushtail possum TNF-α. Immunol Cell Biol 77:28–33.

Wedlock DN, Goh LP, Parlane NA, Buddle BM (1999b) Molecular cloning and physiological effects of brushtail possum interleukin-1β. Vet Immunol Immunopathol 67:359–372.

Wei L-H, Kuo M-L, Chen C-A (2003) Interleukin-6 promotes cervical tumor growth by VEGF-dependent angiogenesis via a STAT3 pathway. Oncogene 22:1517–1527.

Wong ESW, Young LJ, Papenfuss AT, Belov K (2006) In silico identification of opossum cytokine genes suggests the complexity of the marsupial immune system rivals that of eutherian mammals. Immunome Res 2:4.

Xiao Y, Hughes AL, Ando J, Matsuda Y, Cheng JF, Skinner-Noble D, Zhang G (2004) A genome-wide screen identifies a single beta-defensin gene cluster in the chicken: implications for the origin and evolution of mammalian defensins. BMC Genomics 5:56.

Part VIII
Marsupial Genes and Gene Evolution

Paul D. Waters and Janine E. Deakin

All gene families have unique evolutionary histories characterized by duplications, transpositions, pseudogenization and selection for retention. Understanding the often complex evolutionary history of gene families can shed light on their function and organization in humans. However, the goal of completely understanding the evolution of genes and gene families cannot be reached without data from diverse vertebrate (and even invertebrate) representatives. The inclusion of marsupials and the most basal mammals, the monotremes, in comparative studies of gene family organization has led to very interesting and surprising findings which are highlighted in these next three chapters.

In this part, Deeb reviews colour vision and opsin gene evolution in mammals. Most mammals have dichromatic colour vision with the opsin genes short-wave-sensitive 1 (*SWS1*) and either middle-wave-sensitive (*MWS*) or long-wave sensitive (*LWS*); whereas birds, reptiles and fish generally have tri- or even tetra-chromatic vision (reviewed in Collin et al., 2009). It is thought that this was lost in early mammalian evolution in response to selection for more sensitive vision in low light conditions for nocturnal foraging. Trichromatic vision evolved again in the ancestor of old world primates by the duplication of *LWS* to give *MWS*. Interestingly, there has also been an independent duplication of *LWS* in the howler monkey (a new world monkey) (Jacobs et al., 1996). Like the majority of eutherian mammals, most marsupials appear to have dichromatic colour vision, and also like eutherian mammals, some Australian marsupials have evolved trichromatic colour vision. Trichromatic visions appears to have arisen independently several times in therian mammals, and comparing the opsin repertoire of marsupials and therians with monotremes and other vertebrates has allowed for the dating of when these opsin genes arose in different lineages.

Patel and Deakin discuss globin genes and their evolution in Chapter 20. Haemoglobin molecules, consisting of two alpha- and two beta-globin polypeptides, are essential for oxygen transport in vertebrates. These polypeptides are encoded by genes occurring in two distinct clusters in human but occur as a single cluster in fish and amphibians. In the past, several different hypotheses had been put forward to explain the evolution of the two clusters. However, it was the discovery of a novel globin gene in marsupials and monotremes, *HBW,* and closer inspection of regions

flanking the α- and β-globin clusters in representative amniotes that shed light on the evolution of these gene clusters (Patel et al., 2008).

Olfactory receptors, the mediators of olfaction, belong to the largest gene family found in vertebrates with over 1,000 members in some species (reviewed in Kratz et al., 2002), making it one of the most challenging gene families to study. In Chapter 21, Delbridge and colleagues discuss olfactory receptor (OR) genes and their evolution. This review examines the physical structures and the genes involved in olfaction. It compares olfactory receptor gene families in a wide variety of vertebrate species, and explains what marsupials have revealed about lineage specific expansion and contractions of these families in different mammalian groups. There is considerable variation of olfactory receptors between species, and the expansion/contractions of different OR families reflects the niche (i.e how it detects food, evades predators, communicates with other individuals, attracts mates) a particular animal fills in its environment.

These reviews highlight the importance of understanding gene families across all mammals, and indeed all vertebrates, to gain a more complete understanding of their evolution and function. Research on marsupials has led to the discovery of novel genes in the eutherian genome, along with lineage specific genes (and gene organization) that has helped reconstruct the ancestral state and the steps required to reach the patterns observed today in different taxa.

References

Collin SP, Davies WL, Hart NS, Hunt DM (2009) The evolution of early vertebrate photoreceptors. Philos Trans R Soc Lond B Biol Sci 364:2925–2940.

Jacobs GH, Neitz M, Deegan JF, Neitz J (1996) Trichromatic colour vision in New World monkeys. Nature 382:156–158.

Kratz E, Dugas JC, Ngai J (2002) Odorant receptor gene regulation: implications from genomic organization. Trends Genet 18:29–34.

Patel VS, Cooper SJ, Deakin JE, et al. (2008) Platypus globin genes and flanking loci suggest a new insertional model for beta-globin evolution in birds and mammals. BMC Biol 6:34.

Chapter 19
Visual Pigments and Colour Vision in Marsupials and Monotremes

Samir S. Deeb

Abstract The sequences and spectral characteristics of the cone visual pigments of several Australian and a few South American marsupials have been determined. Based on genomic and mRNA sequence analyses, it was found that all of these species have only two cone-specific photopigments: one belonging to the short-wave-sensitive 1 (SWS1 ultraviolet or violet), and the second belonging to either the middle-wave-sensitive (MWS) or long-wave-sensitive (LWS) visual pigments. This predicted that colour vision in these species is dichromatic. However, behavioural studies have indicated that at least some Australian marsupials have trichromatic colour vision. This is consistent with the observation by microspectrophotometry that three classes of cone photoreceptors exist in the retinae of some of the Australian marsupials. Retinal mRNA and genomic DNA sequence analysis approaches did not find a third cone photopigment. These approaches used PCR primers that are specific to cone photopigments. Based on spectral sensitivity, it was suggested that a rod photopigment could be expressed in some cones to establish trichromacy. Sequence analysis of the photopigment genes and corresponding mRNAs in the platypus and echidna indicated the presence of only two opsins: SWS2, typically found in reptiles, birds and fish, instead of the SWS1 pigment typically found in most mammals, and a LWS, suggesting that monotremes also have dichromatic colour vision. Unlike in other mammals, the monotreme *SWS2* and *LWS* genes were tandemly arranged and separated by a locus control region, indicating retention of ancestral reptilian opsin genes. The gene structure and expression of monotreme cone opsin genes bridge the phylogenetic gap between reptiles and other mammals. A proposed evolutionary pathway of the opsin genes is described.

Keywords Colour vision · Genetics · Marsupials · Monotremes · Evolution

---

S.S. Deeb (✉)
Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA, USA
e-mail: sdeeba@u.washington.edu

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_19, © Springer Science+Business Media B.V. 2010

19.1 Introduction

The retina contains two types of photoreceptors: rods and cones. Rods contain rhodopsin and mediate vision in dim light (scotopic vision). Cones mediate vision in daylight (photopic vision) and colour vision. Rods are usually more numerous than cones in mammalian retinae. The rod-to-cone ratio in the retina varies widely among species. Some species have rod-dominated retinae (rats and mice) while others have cone-dominated retinae (ground squirrels and tree-shrews). No colour is perceived under scotopic conditions. Colour vision is achieved by comparing the rates at which photons are absorbed by cone photoreceptors that have different, but overlapping, spectra. The quantity and distribution of cones in the retinae also play an important role in colour perception. The photoreceptors are composed of the 11-*cis*-retinal chromophore that is bound by proteins, called opsins. The number of colour hues that an animal can discriminate is dependent on the number of cone photoreceptors classes in their retinae. For example, humans can discriminate between ∼2.3 million colours using three classes of cones. Most other mammals who have two classes of cones can distinguish between only ∼1,000 colours. The development of different classes of cones relies on mutually exclusive expression of photoreceptor pigments in single cone cells. Animals have evolved regulatory mechanisms that underlie mutually exclusive expression of cone pigment genes. An evolutionarily conserved enhancer motif, referred to as the locus control region (LCR), plays an important role in mutually exclusive expression of opsin genes in single cone cells. The LCR is commonly located upstream of the LWS or MWS pigment genes on the X chromosome of mammals (Wang et al., 1999; Smallwood et al., 2002).

The evolution of cone opsins and colour vision has been thoroughly investigated in many eutherian mammals, reptiles, birds and fish (Yokoyama, 2002; Jacobs, 2008). Monotremes and marsupials have evolutionarily diverged from eutherian mammals about 166 and 148 million years ago (MYA), respectively (Fig. 19.1) (Bininda-Emonds et al., 2007). Characterization of the genes and photoreceptors that predict colour vision in these species would be very important in filling the gap between eutherian mammals and reptiles, birds (diverged about 310 MYA) and fish (∼450 MYA) (Fig. 19.1). This will provide new insights into the evolution of opsin genes and colour vision in vertebrates.

Colour vision in vertebrates is based on visual pigments that exhibit a wide range of spectral sensitivity, ranging from ultraviolet to far red. Adaptations of vertebrates to different environments included selection for visual pigments with the appropriate spectral absorption (Yokoyama, 2002; Jacobs and Rowe, 2004). The evolution of visual pigment tuning towards adaptation to various environments involved gene duplication, translocation to other chromosomes, mutations that altered amino acids that contribute to spectral tuning and acquisition of appropriate gene-regulating motifs. Most eutherian mammals have only two types of cone photoreceptors (SWS1 and either MWS or LWS) in their retinae, and therefore, have dichromatic colour vision. In contrast, reptiles, birds and fish have much richer (trichromatic and perhaps tetrachromatic) colour vision. Their retinae contain four

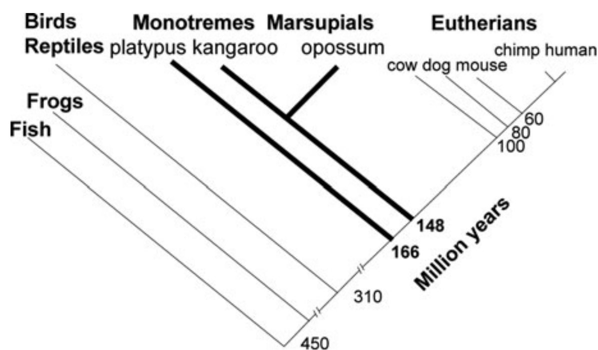


Fig. 19.1 Vertebrate evolution. Monotremes and marsupials diverged from eutherian mammals about 166 and 148 million years ago (MYA), respectively (Bininda-Emonds et al., 2007). There have been extensive studies on colour vision pigment genes and colour vision among eutherian mammals as well as lower vertebrates such as fish and birds. Therefore, information on monotremes and marsupials will bridge the gap between lower vertebrates and eutherian mammals in our understanding of the evolution of colour pigment genes as well as colour vision

or five classes of cone photoreceptors, some of which are orthologous to the mammalian opsins (Yokoyama et al., 2000; Collin et al., 2003). It is hypothesized that this rich colour vision system was lost during early mammalian evolution because of selection for higher degrees of sensitivity in low light (nocturnal foraging). Trichromatic colour vision in mammals re-evolved, in response to the development of diurnal habits, as a result of gene duplication followed by mutations to establish three distinct classes of cone photoreceptors. An important aspect of this evolutionary pathway was to ensure regulation of expression of the three pigment genes in a mutually exclusive fashion (i.e. one pigment in one cone photoreceptor cell), as well as to establish the proper cone mosaic in the retina. Most old world primates have three types (SWS1, MWS and LWS) of cone photoreceptors and rich trichromatic colour vision. This was due to duplication, about 40 MYA, of the *LWS* gene on the X-chromosome to generate *LWS* and *MWS* encoding genes. Most New World primates have two cone pigments because they separated from the old world primates before *LWS* pigment gene duplication. However, *LWS* gene duplication was found to have occurred independently in the South American howler monkey (Jacobs et al., 1996; Dulai et al., 1999). Interestingly, trichromatic colour vision was also found among New World monkey females who are heterozygotes for common genetic variants of the single pigment gene on the X-chromosome that encode pigments with different spectral sensitivities (Rowe and Jacobs, 2004; Jacobs and Deegan, 2005; Jacobs and Williams, 2006). Such females have three classes of cone photoreceptors in their reinae due to X-chromosome inactivation. The selective forces involved in re-evolution of trichromatic colour vision in primates were suggested to include selection for the ability to detect coloured ripe fruits (Regan et al., 2001) or nutritious young leaves (Dominy et al., 2003) against green backgrounds.

This chapter will focus on reviewing molecular genetic, spectrophotometric and behavioural studies conducted on marsupials and monotremes, in order to bridge the

gap in evolution of colour vision between basal vertebrates and mammals. There is evidence suggesting that some of the marsupials may have trichromatic colour vision.

19.2 Marsupial Cone Visual Pigments and Colour Vision

The molecular genetic approaches that were used to determine the sequences of cone visual pigments involved PCR amplification of opsin genes using genomic DNA or retinal mRNA templates with oligonucleotide primers that are highly conserved among many vertebrate species (Deeb et al., 2003; Strachan et al., 2004; Cowing et al., 2008). The spectral characteristics of the pigments were either inferred from the sequence or determined by in vitro expression and spectral measurements. The other approaches that were used to characterize the classes of cones in marsupial retinas included immunohistochemistry, microspectrophotometry, electroretinography (flicker-photometry) and behavioural chromatic discrimination testing.

Most of the data on visual pigments and colour vision so far published have been obtained from studies on the following Australian marsupials: polyprotodonts (fat-tailed dunnart, stripe-faced dunnart and quenda) and diprotodonts (tammar wallaby, honey possum, brush-tailed possum, and quokka). Recently the cone opsin sequences of two species of South American marsupials (the big-eared opossum and the gray short-tailed opossum) have been determined.

19.2.1 Australian Marsupials

The first marsupial cone visual pigments (*SWS* and *MWS*) were sequenced in the tammar wallaby (*Macropus eugenii*) (Deeb et al., 2003). This is consistent with behavioural evidence using colour discrimination tests, the tammar wallaby has dichromatic colour vision (Hemmi, 1999). Subsequently the sequence of the cone pigment genes and mRNAs of the stripe-faced dunnart (*Sminthopsis macroura*), fat-tailed dunnart (*Sminthopsis crassicaudata*) (Strachan et al., 2004), the brush-tailed possum (*Trichosurus vulpecula*) and the long-nosed bandicoot (*Perameles nasuta*) (Anderson et al., in preparation) were determined. The classes and wavelengths of maximum absorption (λ_{max}) of these pigments are shown in Table 19.1. Interestingly, in contrast with the behavioural finding of dichromatic colour vision in the tammar wallaby (Hemmi, 1999), behavioural tests on the fat-tailed dunnart indicated the presence of trichromatic colour vision (Arrese et al., 2006). Only two cone pigment genes were detected in each of the above five species that we analysed: SWS1 and either MWS or LWS. The SWS1 pigments fall into the ultraviolet (360 nm) or the violet (420 nm) classes. Similar findings were made using molecular genetic approaches on the cone visual pigments of the fat-tailed dunnart and the honey possum (*Tarsipes rostratus*) (Table 19.1) (Cowing et al., 2008). The above results suggest that marsupials may generally have dichromatic colour

Table 19.1 Cone Visual pigments of marsupials and monotremes. The rod-like cone pigments are labeled with (*) and were detected by microspectrophotometry (MSP) and immunohistochemistry (IHC), but not by molecular genetic analysis

	λ_{max} of cone pigments (nm)				
	SWS1		MWS	LWS	
Marsupials and monotremes	UVS	VS			Methods of analysis
Australian Marsupials					
Tammar wallaby	–	420	539	–	Mol. Genetics
Fat-tailed dunnart	360	–	509*, 532	–	MSP & Mol. Genetics
Stripe-faced dunnart	360	–	530	–	Mol. Genetics
Honey possum	350	–	509	557	Mol. Genet. & MSP
Brush-tailed possum	360	–	–	556	Mol Genet & MSP
Long-nosed bandicoot	360	–	–	556	Mol. Genetics
Quoka	–	420	505*, 538	–	MSP & IHC
Quenda (Southern-brown bandicoot)	360	–	509*	551	MSP & IHC
South American Marsupials					
Gray short-tailed opossum	362	–	–	557	Mol. Genetics
Big-eared opossum	362	–	–	557	Mol. Genetics & IHC
Australian Monotremes					
Platypus	–	450	–	550	Mol. Genetics
Echidna	–	450	–	560	Mol. Genetics

vision. PCR amplification attempts may have missed some cone pigments that have diverged substantially from eutherian mammalian sequences. This seems unlikely because extensive amplification attempts employed a large number of degenerate primer pairs that are homologous to conserved coding regions of the cone visual pigment genes. What then is the genetic basis of the observed trichromacy in some marsupials?

Microspectrophotometric (MSP) analysis presented evidence that the fat-tailed dunnart and the honey possum have three distinct classes of cone photoreceptors (Table 19.1) in their retinae: SWS1, MWS and LWS (Arrese et al., 2002), indicating the possibility of trichromatic colour vision. In addition, using immunohistochemistry and MSP, the same group showed that two additional Australian marsupials, the quokka (*Setonix brachyurus*) and quenda (Southern brown bandicoot, *Isoodon obesulus*) also have three classes of retinal cones (Table 19.1) (Arrese et al., 2005). It is interesting to note that the MWS cone pigments observed in the above four species have λ_{max} values that are almost identical to those of rods

(500–510 nm) (Fig. 19.2a). Does this suggest that rhodopsin may be expressed in both rods and a specific class of cones? The interesting finding of a potentially duplicated rhodopsin 1 (*RH1*) gene in the fat-tailed dunnart genome (Cowing et al.,

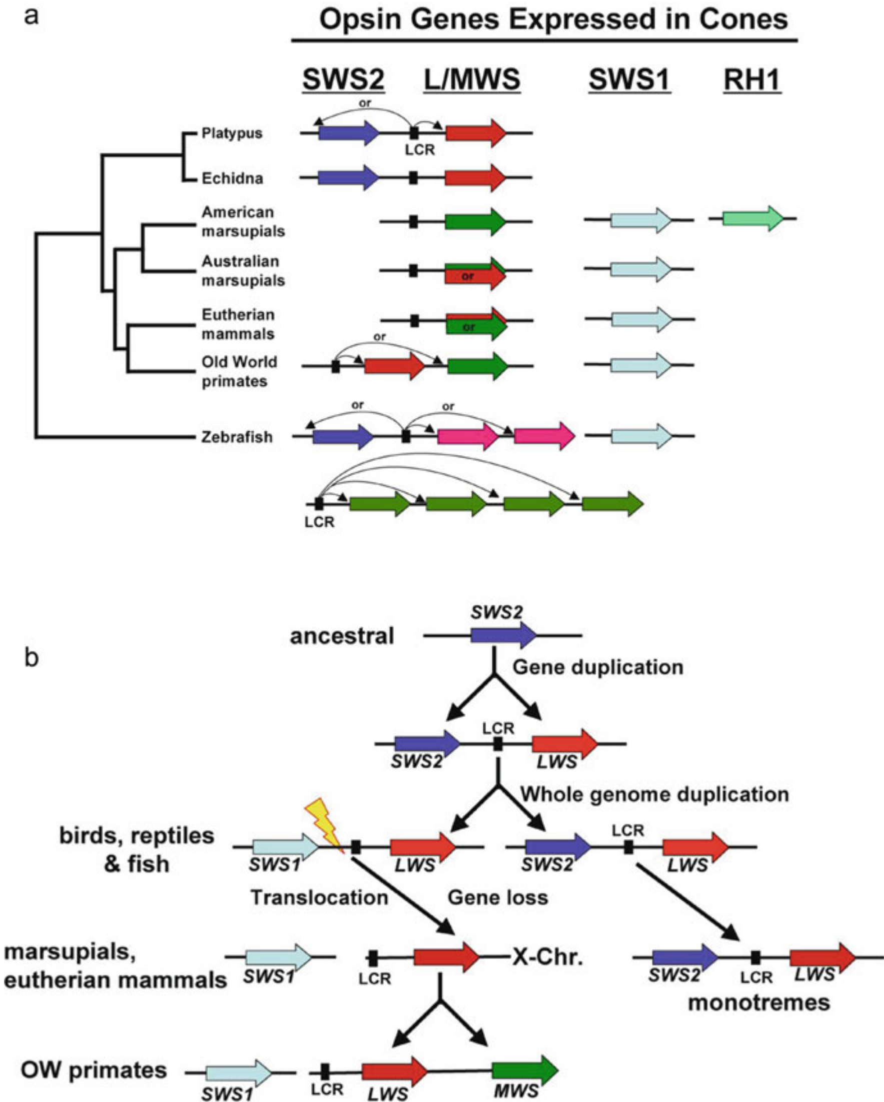


Fig. 19.2 Structure and evolution of the vertebrate cone visual pigment genes. (**a**) Phylogenetic tree of the cone opsin genes (adapted from Wakefield et al., 2008). *Arrows* pointing from the locus control regions (LCR) to the promoters of the opsin genes control expression of only one of the genes per cone photoreceptor cell. The rod RH1 opsin was detected in a class of cones in some marsupials (Arrese et al., 2005). (**b**) Evolutionary pathway of the opsin genes. The ancestral *SWS* pigment gene duplicated to form the *SWS-LCR-LWS* tandem arrangement that existed prior to the

2008) supports the possibility that one of these *RH1* genes may be expressed in rods and the other specifically in cones. The duplicate *RH1* genes encode visual pigments with identical amino acid (but differ in intronic sequence) sequences and λ_{max} value. One would have to assume that some changes had occurred in the regulatory sequences of one of the *RH1* genes to direct its expression, in a mutually exclusive fashion, in cones. As shown in Table 19.1, some of the marsupials have LWS cone pigments with λ_{max} values of 551–557 nm, while others have cone pigments close to the MWS pigments in other mammals with λ_{max} values of 530–54 nm. Does this difference correlate with adaptation to a specific visual behaviour, such as being able to target food with a particular colour? This was explored by studying the behaviour of the honey possum, which has a LWS pigment with a λ_{max} of 557 nm (Sumner et al., 2005). The honey possum targets a specific flower because it feeds almost exclusively on nectar and pollen. It was found that the longer-wavelength cone pigment offered the honey possum an advantage in detecting its target flower against their natural environment. More studies are needed to verify these results.

19.2.2 South American Marsupials

The Australian marsupials separated from the South American marsupials about 80 MYA (Wakefield and Graves, 2005). It would be valuable to investigate differences in the visual pigment sequences and gene structures and locations between these marsupials.

Immunohistochemistry, using monoclonal antibodies that have been proven to consistently label MWS/LWS and SWS1 cones in eutherian mammals, revealed the presence of only two classes of cone photoreceptors in the retina of the South American big-eared opossum (*Didelphis aurita*) (Ahnelt et al., 1996). Molecular analysis of the coding sequences of cone opsins of two South American marsupial species revealed that the big-eared opossum and the gray short-tailed opossum (*Monodelphis domestica*) only had two classes of cone opsin genes that are expressed in the retinae of both species: *SWS1* and *LWS* (Hunt et al., 2009)

Fig. 19.2 (continued) whole-genome duplication in vertebrates. After genome-wide duplication, a large fraction of genes were lost and some evolved to encode different proteins such as the *SWS1* gene. The *SWS2-LCR-LWS* cluster was transmitted to monotremes after loss of the *SWS1* pigment gene. Subsequently, the *LCR-LWS* gene was translocated to the X-chromosome in marsupials and other mammals, while the *SWS1* gene remained on an autosome. Finally, the X-linked *LWS* gene was duplicated, about 40 MYA, to form the LCR-*LWS-MWS* array in the Old World (OW) primates. The LCR sequence is highly conserved and plays a crucial role in mutually exclusive expression of the genes in a cluster. Most New World (NW) primates have only two types of photoreceptors and dichromatic colour vision. However, in the howler monkey gene duplication and evolution have generated three opsin genes and trichromatic colour vision. Interestingly, trichromatic colour vision also exists among female heterozygotes for common alleles of the single pigment gene on the X-chromosome that encode pigments with different absorption spectra

(Table 19.1). Analysis of the short-tailed opossum genome sequence showed that, as in other mammals, the *LWS* gene is located on the X-chromosome, with an upstream (1.75 kb) locus control region (LCR); the *SWS1* gene is located in the same position on chromosome 8 as that in other vertebrates (Hunt et al., 2009). So far, no MSP or behavioural studies have been performed on these species to determine the potential of trichromatic colour vision.

19.2.3 Topography of Retinal Photoreceptors in Marsupials and Colour Vision

The topography and spectral characteristics of photoreceptors correlate with ecological adaptations and evolutionary history of sensory systems. Colour vision requires comparison between signal outputs of signals from different types of photoreceptors and proper neural wiring. The distribution and relative number of cones in the retina contribute to colour vision. In primates, each of the MWS and LWS cones in the fovea privately transmit signals to midget bipolar cells and then to midget ganglion cells. This is thought to be required for comparing signal outputs from these two types of photoreceptors in order to perceive colours (Mollon, 1989; Boycott and Wassle, 1999: Wassle, 2004). However, non-primate mammals do not possess midget bipolar cells, and therefore, it was thought that the addition of extra types of cones into the retina would not result in acquisition of enhanced colour discrimination. This was tested in a genetically engineered mouse that expressed both *SWS, MWS*, along with a *LWS* transgene (Jacobs et al., 2007). It was observed that additional chromatic discrimination was acquired, but only when the MWS: LWS cones were in about equal ratios in the retinae. It was proposed that plasticity in the mammalian visual system allows acquisition of additional colour vision due to acquisition of only visual pigment genes. This is consistent with the observed trichromatic colour vision in some marsupials.

The ratio between rods and cones varies widely among mammals. Mammalian retinea contain relatively low numbers of SWS cones compared to MWS and LWS cones. In general, due to adaptation, nocturnal mammals have a high rod: cone ratio (~100:1), while diurnal mammals have a lower rod: cone ratio (~1:19 in the tree-shrew). The echidna and platypus are crepuscular, primarily active during twilight; the short-billed echidna is nocturnal or diurnal. Most marsupials are nocturnal (bandicoot, wombat, Australian possums and American opossums) (McMenamin, 2007). Some are crepuscular or nocturnal (honey possum, tammar wallaby, red kangaroo, fat-tailed dunnart), and very few are diurnal (the numbat).

Does the cone photoreceptor topography in the retinae of these marsupials and monotremes correlate with the reported trichromatic colour vision? In the fat-tailed dunnart, SWS1 cones are concentrated in the dorso-temporal region of the retina (max of 2,300/mm^2) and decrease in number in a gradient to the periphery (2,300–1,500), while MWS cones form a horizontal streak (31,000–21,000/mm^2). In the honey possum, SWS1 cones are uniformly distributed in the retina (2,700/mm^2); MWS cones form a mid-ventral to peripheral gradient (37,000–26,000/mm^2) (Arrese et al., 2003).

The SWS1 cones in the quokka are located in the dorso-temporal region of the retina, while in the quenda; two peaks were identified in the naso-ventral and dorso-temporal regions. In both species, MWS cones were located essentially all over the retina (Arrese et al., 2005). As in the honey possum, quokka and quenda, high enough density and co-localization of SWS1 and MWS/LWS cones in the central retina would contribute to enhanced colour vision. However, in the fat-tailed dunnart, regions of high density of SWS and MWS/LWS cones do not overlap and this may make it difficult to have trichromatic colour vision.

The retina of the tammar wallaby has three distinct cone regions: the dorsal retina has low cone density but a high percentage of SWS cones; the central horizontal band of the retina contains a high MWS cone density; the ventral retina contains high MWS cone density with few S-cones (Hemmi and Grunert, 1999; Arrese et al., 2003). However, the distribution of SWS1 and MWS cone photoreceptors in the honey possum retina is uniform (Arrese et al., 2003). This cone topography is also not likely to support high levels of chromatic discrimination.

The retina of the South American opossum is rod-dominated ($\sim$400,000/mm^2) and SWS1 and MWS cones ($\sim$3,000/mm^2) (Ahnelt et al., 1995), some of which contained colourless oil droplets, as observed in other marsupials. These oil droplets have no contribution to colour vision, but they function in attracting and focusing light on the retina. The SWS1 and MWS cones (without oil droplets) were evenly distributed all over the retina. However, the MWS cones bearing oil droplets were concentrated in the ventral half of the retina.

19.3 Cone Visual Pigments and Genes of Monotremes

The monotreme lineage is comprised of only the duck-billed platypus (*Ornithorhynchus anatinus*) and two genera of the spiny ant-eaters or echidna: *Tachyglossus* and *Zaglossus*, found in Australia and New Guinea. The platypus and echidna are the only egg-laying mammals and occupy a unique position in the mammalian phylogeny. They are the most basal mammal group, and share phenotypic and genotypic characteristics with more distantly related vertebrates (Daish and Grutzner, 2009), including the colour vision genes (Davies et al., 2007; Wakefield, 2008). Therefore, these species provide an important evolutionary link to decipher the evolution of the mammalian genome. Monotremes diverged from the therian mammals about 166 MYA (Woodburne et al., 2003; Bininda-Emonds et al., 2007) (Fig. 19.1).

Spectrophotometric, MSP, electroretinographic or behavioural studies regarding colour vision phenotype in the monotremes have not been performed. We have performed genomic DNA and retinal RNA analyses to determine the sequence and genomic organization of the genes encoding the cone visual pigments of the platypus and the echidna (Wakefield et al., 2008). We performed PCR amplifications using primers designed in coding sequences that are highly conserved among mammals, including marsupials, to target SWS1, MWS, and LWS. This means that any cone pigment sequences that had diverged from other mammalian sequences

would not have been amplified. Amplification using retinal RNA revealed that LWS pigment mRNA, and the encoded opsin, is homologous in sequence and spectral properties to the primate LWS visual pigments. However, we were unable to amplify the mammalian SWS1 pigment from these two species. At the same time, the platypus genome sequence became available. Examining the genome sequence, we found a LWS pigment gene, but surprisingly we found an adjacent (20 kb) *SWS2* pigment gene in the opposite orientation. Interestingly, an LCR regulatory motif that is homologous to that found in other mammals was found between these two genes (Wakefield et al., 2008). It is predicted that this LCR regulates mutually exclusive expression of the adjacent *LWS* and *SWS2* in cone photoreceptors. We obtained the same arrangement of the *LWS* and *SWS2* pigment genes after sequencing these in the echidna. The SWS2 and LWS pigments of the platypus were expressed in vitro and their spectra had λ_{max} values of 451 and 550 nm, respectively (Davies et al., 2007). The SWS2 pigment is paralogous to the SWS2 pigments typically found in reptiles, birds, and fish but not to the SWS1 of other mammals (Davies et al., 2007; Wakefield, 2008). The *SWS1* gene in the monotremes has been lost, except for exon 5 (Davies et al., 2007). The presence of *SWS1* exon 5 sequence in the genome of both the platypus and echidna dates the deletion of the other exons prior to the echidna and platypus divergence ∼21 MYA. No remnant sequences *SWS2* of the gene were found in the tammar wallaby, opossum or eutherian mammals. This indicates that the therian ancestor lost the *SWS2* gene prior to marsupial and eutherian divergence (Fig. 19.2a) (Wakefield et al., 2008). Based on these data, a pathway of evolution of the visual pigment genes is outlined in Fig. 19.2b. The tandem gene pair of *SWS2-LWS*, with the LCR in between, was formed by gene duplication of the *SWS2* gene followed by mutations in amino acids that interact with the chromophore and changed the λ_{max} value of one of these genes to that of a *LWS*. This arrangement most likely existed before duplication of the vertebrate genome (Ohno, 1970; Dehal and Boore, 2005). After whole-genome duplication the *SWS* genes diverged into *SWS1* and *SWS2*. The *LWS* and LCR were translocated to the X-chromosome. The leftover *LWS* was duplicated in primates to form the LCR-*LWS*-*MWS* gene arrays and the re-evolution of trichromatic colour vision.

19.4 Future Experiments

1. Investigate whether duplication of the *RH1* gene exists in other marsupials that were shown to have three cone opsin classes, and potentially trichromatic colour vision.
2. Analyze genome sequences of marsupials for presence of novel opsin genes or remnants of known genes.
3. Extend behavioural studies to test colour vision of additional marsupials and monotremes.
4. Investigate the relationship between ecological adaptations and spectral characteristics of marsupial visual pigments and their retinal topography.

Acknowledgments This work was supported by National Institutes of Health (National Eye Institute) grant number EY008395.

References

Ahnelt PK, Hokoc J, Rohlich P (1995) Photoreceptors in a primitive mammal, the South American opossum, Didelphis marsupialis aurita: characterization with anti-opsin immunolabeling. Vis Neurosc 12:793–804.

Ahnelt PK, Hokoc J, Rohlich P (1996) The opossum photoreceptors–a model for evolutionary trends in early mammalian retina. Rev Bras Biol 56(Suppl 1(2)):199–207.

Arrese CA, Hart NS, Thomas N, Beazley LD, Shand J (2002) Trichromacy in Australian marsupials. Curr Biol 12:657–660.

Arrese CA, Rodger J, Beazley LD, Shand J (2003) Topographies of retinal cone photoreceptors in two Australian marsupials. Vis Neurosci 20:307–311.

Arrese CA, Oddy AY, Runham PB, Hart NS, Shand J, Hunt DM, Beazley LD (2005) Cone topography and spectral sensitivity in two potentially trichromatic marsupials, the quokka (*Setonix brachyurus*) and quenda (*Isoodon obesulus*). Proc Biol Sci 272:791–796.

Arrese CA, Beazley LD, Neumeyer C (2006) Behavioural evidence for marsupial trichromacy. Curr Biol 16:R.

Bininda-Edmonds OR, Cardillo M, Jones KE, Macphee RD, Beck RM, Grenyer R, Price SA, Vos RA, Gittleman JL, Purvis A (2007) The delayed rise of present-day mammals. Nature 446:507–512.

Boycott B, Wassle H (1999) Parallel processing in the mammalian retina: the Proctor Lecture. Invest Ophthalmol Vis Sci 40:1313–1327.

Collin SP, Knight MA, Davies WL, Potter IC, Hunt DM, Trezise AE (2003) Ancient colour vision: multiple opsin genes in the ancestral vertebrates. Curr Biol 13:R.

Cowing JA, Arrese CA, Davies WL, Beazley LD, Hunt DM (2008) Cone visual pigments in two marsupial species: the fat-tailed dunnart (Sminthopsis crassicaudata) and the honey possum (Tarsipes rostratus). Proc Biol Sci 275:1491–1499.

Daish T, Grutzner F (2009) Location, location, location! Monotremes provide unique insights into the evolution of sex chromosome silencing in mammals. DNA Cell Biol 28: 91–100.

Davies WL, Carvalho LS, Cowing JA, Beazley LD, Hunt DM, Arrese CA (2007) Visual pigments of the platypus: a novel route to mammalian colour vision. Curr Biol 17:R161–R163.

Deeb SS, Wakefield MJ, Tada T, Marotte L, Yokoyama S, Graves JAM (2003) the cone visual pigments of an Australian marsupial, the tammar wallaby (Macropus eugenii): sequence, spectral tuning, and evolution. Mol Biol Evol 20:1642–1649.

Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3:e314.

Dominy NJ, Svenning JC, Li WH (2003) Historical contingency in the evolution of primate color vision. J Hum Evol 44:25–45.

Dulai KS, Von Dornum M, Mollon JD, Hunt DM (1999) The evolution of trichromatic color vision by opsin gene duplication in New World and Old World primates. Genome Res 9:629–638.

Hemmi JM (1999) Dichromatic colour vision in an Australian marsupial, the tammar wallaby. J Comp Physiol A 185:509–515.

Hemmi JM, Grunert U (1999) Distribution of photoreceptor types in the retina of a marsupial, the tammar wallaby (*Macropus eugenii*). Vis Neurosci 16:291–302.

Hunt DM, Chan J, Carvalho LS, Hokoc JN, Ferguson MC, Arrese CA, Beazley LD (2009) Cone visual pigments in two species of South American marsupials. Gene 433:50–55.

Jacobs GH (2008) Primate color vision: a comparative perspective. Vis Neurosci 25:619–633.

Jacobs GH, Deegan JF (2005) Polymorphic New World monkeys with more than three M/L cone types. J Opt Soc Am A Opt Image Sci Vis 22:2072–2080.

Jacobs GH, Williams GA (2006) L and M cone proportions in polymorphic New World monkeys. Vis Neurosci 23:365–370.
Jacobs GH, Neitz M, Deegan JF, Neitz J (1996) Trichromatic colour vision in New World monkeys. Nature 382:156–158.
Jacobs GH, Rowe MP (2004) Evolution of vertebrate colour vision. Clin Exp Optom 87:206–216.
Jacobs GH, Williams GA, Cahill H, Nathans J (2007) Emergence of novel color vision in mice engineered to express a human cone photopigment. Science 315:1723–1725.
McMenamin PG (2007) The unique paired retinal vascular pattern in marsupials: structural, functional and evolutionary perspectives based on observations in a range of species. Br J Ophthalmo 91:1399–1405.
Mollon JD (1989) "Tho' she kneel'd in that place where they grew." The uses and origins of primate colour vision. J Exp Biol 146:21–38.
Ohno S (1970) Evolution of by Gene Duplication. Springer, New York.
Regan BC, Julliot C, Simmen B, Vienot F, Charles-Dominique P, Mollon JD (2001) Fruits, foliage and the evolution of primate colour vision. Philos Trans R Soc Lond B Biol Sci 356:229–283.
Rowe MP, Jacobs GH (2004) Cone pigment polymorphism in New World monkeys: are all pigments created equal? Vis Neurosci 21:217–222.
Smallwood PM, Wang Y, Nathans J (2002) Role of a locus control region in the mutually exclusive expression of human red and green cone pigment genes. Proc Natl Acad Sci U S A 99: 1008–1011.
Strachan J, Chang LY, Wakefield MJ, Graves JAM, Deebs SS (2004) Cone visual pigments of the Australian marsupials, the stripe-faced and fat-tailed dunnarts: sequence and inferred spectral properties. Vis Neurosci 21:223–229.
Sumner P, Arrese CA, Partridge JC (2005) The ecology of visual pigment tuning in an Australian marsupial: the honey possum Tarsipes rostratus. J Exp Biol 208:1803–1815.
Wakefield MJ, Graves JAM (2005) Marsupials and monotremes sort genome treasures from junk. Genome Biol 6:218.
Wakefield MJ, Anderson M, Chang E, Wei KJ, Kaul R, Graves JAM, Grutzner F, Deeb SS (2008) Cone visual pigments of monotremes: filling the phylogenetic gap. Vis Neurosci 25:257–264.
Wang Y, Smallwood PM, Cowan M, Blesh D, Lawler A, Nathans J (1999) Mutually exclusive expression of human red and green visual pigment-reporter transgenes occurs at high frequency in murine cone photoreceptors. Proc Natl Acad Sci U S A 96:5251–5256.
Wassle H (2004) Parallel processing in the mammalian retina. Nat Rev Neurosci 5:747–757.
Woodburne MO, Rich TH, Springer MS (2003) The evolution of tribospheny and the antiquity of mammalian clades. Mol Phylogenet Evol 28:360–385.
Yokoyama S (2002) Molecular evolution of color vision in vertebrates. Gene 300:69–78.
Yokoyama S, Blow NS, Radlwimmer FB (2000) Molecular evolution of color vision of zebra finch. Gene 259:17–24.

Chapter 20
The Evolutionary History of Globin Genes: Insights from Marsupials and Monotremes

Vidushi S. Patel and Janine E. Deakin

Abstract Haemoglobin, which is required for oxygen transportation in the blood, is encoded by members of the alpha (α) and beta (β) globin gene families. They are highly regulated throughout different stages of development in a tissue-specific manner. The number and type of α- and β-globin genes varies between jawed vertebrates. In teleosts and amphibians α- and β-like globin genes are clustered together, but in birds and mammals these genes form distinct clusters on different chromosomes. This chapter reviews how data from marsupials and monotremes have contributed in discovering novel globin genes, which in turn has clarified how globin genes have evolved throughout amniote evolution. We also provide a detailed view of how the duplication of a single primordial globin gene (about 500 million years ago) along with other subsequent events have resulted in more complex α- and β-globin clusters in extant vertebrates. In addition, how this has impacted the current fate and regulation of α- and β-genes in mammals will also be discussed.

Keywords α-globin · β-globin · Evolution · Haemoglobin · Marsupials · Monotremes · Transposition

20.1 Introduction

Haemoglobin is vital for the survival of all species, and originated as an iron-complexed protein at about the time when cellular life originated on earth (Hardison, 2001). It is found in all living organisms including prokaryotes, fungi, plants and animals, and its function ranges from catabolic metabolism in bacteria, yeast and algae, to intercellular oxygen transportation in vertebrates (reviewed in Hardison, 1998). All globin genes are thought to have evolved from a single common ancestral globin via a series of gene duplications, which, due to selective pressures and

V.S. Patel (✉)
ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia
e-mail: vidushi.patel@anu.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_20, © Springer Science+Business Media B.V. 2010

changes in atmospheric oxygen levels, led to the diversity of specialised forms and functions of haemoglobin across all kingdoms of organisms (Hardison, 1998).

In vertebrates, haemoglobin is a major component in blood, but its structure differs from species to species and among isoforms. In agnatha (jawless vertebrates), such as lamprey and hagfish, the haemoglobin molecule is a monomer in the oxygenated form, but changes to heterodimers and heterotetromers in the deoxygenated form (Qiu et al., 2000; Muller et al., 2003). In all gnathostomes (jawed vertebrates), it is always found as a heterotetrameric molecule composed of two alpha (α)-globin and two beta (β)-globin polypeptides, with a heme tightly bound to a pocket in each globin monomer. In highly oxygenated surroundings, i.e. in lungs or gills, the interaction between each globin polypeptide chain allows heme to cooperatively bind to oxygen and unloads it efficiently in the peripheral respiring tissues that need oxygen for efficient metabolism.

Most gnathostomes possess multiple variants of α-like and β-like globin genes that are located in clusters, and upon activation constitute different forms of haemoglobin during different stages of ontogeny. The multiplicity of these genes has the advantage of producing different haemoglobin molecules with distinct oxygen binding affinity during different stages of development and in different tissues. This enables the organism to cope with environmental changes in oxygen tension (Coates, 1975). For example, early expressed haemoglobins in embryos lack cooperative oxygen binding and have a higher oxygen affinity than adult haemoglobin. It allows the embryo to extract more oxygen with greater efficiency from their surroundings or maternal circulation (Hofmann et al., 1997; Brownlie et al., 2003). As the gas exchange structures develop and environmental conditions improve around fishes and amphibians, or placental barriers are reduced in eutherian mammals, these high oxygen affinity haemoglobins are then slowly replaced by lower affinity haemoglobins with greater cooperative binding (Iuchi, 1973; Brittain et al., 1997). Thus, there is a huge demand for coordinating and regulating these α- and β-globin genes during different stages of ontogeny and tissues in gnathostomes for their balanced production in forming proper haemoglobin molecules, which otherwise if altered could lead to haemoglobinopathies such as sickle cell amenia and thalassemia (Wells, 1999).

20.2 Unique Globin Properties in Marsupial Newborns

In marsupial embryos and newborns, the red blood cells and haemoglobin properties are quite different to those in older pouch young or adult marsupials. At birth, the red blood cells of marsupials are all nucleated and larger, whereas in more mature pouch young and adults the red blood cells lack nuclei and are smaller (Holland and Gooley, 1997). Haemoglobin in marsupial embryos from species such as the tammar wallaby (*Macropus eugenii*) and brush-tail possum (*Trichosurus vulpecula*) have lower affinity for oxygen than other vertebrate embryos and foetuses (Tibben et al., 1991; Calvert et al., 1994). This allows marsupial embryos to extract less oxygen from maternal circulation (Holland and Gooley, 1997). The lower affinity

for oxygen is really interesting, particularly given that marsupials have a short period of gestation and are born at a very pre-mature stage of development with an underdeveloped respiratory system. In the tammar wallaby neonate, there are four different types of haemoglobin molecules in embryonic blood, which are different from those found in adults (Holland et al., 1998). These molecules are also larger than the adult haemoglobin molecules, containing probably eight sub-units rather than the normal four sub-units (Holland and Gooley, 1997; Holland et al., 1998). These specialised forms of haemoglobin may play a role in the transition from placental respiration in utero to pulmonary respiration in the pouch, since they disappear a few days after birth once the nucleated red blood cells are replaced by adult-like non-nucleated red blood cells (Tyndale-Biscoe, 2005).

These unique properties make studying marsupial globin genes of particular interest. Genomic studies on genes encoding haemoglobin molecules have led to the discovery of novel globin genes in marsupials, which then sparked new interest in determining the evolutionary history of jawed vertebrate α- and β-globin genes.

20.3 How Did α- and β-Globin Clusters Evolve in Jawed Vertebrates?

20.3.1 History of Globin Gene Evolution

Globin gene evolution can be traced by studying organisational gene structure and function of α- and β-globin genes in different jawed vertebrate lineages (i.e. from cartilaginous fishes to humans). In cartilaginous fishes such as sharks and ray fishes, although the whole organisational gene structure has not yet been studied, the structure of their haemoglobin molecules is similar to that of other fishes and tetrapods, being composed of two α and two β globin polypeptide chains (Chong et al., 1999; Naoi et al., 2001; Verde et al., 2005). These α and β chains have many structural similarities with human α- and β-globin genes respectively (Chong et al., 1999; Naoi et al., 2001), and therefore the sequence and structure of these chains have been conserved across all jawed vertebrates and evolved from a common ancestral globin gene before fish-tetrapod divergence (∼500 MYA).

In teleosts and amphibians, α- and β-globin gene families, each containing a set of differentially expressed genes, are closely linked. The main difference between these two species are that in teleosts the globin genes are linked in a "head to head" orientation (α, 3′–5′; β, 5′–3′), so that they are transcribed in opposite directions (Wagner et al., 1994; Chan et al., 1997; Miyata and Aoki, 1997; Brownlie et al., 2003), whereas in amphibians the globin genes are arranged in a "head to tail" orientation (α, 5′–3′, β, 5′–3′) (Jeffreys et al., 1980; Patient et al., 1980; Hosbach et al., 1983; Fuchs et al., 2006). Also, in the amphibians (*Xenopus laevis* and *X. tropicalis*), a new member of the globin superfamily was recently discovered adjacent (3′) to the α-β cluster. This gene called Globin Y (*GBY*) encoded a *bona fide* functional globin peptide of 156 amino acid and bore sequence hallmarks of a functional respiratory

protein (i.e. contains three-exons/two-introns with conserved splicing sites), but the precise function is currently unknown (Fuchs et al., 2006).

Unlike teleostan fishes and amphibians, birds and eutherian mammals possess two unlinked clusters for α- and β-genes that are located on different chromosomes. The number of globin genes in each of the α- and β-globin clusters are also different. For example, the chicken (*Gallus gallus*) α-globin cluster contains three α-like globin genes (*HBP, HBK, HBA*) located on chromosome 14 (Engel and Dodgson, 1980; Dodgson et al., 1981), whereas their β-globin clusters contain four β-globin genes (*HBR, HBB-T1, HBB-T2, HBE*) located on chromosome 1 (Dolan et al., 1981; Villeponteau and Martinson, 1981) (Fig. 20.1). In eutherian mammals,

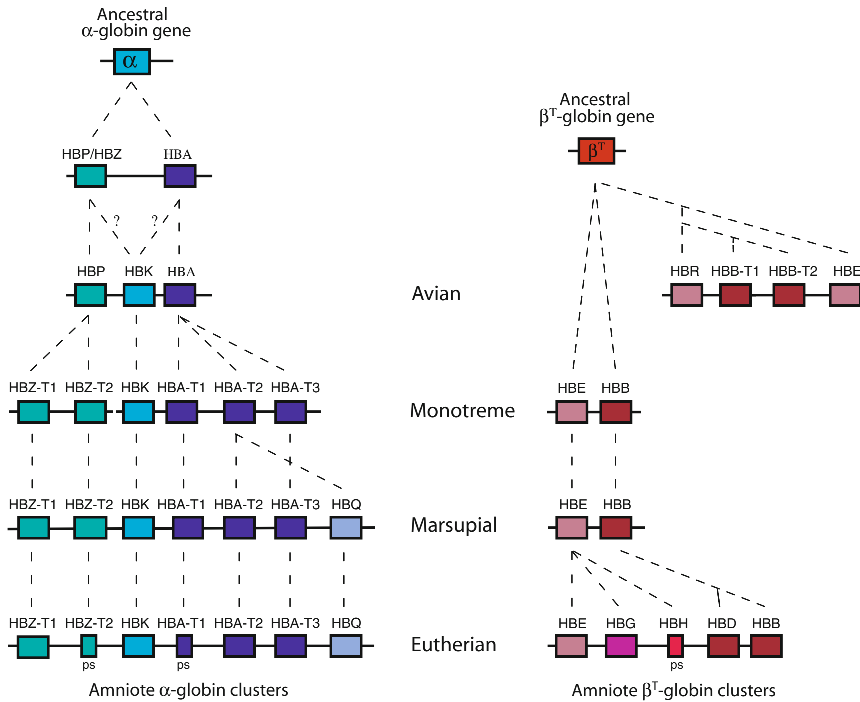


Fig. 20.1 The gene organisation, structure and evolution of α- and β-like globin genes in different amniote lineages. Globin genes are named according to the new standard nomenclature system specified by Aguileta et al. (2006b), where they should start with *HB* representing haemoglobin and the third letter representing their generic names. For example, *HBA* is for α-globin gene in haemoglobin, *HBB* is for β-globin gene in haemoglobin, E is for epsilon (ε), P for pi (π), K for alphaD (α^D), R for rho (ρ), and so forth. The symbol "-T" followed by a number indicates that the gene is a part of a known tandemly duplicated gene block and the number corresponds to the linkage order within the block according to the 5′–3′ orientation. The lowercase letter "ps" at the end of the symbol (in the text) or below the *squares* (in the figure) designates a psuedogene (Aguileta et al., 2006b). This figure not only shows the gene organization in various amniote species, but also how they evolved from a common ancestral α- and β^T-globin genes and their orthologous relationship

all of these genes are present but their numbers have expanded due to many tandem duplications within the cluster. For example, the human (*Homo sapiens*) α-globin cluster contains seven genes arranged in the order *HBZ-T1, HBZ-T2ps, HBK, HBA-T1ps, HBA-T2, HBA-T3, HBQ* and located on chromosome 16 (Orkin, 1978; Lauer et al., 1980; Proudfoot and Maniatis, 1980; Goh et al., 2005), whereas their β-globin cluster (*HBE, HBG-T1, HBG-T2, HBHps, HBD, HBB*) is located on chromosome 11 (Bernard et al., 1979; Fritsch et al., 1980). Irrespective of their chromosomal locations, the amino acid sequences of α and β-globin genes display at least 50% similarity, with similar structures (three-exons/two-introns) to that of fishes and amphibians, suggesting that the structure of globin genes has been conserved since prior to fish-tetrapod divergence.

There are a lot of sequence and structural similarities between the avian and eutherian α-globin clusters, but not between their β-globin clusters. The genes in the avian α-globin cluster show high sequence similarities to their orthologous counterparts in the mammalian α-globin cluster (Proudfoot et al., 1982). For example, *HBP* from chicken is orthologous to *HBZ* from mammals (Fig. 20.1) with 71% sequence identity at protein level. Moreover, the genes in both clusters are arranged in their order of expression, that is, *HBP/HBZ* are first expressed during embryonic stages, followed by *HBK* during late embryonic/fetal and adult stages, and finally *HBA* is dominantly expressed during adult life (Bruns and Ingram, 1973; Proudfoot et al., 1980). However, these properties are not seen for their β-globin clusters. The alignment score between the avian and mammalian β-globin sequences are poor (Reitman et al., 1993). For instance, the avian *HBE* is as equally similar to its eutherian orthologue (*HBE*) as it is to other members of the eutherian β-globin cluster. The genes are also not arranged in their order of expression in aves (Dodgson et al., 1981) but are in mammals (Lacy et al., 1979; Lauer et al., 1980; Maniatis et al., 1980). These differences are believed to be a result of independent duplications of an ancestral β-globin gene in the avian and mammalian lineages (Proudfoot et al., 1982; Reitman et al., 1993).

The fact that α- and β-genes are closely linked in teleosts and amphibians, their structures are similar to eutherian globin genes, and the eutherian α- and β-globin amino acid sequences are homologous, suggests that these genes evolved by a tandem duplication of the ancestral primordial globin gene some time prior to the divergence of gnathostomes ~500 MYA (Fig. 20.2; Dayhoff et al., 1972; Jeffreys et al., 1980; Goodman et al., 1987). Therefore, the duplicated ancestral proto α-β globin genes were closely linked in fishes and frogs, but became separated in the common ancestor of all amniotes (315 MYA).

Jeffreys et al. (1980) were the first to hypothesize two possible models for the evolution of separate α- and β-globin clusters in amniotes. One model suggested that the separation of these genes occurred by translocation between the α- and β-globin genes, possibly by chromosomal rearrangements after amphibian-amniote divergence, resulting in two unlinked clusters (Pisano et al., 2003). The alternative model suggested that two distinct clusters arose via chromosomal duplication, where the duplicate copies would then have evolved to the present arrangement found in birds and mammals by tandem duplications and silencing of

the linked α- or β-globin genes within each cluster. According to this model, some evolutionary remnants of inactivated globin genes would be present near contemporary amniote globin gene clusters. For instance, a "fossil" β-like globin gene would be present within the amniote α-globin cluster and likewise, a "fossil" α-like globin gene would be present within the amniote β-globin cluster.

20.3.2 The Discovery of a Novel Marsupial Globin Gene: Implications for Globin Gene Evolution

Early work focused on studying globin genes in birds and eutherian mammals leaving a void in our knowledge of globin genes from lineages falling between these two divergent groups of animals. Studies of α- and β-globin genes in marsupials, thus, provided useful information not only about the organization of α- and β-globin clusters, but also led to the discovery of novel globin genes. This has permitted the reconstruction of the ancestral α- and β-globin clusters in the therian ancestor.

Characterisation of marsupial globin genes showed that their α-globin clusters were similar to humans (i.e. *HBZ-T1, HBZ-T2, HBK, HBA-T1, HBA-T2, HBA-T3, HBQ*; Fig. 20.1) and located on chromosome 1 in the tammar wallaby and striped-faced dunnart (*Sminthopsis macroura*) (De Leo et al., 2005; Cooper et al., 2006). The *HBK* (known as α^{D}-globin or μ-globin) was previously thought only to be present in birds and reptiles, and lost from the genomes of mammals. However, the discovery of *HBK* in marsupials led to the discovery of *HBK* in eutherian mammals (Goh et al., 2005; Cooper et al., 2006; Hoffmann and Storz, 2007). On the other hand, marsupials have much simpler β-globin clusters than do birds and eutherian mammals, being comprised of only two β-globin genes (*HBE, HBB*; Fig. 20.1) (Koop and Goodman, 1988; Cooper et al., 1996) located on chromosomes 5 and 3 in the tammar wallaby and dunnart respectively (De Leo et al., 2005; Deakin et al., 2008). Since both the α- and β-globin clusters were located on different chromosomes and were orthologous to their respective eutherian α- and β-globin counterparts, it provided support for the Jeffreys et al. (1980) translocation hypothesis where the proto α-β separated from each other by a possible split between them sometime before the radiation of amniotes (>315 MYA).

The recent discovery of a third β-like globin gene called *HBW* in marsupials sparked new interested into determining the evolutionary history of globin genes. The marsupial-specific *HBW* was an extremely interesting discovery since *HBW* was not linked to the main β-globin cluster, but instead was found at the 3′ end of the α-globin cluster (Wheeler et al., 2001, 2004). This "fossil" β-like gene, like other marsupial β-globin genes, was expressed in tammar wallaby neonates and bound with α globin polypeptides to form functional haemoglobin (Holland and Gooley, 1997; Holland et al., 1998). However, the phylogenetic analysis surprisingly indicated that the *HBW* was more closely related to bird β-globin genes rather than mammalian β-globin genes (Wheeler et al., 2001, 2004).

Since *HBW* was seen as an "orphan" (i.e. separated from the other β-globin genes), it provided support for the Jeffreys et al. (1980) chromosome duplication

hypothesis where the whole proto α-β cluster duplicated to another region of the genome and underwent differential gene silencing in different amniote lineages. One of the two globin gene clusters became the β-globin cluster in birds, but the β-globin genes within this cluster became redundant and inactive, and finally deleted in mammals, except for the marsupial *HBW*, which was seen as the last mammalian relic β-globin gene in this cluster (Wheeler et al., 2001, 2004). The other cluster diverged and duplicated to form the β-globin cluster of marsupials and eutherian mammals. Thus, under this model, the avian β-globin genes were orthologous to marsupial *HBW*, and both were paralogous to the mammalian β globin genes (Wheeler et al., 2001, 2004), contradicting previous claims that the β-globin clusters of birds and mammals were orthologous.

The chromosome duplication hypothesis answered some questions but ignited new ones. It made it clear why a "fossil" and "orphaned" β-like *HBW* gene was found adjacent (3′) to the α-globin cluster. This also partially explained why chicken β-globin genes did not show high sequence similarities to mammalian β-globin genes (Reitman et al., 1993). However, the *HBW* gene was not a pseudogene, but a transcriptionally active gene that encoded a functional haemoglobin molecule (reviewed in Hardison, 2008). Also, no evidence has been found for a fossil or functional α-like globin gene adjacent to any amniote β-globin cluster, nor have orthologues of the chicken β-globin gene cluster and marsupial *HBW* gene been found in eutherian mammals. It also failed to explain why the regions spanning the duplicated globin clusters are so different, i.e. the mammalian α-globin clusters reside in a G+C-rich region that contains a CpG island, whereas their β-globin clusters are in an A+T-rich region (reviewed in Hardison, 1998). If the whole α-β cluster duplicated, then one would expect to see some flanking genes and regulatory regions common to both clusters. Therefore, the model proposed by Wheeler et al. (2001, 2004) explained the presence of the *HBW* gene beside the marsupial α-globin cluster, but failed to explain the differences between α- and β-globin clusters in amniotes.

20.3.3 A New Model for Globin Gene Evolution: Insights from Monotremes

The discovery of the marsupial *HBW* gene and the poor explanation of globin evolution by chromosome duplication sparked further interest in solving the mystery of globin gene evolution by looking at the most basal lineage of mammals, the monotremes. This particular mammalian lineage has unique characteristics distinct from other mammals, such being the only egg-laying mammals.

Studies of globin genes in the platypus (*Ornithorhynchus anatinus*) revealed that, like marsupials, the α-globin cluster contains the *HBW* gene, with the entire cluster consisting of a similar arrangement to other therian mammals i.e. *HBZ-T1, HBZ-T2, HBK, HBA-T1, HBA-T2, HBA-T3, HBW, GBY* (Fig. 20.1; Patel et al., 2008). The only two differences are that it does not possess *HBQ* between *HBA-T3* and *HBW*, suggesting that it evolved by a duplication of *HBA-T2* after the divergence of

monotremes and therian mammals (Fig. 20.1). The platypus α-globin cluster also contained a homolog of *GBY*, a gene recently discovered in amphibians by Fuchs et al. (2006). The presence of *GBY* in both amphibians and monotremes at a similar position within the cluster indicated that it was present in the last common ancestor of tetrapods (>354 MYA).

Furthermore, like marsupials, the platypus had three β-globin genes; two in the main β-globin cluster (*HBE, HBB*; Fig. 20.1) located on chromosome 2, and the third (*HBW*) residing adjacent (3′) to the α-globin cluster on chromosome 21 (Opazo et al., 2008; Patel et al., 2008). The finding of *HBW* in monotremes was significant but the phylogenetic analyses presented by Patel et al. (2008) was different to that of Wheeler et al. (2001, 2004). The *HBW* lineage grouped separately from the bird and other mammalian β-globin genes (Patel et al., 2008) (Fig. 20.2). The position of *HBW* in the phlogenetic tree suggested that the *HBW* lineage was more ancient than other β-globin genes in the avian and mammalian β-globin clusters, and that the avian and mammalian β-globin genes were orthologous to each other (Aguileta et al., 2006a; Patel et al., 2008), but paralogous to fish and amphibian β-globin genes and to the marsupial and monotreme β-like (*HBW*) gene.

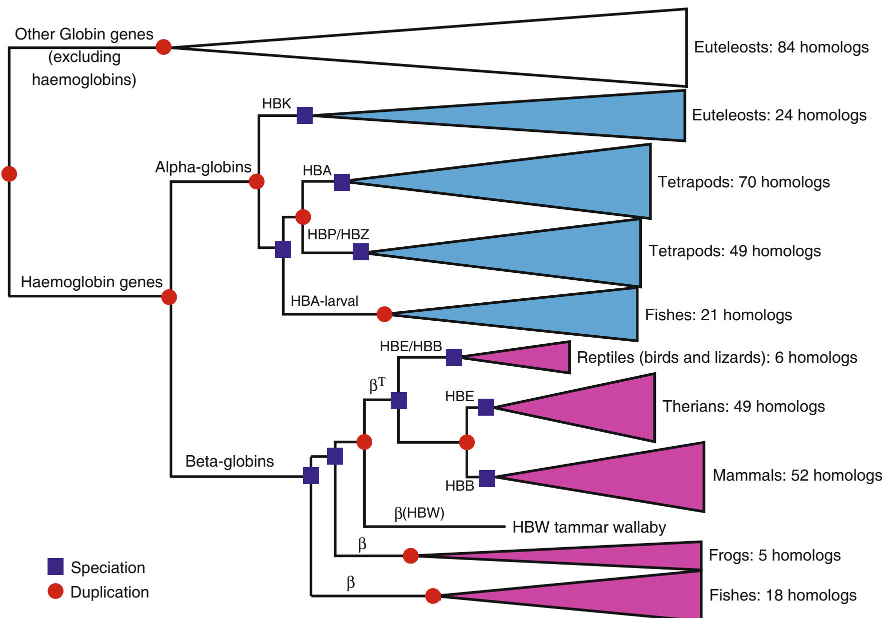


Fig. 20.2 The phylogeny of all globin genes. The globin phylogeny shows how an ancestral globin gene duplicated into ancestral α- and β-globin genes, followed by further duplications and divergences in various jawed vertebrates. This “Gene Tree” was extracted from Ensembl v56 (15 September 2009; Vilella et al., 2009)

Further evidence for this came through flanking analysis of α- and β-globin clusters in other vertebrates. The avian and mammalian α-globin clusters were flanked by *MPG* and *C16orf35* on the 5′ end, and *GBY* and *LUC7L* on the 3′ end. These 5′ flanking markers were also present adjacent to the α-β cluster of teleosts and amphibians (including GBY), supporting its orthology to the α-globin cluster of all amniotes. Therefore, this ancient α-β cluster (and flanking regions) has been conserved for more than 500 MYA, since before the divergence of gnathostomes (Patel et al., 2008).

After the evolution of gnathostomes and the duplication of an ancestral single primordial globin gene into the ancestral proto α-β globin genes (Fig. 20.3), the proto-α globin gene underwent further tandem duplications, subsequent divergence and specialisation of embryonic and adult globin complexes in different vertebrate lineages. It first duplicated sometime before the teleost-tetrapod divergence more than 450 MYA into the progenitors of embryonic globin genes *HBP/HBZ* and adult globin genes *HBA* (Czelusniak et al., 1982; Goodman et al., 1987; Fig. 20.1). In teleosts and amphibians, further tandem duplications, followed by sequence divergence, expanded the embryonic and adult α- and β-globin genes. Then after amphibian-amniote divergence, there was an additional tandem duplication of either *HBP/HBZ* or *HBA* to result in the *HBK* lineage (Fig. 20.1), since all three genes are present in aves (Engel and Dodgson, 1978; Dodgson et al., 1981; Alev et al., 2009) and mammals (Goh et al., 2005; Cooper et al., 2006; Hoffmann and Storz, 2007). The order, timing and origin of these duplications are still unknown. For example, *HBK* may have evolved by duplication of an adult *HBA* (Cooper et al., 2006) or of an embryonic *HBP/HBZ* and this duplication may have occurred before (and not after) tetrapod divergence with *HBK* lost secondarily in amphibians (Hoffmann and Storz, 2007).

After the divergence of birds and mammals, there were further tandem duplications of embryonic and adult α-like globin genes to produce a six-gene cluster (*HBZ-T1, HBZ-T2, HBK, HBA-T1, HBA-T2, HBA-T3*; Fig. 20.1) in monotremes, with ongoing gene conversion events homogenising the two embryonic (*HBZ-T1* and *HBZ-T2*) and three adult (*HBA-T1,2,3*) genes (Patel et al., 2008). There was an additional tandem duplication of *HBA-T2* to generate *HBQ* after monotreme and therian divergence, thus giving rise to a seven-gene cluster (*HBZ-T1, HBZ-T2, HBK, HBA-T1, HBA-T2, HBA-T3, HBQ*; Fig. 20.1) in marsupial (Cooper et al., 2006) and eutherian mammals (Orkin, 1978; Lauer et al., 1980; Proudfoot and Maniatis, 1980; Goh et al., 2005).

The amniote β-globin cluster resides in a completely different genomic context to that of the α-globin cluster, sharing none of its flanking markers with the amniote α-globin cluster, or even the α-β cluster of teleosts and amphibians. Instead, they are surrounded by many olfactory receptor (*OR*) genes (Bulger et al., 1999). Since *OR* genes are a huge superfamily in vertebrates (reviewed in Delbridge et al., Chapter 21; Niimura and Nei, 2005), it was necessary to look at more distant flanking regions for single-copy genes to make genomic context comparisons. In this case, Patel et al. (2008) used the markers *ILK* and *CCKBR,* which were located 5′ of the β-globin cluster, and *RRM1* that was located 3′. These three flanking markers were

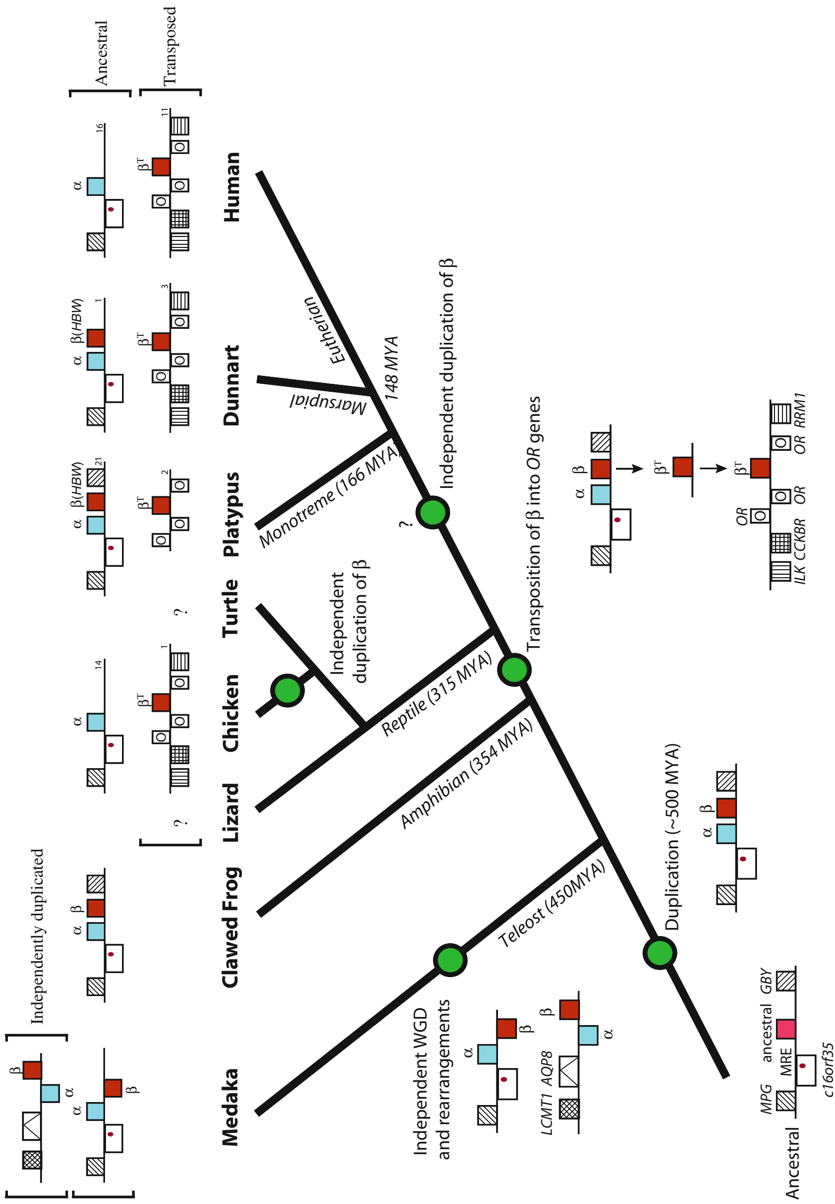


Fig. 20.3 (continued)

always present in their correct orientation near the avian and mammalian β-globin clusters, but not near to any teleost and amphibian α-β globin clusters, indicating that the amniote β-globin cluster evolved recently after the amphibians and amniotes diverged (Proudfoot et al., 1982; Patel et al., 2008; Hardison, 2008). A copy of an ancestral β-like globin gene was transposed (henceforth referred to as β^T-globin gene) from the ancient α-β globin cluster and inserted into the array of *OR* genes flanked by *ILK*, *CCKBR* and *ILK* before avian-mammalian divergence (315 MYA), which then duplicated and diverged further to result in modern β-globin clusters (henceforth referred to as β^T-globin clusters) in contemporary reptilian, avian and mammalian species (Fig. 20.1; Patel et al., 2008).

In the avian lineage, the β^T-globin gene tandemly duplicated independent to that of mammals giving rise to the four-gene cluster *HBP, HBB-T1, HBB-T2, HBE* (Fig. 20.1), followed by some gene conversions and rearrangements (Czelusniak et al., 1982; Reitman et al., 1993). The β^T-globin gene also tamdemly duplicated independently in the mammalian lineage, but the number and timing of duplications are still debatable at this stage since there is an uncertainty in the identification of the monotremes *HBE*. Although the platypus possesses two genes in their β^T-globin cluster (*HBE, HBB*) just like marsupials, its *HBE* gene phylogenetically grouped closely to adult *HBB* genes rather than other mammalian embryonic *HBE* genes (Opazo et al., 2008; Patel et al., 2008). Since platypus *HBE* was located 5′ to *HBB* and had an expression profile similar to other embryonic α-like globin genes (*HBZ-T1* and *HBZ-T2*), Patel et al. (2008) argued that the β^T-globin gene duplicated into progenitors of embryonic *HBE* and adult *HBB* just prior to monotreme-therian divergence (Fig. 20.1), but the ongoing gene conversions homogenised the platypus *HBE* to group with monotreme adult *HBB*.

Alternatively, Opazo et al. (2008) suggested that there had been two independent duplications of the β^T-globin gene in the mammalian lineage; one in the monotreme lineage and the other after monotreme-therian divergence, thereby resulting in a total of three independent duplications of the ancestral β^T-globin gene in amniotes. However, it is not clearly understood how three independent duplications in different lineages would have resulted in genes having a similar pattern of expression during ontogeny. This perhaps could be a consequence of convergent evolution. In either

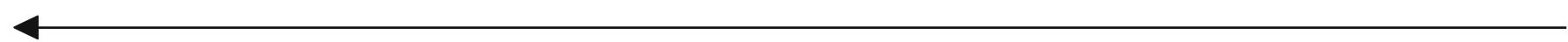


Fig. 20.3 Recent model for the evolution of α- and β-globin clusters in vertebrate lineages. The *circles* drawn on the phylogeny tree indicate major events that have occurred during the evolution of globin genes (i.e. from a duplication of a single primordial globin gene to the transposition event followed by complex tandem/independent duplications and subsequent divergence in different lineages). On the *top panel* of each species are the appearance of the clusters and the flanking genes in them. Note: Generic names (α and β) are used in this context to represent the α and β globin clusters and not the individual genes contained in them. The only exception is *HBW*, which specifically represents the β-like globin gene *HBW*, adjacent to the monotreme and marsupial α-globin clusters. The β^T represents the transposed β-globin gene found within a cluster of olfactory genes (marked as "O" inside the *box*). For individual genes contained in the α, β and β^T globin clusters of various gnathosomes, refer to the text and Fig. 20.1

case, further research into the gene structure and function of β-like globin genes in other birds, reptiles and monotremes would clarify this evolutionary step.

Data from marsupials gave strong evidence that the duplication of the β^T-globin gene into progenitors of embryonic *HBE* and adult *HBB* occurred at least before the divergence of marsupials and eutherian mammals since these two genes are present in all therian mammals studied to date (Koop and Goodman, 1988; Cooper and Hope, 1993; Cooper et al., 1996; De Leo et al., 2005). After marsupial-eutherian divergence, the progenitors of *HBE* and *HBB* tandemly duplicated to produce a five-gene ancestral cluster (*HBE, HBG, HBH, HBD, HBB*; Fig. 20.1) in eutherian mammals (Goodman et al., 1984). In some mammals, additional events such as block duplications, gene conversions and transpositions have resulted in more complex β-globin clusters.

20.3.4 What Was the Mechanism Behind Transposition?

The evolution of the β^T-globin cluster in amniotes was most likely due to a single transposition event that moved a copy of the β-globin gene to a new site containing *OR* genes. The exact mechanism for this recruitment remains unclear at this stage but it is likely to have involved transposable elements for the gene's survival. It also seems possible that the region in which the β^T-globin cluster became embedded in amniotes was a "hot-spot" for transposition events. In marsupials, such as the tammar wallaby and South American opossum (*Monodelphis domestica*), the *FTSJ1* gene, found on the X chromosome in humans, has also transposed into this region beside the β^T-globin cluster on opossum chromosome 4 and tammar wallaby chromosome 5 (Deakin et al., 2008). Therefore, this region appears to have recruited genes from various parts of the genome.

20.4 Regulation of α- and β-Globin Genes

The switching of the α and β globin genes at various developmental stages is common across gnathostomes, indicating a highly conserved regulatory system. The α-globin clusters of birds and therian mammals and the β^T-globin clusters of therian mammals have distal cis-regulatory elements called the major regulatory region (MRE) located in the intron of *C16orf35* for α-globin clusters (Higgs et al., 1990), and locus control region (LCR) for β^T-globin clusters (Forrester et al., 1986; Grosveld et al, 1987). These regulatory regions contain conserved binding sites for transcription factors that control the high level, copy number-dependent, position-independent expression of associated globin gene in transgenic mice (Grosveld et al., 1987; Higgs et al., 1990).

Both these regulatory regions differ from each other in many ways: the MRE is about 0.4 kb long, contains only one DNase hypersensitive site and is located about 20–60 kb upstream of the α-globin loci; whereas the LCR comprises five DNase hypersensitive sites (HS 1–5), is located about 6–20 kb upstream of the

β^T-globin locus and is about 20–25 kb long. Each HS site in the LCR has a core sequence of ~250 bp and contain binding sites for the transcription factors (Stamatoyannopoulos, 2005). The β^T-globin LCR is also required for tissue-specific chromosomal domain opening, whereas no such function has been defined for α-globin MRE (reviewed in Hardison et al., 1997; Higgs and Wood, 2008). The only similarity between them is that α-globin MRE shares many of the structural features of the HS2 site of the β^T-globin LCR (Jarman et al., 1991). Furthermore, the tissue-specific and developmental-specific expression of α- and β-like globin genes are not only controlled by MRE and LCR, but also by other promoters and local enhancers located near them (Stamatoyannopoulos, 1991).

The α-globin MRE is well conserved in amniotes but the β^T-globin LCR is not. The MRE sequences of eutherian mammals show some similarities to those of marsupials (Hughes et al., 2005) and birds, and even to fish, acting as active enhancers (Flint et al., 2001). Although MRE has not been studied in monotremes to date, it is expected that their MRE would be conserved with other mammalian and avian MREs. It has therefore been conserved since the evolution of α-globin genes, prior to the divergence of fish and tetrapods (>450 MYA).

On the other hand, the LCR of eutherian mammals shows conservation only with marsupials (De Leo et al., 2005) and not chickens (Reitman et al., 1993). Data from marsupials show that, although they contain a much simpler β^T-globin cluster than eutherian mammals, their LCR (HS 1–5) is equivalent to that of eutherians. It therefore supports the holocomplex theory that the LCR acts as an integral unit to regulate the β-globin genes, by activating the transcriptional apparatus at the globin gene promoters (reviewed in Engel and Tanimoto, 2000; De Leo et al., 2005). The conservation of these regulatory regions across taxa indicates that their function has also been conserved, even in marsupials whose haemoglobin molecules have different properties in embryos and newborns. Chickens, on the other hand, although they have four HS sites located upstream of their β-globin loci, they do not have major enhancing activities to confer copy number-dependent and high-expression of the β-globin gene in transgenic mice nor the sequences show any conservation with therian HS sites (Reitman and Felsenfeld, 1990; Reitman et al., 1993; Abruzzo and Reitman, 1994; Mason et al., 1995). Instead, chickens have a strong tissue-specific enhancer located within its β-globin loci called β^A/ε enhancer (Choi and Engel, 1986; Hesse et al., 1986) that together with upstream HS sites and local promoters regulate the high expression of β-globin genes in them (Mason et al., 1995). The lack of conservation between the chicken and therian β-globin regulatory regions (Reitman et al., 1993; Abruzzo and Reitman, 1994) suggests that they evolved independently in the avian and mammalian lineages.

20.4.1 How Did Transposition Affect the Regulation of the Globin Clusters?

In amniotes, the α- and β-globin genes are expressed specifically in erythroid cells and require a complex coordinated regulatory network for the balanced expression

of α and β globin genes to form functional haemoglobin. Given that a β^T-globin gene transposed into a region containing *OR* genes that are expressed only in nasal epithelium, raises the question of how has it maintained its erythroid-specific expression. All amniote species studied to date (except monotremes) have a *cis*-regulatory element (LCR for mammals and β^A/ε enhancer for chickens) regulating the expression of genes in their β^T-globin clusters. How this has evolved is a challenging question. Perhaps during transposition of the β^T-globin gene sufficient regulatory sequences from the α-globin MRE were also transposed, which subsequently duplicated and diverged further in amniotes, thus making it quite different to the α-globin MRE. Alternatively, amniotes evolved their own regulatory regions for β^T-globin loci after the transposition event followed by duplications and divergence (reviewed in Hardison, 2008).

In either case, it would be predicted that β^T-globin regulatory regions of aves and mammals would be conserved but indeed they are not. Eutherian LCR share similarities with marsupials only (De Leo et al., 2005), but not with aves (Reitman et al., 1993). The status of the monotreme LCR is debateable at this stage. De Leo et al. (2005) claim to have isolated the platypus HS2 and HS3 sites by amplification using the primers that have been conserved with all other therian mammals. According to De Leo et al. (2005), the sequences of these two hypersensitive sites show high similarities to other marsupial and eutherian HS sites but not to chicken HS sites. Now that the platypus genome assembly (OrnAna1 – March 2007; http://genome.ucsc.edu/cgi-bin/hgGateway) and the bacterial artificial chromosome clones containing the β^T-globin cluster (Patel et al., 2008) have been sequenced, one would assume that these platypus HS2 and HS3 sites would clearly be present. However, the HS2 and HS3 sequences isolated by De Leo et al. (2005) could not be found in the proximity of the platypus β^T-globin cluster, or anywhere else in the platypus genome assembly (OrnAna1). Even attempts to re-isolate these two regions using the same primer sequences and the same method described by De Leo et al. (2005) have failed to amplify any fragments (Patel et al., in preparation). These recent results suggest that there was potentially a contamination issue with the experiments described in De Leo et al. (2005). It is therefore, uncertain if platypus does contain conserved HS2 and HS3 sites or contains a different LCR region that controls the expression of β^T-globin genes.

Given that the transposition event occurred before avian-mammalian divergence (315 MYA), it is possible that the regulatory regions would have diverged significantly, resulting in the low similarity observed between aves and mammals (reviewed in Hardison, 2008). Alternatively, since there has been independent duplications of the β^T-globin gene in avian and mammalian lineages, it is equally likely that their regulatory regions would have evolved independently with them as well, thereby resulting in different HS sites, enhancers and promoters. Further research into the LCR of monotremes and reptilian species would be useful in clarifying the evolution of β^T-globin regulatory regions in amniotes.

20.5 Unsolved Questions and Future Work

In conclusion, many studies into the gene structures of α- and β-globin clusters in various vertebrates have provided clues about their evolution and regulation. It wasn't until the discovery of the marsupial *HBW* gene (Wheeler et al., 2001, 2004), and the subsequent genomic context analysis in monotremes and other species (Patel et al., 2008), that a simplified model for globin gene evolution was proposed. In the ancestor of jawed vertebrates (500 MYA), a single primordial globin gene duplicated into proto α-β globin genes linked together, followed by further tandem duplications, subsequent sequence divergence and specialisation to result in the modern α-β globin cluster in contemporary teleosts and amphibians, and the α-globin cluster in all amniotes (Fig. 20.3).

In amniotes, a distinct and unlinked β-globin cluster is found on a separate chromosome from that of α-globin cluster. Of all proposed models, the transposition model proposed by Patel et al. (2008) provided the most parsimonious explanation of how a copy of β-globin gene, was transposed into a sea of olfactory receptor genes before the evolution of amniotes. The transposed β^T-globin gene duplicated and diverged independently in contemporary aves and mammals to form a dominant β^T-globin cluster. As a consequence, the avian and mammalian β^T-globin clusters are quite different to each other. Furthermore, this transposition event also explained why both α- and β^T-globin clusters in amniotes were so different to each other in respect to their flanking regions, regulatory regions and locations. Research into marsupials and monotremes globin genes have been instrumental in this study, as it has clarified the timing of when many globin genes arose in mammals.

We have come a step closer to understanding globin genes, their evolution and regulation in jawed vertebrates. Based on the studied species, models are proposed to gain insight into what might have happened over the course of evolution. However, all models, including the transposition model demonstrated here, need further validation from other species that have not yet been studied. Sequencing the genomes of various species, including the Australian model tammar wallaby, have advanced and could facilitate research into globin studies. The genome of the tammar wallaby could answer questions such as: (1) Does it contain *GBY*? (2) What are the structures of the MRE and LCR of its α- and β^T-globin clusters respectively? (3) and does the region in close proximity to the β^T-globin cluster show signs of transposable elements? The data from marsupials also need to be compared to other species such as monotremes and reptiles to provide constructive evaluation of the data, and potentially answer how and why the transposition event took place and the consequences of such movement.

In addition, more intensive research into the expression and regulation of globin genes in marsupials, especially in the embryos and newborns is needed. This would perhaps hold the answer to why some marsupial embryos and newborns have unique haemoglobin properties compared to other amniote embryos. Research into globin genes in marsupials has been, and will continue to be, very valuable not only to study globin gene evolution, but also to study their unique properties and regulation.

Acknowledgments We thank Hardip Patel for his assistance with drawing Fig. 20.2.

References

Abruzzo LV, Reitman M (1994) Enhancer activity of upstream hypersensitive site 2 of the chicken beta-globin cluster is mediated by GATA sites. J Biol Chem 269:32565–32571.

Aguileta G, Bielawski JP, Yang Z (2006a) Evolutionary rate variation among vertebrate beta globin genes: implications for dating gene family duplication events. Gene 380:21–29.

Aguileta G, Bielawski JP, Yang Z (2006b) Proposed standard nomenclature for the alpha- and beta-globin gene families. Genes Genet Syst 81:367–371.

Alev C, Shinmyozu K, McIntyre BA, et al. (2009) Genomic organization of zebra finch alpha and beta globin genes and their expression in primitive and definitive blood in comparison with globins in chicken. Dev Genes Evol 219:353–360.

Bernard R, Little PFR, Annison G, et al. (1979) Structure of the human gammaG-gammaA-delta-beta-globin gene locus. Proc Natl Acad Sci U S A 76:4827–4831.

Brittain T, Hofmann OM, Watmough NJ, et al. (1997) A two-state analysis of co-operative oxygen binding in the three human embryonic haemoglobins. Biochem J 326:299–303.

Brownlie A, Hersey C, Oates AC, et al. (2003) Characterization of embryonic globin genes of the zebrafish. Dev Biol 255:48–61.

Bruns GA, Ingram VM (1973) Erythropoiesis in the developing chick embryo. Dev Biol 30:455–459.

Bulger M, van Doorninck JH, Saitoh N, et al. (1999) Conservation of sequence and structure flanking the mouse and human beta-globin loci: the beta-globin genes are embedded within an array of odorant receptor genes. Proc Natl Acad Sci U S A 96:5129–5134.

Calvert SJ, Holland RAB, Gemmell RT (1994) Respiratory Properties of the Neonatal Blood of the Common Brushtail Possum (*Trichosurus vulpecula*). Physiol Zool 67:407–417.

Chan FY, Robinson J, Brownlie A, et al. (1997) Characterization of adult alpha- and beta-globin genes in the zebrafish. Blood 89:688–700.

Choi OR, Engel JD (1986) A 3′ enhancer is required for temporal and tissue-specific transcriptional activation of the chicken adult beta-globin gene. Nature 323:731–734.

Chong KT, Miyazaki G, Morimoto H, et al. (1999) Structures of the deoxy and CO forms of haemoglobin from *Dasyatis akajei*, a cartilaginous fish. Acta Crystallogr D Biol Crystallogr 55:1291–1300.

Coates ML (1975) Hemoglobin function in the vertebrates: an evolutionary model. J Mol Evol 6:285–307.

Cooper SJ, Hope RM (1993) Evolution and expression of a beta-like globin gene of the Australian marsupial *Sminthopsis crassicaudata*. Proc Natl Acad Sci U S A 90:11777–11781.

Cooper SJ, Murphy R, Dolman G, et al. (1996) A molecular and evolutionary study of the beta-globin gene family of the Australian marsupial *Sminthopsis crassicaudata*. Mol Biol Evol 13:1012–1022.

Cooper SJ, Wheeler D, De Leo A, et al. (2006) The mammalian alphaD-globin gene lineage and a new model for the molecular evolution of alpha-globin gene clusters at the stem of the mammalian radiation. Mol Phylogenet Evol 38:439–448.

Czelusniak J, Goodman M, Hewett-Emmett D, et al. (1982) Phylogenetic origins and adaptive evolution of avian and mammalian haemoglobin genes. Nature 298:297–300.

Dayhoff MO, Hunt LT, McLaughlin PJ, et al. (1972) Gene duplications in evolution. In: Dayhoff MO (ed) Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Washington, DC, pp. 17–30.

De Leo AA, Wheeler D, Lefevre C, et al. (2005) Sequencing and mapping hemoglobin gene clusters in the Australian model dasyurid marsupial *Sminthopsis macroura*. Cytogenet Genome Res 108:333–341.

Deakin JE, Koina E, Waters PD, et al. (2008) Physical map of two tammar wallaby chromosomes: a strategy for mapping in non-model mammals. Chromosome Res 16:1159–1175.

Dodgson JB, McCune KC, Rusling DJ, et al. (1981) Adult chicken alpha-globin genes alpha A and alpha D: no anemic shock alpha-globin exists in domestic chickens. Proc Natl Acad Sci U S A 78:5998–6002.

Dolan M, Sugarman BJ, Dodgson JB, et al. (1981) Chromosomal arrangement of the chicken beta-type globin genes. Cell 24:669–677.

Engel JD, Dodgson JB (1978) Analysis of the adult and embryonic chicken globin genes in chromosomal DNA. J Biol Chem 253:8239–8246.

Engel JD, Dodgson JB (1980) Analysis of the closely linked adult chicken alpha-globin genes in recombinant DNAs. Proc Natl Acad Sci U S A 77:2596–2600.

Engel JD, Tanimoto K (2000) Looping, linking, and chromatin activity: new insights into beta-globin locus regulation. Cell 100:499–502.

Flint J, Tufarelli C, Peden J, et al. (2001) Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster. Hum Mol Genet 10:371–382.

Forrester WC, Thompson C, Elder JT, et al. (1986) A developmentally stable chromatin structure in the human beta-globin gene cluster. Proc Natl Acad Sci U S A 83:1359–1363.

Fritsch EF, Lawn RM, Maniatis T (1980) Molecular cloning and characterization of the human beta-like globin gene cluster. Cell 19:959–972.

Fuchs C, Burmester T, Hankeln T (2006) The amphibian globin gene repertoire as revealed by the *Xenopus* genome. Cytogenet Genome Res 112:296–306.

Goh SH, Lee YT, Bhanu NV, et al. (2005) A newly discovered human alpha-globin gene. Blood 106:1466–1472.

Goodman M, Koop BF, Czelusniak J, et al. (1984) The eta-globin gene. Its long evolutionary history in the beta-globin gene family of mammals. J Mol Biol 180:803–823.

Goodman M, Czelusniak J, Koop BF, et al. (1987) Globins: a case study in molecular phylogeny. Cold Spring Harb Symp Quant Biol 52:875–890.

Grosveld F, van Assendelft GB, Greaves DR, et al. (1987) Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell 51:975–985.

Hardison R (1998) Hemoglobins from bacteria to man: evolution of different patterns of gene expression. J Exp Biol 201:1099–1117.

Hardison R, Slightom JL, Gumucio DL, et al. (1997) Locus control regions of mammalian beta-globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights. Gene 205:73–94.

Hardison RC (2001) Organization, evolution and regulation of the globin genes. In: Steinberg MC, Forget BG, Higgs D, et al. (eds) Disorders of Hemoglobin: Genetics, Pathophysiology, Clinical Management. Cambridge University Press, Cambridge.

Hardison RC (2008) Globin genes on the move. J Biol 7:35.

Hesse JE, Nickol JM, Lieber MR, et al. (1986) Regulated gene expression in transfected primary chicken erythrocytes. Proc Natl Acad Sci U S A 83:4312–4316.

Higgs DR, Wood WG (2008) Long-range regulation of alpha globin gene expression during erythropoiesis. Curr Opin Hematol 15:176–183.

Higgs DR, Wood WG, Jarman AP, et al. (1990) A major positive regulatory region located far upstream of the human alpha-globin gene locus. Genes Dev 4:1588–1601.

Hoffmann FG, Storz JF (2007) The αD-globin gene originated via duplication of an embryonic α-like globin gene in the ancestor of tetrapod vertebrates. Mol Biol Evol 24:1982–1990.

Hofmann OM, Brittain T, Wells RM (1997) The control of oxygen affinity in the three human embryonic haemoglobins by respiration linked metabolites. Biochem Mol Biol Int 42:553–566.

Holland RA, Gooley AA (1997) Characterization of the embryonic globin chains of the marsupial Tammar wallaby, *Macropus eugenii*. Eur J Biochem 248:864–871.

Holland RA, Gooley AA, Hope RM (1998) Embryonic globins of the marsupial the Tammar Wallaby (*Macropus eugenii)*: bird like and mammal like. Clin Exp Pharmacol Physiol 25:740–744.

Hosbach HA, Wyler T, Weber R (1983) The *Xenopus laevis* globin gene family: chromosomal arrangement and gene structure. Cell 32:45–53.

Hughes JR, Cheng JF, Ventress N, et al. (2005) Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. Proc Natl Acad Sci U S A 102:9830–9835.

Iuchi I (1973) The post-hatching transition of erythrocytes from larval to adult type in the rainbow trout, *Salmo gairdnerii irideus*. J Exp Zool 184:383–396.

Jarman AP, Wood WG, Sharpe JA, et al. (1991) Characterization of the major regulatory element upstream of the human alpha-globin gene cluster. Mol Cell Biol 11:4679–4689.

Jeffreys AJ, Wilson V, Wood D, et al. (1980) Linkage of adult alpha- and beta-globin genes in *X. laevis* and gene duplication by tetraploidization. Cell 21:555–564.

Koop BF, Goodman M (1988) Evolutionary and developmental aspects of two hemoglobin beta-chain genes (epsilon M and beta M) of opossum. Proc Natl Acad Sci U S A 85:3893–3897.

Lacy E, Hardison RC, Quon D, et al. (1979) The linkage arrangement of four rabbit beta-like globin genes. Cell 18:1273–1283.

Lauer J, Shen CK, Maniatis T (1980) The chromosomal arrangement of human alpha-like globin genes: sequence homology and alpha-globin gene deletions. Cell 20:119–130.

Maniatis T, Fritsch EF, Lauer J, et al. (1980) The molecular genetics of human hemoglobins. Annu Rev Genet 14:145–178.

Mason MM, Lee E, Westphal H, et al. (1995) Expression of the chicken beta-globin gene cluster in mice: correct developmental expression and distributed control. Mol Cell Biol 15:407–414.

Miyata M, Aoki T (1997) Head-to-head linkage of carp alpha- and beta-globin genes. Biochim Biophys Acta 1354:127–133.

Muller G, Fago A, Weber RE (2003) Water regulates oxygen binding in hagfish (*Myxine glutinosa*) hemoglobin. J Exp Biol 206:1389–1395.

Naoi Y, Chong KT, Yoshimatsu K, et al. (2001) The functional similarity and structural diversity of human and cartilaginous fish hemoglobins. J Mol Biol 307:259–270.

Niimura Y, Nei M (2005) Evolutionary dynamics of olfactory receptor genes in fishes and tetrapods. Proc Natl Acad Sci U S A 102:6039–6044.

Opazo JC, Hoffmann FG, Storz JF (2008) Genomic evidence for independent origins of beta-like globin genes in monotremes and therian mammals. Proc Natl Acad Sci U S A 105:1590–1595.

Orkin SH (1978) The duplicated human alpha globin genes lie close together in cellular DNA. Proc Natl Acad Sci U S A 75:5950–5954.

Patel VS, Cooper SJ, Deakin JE, et al. (2008) Platypus globin genes and flanking loci suggest a new insertional model for beta-globin evolution in birds and mammals. BMC Biol 6:34.

Patient RK, Elkington JA, Kay RM, et al. (1980) Internal organization of the major adult alpha- and beta-globin genes of *X. laevis*. Cell 21:565–573.

Pisano E, Cocca E, Mazzei F, et al. (2003) Mapping of alpha- and beta-globin genes on Antarctic fish chromosomes by fluorescence in-situ hybridization. Chromosome Res 11:633–640.

Proudfoot NJ, Gil A, Maniatis T (1982) The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. Cell 31:553–563.

Proudfoot NJ, Maniatis T (1980) The structure of a human alpha-globin pseudogene and its relationship to alpha-globin gene duplication. Cell 21:537–544.

Proudfoot NJ, Shander MH, Manley JL, et al. (1980) Structure and in vitro transcription of human globin genes. Science 209:1329–1336.

Qiu Y, Maillett DH, Knapp J, et al. (2000) Lamprey hemoglobin. Structural basis of the bohr effect. J Biol Chem 275:13517–13528.

Reitman M, Felsenfeld G (1990) Developmental regulation of topoisomerase II sites and DNase I-hypersensitive sites in the chicken beta-globin locus. Mol Cell Biol 10:2774–2786.

Reitman M, Grasso JA, Blumenthal R, et al. (1993) Primary sequence, evolution, and repetitive elements of the *Gallus gallus* (chicken) beta-globin cluster. Genomics 18:616–626.

Stamatoyannopoulos G (1991) Human haemoglobin switching. Science 252:383.

Stamatoyannopoulos G (2005) Control of globin gene expression during development and erythroid differentiation. Exp Hematol 33:259–271.

Tibben EA, Holland RA, Tyndale-Biscoe CH (1991) Blood oxygen carriage in the marsupial, tammar wallaby (*Macropus eugenii*), at prenatal and neonatal stages. Respir Physiol 84:93–104.

Tyndale-Biscoe H (2005) Life Of Marsupials. CSIRO Publishing, Collingwood.

Verde C, De Rosa MC, Giordano D, et al. (2005) Structure, function and molecular adaptations of haemoglobins of the polar cartilaginous fish *Bathyraja eatonii* and *Raja hyperborea*. Biochem J 389:297–306.

Vilella AJ, Severin J, Ureta-Vidal A, et al. (2009) EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. Genome Res 19:327–335.

Villeponteau B, Martinson H (1981) Isolation and characterization of the complete chicken beta-globin gene region: frequent deletion of the adult beta-globin genes in lambda. Nucleic Acids Res 9:3731–3746.

Wagner A, Deryckere F, McMorrow T, et al. (1994) Tail-to-tail orientation of the Atlantic salmon alpha- and beta-globin genes. J Mol Evol 38:28–35.

Wells RMG (1999) Evolution of haemoglobin function: Molecular adaptations to environment. Clinical and Experimental Pharmacology and Physiology, 26:591–595.

Wheeler D, Hope R, Cooper SB, et al. (2001) An orphaned mammalian beta-globin gene of ancient evolutionary origin. Proc Natl Acad Sci U S A 98:1101–1106.

Wheeler D, Hope RM, Cooper SJ, et al. (2004) Linkage of the beta-like omega-globin gene to alpha-like globin genes in an Australian marsupial supports the chromosome duplication model for separation of globin gene clusters. J Mol Evol 58:642–652.

Chapter 21
The Olfactory Receptor Gene Family of Marsupials

Margaret L. Delbridge, Amir Mohammadi, and Jennifer A. Marshall Graves

Abstract Olfaction in vertebrates is mediated mainly by a large family of olfactory receptors in the olfactory epithelium that belong to the superfamily of G protein-coupled receptors. Olfactory systems are well conserved among vertebrates, including marsupials, but there is a large variation in the numbers of olfactory genes in different animals. Most marsupials are nocturnal so depend on their sense of smell to locate food, avoid predators and identify potential mates in similar ways to other mammals. The olfactory bulbs are quite large in adult marsupials, suggesting that the sense of smell is very important in these animals. In addition, very undeveloped newborn marsupials have the special challenge of locating the pouch unassisted. It is likely that these newborns utilise their sense of smell for this unique pouch-finding task. The olfactory system is one of the few systems that is sufficiently developed in newborn marsupials to accomplish the task of finding the pouch. The opossum OR repertoire of one marsupial, the American opossum, is one of the largest characterised in mammals so far, containing over 1,500 genes. Interestingly comparisons of the opossum OR repertoire with the repertoire in an Australian marsupial, the tammar wallaby, suggests that a large conserved OR repertoire may be a feature of marsupials. The OR repertoires of the two marsupials show a high degree of similarity in total gene numbers and range of genes. This is unlike placental mammals, where the OR repertoires show a greater range. Results from these comparisons provide evidence for both the major forces (adaptation and genomic drift) behind the "birth-and-death" theory for the evolution of OR genes.

Keywords Olfactory receptor genes · Marsupials · Olfaction

M.L. Delbridge (✉)
ARC Centre of Excellence for Kangaroo Genomics, Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia
e-mail: margaret.delbridge@anu.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_21, © Springer Science+Business Media B.V. 2010

21.1 Introduction

21.1.1 The Importance of Olfaction

The sense of smell is so sensitive that it can detect and discriminate thousands of odours at very low concentration. Most animals rely on their sense of smell to find food and water, and avoid environmental dangers such as fire and predators. Animals use body odours to convey information about their identity, sex, health, and reproductive status to other individuals, and signal their home territories to other animals of the same and different species. Mammals also utilize their sense of smell to identify individuals and the mating readiness of other animals (Knapp et al., 2006). Olfaction therefore, is one of the most important senses animals utilize to learn about their environment.

21.1.2 Structure of the Olfactory Epithelium

Olfaction is mediated by the olfactory receptors that are located in the cilia of olfactory sensory neurons in the olfactory epithelium of mammals. Olfactory sensory neurons are bipolar neurons that span the olfactory epithelium and terminate externally in a number of fine cilia. Like other epithelial cells, olfactory sensory neurons have only a very short life and are constantly being regenerated from the basal cells of the olfactory epithelium approximately every 30 days (Moulton, 1974).

Odorant molecules become immersed in a layer of mucus, then bind to olfactory receptor proteins on the surface of the cilia. Binding of an odour molecule to the olfactory receptor activates the coupled G protein inside the olfactory sensory neuron. This event activates the conversion of ATP to the second messenger, cyclic AMP (cAMP), in the cytosol. cAMP opens up ligand-gated sodium (Na^+) channels, reduces the potential across the plasma membrane, and produces an action potential to be sent to the olfactory bulb in forebrain (reviewed by Menini et al., 2004).

21.1.3 The Organization of the Olfactory System

One of the features of the olfactory system in all animals from insects to mammals is the specificity of the olfactory receptor cells. It is a widely held view that most olfactory neurons express only one olfactory receptor gene, whereas all the other olfactory receptor genes in that neuron (over 1,000 in some mammalian species) are silent. Initial mapping of receptors in the olfactory epithelium suggested that olfactory neurons expressing a single type of receptor were distributed in one of four anterior-posterior bands or expression “zones” in the olfactory epithelium (Vassar et al., 1993). More recently it has been shown that single types of olfactory receptors are not confined to these four zones, but are found in multiple overlapping zones in the olfactory epithelium. Therefore a given olfactory receptor type is generally

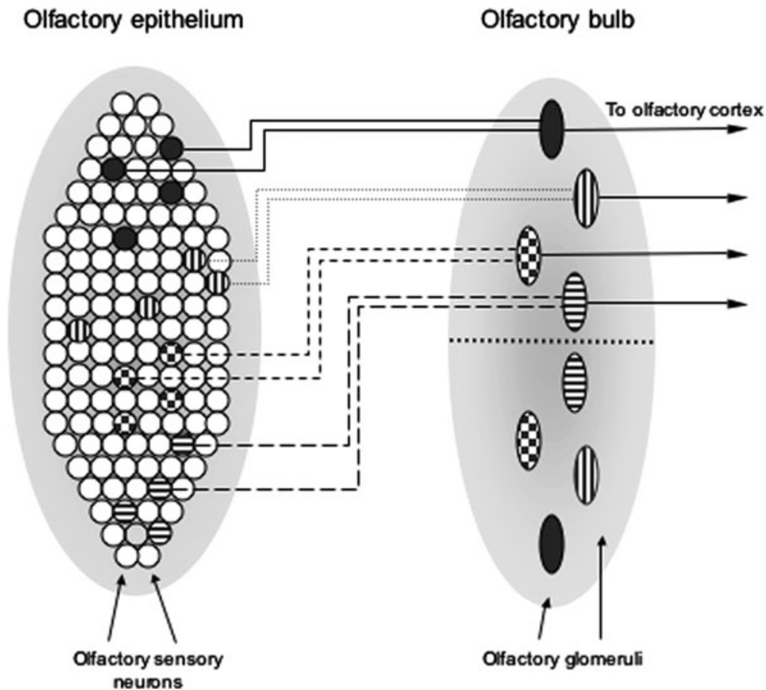


Fig. 21.1 Spatial topography of olfactory signalling. Within each zone of the olfactory epithelium, olfactory sensory neurons that express the same olfactory receptor are randomly distributed. The axons of the olfactory receptor neurons converge onto one of two glomeruli positioned approximately bilaterally in the olfactory bulb, before the signal is transported to numerous regions in the olfactory cortex of the brain by mitral cells

confined to a zone of expression, but many different olfactory receptor types are expressed in overlapping areas within a zone (Iwema et al., 2004; Miyamichi et al., 2005) (Fig. 21.1).

The olfactory bulb is organised into discrete regions called glomeruli, of which there are approximately 2,000 in mice (Menini et al., 2004). The position of each glomerulus in the olfactory bulb is roughly conserved between individuals, and each glomerulus represents the gathering together of the axons of all olfactory sensory neurons that express the same kind of olfactory receptor (Fig. 21.1) (Menini et al., 2004). The olfactory bulb is symmetrical around the midline, and so each olfactory sensory neuron converges to one or two glomeruli on each side of the olfactory bulb (Mombaerts, 2006). Therefore the olfactory receptor types are represented spatially by the organization of the glomeruli in the olfactory bulb. Mitral cells deliver olfactory signals from one glomerulus to multiple clusters of cortical neurons in several areas of the olfactory cortex, so that the spatial organization of the olfactory epithelium and olfactory bulb is not reflected in the organization of the olfactory cortex. The olfactory receptor can detect multiple odour molecules with different affinities,

and a single odour is detected by multiple olfactory receptors. The combination of olfactory receptors reacting to a single odour and the transmission of signals from multiple glomeruli to similar areas in the olfactory cortex results in the processing and recognition of complex odour signals (Malnic et al., 1999; Menini et al., 2004).

The olfactory receptor expressed by the neuron has been shown to have a role in the specification of the neurons for particular glomeruli. Rather than a particular type of olfactory receptor-containing neuron "targeting" a specific glomerulus, one model proposes that during development it is the coalescence of the axons of olfactory sensory neurons in a self-sorting mechanism that results in the formation of glomeruli. This self-sorting mechanism is perhaps mediated by interactions between olfactory receptor protein fragments located within the axon growth cones. This model is not inconsistent with a mechanism that would allow constant renewal of the olfactory neurons in the olfactory epithelium. The processes by which the complex connections between the olfactory neurons, glomeruli and regions of the olfactory cortex are laid down during development and throughout adult life are still not well understood (reviewed by Mombaerts, 2006).

In addition the evidence in support of the "one receptor-one neuron" hypothesis is not rigorously conclusive (Mombaerts, 2004), but there are only a limited number of exceptions that have been reported (Rawson et al., 2000; Sato et al., 2007; Tian and Ma, 2008). A variation on this hypothesis suggests that each neuron expresses just a few olfactory receptors, perhaps zero, one or two (Mombaerts, 2004). Neurons that express zero receptors would be negatively selected, neurons expressing one or more olfactory neurons would be positively selected. Neurons that express two or more olfactory receptors may persist only if there is no conflict between the expressed receptors; otherwise they will be selected against. This hypothesis proposes that immature olfactory sensory neurons would be likely to express more than one olfactory receptor with differing levels of expression, and that down-regulation of all other olfactory receptors would occur following coalescence of an axon into a glomerulus. This hypothesis also proposes that an olfactory sensory neuron can switch between the olfactory receptors that it expresses, so that it expresses different olfactory receptors sequentially during its lifespan (reviewed in Mombaerts, 2004). The choice of olfactory receptor expression in the olfactory sensory neurons is therefore proposed to be governed by selective forces similar to those proposed for the choice of alleles expressed by T and B lymphocytes (reviewed in Cedar and Bergman, 2008).

21.2 Olfactory Receptor Genes

The olfactory epithelium of vertebrates has a well conserved structure (Stoddart, 1980; Ache and Young, 2005). Olfactory receptors are a subfamily of the G protein-coupled receptor (GPCR) family. Olfactory receptors were first identified by their expression in the olfactory epithelium of the rat (*Rattus norvegicus*) (Buck and Axel,

1991). Like other members of the GPCR family, they are transmembrane proteins, with seven α-helical transmembrane domains composed of 20–28 hydrophobic amino acids each (Feingold et al., 1999). The third, fourth and fifth transmembrane domains, and the third extracellular loop are the most variable regions of the olfactory receptor protein, and are suggested to be the regions that play a role in binding to external odour molecules (Ngai et al., 1993; Pilpel and Lancet, 1999; Irie-Kushiyama et al., 2004). The first, second, sixth and seventh transmembrane domains, and the second intracellular loop of the olfactory receptors contain highly conserved motifs, that are proposed to be involved in protein-membrane interactions, and may initiate the intracellular signaling pathways to the brain (Malnic et al., 2004).

Olfactory receptor proteins are encoded by genes of the olfactory receptor (OR) gene family. This gene family is the largest in vertebrates, having between 700 and 1,500 members in different mammals, and accounting for 3–5% of all genes in the genome (Ben-Arie et al., 1993; Glusman et al., 1996). The vertebrate OR gene family has no homology to the olfactory genes of invertebrates, despite the similarity in the nervous pathways to the brain and similarities in the patterns of evolution of each gene superfamily (reviewed in Sanchez-Gracia et al., 2009). Vertebrate OR genes are short, intronless sequences, approximately 1 kb in length, that are organized in clusters and are present on almost all chromosomes. There is a large variation in the number of OR genes and the fraction of pseudogenes identified within a family in different species, and up to 50% of OR genes are non-functional in some species (Table 21.1).

21.2.1 Regulation of OR Gene Expression

The expression of olfactory receptors in the olfactory epithelium is not restricted to functional genes. There is widespread expression of human olfactory receptor pseudogenes in the olfactory epithelium (Zhang et al., 2007). It has been suggested that this indicates that the choice of olfactory receptor to be expressed in a given neuron is largely stochastic, whereby different OR genes are expressed and the expression is stabilised only once a functional gene is expressed (reviewed in Fuss and Ray, 2009).

It has also been demonstrated that OR genes are mono-allelically expressed. The expression of the paternally and maternally inherited allele is random. As for the mammalian X chromosome, which also undergoes random epigenetic silencing, active and inactive alleles replicate asynchronously (Singh et al., 2003). The mammalian OR gene loci are enriched in long interspersed nuclear elements (LINEs) to a level similar to that of the LINE-rich X chromosome, and it has been suggested that the LINE elements at the OR loci contribute to allelic exclusion, in the same way they are proposed to act as way-stations for spreading inactivation along the X chromosome (Kambere and Lane, 2009). However, unlike X inactivation or genomic imprinting (Goldmit and Bergman, 2004) silencing of OR alleles does not

Table 21.1 Numbers of OR genes in 20 vertebrate species

Common name	Species name	T[a]	F[b]	P/T[c]	References
Amphioxus	*Branchiostoma floridae*	43	31	12	Niimura (2009)
Zebrafish	*Danio rerio*	176	154	22	Niimura (2009)
Pufferfish	*Tetraodon nigroviridis*	34	11	23	Niimura (2009)
Elephant shark	*Callorhinchus milii*	2	1	1	Niimura (2009)
Medaka	*Oryzias latipes*	98	68	30	Niimura (2009)
Stickleback	*Gasterosteus aculeatus*	159	102	57	Niimura (2009)
Fugu	*Takifugu rubripes*	125	47	78	Niimura (2009)
Frog	*Xenopus tropicalis*	1,638	824	814	Niimura (2009)
Lizard	*Anolis carolinensis*	146	112	34	Niimura (2009)
Chicken	*Gallus gallus*	433	211	222	Niimura (2009)
Platypus	*Ornithorhynchus anatinus*	718	265	453	Niimura and Nei (2007)
Opossum	*Monodelphis domestica*	1,492	1,188	619	Niimura and Nei (2007)
Tammar wallaby	*Macropus eugenii*	1,753	601	1,152	Mohammadi et al. (in preparation)
Cow	*Bos taurus*	2,129	970	1,159	Niimura (2009)
Dog	*Canis lupus familiaris*	1,100	811	289	Niimura (2009)
Mouse	*Mus musculus*	1,391	1,037	354	Niimura and Nei (2005)
Rat	*Rattus norvegicus*	1,765	1,205	560	Niimura (2009)
Macaque	*Macaca mulatta*	606	309	297	Niimura and Nei (2007)
Chimpanzee	*Pan troglodytes*	813	380	433	Niimura (2009)
Human	*Homo sapiens*	802	387	415	Niimura and Nei (2007)

[a]Total number of genes
[b]Number of functional genes
[c]Number of pseudogenes and truncated genes

appear to be stable, for if the expressed allele is non-functional, then the other allele is subsequently expressed (Feinstein et al., 2004).

Local regulation of the *MOR28* OR gene cluster on mouse chromosome 14 by the *cis*-acting H element has also been investigated as a *trans*-acting mechanism for the widespread regulation of OR gene expression. The H element has been proposed to interact with only one *MOR28* OR gene at a time, through a physical

association of the H element with an OR gene promoter (Serizawa et al., 2003). This association then regulates the expression of the *MOR28* gene cluster. This mechanism does not appear to be transferable as the H element on chromosome 14 does not regulate any other OR genes, either on chromosome 14 outside the *MOR28* cluster or other chromosomes (Fuss et al., 2007; Nishizumi et al., 2007; reviewed by Fuss and Ray, 2009).

The most likely model for the determination of which OR gene will be expressed by any particular neuron is probably stochastic. As all neurons express a functional OR gene, there is the possibility that immature olfactory neurons switch between expression of OR genes before the neuron has established a stable connection to the brain. A negative feedback signal has been proposed to prevent continuation of the olfactory receptor choice process as soon as a functional olfactory receptor is expressed. This permits a second choice if, for example, a non-functional pseudogene was expressed. Interestingly it also appears that the second choice of receptor for expression is not random, but is restricted to the class of gene from which the first OR gene was expressed. The negative feedback of an expressed functional OR gene prevents the further expression of other OR genes and ensures that only a single gene will be expressed from each olfactory sensory neuron (Shykind et al., 2004, reviewed in Fuss and Ray, 2009).

It is not known if the *cis*-regulation of the *MOR28* cluster by a locus control region (LCR) such as the H element, will be generally applicable to the regulation of other OR gene clusters, or if this type of regulation is a consequence of the proximity of this cluster to the T-cell receptor α locus. Multiple LCRs may be present in a zebrafish OR gene cluster (Nishizumi et al., 2007), but the H element is the first LCR identified so far in mammals. It remains to be seen whether multiple LCRs and this type of *cis*-acting regulation also exist in other OR gene clusters. It is not yet known whether mechanisms of silencing such as methylation and/or histone modification have a role in the regulation of olfactory receptor choice, but it may be that modifications such as these are involved in the feedback mechanism to stabilise gene expression rather than be involved in OR gene switching (reviewed in Fuss and Ray, 2009).

21.2.2 Vertebrate Olfactory Receptor Gene Families

There are four major families of GPCR genes that encode olfactory receptor proteins. The neurons of the olfactory epithelium in the nose express two of the gene families; the OR and trace amine-associated receptor (TAAR) gene families. The vomeronasal organ is a specialised region of the nasal epithelium in mammals that is thought to be largely responsible for the detection of pheromones, and it expresses the other two gene families; the vomeronasal receptor type 1 and 2 (V1R and V2R) gene families. The four families of olfactory receptors are expressed in functionally different types of neurons, and use different signal transduction pathways (Nei et al., 2008). The TAAR receptors share sequence similarity with the OR receptors, and also are expressed in different zones within the olfactory epithelium, and so have

been classified as a second class of olfactory receptors (Liberles and Buck, 2006). TAARs may have a role in detecting social cues and may be a separate group of receptors in the nasal epithelium responsible for detecting pheromones (Stephen, 2009).

Olfactory gene repertoires are characterised by species-specific and lineage-specific expansion and reduction of specific gene families. The mammalian vomeronasal organ seems to be specialised for the detection of pheromones, and analysis of the morphology of the vomeronasal organ and the functional repertoire of the V1R and V2R gene families suggests that the function of this organ has been greatly enhanced in the rodents. V1R genes are intronless like OR genes, but the V2R genes contain introns and their proteins have a long extracellular N-terminus (Nei et al., 2008). The mouse expresses over 150 functional V1R genes, whilst the rat genome contains over 100 V1R genes. In contrast, humans have only five V1R genes, chimpanzees have none, and the dog and cow have only 8 and 32 V1R genes respectively. Like the rodent V1R repertoire, the V2R repertoire also seems to be enhanced particularly in the rodents, consisting of over 100 intact genes in both mouse and rat (Hasin-Brumshtein et al., 2009).

In fish, which have no vomeronasal organ, there is a single V1R-like gene that is expressed in the main olfactory epithelium (Pfister and Rodriguez, 2005), but there has been a lineage-specific expansion of V2R-like genes (Hashiguchi and Nishida, 2006). This suggests that the V1R and V2R gene repertoires have undergone specific expansions in the rodent lineage, along with the development of the vomeronasal organ. Interestingly, the South American grey, short-tailed opossum (*Monodelphis domestica*) also has well developed V1R and V2R repertoires, consistent with having a well-developed vomeronasal organ in this marsupial mammal (Grus et al., 2005; Young and Trask, 2007). Another mammal in which there has been a lineage specific expansion of the V1R gene repertoire is the platypus. This semi-aquatic monotreme has undergone the largest lineage-specific V1R gene expansion of any mammal so far investigated, and it also has a large V2R repertoire. In contrast, the main olfactory system of the platypus seems to be somewhat reduced. This suggests that the vomeronasal system is a more important sensory organ than the main olfactory system of the nose in the platypus, which forages underwater with the eyes, ears and nostrils closed (Grus et al., 2007).

Olfactory receptors in the main nasal epithelium are encoded by the OR gene family that is found in all vertebrates. In eutherian mammals, the OR gene family can contain between 700 and 1,500 genes in different species. In contrast, the OR gene family is much smaller in fish, ranging from about 40–140 genes in different fish species. The fish OR gene family is more divergent than the mammalian OR gene repertoire, suggesting that fish might be able to detect a wider range of odorants, but that the mammalian OR repertoire might be capable of discriminating between more structurally similar odorants (Niimura and Nei, 2005). It is unlikely that the number of OR genes is directly related to the sensitivity of the sense of smell. Comparisons of olfactory abilities must also take into account the size of the olfactory epithelium, the fraction of the brain devoted to processing olfactory information, and the ability to memorize the signals from new odours.

The OR gene family is classified into two major classes: Class I and Class II. In mammals, Class I olfactory receptors represent only 10% of the total olfactory receptor repertoire, whereas Class II contains approximately 90% of the total repertoire. Class I receptors have been proposed to be responsible for detecting water-soluble odorants, whereas Class II receptors are thought to bind air-borne odorants. In support of this hypothesis is evidence that in goldfish, the only functional OR genes have homology to Class I genes, and all Class II genes are non-functional pseudogenes (Freitag et al., 1998). In amphibians olfaction occurs in two anatomically separate nasal cavities, and both Class I and Class II olfactory receptors are expressed. Class I OR genes are expressed in a water-filled nasal cavity and Class II OR genes are expressed in the air-filled nasal cavity (Freitag et al., 1995). Further evidence for the distinction between the functions of the two classes is provided by studies of the dolphin, in which a large array of Class II OR genes were found to be nonfunctional pseudogenes. It was suggested that the Class II OR genes were lost in these aquatic mammals when they evolved from land mammals and lost the ability to detect air-borne odorants (Freitag et al., 1998). It is unlikely however that there is a functional distinction between Class I and Class II OR genes. High numbers of functional Class I OR genes were found in the human and mouse genomes (Glusman et al., 2001; Zhang and Firestein, 2002), suggesting that Class I OR genes are likely to respond to additional ligands in these species, as well as water-soluble odour molecules in aquatic animals.

The blurring of the distinction between Class I and Class II OR genes has prompted a reclassification of the OR gene family in vertebrates, although the well established Class I and Class II nomenclature is still in common usage. The availability of genome sequence data and its use for phylogenetic analysis of vertebrate OR gene families led to classification into nine groups (α, β, γ, δ, ϵ, ζ, η, θ), each of which had a single or small number of founding genes prior to the divergence of fish and tetrapods. The α and γ groups are the only groups that have been retained by, and greatly expanded in, the tetrapods, and these two groups represent the Class I and Class II genes respectively in mammals (Niimura and Nei, 2005; Niimura and Nei, 2006; Nei et al., 2008; Niimura, 2009).

21.2.3 Ectopic Expression of OR Genes

The discovery of olfactory receptor genes was based partly on the observation that they were specifically expressed in the olfactory epithelium (Buck and Axel, 1991). They are also expressed in the olfactory bulb, where they are likely to play a role in the coalescence of olfactory sensory neurons into glomeruli (Mombaerts, 2006). Several studies have shown that OR genes can be expressed in tissues other than the olfactory epithelium, which may suggest other roles for these receptors. Expression of various OR genes has been reported in the testis and germ cells, as well as a wide range of other tissues such as the tongue, erythroid cells, prostate, placenta, brain, and peripheral nervous system (see Feldmesser et al., 2006). Expression of OR genes in the testis has been proposed to play a role in sperm chemotaxis or

olfaction-driven mate choice, whereas expression of OR genes in a wide range of tissues has been suggested to indicate a role for the OR genes in cell-cell recognition (Ziegler et al., 2002; Dreyer, 1998; Fukuda et al., 2004).

Database mining and microarray experiments (Feldmesser et al., 2006; Zhang et al., 2007) have now shown that widespread expression of OR genes in non-olfactory tissues may be a consequence of transcription from leaky promoters during the decondensation of chromatin that is required for transcriptional regulation. Leaky expression of genes may be evolutionarily important by generating potential diversity (Rodriguez-Trelles et al., 2005; Feldmesser et al., 2006). However, microarray analysis of OR gene expression shows that ectopic patterns of expression of orthologous OR genes are conserved between chimpanzee and human, suggesting that they evolved more slowly because they serve a selectable function (Branscomb et al., 2000; De la Cruz et al., 2009). This is circumstantial evidence that at least some OR genes may have gained new functions in non-olfactory tissues.

21.2.4 Evolution of OR Genes

The OR gene family in vertebrates is characterized by lineage-specific expansions and contractions in the numbers of OR genes, described as a "birth-and-death" pattern of evolution (reviewed in Nei and Rooney, 2005). This model predicts that the composition of a gene family may be unique in different lineages. For example, in one species, particular genes may be duplicated one or more times, resulting in expansion of the family, whereas in another species the same genes may be lost or become pseudogenes, resulting in contraction of the family. For example, the most recent common ancestor of human and mouse is estimated to have had 754 functional OR genes, but there has been a lineage-specific contraction in the OR repertoire to 388 functional genes in humans and a lineage-specific expansion of the repertoire to 1,037 functional genes in the mouse. Of the 754 ancestral OR genes, 193 and 267 genes have become new non-functional pseudogenes in human and mouse, respectively, whilst humans have lost 249 of the ancestral pseudogenes that have been proposed to be part of the OR repertoire of their most recent common ancestor with mice (Nei and Rooney, 2005).

Class I and Class II OR genes are further subdivided based on protein sequence similarity. OR gene families encode proteins with more than 40% protein similarity, and OR gene subfamilies encode proteins with more than 60% protein similarity. Currently there are 18 gene families and 300 subfamilies that have been characterised in humans (Olender et al., 2008), but the publication of whole genome sequences has made it possible to characterise OR gene families bioinformatically in a number of vertebrate species. This has shown that there are large numbers of OR genes that vary widely between different species. Another hallmark of the OR gene families in vertebrates is that in addition to functional gene members, they all contain significant numbers of disrupted genes, the numbers of which also

vary widely between species. The disrupted genes have interruptions to their coding region, suggesting that they are non-functional pseudogenes.

Another feature of the OR gene family is that the genes are located in a number of clusters on nearly all chromosomes in the genome. Fish have one of the smallest OR repertoires, and it has been proposed that one of the major events in the evolution of the OR gene families was the adaptation of the olfactory system to a non-aquatic environment (Glusman et al., 2001). Early duplications in the land mammals gave rise to the major families of Class II OR genes found in all mammals. Subsequently there have been local expansions of different OR families and subfamilies to form large gene clusters (Glusman et al., 2001). This pattern of duplication means that OR genes are found at many chromosomal locations, and related OR genes are often present in the same clusters (Sullivan et al., 1996; Rouquier et al., 1998; Zhang and Firestein, 2002), many of which are conserved across species (Aloni et al., 2006).

Segmental duplications between chromosomes have also contributed to the wide distribution of OR gene clusters, resulting in the location of members of different subfamilies on different chromosomes. Other mechanisms that are likely to have contributed to the evolution of this gene family are tandem duplication, which has generated long tandem arrays of closely related genes, gene conversion and recombination which may allow conversion of non-functional to functional genes or *vice versa*, and retrotransposition (reviewed in Kambere and Lane, 2007). Some OR gene clusters are broken up by many interspersed non-OR genes (Olender et al., 2008), whereas other clusters contain OR genes exclusively (Glusman et al., 2001; Lane et al., 2002). The presence of related OR genes in single clusters suggests that different clusters may largely represent different subfamilies of OR genes and therefore might be responsible for detecting different types of odorant molecules. It may be that gene clusters are preserved so that all the genes within a cluster can be transcriptionally regulated by common control elements (Krumlauf, 1992).

Variation in the total numbers of functional and non-functional OR genes in the repertoires of different animals has been suggested to reflect an adaptive response to the environment. It has been suggested that there have been three major events in the evolution of the vertebrate olfactory system; first, the expansion of the olfactory system as a response to the adaptation to a terrestrial environment, secondly, the reduction in the olfactory system in the primate lineage along with the development of trichromatic vision, and thirdly the development of the pheromone system in the rodent lineage (Kambere and Lane, 2007; Olender et al., 2008). This theory of adaptive evolution is supported by the relatively few genes in the OR repertoire of all fish species (approximately 100 OR genes) compared to a 8–10 times larger number of genes in the OR repertoires of terrestrial mammals (between approximately 800 and 1,500), suggesting an expansion of the OR repertoire has occurred during the adaptation of vertebrates to a terrestrial existence (Niimura and Nei, 2006; Saraiva and Korsching, 2007; Nei et al., 2008). In addition, there are many fewer OR genes in aquatic mammals (dolphins, whales), consistent with loss of OR genes following the readaptation of these mammalian species to the water (Freitag et al., 1998). This does not explain, though, why a smaller number of fish OR genes can detect the

same range of odorants as the larger mammalian OR repertoire (Niimura and Nei, 2005).

In primates, lower numbers of functional OR genes and higher numbers of pseudogenes seems to have resulted from increased pseudogenisation of OR genes. This may have occurred when the olfactory system became relatively less important with the increasing reliance of primates on the development of trichromatic colour vision, or the development of higher brain mechanisms such as memory. However, this does not explain why cattle with dichromatic colour vision have a high proportion of pseudogenisation in their OR repertoires, or why the numbers of functional OR genes varies so much among terrestrial mammals (Nei et al., 2008).

Diversity of the OR repertoire in different animals was clearly generated by gene duplication and deletion events, then adaptation of species to different environmental niches with different requirements for olfaction. Such gene duplication and deletion events occur more or less randomly, so genomic drift may play an important role in the generation of diverse OR gene repertoires, and may wholly account for instances in which species-specific expansions and contractions are difficult to explain by adaptation (Nei et al., 2008). Random gene duplication can increase the number of genes which are then fixed by positive selection or chance (Nei et al., 2008).

If genetic drift has a significant effect on the OR gene repertoire of mammals, it would be expected that copy number variation should exist among individuals of the same species. Indeed, up to 30% of functional genes in the human and mouse OR repertoires are polymorphic for copy number (Young et al., 2008; Nozawa et al., 2007). There is also a high degree of allelic variation in human and mouse OR genes, which may account for the variation observed in sensitivity to odours observed between different individuals. Selection for a diverse repertoire can explain why the exceptionally large numbers of functional OR genes and pseudogenes are tolerated in vertebrate genomes. Gene conversion between different functional and non-functional OR copies could rapidly generate new functional genes (reviewed in Keller and Vosshall, 2008; Hasin-Brumshtein et al., 2009).

The size of OR gene families, as well as the proportions of intact OR genes and pseudogenes, varies widely among vertebrate species (Table 21.1). Different subsets of pseudogenes are found even between different human populations, or different dog breeds (Tacher et al., 2005; Rouquier and Giorgi, 2007), implying very rapid evolution.

21.3 The Olfactory System in Marsupials

Most marsupials are nocturnal, and forage alone or in small groups at night, although some marsupials, such as the kangaroos, may continue to forage at either end of the day. During the night therefore, the auditory and olfactory senses are more important than vision for these animals. The relative importance of the sense of smell for the marsupials is confirmed by the observation that all marsupials have

a relatively large olfactory bulb. Like most other vertebrates, marsupials use their sense of smell to find food, avoid predators, and to identify animals and marked territories of their own and other species. This is particularly important for species that live a solitary existence and meet only to mate (reviewed by Croft and Eisenberg, 2006).

A unique feature of marsupial reproduction is the birth of the young at a very early developmental stage. The tiny newborn marsupial young manages to move unaided from the urogenital opening to the pouch some distance away, where it attaches to a teat (reviewed in Pask and Renfree, Chapter 14). At birth, the eyes and ears of the pouch young are very undeveloped, but the olfactory system is relatively well developed, suggesting that both gravity and the sense of smell guide the newborn marsupials to the pouch (Gemmell and Rose, 1989; Hughes et al., 1989).

21.4 Olfactory Apparatus of Marsupials

Exactly which signals guide the newborn marsupial to the pouch is still unclear, and the relative importance of different stimuli may be different for different marsupial species. At birth, the olfactory epithelium of a number of marsupials, including the brush-tailed possum (*Trichosurus vulpecula*), rat-kangaroo (*Potorous tridactylus*) and quoll (*Dasyurus hallucatus*), is well developed. Olfactory receptor neurons with cilia are observed in the olfactory epithelium of the newborn possum, and may be capable of responding to odours. However, the apparent maturity of the olfactory epithelium contrasts with the morphological immaturity of the olfactory bulb in these species (Gemmell and Rose, 1989; Gemmell and Nelson, 2004). Indeed in both the opossum (*M. domestica*) and the tammar wallaby (*Macropus eugenii*), the olfactory bulb appears to be relatively undeveloped at birth, and there is no evidence for the presence of glomeruli (Ashwell et al., 2008; Schneider et al., 2009) until approximately 15 days after birth in the opossum and 25 days after birth in the tammar wallaby. It is not until approximately 54 days after birth that the olfactory bulb seems to be fully developed in the tammar wallaby (Ashwell et al., 2008). In the tammar wallaby, the vomeronasal organ, which is responsible for the detection of pheromones, is very immature in appearance at birth and does not have neurons extending to the olfactory bulb until day 5 after birth (Ashwell et al., 2008).

The olfactory system of most marsupials undergoes further development and differentiation during the first 25–40 days of pouch life. Although the olfactory epithelium of marsupials appears to be functional at birth, it has been suggested that the olfactory bulb and signaling pathway to the brain is too immature at birth to allow the newborn to detect odour and use this sense to locate the pouch (Gemmell and Rose, 1989; Ashwell et al., 2008). However, behavioural studies have shown that newborn tammar wallabies are attracted toward the smell of the mother's pouch, indicating that these newborns have a functional olfactory system (Schneider et al., 2009). This suggests that although the olfactory system is not fully developed at birth, it may be sufficiently developed to contribute to finding the pouch.

21.5 Marsupial OR Repertoire

The koala (*Phascolarctos cinereus)* was the first marsupial from which OR genes were isolated (Glusman et al., 2000). Twenty-five OR genes were identified, all of which belonged to the class II OR genes, and five of which were pseudogenes. From these limited numbers, it was estimated that marsupials must have a very small OR repertoire of up to 50 genes (Glusman et al., 2000).

The availability of genome sequence of the opossum (*M. domestica*) provided the first opportunity to isolate almost the whole OR repertoire from a marsupial. Surveying the first genome assembly of the opossum resulted in the detection of 1,518 OR genes, more than 60% of which are functional (Aloni et al., 2006). Somewhat different estimates have since been reported for the opossum OR gene repertoire, ranging from 1,492 genes of which 80% are functional (Niimura and Nei, 2007), to 1,548 genes of which 77% are functional (Kishida, 2008). However these differences are small and mostly attributable to different methodologies and/or different versions of the assembled genome being used.

The availability of genome data from the opossum and the platypus enables a more complete account of the evolution of OR gene repertoire. Examination of the OR repertoires of several mammalian species reveals several major duplications during vertebrate evolution that led to differences in the total number of OR genes in different lineages. In the monotremes, as well as amphibians, birds and fish, there are fewer than 1,000 OR genes, suggesting that the OR repertoire was relatively small prior to the mammalian radiation. The large OR repertoire in the opossum implies that there was a sudden burst of OR gene duplication and diversification between the divergence of the monotremes from Theria (marsupials and eutherians), and the divergence of marsupials from eutherians. During this short period (between 166 and 148 million years ago, Bininda-Emonds et al., 2007), there was a gain of approximately 350 genes to increase the total number of OR genes in the most recent common ancestor of the therian mammals from 152 to 492. A further 350 genes were gained in the eutherian lineage between 145 and 105 million years ago, between the divergence of the marsupials from the Eutheria and prior to the radiation of eutherian mammals. There was a remarkable gain of approximately 750 genes in the marsupial lineage, and further significant increases (up to 1,500) in the number of genes then occurred in the OR repertoires of the rodent lineages, whilst significant decreases (down to about 300) in the number of genes occurred in the primate lineage (Niimura and Nei, 2007) (Fig. 21.2).

Most of the expansion in the marsupial lineage can be accounted for by expansion in a small number (four or five) of gene lineages. The majority (80–90%) of the mammalian OR repertoire is made up of Class II OR genes, leaving only 10–20% represented by the Class I OR genes. Niimura and Nei (2007) divided the Class II genes into 34 subgroups by constructing phylogenetic trees, so that each clade contained genes that had the greatest sequence similarity both within and between species. They found that the representation of the Class II subgroups can be quite different between species (Niimura and Nei, 2007). Some Class II subgroups appeared to be quite stable throughout mammalian evolution,

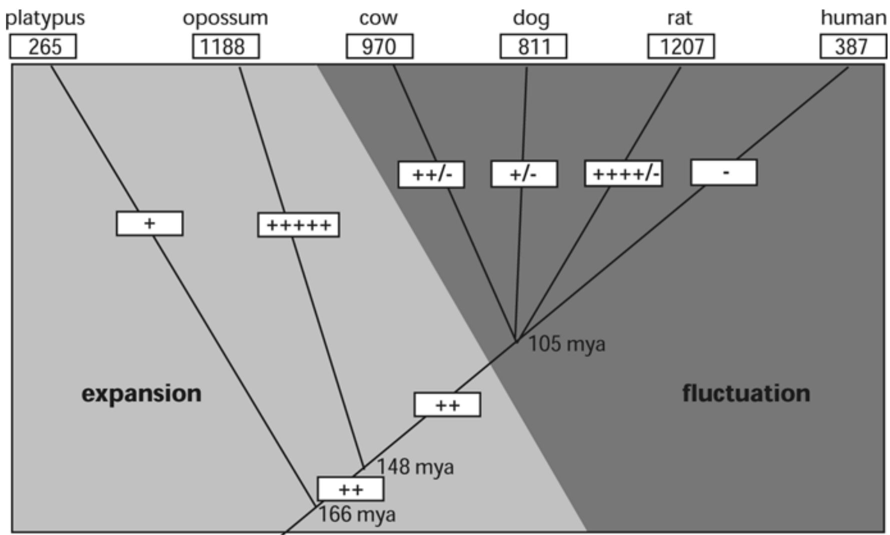


Fig. 21.2 Evolutionary changes in functional gene numbers in mammalian OR gene repertoires. The numbers of OR genes in each species is given in each box. "+" represents a gain of approximately 150 genes and "–" represents a loss of approximately 150 genes. Divergence times are estimated in millions of years ago (mya) (Bininda-Emonds et al., 2007). There has been significant expansion in OR gene numbers in the most recent common ancestor of the Theria and the Eutheria, as well as lineage specific expansion in the marsupials and rodents. Generally, there have been fluctuations in the numbers of OR genes in separate lineages following the eutherian radiation, including significant amounts of both gain and loss of genes from the OR gene repertoires. Adapted from Niimura and Nei (2007)

containing approximately the same numbers of genes in different species. However, other Class II subgroups contained vastly different numbers of genes in different species. The expansion in the opossum lineage can be accounted for mostly by expansion in four OR Class II groups, presumably because these OR genes encode receptors that detect odours that are important for opossums. Interestingly, some of the same subgroups are expanded also in the rodent lineages. This analysis provides valuable insight into the gains and losses of OR genes that have occurred during mammalian evolution, but the function of each of the subgroups remains unclear. Further investigation is needed to ascertain whether there is a relationship between these subgroups and their function and/or genomic location. Additionally there is limited information about which ligands elicit a response from which OR gene families, so it cannot be determined, for example, whether the same OR gene family in marsupials and rodents is responding to the same ligand (Niimura and Nei, 2007).

Information from marsupials also makes it possible to interpret the evolution of the TAAR family of OR receptors in the main olfactory epithelium, which detect amines in the urine that function as sex pheromones in the mouse (Liberles and Buck, 2006). In all tetrapods the TAAR family is located in a conserved cluster on

human chromosome 6, mouse chromosome 10 and opossum chromosome 2. Only one or two TAAR genes are found in the frog and the chicken, suggesting that these genes encode less important odour receptors in amphibians and birds. A similar pattern of lineage specific gain and loss of genes is observed in this gene family. Subfamily I of the TAAR family is expanded in the opossum as well as the mouse (Hashiguchi and Nishida, 2007), which is consistent with the detection of amines in the urine being important for the identification of oestrus in female opossums (Zuri et al., 2003).

The repertoire of opossum OR genes also shows that this pattern of lineage-specific gain and loss of genes has occurred to a much greater degree in the other pheromone receptor genes, the V1R and V2R genes of the vomeronasal epithelium. The numbers of V1R and V2R genes in opossum is similar to those of the rodents, whereas there has been a loss of these genes in primates, dog and cow (Grus et al., 2007; Young and Trask, 2007). This is consistent with the well-developed vomeronasal system in the opossum, which has been suggested to relate to the reliance of marsupials, as well as rodents, on the detection of pheromones, and the loss of this ability in other mammals. The platypus has the largest V1R and V2R repertoires of any mammal, which may have been selected for because of its reliance on these receptors underwater, and the orientation of its vomeronasal organ to open into the oral cavity rather than the nasal cavity as is usual for most other mammals (Grus et al., 2007).

Another remarkable feature of the V1R and V2R repertoires is their rapid and extreme expansions within mammal lineages. This is evident from phylogenies of V1R and V2R genes that form lineage-specific clades, in which genes within a species are most closely related and there are very few genes with one-to-one orthology between different species. This means that the evolution of these genes has been much more rapid than that of the OR genes of the main olfactory epithelium (Grus et al., 2007; Grus and Zhang, 2008), in which orthologous clusters of OR genes can still be recognised in different species (Aloni et al., 2006).

Data from the opossum and tammar genomes also shows that as well as an increase in the total number of genes in the marsupial lineage, there has been remarkable conservation of the marsupial OR repertoire. In almost all terrestrial marsupials and eutherians whose genomes have been sequenced, the Class I genes consistently represent between 12 and 18% of the total OR repertoire, and most of them are functional. This implies that the Class I genes have been functional since before the divergence of the marsupials from the eutherians 148 million years ago. The size and the composition of the OR repertoire is well conserved between the tammar wallaby and opossum, indicating that the marsupial OR repertoire has been conserved for more than 80 million years (Fig. 21.3a). This extent of conservation is unusual; it is not found between any two eutherian mammals and might be the result of some functional constraints in marsupials resulting from similar biological needs.

The OR51 family is an example of a Class I family that appears to have undergone lineage-specific expansion in the therians. Within family OR51 there is almost a one-to-one relationship between the tammar and opossum OR genes of this family (Fig. 21.3b). This suggests that the expansion of this family occurred in the

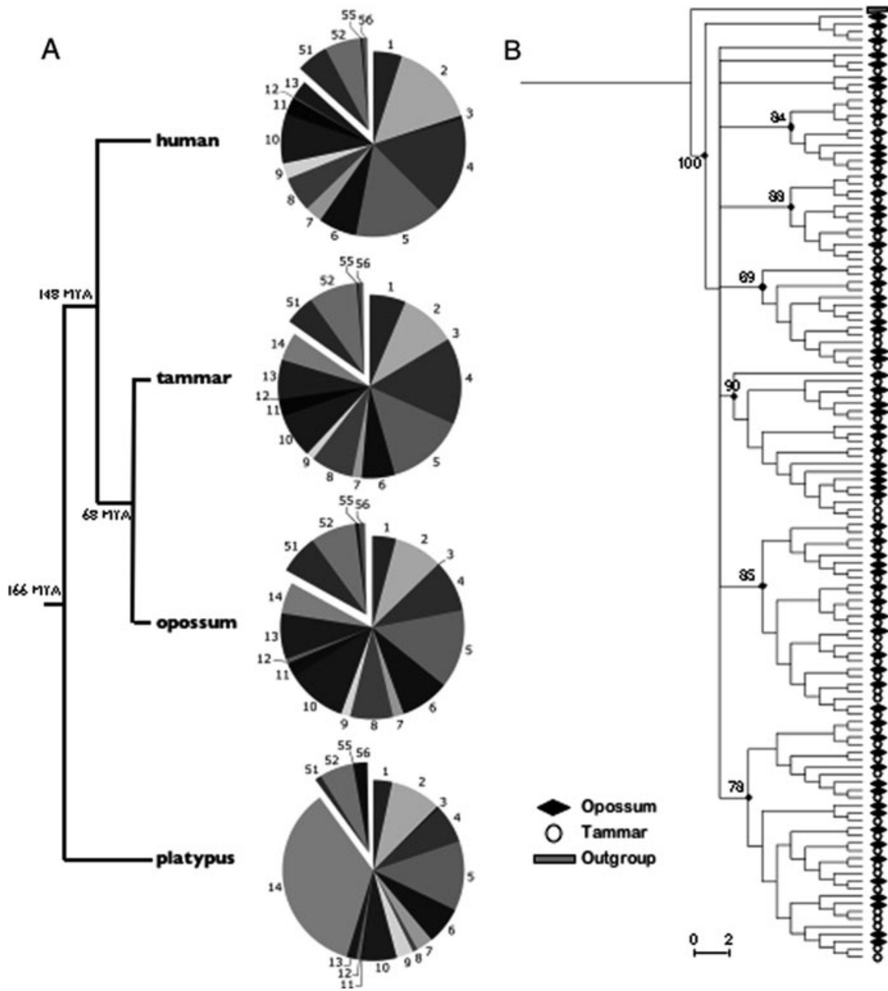


Fig. 21.3 **(a)** A comparison of the composition of the total OR gene repertoires in mammals. Numbers next to the sections represent the family number. Class I families are represented in the excised portion of the pie charts. Comparisons between species suggest there has been a loss of family OR14 genes in eutherians, expansion of class I genes (especially family OR51) prior to the divergence of the therians, and a high level of similarity between two marsupial species. **(b)** The phylogenetic relationship between 125 genes of the OR51 family from the opossum and tammar wallaby. A class II OR gene is used as outgroup. Bootstrap values are shown for important nodes. In most cases there is a one-to-one relationship between the opossum and tammar wallaby genes. Absence of species-specific expansions and/or contractions in this family indicates an important function for this family in marsupials. Similar sized human and marsupial OR51 families suggests that there was an expansion of this family between the divergence of the monotremes 166 mya, and the divergence of the marsupials from the eutherians 148 mya (Mohammadi et al., in preparation)

marsupial lineage prior to the divergence of these two marsupials approximately 80 million years ago (Mohammadi et al., in preparation). Similar numbers of OR51 genes are found in human and marsupials, whereas low numbers of family OR51 are found in platypus, suggesting that the expansion of the family occurred in the therian lineage following its divergence from the monotremes, between 148 and 166 million years ago. Comparisons with an outgroup, such as the counterpart of the OR51 family in another vertebrate, would be necessary to confirm the expansion of this lineage in therians. The analysis of therian Class I OR genes such as this may be an opportunity to examine the role of olfactory receptors in detecting airborne molecules and/or gaining new functions in land-living mammals.

Data from the opossum and tammar genomes also shows that as well as an increase in the total number of genes, the diversity of the OR genes has increased since the divergence of Theria from the monotremes. There are only 130 subfamilies in the platypus genome, compared to 240 OR gene subfamilies in the opossum and 286 in the tammar wallaby. These duplications have affected mostly the gene families OR8, OR11, OR13 and OR51, which have expanded up to 20 times compared to their counterparts in platypus. Further analysis including a vertebrate outgroup will confirm whether these expansions are specific to therians, or whether the contrast between platypus and therian OR family sizes are the result of lineage specific loss of genes in the platypus (Mohammadi et al., in preparation).

In contrast, the Class II OR14 gene family may have undergone a significant loss of genes during therian evolution; the platypus genome contains more than 250 members of the OR14 gene family, but only one third of them have been retained in marsupials, and this family has been completely lost from the eutherian genome. As all the OR14 genes of platypus belong to only six subfamilies, this implies coincidental tandem duplications of a few ancestral genes. Further analysis using a vertebrate outgroup will confirm whether this is the case, or if there has been an expansion of these Class II genes in platypus (Mohammadi et al., in preparation) (Fig. 21.3a).

21.6 Conclusion

The sense of smell is an important means by which vertebrates receive information from the environment and communicate with each other. The olfactory system responds to airborne odour molecules and pheromones. One important area of research in this field examines how the connections between the olfactory epithelium, olfactory bulb and olfactory cortex of the brain are established and regulated. Olfactory neurons contain positional information within in the epithelial layer, they express one or a few olfactory receptors, and their axons can coalesce with other neurons of the same type to form glomeruli in the olfactory bulb. Genomic imprinting mechanisms may play a role in the regulation of olfactory gene expression influencing the construction of the olfactory system as well as detection of odour molecules, so the investigation of epigenetic regulation in marsupials is likely to

provide valuable insight into the regulation of olfactory genes as it has other gene systems in mammals (reviewed in Hore et al., Chapter 12). Marsupials are ideal organisms in which to study the development of the olfactory connections through the olfactory bulb to the olfactory cortex, as they are born at an undeveloped stage and a significant proportion of the complexity of olfactory bulb development takes place in the pouch.

The marsupial genome harbours one of the largest mammalian OR repertoires, containing up to 1,500 olfactory receptor genes, that respond in combinatorial codes to odour molecules to allow marsupials to process and recognise an almost unlimited range of odours. Comparisons between the repertoires of different animals has suggested that different families and subfamilies of OR genes have been expanded in response to the adaptation of animals to different environments. In marsupials there has been expansion of both the OR repertoire of the main olfactory epithelium and the V1R and V2R repertoires of the vomeronasal organ, similar to that observed in rodents. This suggests that the sense of smell and detection of pheromone signals are at least as important to marsupials as they are to rodents, perhaps due to marsupials nocturnal habits and unique reproductive biology. These unique evolutionary adaptations of the marsupials make the characterisation of the marsupial OR repertoire essential to any comprehensive study of the olfactory system.

References

Ache BW, Young JM (2005) Olfaction: diverse species, conserved principles. Neuron 48: 417–430.

Aloni R, Olender T, Lancet D (2006) Ancient genomic architecture for mammalian olfactory receptor clusters. Genome Biol 7:R88.

Ashwell KWS, Marotte LR, Cheng G (2008) Development of the olfactory system in a wallaby (*Macropus eugenii*). Brain Behav Evol 71:216–230.

Ben-Arie N, Khen M, Lancet D (1993) Glutathione S-transferases in rat olfactory epithelium: purification, molecular properties and odorant biotransformation. Biochem J 292: 379–384.

Bininda-Emonds OR, Cardillo M, Jones KE, et al. (2007) The delayed rise of present-day mammals. Nature 446:507–512.

Branscomb A, Seger J, White RL (2000) Evolution of odorant receptors expressed in mammalian testes. Genetics 156:785–797.

Buck L, Axel R (1991) A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell 65:175–187.

Cedar H, Bergman Y (2008) Choreography of Ig allelic exclusion. Curr Opin Immunol 20: 308–317.

Croft DB, Eisenberg JF (2006) Behaviour. In: Armati PJ, Dickman CR, Hume ID (eds) Marsupials. Cambridge University Press, Cambridge.

De la Cruz O, Blekhman R, Zhang XH, et al. (2009) A signature of evolutionary constraint on a subset of ectopically expressed olfactory receptor genes. Mol Biol Evol 26:491–494.

Dreyer WJ (1998) The area code hypothesis revisited: olfactory receptors and other related transmembrane receptors may function as the last digits in a cell surface code for assembling embryos. Proc Natl Acad Sci USA 95:9072–9077.

Feingold EA, Penny LA, Nienhuis AW, Forget BG (1999) An olfactory receptor gene is located in the extended human beta-globin gene cluster and is expressed in erythroid cells. Genomics 61:15–23.

Feinstein P, Bozza T, Rodriguez I, Vassalli A, Mombaerts P (2004) Axon guidance of mouse olfactory sensory neurons by odorant receptors and the beta2 adrenergic receptor. Cell 117:833–846.

Feldmesser E, Olender T, Khen M, et al. (2006) Widespread ectopic expression of olfactory receptor genes. BMC Genomics 7:121.

Freitag J, Krieger J, Strotmann J, Breer H (1995) Two classes of olfactory receptors in *Xenopus laevis*. Neuron 15:1383–1392.

Freitag J, Ludwig G, Andreini I, Rossler P, Breer H (1998) Olfactory receptors in aquatic and terrestrial vertebrates. J Comp Physiol A 183:635–650.

Fukuda N, Yomogida K, Okabe M, Touhara K (2004) Functional characterization of a mouse testicular olfactory receptor and its role in chemosensing and in regulation of sperm motility. J Cell Sci 117:5835–5845.

Fuss SH, Omura M, Mombaerts P (2007) Local and *cis* effects of the H element on expression of odorant receptor genes in mouse. Cell 130:373–384.

Fuss SH, Ray A (2009) Mechanisms of odorant receptor gene choice in *Drosophila* and vertebrates. Mol Cell Neurosci 41:101–112.

Gemmell RT, Nelson J (2004) The birth and subsequent development of the common brushtail possum *Trichosurus vulpecula*. In: Goldingay RL, Jackson SM (eds) The Biology of Australian Possums and Gliders. Surrey Beatty & Sons, Chipping Norton.

Gemmell RT, Rose RW (1989) The senses involved in movement of some newborn Macropodidae and other marsupials from cloaca to pouch. In Grigg G, Jarman P, Hume I (eds) Kangaroos, Wallabies and Rat-Kangaroos. Surrey Beatty & Sons, Chipping Norton.

Glusman G, Bahar A, Sharon D, et al. (2000) The olfactory receptor gene superfamily: data mining, classification, and nomenclature. Mamm Genome 11:1016–1023.

Glusman G, Clifton S, Roe B, Lancet D (1996) Sequence analysis in the olfactory receptor gene cluster on human chromosome 17: recombinatorial events affecting receptor diversity. Genomics 37:147–160.

Glusman G, Yanai I, Rubin I, Lancet D (2001) The complete human olfactory subgenome. Genome Res 11:685–702.

Goldmit M, Bergman Y (2004) Monoallelic gene expression: a repertoire of recurrent themes. Immunol Rev 200:197–214.

Grus WE, Shi P, Zhang J (2007) Largest vertebrate vomeronasal type 1 receptor gene repertoire in the semiaquatic platypus. Mol Biol Evol 24:2153–2157.

Grus WE, Shi P, Zhang YP, Zhang J (2005) Dramatic variation of the vomeronasal pheromone receptor gene repertoire among five orders of placental and marsupial mammals. Proc Natl Acad Sci USA 102:5767–5772.

Grus WE, Zhang J (2008) Distinct evolutionary patterns between chemoreceptors of 2 vertebrate olfactory systems and the differential tuning hypothesis. Mol Biol Evol 25:1593–1601.

Hashiguchi Y, Nishida M (2006) Evolution and origin of vomeronasal-type odorant receptor gene repertoire in fishes. BMC Evol Biol 6:76.

Hashiguchi Y, Nishida M (2007) Evolution of trace amine associated receptor (TAAR) gene family in vertebrates: lineage-specific expansions and degradations of a second class of vertebrate chemosensory receptors expressed in the olfactory epithelium. Mol Biol Evol 24:2099–2107.

Hasin-Brumshtein Y, Lancet D, Olender T (2009) Human olfaction: from genomic variation to phenotypic diversity. Trends Genet 25:178–184.

Hughes RL, Hall LS, Tyndale-Biscoe CH, Hinds LA (1989) Evolutionary implication of macropodid organogenesis. In Grigg G, Jarman P, Hume I (eds) Kangaroos, Wallabies and Rat-Kangaroos. Surrey Beatty & Sons, Chipping Norton.

Irie-Kushiyama S, Asano-Miyoshi M, Suda T, Abe K, Emori Y (2004) Identification of 24 genes and two pseudogenes coding for olfactory receptors in Japanese loach, classified into four subfamilies: a putative evolutionary process for fish olfactory receptor genes by comprehensive phylogenetic analysis. Gene 328:203–204.

Iwema CL, Fang H, Kurtz DB, Youngentob SL, Schwob JE (2004) Odorant receptor expression patterns are restored in lesion-recovered rat olfactory epithelium. J Neurosci 24:356–369.

Kambere MB, Lane RP (2007) Co-regulation of a large and rapidly evolving repertoire of odorant receptor genes. BMC Neurosci 8(Suppl 3):S2.

Kambere MB, Lane RP (2009) Exceptional LINE density at V1R loci: the Lyon repeat hypothesis revisited on autosomes. J Mol Evol 68:145–159.

Keller A, Vosshall LB (2008) Better smelling through genetics: mammalian odor perception. Curr Opin Neurobiol 18:364–369.

Kishida T (2008) Pattern of the divergence of olfactory receptor genes during tetrapod evolution. PLoS One 3:e2385.

Knapp LA, Robson J, Waterhouse JS (2006) Olfactory signals and the MHC: a review and a case study in *Lemur catta*. Am J Primatol 68:568–584.

Krumlauf R (1992) Evolution of the vertebrate *Hox* homeobox genes. Bioessays 14:245–252.

Lane RP, Roach JC, Lee IY, et al. (2002) Genomic analysis of the olfactory receptor region of the mouse and human T-cell receptor alpha/delta loci. Genome Res 12:81–87.

Liberles SD, Buck LB (2006) A second class of chemosensory receptors in the olfactory epithelium. Nature 442:645–650.

Malnic B, Godfrey PA, Buck LB (2004) The human olfactory receptor gene family. Proc Natl Acad Sci USA 101:2584–2589.

Malnic B, Hirono J, Sato T, Buck LB (1999) Combinatorial receptor codes for odours. Cell 96: 713–723.

Menini A, Lagostena L, Boccaccio A (2004) Olfaction: from odorant molecules to the olfactory cortex. News Physiol Sci 19:101–104.

Miyamichi K, Serizawa S, Kimura H M, Sakano H (2005) Continuous and overlapping expression domains of odorant receptor genes in the olfactory epithelium determine the dorsal/ventral positioning of glomeruli in the olfactory bulb. J Neurosci 25:3586–3592.

Mombaerts P (2004) Odorant receptor gene choice in olfactory sensory neurons: the one receptor-one neuron hypothesis revisited. Curr Opin Neurobiol 14:31–36.

Mombaerts P (2006) Axonal wiring in the mouse olfactory system. Annu Rev Cell Dev Biol 22:713–737.

Moulton DG (1974) Dynamics of cell populations in the olfactory epithelium. Ann NY Acad Sci 237:52–61.

Nei M, Niimura Y, Nozawa M (2008) The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. Nat Rev Genet 9:951–963.

Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. Annu Rev Genet 39:121–152.

Ngai J, Chess A, Dowling MM, et al. (1993) Coding of olfactory information: topography of odorant receptor expression in the catfish olfactory epithelium. Cell 72:667–680.

Niimura Y (2009) On the origin and evolution of vertebrate olfactory receptor genes: comparative genome analysis among 23 chordate species. Genome Biol Evol 2009:34–44.

Niimura Y, Nei M (2005) Evolutionary dynamics of olfactory receptor genes in fishes and tetrapods. Proc Natl Acad Sci USA 102:6039–6044.

Niimura Y, Nei M (2006) Evolutionary dynamics of olfactory and other chemosensory receptor genes in vertebrates. J Hum Genet 51:505–517.

Niimura Y, Nei M (2007) Extensive gains and losses of olfactory receptor genes in Mammalian evolution. PLoS One 2:e708.

Nishizumi H, Kumasaka K, Inoue N, Nakashima A, Sakano H (2007) Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in *cis*. Proc Natl Acad Sci USA 104:20067–20072.

Nozawa M, Kawahara Y, Nei M (2007) Genomic drift and copy number variation of sensory receptor genes in humans. Proc Natl Acad Sci USA 104:20421–20426.

Olender T, Lancet D, Nebert DW (2008) Update on the olfactory receptor (OR) gene superfamily. Hum Genomics 3:87–97.

Pfister P, Rodriguez I (2005) Olfactory expression of a single and highly variable V1R pheromone receptor-like gene in fish species. Proc Natl Acad Sci USA 102:5489–5494.

Pilpel Y, Lancet D (1999) The variable and conserved interfaces of modeled olfactory receptor proteins. Protein Sci 8:969–977.

Rawson NE, Eberwine J, Dotson R, et al. (2000) Expression of mRNAs encoding for two different olfactory receptors in a subset of olfactory receptor neurons. J Neurochem 75:185–195.

Rodriguez-Trelles F, Tarrio R, Ayala FJ (2005) Is ectopic expression caused by deregulatory mutations or due to gene-regulation leaks with evolutionary potential? Bioessays 27:592–601.

Rouquier S, Giorgi D (2007) Olfactory receptor gene repertoires in mammals. Mutat Res 616: 95–102.

Rouquier S, Taviaux S, Trask BJ, et al. (1998) Distribution of olfactory receptor genes in the human genome. Nat Genet 18:243–250.

Sanchez-Gracia A, Vieira FG, Rozas J (2009) Molecular evolution of the major chemosensory gene families in insects. Heredity 103:208–216.

Saraiva LR, Korsching SI (2007) A novel olfactory receptor gene family in teleost fish. Genome Res 17:1448–1457.

Sato Y, Miyasaka N, Yoshihara Y (2007) Hierarchical regulation of odorant receptor gene choice and subsequent axonal projection of olfactory sensory neurons in zebrafish. J Neurosci 27:1606–1615.

Schneider N, Fletcher T, Shaw G, Renfree MB (2009) The olfactory system of the tammar wallaby is developed at birth and directs the neonate to its mother's pouch odours. Reproduction 138:849–857.

Serizawa S, Miyamichi K, Nakatani H, et al. (2003) Negative feedback regulation ensures the one receptor-one olfactory neuron rule in mouse. Science 302:2088–2094.

Shykind BM, Rohani SC, O'Donnell S, et al. (2004) Gene switching and the stability of odorant receptor gene choice. Cell 117:801–815.

Singh N, Ebrahimi FA, Gimelbrant AA, et al. (2003) Coordination of the random asynchronous replication of autosomal loci. Nat Genet 33:339–341.

Stephen DL (2009) Trace amine-associated receptors are olfactory receptors in vertebrates. Ann NY Acad Sci 1170:168–172.

Stoddart DM (1980) The Ecology of Vertebrate Olfaction. Chapman and Hall, London and New York.

Sullivan SL, Adamson MC, Ressler KJ, Kozak CA Buck LB (1996) The chromosomal distribution of mouse odorant receptor genes. Proc Natl Acad Sci USA 93:884–888.

Tacher S, Quignon P, Rimbault M, et al. (2005) Olfactory receptor sequence polymorphism within and between breeds of dogs. J Hered 96:812–816.

Tian H, Ma M (2008) Activity plays a role in eliminating olfactory sensory neurons expressing multiple odorant receptors in the mouse septal organ. Mol Cell Neurosci 38:484–488.

Vassar R, Ngai, J, Axel R (1993) Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell: 74:309–318.

Young JM, Endicott RM, Parghi SS, et al. (2008) Extensive copy-number variation of the human olfactory receptor gene family. Am J Hum Genet 83:228–242.

Young JM, Trask BJ (2007) V2R gene families degenerated in primates, dog and cow, but expanded in opossum. Trends Genet 23:212–215.

Zhang X, De la Cruz O, Pinto JM, et al. (2007) Characterizing the expression of the human olfactory receptor gene family using a novel DNA microarray. Genome Biol 8:R86.

Zhang X, Firestein S (2002) The olfactory receptor gene superfamily of the mouse. Nat Neurosci 5:124–133.

Ziegler A, Dohr G, Uchanska-Ziegler B (2002) Possible roles for products of polymorphic MHC and linked olfactory receptor genes during selection processes in reproduction. Am J Reprod Immunol 48:34–42.

Zuri I, Su W, Halpern M (2003) Conspecific odor investigation by gray short-tailed opossums (*Monodelphis domestica*). Physiol Behav 80:225–232.

Part IX
Marsupial Conservation Genetics

Jennifer A. Marshall Graves

The preceding parts have detailed what marsupials can do for genetics. In this part we address the question of what genetics can do for marsupials (Sherwin and Graves, 2006).

There are about 270 extant species of marsupials, most living in Australia. Many are endangered or vulnerable, including such icons as the bilby (the Australian "Easter bunny"), the fabulous koala and gliding possums, and the Tasmanian devil. In contrast, several species are in danger from overbreeding, for instance the koala lays waste to limited tree habitat and the population would crash if there were no intervention. Australia has a truly horrible record for marsupial extinctions (Johnson, 2006).

Molecular markers have armed conservation geneticists with powerful new tools to study marsupial population structure, numbers and subdivisions. Starting from restriction fragment length polymorphisms (RFLPs) in the 1980s (Sherwin and Murray, 1990) even limited information was of use in the management of koalas, which, though numerous, were almost invariant in the south, where they were almost extincted by fire and trapping. Mitochondrial sequence variation, particularly in the D-loop region also provided useful information about the maternal lineage, and limited use of Y chromosome markers has been made to track the contribution of individual males. These techniques produced useful data, but were not easy to apply to many marsupial species because of the large samples that were required.

Conservation genetic research has become much more accessible and powerful since Sherwin and Murray's contribution to the 1990 Boden Conference (Sherwin and Murray, 1990). The development of PCR-based assays completely changed the ease and scope of mammal population studies. Almost unlimited supplies of useful markers can be easily attained by high throughput screening methods. Non-invasive sampling of populations is now a possibility with the increased reliability of detection methods. These methods are now being applied to marsupials to produce a great spurt in population studies. The availability of genomic sequences from two marsupials also makes it possible to target regions of the genome for variation.

In Chapter 22 Eldridge reviews the types of marker now available, their isolation and use. He summarises the literature on frequency of polymorphisms in marsupial populations, and provides some case studies attesting to their usefulness in

marsupials and their potential for filling in the considerable gaps. Eldridge notes that only 15% of marsupial species have so far been the subject of molecular population studies to date, and there is a dearth of knowledge, particularly about large and secure populations, which would be a useful backdrop for studies of endangered species.

The necessity for studying the populations of apparently secure and widespread species of marsupials has been thrown into stark relief by the calamitous decline of the Tasmanian devil population due to the ravages of a unique facial cancer that is transmitted from animal to animal by biting (Bender, Chapter 23). The population structure, and even the numbers of this species, was almost unknown until the scale of the decline became obvious just a few years ago.

Using new genetic technology it has been possible rapidly to screen the genome for microsatellite markers, and to deploy these to examine genetic variation within and between populations all over Tasmania. In particular, the major histocompatibility locus MHC, usually the most variable locus in many mammals, has few alleles, suggesting that the population bottlenecked, probably multiple times. The lack of genetic variation underlies the potential of cancer cells to colonise unrelated animals, since they will not mount an immune response (Siddle et al., 2007).

These findings substantiate the long-held view (O'Brien et al., 1983) that limited genetic variability renders a species vulnerable to pathogens and declining fertility. While holding out little hope for natural immunity of devils to the tumour, the MHC and microsatellite typing is proving crucial for maximising genetic variability in insurance populations of devils that are now being set up in case the species becomes extinct in the wild.

High throughput genomics may yet come to the rescue of the beleaguered devil. Very recently, the sequences of all transcripts ("transcriptome") in normal devil tissues, and of the tumour, have been determined and compared (Murchison et al., 2010). The expression profile of the tumour resembles that of Schwann cells, enabling us, for the first time, to determine that the tumour cell that entered a life of its own originally derived from a Schwann cell in a single animal. Not only does this information help to know the enemy, potentially with the aim of developing a vaccine directed against some unique protein of the tumour, but also delivers us a valuable diagnostic marker. We have truly entered the era of genomic and evolutionary approaches to animal, as well as human health.

References

Johnson C (2006) Australia's Mammal Extinctions. Cambridge University Press, Cambridge.

Murchison EP, Tovar C, Hsu A, et al. (2010) The Tasmanian devil transcriptome reveals Schwann cell origins of a clonally transmissible cancer. Science 327:84–87.

O'Brien SJ, Wildt DE, Goldman D, Merril CR, Busch M (1983) The cheetah is depauperate in genetic variation. Science 221:459–462.

Sherwin WB, Graves JAM (2006) What Marsupials can do for genetics and what genetics can do for marsupials. In: Armati P, Dickman CR, Hume ID (eds) Marsupials. Cambridge University Press, Cambridge.

Sherwin WB, Murray ND (1990) Population and conservation genetics of marsupials. Aust J Zool 37:161–180.

Siddle HV, Kreiss A, Eldridge MD, et al. (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA 104:16221–16226.

Chapter 22
Marsupial Population and Conservation Genetics

Mark D.B. Eldridge

Abstract Recent technological advances have resulted in a dramatic renaissance of population genetics and its application to species ecology and conservation. This review summarizes the progress made in applying these new techniques, notably hypervariable genetic markers (microsatellite loci and mitochondrial control region), to the study of marsupials. Since 1990, population genetic studies have overwhelmingly been of marsupial species from eastern and southern Australia, largely focusing on threatened species or those with restricted distributions. To date, over 500 polymorphic microsatellite loci have been isolated from 38 marsupial species (six American, 32 Australian), including representatives from 13 of the 18 extant marsupial families. Levels of microsatellite diversity identified within the 209 marsupial populations (43 species) so far examined have varied greatly, although the range is similar to those reported from eutherians and other vertebrates. Marsupial populations with high levels of genetic diversity tend to be those from relatively abundant or widespread species, while those with lower levels are typically species with restricted distributions, that are threatened, found on islands or have been established via translocation. Although data for most families are still limited, bandicoots, koalas and wombats appear less and phalangerids more diverse than the marsupial average. The application of these hypervariable genetic markers to investigate marsupial species ecology has substantially improved our understanding of population biology, behaviour and reproduction in many species, as well as informing conservation initiatives and management plans for many threatened marsupial taxa. It is hoped that in the future, marsupial population genetic studies can be expanded to include a larger number of South American and New Guinean species, as well as a better representation of arid, tropical, widespread and abundant Australian species.

Keywords Marsupial · Microsatellite · Mitochondrial DNA · Genetic diversity · Wildlife management

M.D.B. Eldridge (✉)
Australian Museum, Sydney, NSW 2010, Australia
e-mail: mark.eldridge@austmus.gov.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_22, © Springer Science+Business Media B.V. 2010

22.1 Introduction

Population genetics is the study of the distribution and abundance of genetic diversity and how it is impacted through time by gene flow, genetic drift, mutation and selection. It is therefore able to offer unique insights into the ecological and evolutionary processes that have shaped populations and species, as well as providing a means for examining the influence of contemporary forces and predicting the impact of future changes. As such, population genetics is able to increase our understanding of fundamental biological processes, as well as species behaviour, reproduction and ecology, and increasingly is being used to inform wildlife management.

The unique biology of marsupials has made them of enduring international scientific interest and the focus of ongoing research. Population genetic studies of marsupials have historically been hampered, as for other mammals, by difficulties in detecting sufficient genetic variation. Throughout the 1970s and 1980s a series of allozyme electrophoresis studies revealed that marsupials contained the relatively low levels of genetic diversity (average heterozygosity $He = 0.040 \pm 0.012$ from over 25 species) typically detected with this technique (reviewed in Sherwin and Murray, 1990). These studies, however, made some progress in clarifying species boundaries and determining broad-scale population structure (e.g., Johnston and Sharman, 1979; Seebeck and Johnston, 1980; Baverstock et al., 1984; Dickman et al., 1988).

Since the last major review of marsupial population and conservation genetics (Sherwin and Murray, 1990) a series of technological discoveries and innovations have transformed molecular genetics and lead to a dramatic renaissance in population genetics. Central to this revolution was the development of the polymerase chain reaction (PCR) (reviewed in Avise, 1994) and the identification of hypervariable segments of the genome (both nuclear and mitochondrial) which could be accessed, and variation assessed, via PCR (reviewed in Sunnucks, 2000). Coupled with this dramatic increase in the information content of genetic data, which enabled individuals to be routinely and unequivocally identified for the first time, were some equally innovative developments in methods of data analysis (reviewed in Excoffier and Heckel, 2006; Majoram and Tavare, 2006), allowing the potential of these data to be more fully realised.

This review seeks to summarise the progress made to date in applying the most commonly used of these new generation of hypervariable genetic markers (microsatellites and the mtDNA control region) to marsupial populations and to illustrate how these studies have enhanced our knowledge of marsupial biology and conservation. For a summary of Major Histocompatibility Complex (MHC) diversity studies in marsupials see Belov (Chapter 16).

22.2 Microsatellites

Microsatellites are tandem repeats of DNA sequences 2–5 base pairs in length that have become the marker of choice in population genetics studies. They are hypervariable in repeat length, widespread in the genome and mostly show codominant

Mendelian inheritance and selective neutrality (reviewed in Bruford et al., 1996; Oliveira et al., 2006).

To date, 506 polymorphic microsatellite loci have been isolated from 38 marsupial species (six American, 32 Australian), including representatives from 13 of the 18 extant marsupial families (Table 22.1). Loci are yet to be published for any representative of the South American families Microbiotheridae and Caenolestidae, as well as the Australian families Myrmecobiidae, Notoryctidae and Hypsiprymnodontidae. Almost half (205/506, 41%) of published loci have been isolated from just two species, the South American gray short-tailed opossum (*Monodelphis domestica*), and the Australian tammar wallaby (*Macropus eugenii*) (Table 22.1). This bias is unsurprising as both species are model marsupials for genetic and genomic research (Deakin, Chapter 5; Papenfuss et al., Chapter 6; Samollow, Chapter 4). Of the remaining loci characterised, most are from Australian species although the overall coverage within families remains quite patchy and uneven.

Although microsatellite loci have been isolated from only 38 of the ~320 recent marsupial species, many loci exhibit remarkable cross-species utility, successfully amplifying homologous polymorphic loci in congeneric as well as inter-generic species within families (Table 22.1) and occasionally extending to other families (Table 22.1). Even though a decrease in successful cross-species amplification with increasing genetic distance has been observed, as well as some reduction in diversity in non-source species (Zenger et al., 2003c), the loci currently available appear sufficiently versatile to enable population genetic studies for most living marsupial species. Microsatellite-based population genetic studies have so far been conducted on 43 marsupial species (Table 22.2); all from Australia (almost entirely eastern and southern) and predominantly (44%) macropodoids. Comparatively few widespread or common species have been examined, with most studies focusing on threatened species or those with relatively restricted distributions (Table 22.2).

22.2.1 Genetic Diversity at Microsatellite Loci

The levels of genetic diversity identified within the 209 marsupial populations examined to date using microsatellite loci have varied enormously (Table 22.2). Allelic diversity (A), that is the average numbers of alleles per locus, was found to vary from 1.2 to 18.2 (average ± se = 7.0±0.2); while reported values of average heterozygosity (He) have ranged from 0.05 to 0.90 (average ± se = 0.66±0.01). This range of values is similar to those reported from eutherians (Garner et al., 2005) and other vertebrates (DeWoody and Avise, 2000). Garner et al. (2005) reported that on average marsupials (0.765±0.002, n=11) had significantly higher heterozygosity values than eutherians (0.677±0.01, n=74). However, this reported difference appears to be an artefact of the small number of marsupial populations included in their study.

As might be predicted, marsupial populations with the highest microsatellite diversity tend to be those from widespread and abundant or high-density species (e.g., medium-large macropodoids, small dasyurids, one phalangerid:

Table 22.1 Polymorphic microsatellite loci isolated from marsupials and their cross-species utility

Family (number of recent species[a])	Source taxon	No. of loci	References	Non-source taxa in which loci are polymorphic	References
Didelphidae (87)	*Didelphis marsupialis*	8	Lavergne et al. (1999), Guillemin et al. (2000)	Other *Didelphis, Philander*	Lavergne et al. (1999), Guillemin et al. (2000)
	Didelphis viginiana	21	Fike et al. (2009)		
	Micoureus paraguayanus	5	Rodrigues et al. (2006)		
	Micoureus demerarae	8	Dias et al. (2008)	Other *Micoureus*	Dias et al. (2008)
	Monodelphis domestica	150	Samollow et al. (2004), Gouin et al. (2005a, b), Samollow et al. (2007)		
	Philander opossum	4	Guillemin et al. (2000)		
Dasyuridae (69)	*Antechinus agilis*	16	Kraaijeveld-Smit et al. (2002c), Banks et al. (2005a), Lada et al. (2007)	Other *Antechinus*	Holleley et al. (2006), Lada et al. (2007)
	Dasyurus geoffroii	5	Spencer et al. (2007)	Other *Dasyurus, Phascogale*	Spencer et al. (2007)
	Dasyurus maculatus & *D. viverrinus*	5	Firestone (1999)	Other *Dasyurus, Phascogale*	Spencer et al. (2007), Cardoso et al. (2009)
	Parantechinus apicalis	7	Mills and Spencer (2003)		
	Sarcophilis harrisii	11	Jones et al. (2003)	*Dasyurus, Phascogale*	Spencer et al. (2007)
	Sminthopsis douglasi	7	Spencer et al. (2003)	Other *Sminthopsis, Planigale, Ningaui*	Spencer et al. (2003)
Peramelidae (18)	*Isoodon obesulus*	8	Zenger and Johnston (2001)	Other *Isoodon, Perameles, Echymipera, Sminthopsis, Dasyurus, Phascolarctos*	Zenger and Johnston (2001), Smith and Hughes (2008)
	Perameles bougainville	1	Smith and Hughes, (2008)		
Thylacomyidae (2)	*Macrotis lagotis*	10	Moritz et al. (1997), Smith et al. (2009)		

Table 22.1 (continued)

Family (number of recent species[a])	Source taxon	No. of loci	References	Non-source taxa in which loci are polymorphic	References
Vombatidae (3)	*Lasiorhinus krefftii*	20	Beheregaray et al. (2000), Walker et al. (2009)	Other *Lasiorhinus, Vombatus*	Beheregaray et al. (2000), Banks et al. (2002b), Walker et al. (2009)
	Lasiorhinus latifrons	12	Beheregaray et al. (2000)	Other *Lasiorhinus, Vombatus*	Beheregaray et al. (2000), Banks et al. (2002b)
Phascolarctidae (1)	*Phascolarctos cinereus*	17	Houlden et al. (1996a), Cristescu et al. (2009)		
Phalangeridae (27)	*Trichosurus vulpecula*	10	Taylor and Cooper (1998a), Lam (2000), Martin et al. (2007b)	Other *Trichosurus*	Martin et al. (2007b)
Burramyidae (5)	*Burramys parvus*	8	Mitrovski et al. (2005)		
Tarsipedidae (1)	*Tarsipes rostratus*	5	Spencer and Bryant (2000)		
Petauridae (11)	*Dactylopsila trivirgata*	9	Hansen et al. (2003)	*Petaurus, Gymnobelideus*	Hansen et al. (2003)
	Gymnobelideus leadbeateri	17	Hansen et al. (2005)	*Dactylopsila, Petaurus*	Hansen et al. (2005)
	Petaurus australis	3	Brown et al. (2004)	Other *Petaurus*	Brown et al. (2004)
	Petaurus breviceps	6	Brown et al. (2004)	Other *Petaurus*	Brown et al. (2004)
	Petaurus norfolcensis	2[b]	Millis (2000)	Other *Petaurus*	Millis (2000)
Pseudocheiridae (17)	*Petauroides volans*	12	Taylor et al. (2002)		
	Pseudocheirus occidentalis	14	Wilson et al. 2009		
	Pseudocheirus peregrinus	11	Lancaster et al. (2009)	*Pseudochirops, Pseudochirulus*	Lancaster et al. (2009)
Potoroidae (10)	*Potorous longipes*	6	Luikart et al. (1997)	*Bettongia*, other *Potorous*	Luikart et al. (1997), Donaldson and Vercoe (2008)

Table 22.1 (continued)

Family (number of recent species[a])	Source taxon	No. of loci	References	Non-source taxa in which loci are polymorphic	References
	Bettongia tropica	3	Pope et al. (2000)	*Aepyprymnus*, other *Bettongia*, *Hypsiprymnodon*	Johnson et al. (2005), Pope et al. (2005), Donaldson and Vercoe (2008)
Macropodidae (65)	*Macropus eugenii*	56[c]	Taylor and Cooper (1998b), Hawken et al. (1999) Zenger and Cooper (2001a), MacDonald et al. (2006, 2007), Cheng et al. (2009a, b)	*Bettongia, Dendrolagus, Hypsiprymnodon, Lagorchestes*, other *Macropus, Petrogale, Thylogale, Wallabia*	Heckenberg (1997), Taylor and Cooper (1998b), Eldridge et al. (1999), Bowyer et al. (2002), Eldridge et al. (2004a), Johnson et al. (2005), Donaldson and Vercoe (2008), Paplinska et al. (2009)
	Macropus giganteus	9	Zenger and Cooper (2001b)	*Hypsiprymnodon*, other *Macropus, Petrogale, Thylogale, Wallabia*	Eldridge et al. (2001b), Zenger et al. (2003a), Johnson et al. (2005), Macqueen et al. (2009), Paplinska et al. (2009)
	Macropus parma	3	Ivy et al. (2009)		
	Petrogale xanthopus	7	Pope et al. (1996), Pope (personal communication)[d]	*Aepyprymnus, Bettongia, Dendrolagus, Lagorchestes, Macropus, Onychogalea, Thylogale*, other *Petrogale*	Eldridge et al. (1999), Pope et al. (2000), Bowyer et al. (2002), Zenger et al. (2003a), Eldridge et al. (2004a), Pope et al. (2005), Sigg et al. (2005),

Table 22.1 (continued)

Family (number of recent species[a])	Source taxon	No. of loci	References	Non-source taxa in which loci are polymorphic	References
					Donaldson and Vercoe (2008), Macqueen et al. (2009)
	Petrogale assimilis	6	Spencer et al. (1995), Spencer (1996)	*Aepyprymnus, Dendrolagus, Macropus,* other *Petrogale, Thylogale, Wallabia*	Heckenberg (1997), Eldridge et al. (1999), Bowyer et al. (2002), Zenger et al. (2003a), Pope et al. (2005), Paplinska et al. (2009)
	Onychogalea fraenata	4	Pope et al. (2000), Moritz (personal communication)[d]	*Aepyprymnus, Bettongia, Dendrolagus, Hypsiprymnodon, Thylogale, Macropus*	Pope et al. (2000), Bowyer et al. (2002), Zenger et al. (2003a), Johnson et al. (2005), Pope et al. (2005), Macqueen et al. (2009)

[a]See Wilson and Reeder (2005)
[b]See Brown et al. (2004)
[c]An additional 9 monomorphic loci have been described (MacDonald et al., 2006; 2007)
[d]Details of loci given in Zenger et al. (2002) or Sigg et al. (2005)

Table 22.2 Allelic diversity (*A*) and average heterozygosity (*He*) at polymorphic microsatellite loci in 209 populations of 43 marsupial species

Taxon	*N*	No. of loci	*A*	*He*	References
Antechinus agilis (Mt Donna Buang)	206	5	18.2	0.89	Kraaijeveld-Smit et al. (2002a)
Trichosurus vulpecula (Guy Fawkes NP-west)	83	6	17.8	0.90	Clinchy et al. (2004)
Antechinus agilis (Mt Disappointment 2000)	313	5	17.8	0.89	Kraaijeveld-Smit et al. (2002a)
Trichosurus vulpecula (Guy Fawkes NP-east)	71	6	16.8	0.90	Clinchy et al. (2004)
Macropus fuliginosus (Congelin)	38	9	15.6[a]	0.90	Neaves et al. (2009)
Antechinus stuartii (Chichester SF)	401	5	14.8	0.89	Holleley et al. (2006)
Macropus fuliginosus (Geraldton)	24	9	14.3[a]	0.88	Neaves et al. (2009)
Macropus fuliginosus (Esperance)	29	9	14.2[a]	0.89	Neaves et al. (2009)
Antechinus agilis (Buccleuch SF)	170	15	14.1	0.86	Banks et al. (2005a)
Trichosurus vulpecula (nw Sydney)	64	7	14.1	0.85	Stow et al. (2006)
Macropus fuliginosus (Wongan Hills)	30	9	13.9[a]	0.88	Neaves et al. (2009)
**Macropus eugenii* (Tutanning NR)	63	7	13.9	0.86	Eldridge et al. (2004a)
Macropus fuliginosus (Albany)	30	9	13.8[a]	0.89	Neaves et al. (2009)
Macropus fuliginosus (Perth)	27	9	13.7[a]	0.87	Neaves et al. (2009)
Antechinus agilis (Mt Disappointment 1999)	67	5	13.0	0.88	Kraaijeveld-Smit et al. (2002b)
Macropus fuliginosus (Coffin Bay)	36	9	12.9[a]	0.85	Neaves et al. (2009)
Antechinus flavipes (Koondrook-Gunbower)	360	11	12.8[a]	0.83	Lada et al. (2008b)
Aepyprymnus rufescens (Mt Fox)	134	6	12.3	0.83	Pope et al. (2005)
Macropus fuliginosus (Witchcliffe)	19	9	11.8[a]	0.88	Neaves et al. (2009)
Trichosurus vulpecula (Hautapu, NZ)[b]	187	6	11.8	0.83	Taylor et al. (2000)
[T]**Onychogalea fraenata* (Taunton)	100	7	11.7	0.81	Sigg (2006)
Petrogale assimilis (Black Rock)	128	5	11.6	0.86	Spencer and Marsh (1997)
Macropus fuliginosus (Madura)	34	9	11.6[a]	0.84	Neaves et al. (2009)
Antechinus swainsonii (Mt Donna Buang)	166	5	11.6	0.80	Kraaijeveld-Smit et al. (2007)
Antechinus flavipes (Chilton)	156	11	11.5[a]	0.83	Lada et al. (2008b)
Macropus fuliginosus (Hillston)	25	9	11.2[a]	0.87	Neaves et al. (2009)
Macropus fuliginosus (Moonera)	21	9	11.2[a]	0.87	Neaves et al. (2009)
[T]*Gymnobelideus leadbeateri* (Lake Mountain)	159	15	11.2	0.79	Hansen and Taylor (2008)
Macropus fuliginosus (Kalgoorlie)	30	9	11.1[a]	0.86	Neaves et al. (2009)

Table 22.2 (continued)

Taxon	N	No. of loci	A	He	References
Macropus eugenii (Kangaroo Island)	46	7	11.1	0.84	Taylor and Cooper (1999)
Trichosurus vulpecula (Orongorongo, NZ)[b]	69	5	11.0	0.76	Taylor et al. (2004)
Trichosurus vulpecula (Hawke's Bay, NZ)[b]	155	5	10.8	0.83	Taylor et al. (2004)
Trichosurus vulpecula (Healesville)	39	5	10.6	0.83	Taylor et al. (2004)
Macropus giganteus (Tidbinbilla)	48	8	10.4	0.81	Zenger et al. (2003a)
Antechinus flavipes (Millewa-Barmah)	73	11	10.4[a]	0.80	Lada et al. (2008b)
Macropus giganteus (Breadalbane)	29	8	10.1	0.80	Zenger et al. (2003a)
Wallabia bicolor (Maryvale)	178	7	10.1	0.75	Paplinska et al. (2009)
Macropus rufogriseus (Grenta, Tasmania)	44	5	10.0	0.76	Le Page et al. (2000)
Aepyprymnus rufescens (School House)	58	6	9.8	0.82	Pope et al. (2005)
Aepyprymnus rufescens (Dingo Creek)	39	6	9.7	0.81	Pope et al. (2005)
Antechinus flavipes (Rushworth-Reedy)	64	11	9.7[a]	0.77	Lada et al. (2008b)
Macropus giganteus (South Bathurst)	35	8	9.6	0.81	Zenger et al. (2003a)
Thylogale billardierii (northcentral Tasmania)	25	10	9.6[a]	0.75	Macqueen et al. (2009)
Thylogale billardierii (southcentral Tasmania)	25	10	9.5[a]	0.75	Macqueen et al. (2009)
Thylogale billardierii (nw Tasmania)	25	10	9.5[a]	0.73	Macqueen et al. (2009)
Petauroides volans (Buccleuch SF 1966)[c]	38	12	9.4[a]	0.74	Taylor et al. (2007)
Antechinus flavipes (Tumut)	24	9	9.3	0.77	Beckman et al. (2007)
Vombatus ursinus (Glenburn)	49	8	9.3	0.75	Banks et al. (2002b)
[T]*Dasyurus geoffroii* (Batalling SF)[d]	35	6	9.2	0.86	Firestone et al. (2000)
Macropus rufogriseus (Deloraine, Tasmania)	26	5	9.2	0.79	Le Page et al. (2000)
Trichosurus vulpecula (Shannon, NZ)[b]	45	5	9.2	0.78	Taylor et al. (2004)
Trichosurus vulpecula (Stonehenge, Tasmania)	45	5	9.2	0.77	Taylor et al. (2004)
Macropus fuliginosus (Peterborough)	20	9	9.1[a]	0.83	Neaves et al. (2009)
Vombatus ursinus (Tonimbuk)	43	8	9.1	0.76	Banks et al. (2002b)
[T]*Dasyurus hallucatus* (Kapalaga, Kakadu NP)	26	6	9.0	0.68	Firestone et al. (2000)
[T]*Macrotis lagotis* (NT)	19	9	8.9	0.86	Moritz et al. (1997)
Macropus giganteus (Marulan)	15	8	8.9	0.85	Zenger et al. (2003a)
Trichosurus vulpecula (Scotland Island)	28	7	8.9	0.81	Stow et al. (2006)

Table 22.2 (continued)

Taxon	*N*	No. of loci	*A*	*He*	References
Macropus rufogriseus (se mainland)[#]	15	5	8.8	0.87	Le Page et al. (2000)
[T]*Dasyurus geoffroii* (Perth)[#,e]	23	6	8.8	0.80	Firestone et al. (2000)
Thylogale s. stigmatica[#]	24	8	8.8	0.80	Heckenberg (1997)
Macropus fuliginosus (Para Wirra)	32	9	8.8[a]	0.78	Neaves et al. (2009)
Thylogale billardierii (ne Tasmania)	24	10	8.8[a]	0.72	Macqueen et al. (2009)
Macropus fuliginosus (Shark Bay)	20	9	8.7[a]	0.83	Neaves et al. (2009)
Petrogale assimilis (Little Black Rock)	15	5	8.6	0.85	Spencer et al. (1995)
Antechinus flavipes (Warby-Ovens)	45	11	8.6[a]	0.78	Lada et al. (2008b)
Antechinus flavipes (Campbells-Guttrum)	57	11	8.6[a]	0.77	Lada et al. (2008b)
Thylogale thetis[#]	20	9	8.6	0.76	Heckenberg (1997)
Thylogale billardierii (n Tasmania)	26	10	8.6[a]	0.70	Macqueen et al. (2009)
Thylogale billardierii (King Island)	25	10	8.3[a]	0.74	Macqueen et al. (2009)
Thylogale billardierii (se Tasmania)	25	10	8.3[a]	0.74	Macqueen et al. (2009)
Aepyprymnus rufescens (Quartz Gum)	17	6	8.2	0.84	Pope et al. (2005)
Antechinus flavipes (Mt Donna Buang)	15	9	8.1	0.82	Beckman et al. (2007)
Phascolarctus cinereus (Nowendoc)	27	6	8.0	0.83	Houlden et al. (1996b)
Trichosurus vulpecula (Pigeon Flat, NZ)[b]	106	5	8.0	0.72	Taylor et al. (2004)
[T]*Petrogale x. xanthopus* (Flinders Ranges)[#]	69	7	7.8	0.75	Zenger et al. (2003c)
[T*]*Macrotis lagotis* (Qld)	27	9	7.7	0.75	Moritz et al. (1997)
[T*]*Onychogalea fraenata* (Townsville)[d]	34	7	7.6	0.79	Sigg (2006)
Thylogale stigmatica wilcoxi[#]	18	8	7.6	0.76	Heckenberg (1997)
[T]*Burramys parvus* (Mt Higginbotham)	59	8	7.6	0.67	Mitrovski et al. (2007)
Aepyprymnus rufescens (Loading Ramp)	21	6	7.5	0.78	Pope et al. (2005)
Petrogale xanthopus celeris (Hill of Knowledge)	50	4	7.5	0.69	Pope et al. (1996)
[T]*Isoodon obesulus* (East Gippsland)	22	8	7.4	0.79	Zenger et al. (2005)
[T]*Burramys parvus* (Bundara)	78	8	7.3	0.64	Mitrovski et al. (2007)
Phascolarctus cinereus (Mutdapilly)	20	6	7.2	0.81	Houlden et al. (1996b)
Trichosurus vulpecula (Waipoua, NZ)[b]	33	5	7.2	0.78	Taylor et al. (2004)

Table 22.2 (continued)

Taxon	*N*	No. of loci	*A*	*He*	References
Lasiorhinus latifrons (Brookfield CP)	102	15	7.2	0.76	Walker (2004)
Antechinus flavipes (Seaview, Otway)	20	9	7.2	0.76	Beckman et al. (2007)
[T]**Macrotis lagotis* (Astrebla Downs NP)	19	8	7.1[a]	0.76	Smith et al. (2009)
Antechinus flavipes (Charley's Cr, Otway)	26	9	7.1	0.70	Beckman et al. (2007)
[T]*Bettongia tropica* (Davies Creek)	94	7	7.1	0.73	Pope et al. (2000)
Dasyurus viverrinus (NSW)[#,c,e]	25	6	7.0	0.80	Firestone et al. (2000)
[T]**Parantechinus apicalis* (Fitzgerald River NP)	17	7	7.0	0.75	Mills et al. (2004)
Trichosurus vulpecula (Kawau Is., NZ)[b]	37	5	7.0	0.72	Taylor et al. (2004)
Macropus rufogriseus (Flinders Island)	28	5	7.0	0.71	Le Page et al. (2000)
Antechinus flavipes (Kaanglang, Otway)	22	9	7.0	0.70	Beckman et al. (2007)
[T]*Burramys parvus* (Pretty Valley West)	69	8	7.0	0.68	Mitrovski et al. (2007)
[T]*Burramys parvus* (Mt Loch)	93	8	7.0	0.67	Mitrovski et al. (2007)
Trichosurus cunninghami (Boho South-road)	17	7	7.0	0.64	Martin et al. (2007a)
[T]*Dasyurus hallucatus* (Astell Island 2006)[b]	77	5	7.0	0.57	Cardoso et al. (2009)
Antechinus flavipes (Bambra, Otway)	20	9	6.9	0.67	Beckman et al. (2007)
[T]*Burramys parvus* (Timm Spur)	120	8	6.9	0.65	Mitrovski et al. (2007)
[T]*Gymnobelideus leadbeateri*[d]	42	15	6.8	0.74	Hansen and Taylor (2008)
Aepyprymnus rufescens (Mt Claro)	17	6	6.7	0.82	Pope et al. (2005)
Lasiorhinus latifrons (Wauraltee 8B)	17	6	6.7	0.75	Walker (2004)
Wallabia bicolor (Coranderrk)[d]	41	7	6.7	0.73	Paplinska (2005)
Antechinus flavipes (Goldhole, Otway)	23	9	6.7	0.72	Beckman et al. (2007)
Thylogale billardierii (Flinders Island)	25	10	6.7[a]	0.65	Macqueen et al. (2009)
Petrogale assimilis (Mt Stuart)	20	5	6.6	0.89	Spencer et al. (1995)
Lasiorhinus latifrons (Far West Coast)	16	8	6.6	0.78	Alpers et al. (1998)
[T]*Burramys parvus* (Falls Cr)	35	8	6.6	0.66	Mitrovski et al. (2007)
[T]*Burramys parvus* (Paralyser)	40	8	6.6	0.62	Mitrovski et al. (2007)
[T]*Burramys parvus* (Mt Bogong)	42	8	6.5	0.60	Mitrovski et al. (2007)
Lasiorhinus latifrons (Swan Reach)	39	8	6.4	0.75	Alpers et al. (1998)

Table 22.2 (continued)

Taxon	*N*	No. of loci	*A*	*He*	References
Trichosurus vulpecula (Wanganui, NZ)[b]	49	5	6.4	0.72	Taylor et al. (2004)
Trichosurus vulpecula (Hokitika, NZ)[b]	36	5	6.4	0.65	Taylor et al. (2004)
**Petauroides volans* (Buccleuch SF-c3)	18	12	6.3[a]	0.72	Taylor et al. (2007)
**Macropus eugenii* (Kawau Is, NZ)[b]	37	7	6.3	0.62	Taylor and Cooper (1999)
Petaurus australis (Rennick SF)	36	5	6.2	0.69	Brown et al. (2004)
Antechinus flavipes (Wonga, Otway)	29	9	6.1	0.77	Beckman et al. (2007)
[T]*Isoodon obesulus* (Mt Gambier)	22	8	6.1	0.69	Zenger et al. (2005)
[T]*Burramys parvus* (Summit Rd)	43	8	6.1	0.60	Mitrovski et al. (2007)
[T]**Onychogalea fraenata* (Idalia)[b]	75	7	6.0	0.78	Sigg (2006)
Macropus rufogriseus (Kempton, Tasmania)	16	5	6.0	0.74	Le Page et al. (2000)
[T]*Burramys parvus* (Charlottes Pass)	44	8	6.0	0.61	Mitrovski et al. (2007)
Lasiorhinus latifrons (Kulpara)	73	14	5.9	0.71	Walker (2004)
Lasiorhinus latifrons (Nullarbor)	25	15	5.7	0.66	Walker (2004)
[T]*Bettongia tropica* (Emu Creek)	21	7	5.7	0.79	Pope et al. (2000)
Trichosurus cunninghami (Boho South-forest)	29	7	5.6	0.63	Martin et al. (2007b)
Phascolarctus cinereus (Gold Coast)	23	6	5.6	0.68	Houlden et al. (1996b)
Trichosurus vulpecula (Banks Peninsula, NZ)[b]	36	5	5.6	0.59	Taylor et al. (2004)
[T]**Macrotis lagotis* (Charleville)[d]	18	8	5.5[a]	0.68	Smith et al. (2009)
Trichosurus vulpecula (Riverton, NZ)[b]	46	5	5.4	0.53	Taylor et al. (2004)
Trichosurus vulpecula (Nelson, NZ)[b]	36	5	5.4	0.69	Taylor et al. (2004)
Dendrolagus lumholtzi[#]	45	10	5.3	0.60	Bowyer et al. (2002)
**Macropus parma* (Kawau Island, NZ)[b,d]	75	9	5.3	0.60	Ivy et al. (2009)
[T]*Petrogale penicillata* (Ingles Rd)	28	11	5.2	0.70	Eldridge et al. (2004b)
[T]*Bettongia tropica* (Tinaroo)	20	7	5.2	0.68	Pope et al. (2000)
[T]*Petrogale penicillata* (Bowmans Rd)	20	11	5.1	0.69	Eldridge et al. (2004b)
[T]*Petrogale penicillata* (Crocodile Rock)	15	7	5.1	0.65	Piggott et al. (2006b)
[T]*Petrogale penicillata* (Rocky Creek)	17	7	5.1	0.65	Piggott et al. (2006b)
[T]*Dasyurus hallucatus* (East Alligator, Kakadu NP)	15	5	5.0	0.61	Cardoso et al. (2009)

Table 22.2 (continued)

Taxon	*N*	No. of loci	*A*	*He*	References
Vombatus ursinus (Mt Lofty)	17	5	5.0	0.51	Banks et al. (2002a)
[T]*Petrogale penicillata* (Wolgan, Main)	44	7	4.9	0.65	Piggott et al. (2006b)
Vombatus ursinus (Comaum)	36	11	4.8	0.65	Banks et al. (2002b)
[T]*Dasyurus hallucatus* (Pobassoo Island 2006)[b]	34	5	4.8	0.59	Cardoso et al. (2009)
Dasyurus viverrinus (Gladstone, Tas)[c]	58	6	4.8	0.51	Firestone et al. (2000)
[T]*Dasyurus maculatus* (Barrington Guest House)	16	6	4.7	0.67	Firestone et al. (1999)
[*]*Phascolarctus cinereus* (Strzelecki Ranges)	33	6	4.7	0.48	Seymour et al. (2001)
Lasiorhinus latifrons (Scrubby Peak)	24	14	4.6	0.65	Walker (2004)
[*]*Macropus rufogriseus* (New Zealand)[b]	44	5	4.6	0.59	Le Page et al. (2000)
[T]*Petrogale penicillata* (Hurdle Creek)	54	12	4.5	0.67	Hazlitt et al. (2006)
Lasiorhinus latifrons (Brookfield CP)	16	15	4.4	0.65	Taylor et al. (1994)
[*]*Macropus fuliginosus* (Kangaroo Island)	25	9	4.4	0.60	Neaves et al. (2009)
[T]*Dasyurus maculatus* (Wynyard, Tasmania)[c]	32	6	4.3	0.61	Firestone et al. (2000)
[T*]*Lagorchestes hirsutus* (Tanami Desert)	15	8	4.3	0.61	Eldridge et al. (2004a)
[T]*Dasyurus maculatus* (Marengo)	52	6	4.2[a]	0.67[a]	Glen et al. (2009)
[*]*Phascolarctus cinereus* (South Gippsland)	47	6	4.2	0.48	Houlden et al. (1996b)
[*]*Macropus eugenii* (Rotorua, NZ)[b]	61	7	4.1	0.60	Taylor and Cooper (1999)
[T]*Petrogale penicillata* (Broke)	17	11	4.0	0.64	Eldridge et al. (2004b)
[*]*Macropus parma* (Kawau Is., NZ)[b]	51	2	4.0	0.55	Taylor et al. (1999)
[T*]*Petrogale l. lateralis* (Tutakin)	19	11	4.0	0.62	Eldridge et al. (2004a)
[T*]*Isoodon obesulus* (Sydney)	30	8	3.8	0.57	Zenger et al. (2005)
[*]*Phascolarctus cinereus* (French Island)[b]	49	15	3.8	0.48	Cristescu et al. (2009)
Macropus eugenii (Garden Island)	29	7	3.7	0.44	Eldridge et al. (2004a)
Dasyurus viverrinus (Vale of Belvoir, Tasmania)	21	6	3.5	0.59	Firestone et al. (2000)
[T*]*Petrogale l. lateralis* (Sales Rock)	31	11	3.5	0.55	Eldridge et al. (2004a)
[*]*Phascolarctus cinereus* (Philip Island)[b]	17	6	3.5	0.48	Houlden et al. (1996b)
[T]*Petrogale penicillata* (Perseverance Dam)	15	12	3.4	0.57	Hazlitt et al. (2006)
[T*]*Petrogale l. lateralis* (Exmouth)	15	10	3.4	0.62	Eldridge et al. (1999)

Table 22.2 (continued)

Taxon	*N*	No. of loci	*A*	*He*	References
T* *Gymnobelideus leadbeateri* (Yellingbo)	198	15	3.4	0.55	Hansen and Taylor (2008)
T* *Burramys parvus* (Mt Buller)	66	8	3.4	0.28	Mitrovski et al. (2007)
T* *Petrogale l. lateralis* (Mt Caroline East)	32	11	3.3	0.51	Eldridge et al. (2004a)
T *Dasyurus hallucatus* (Boongaree Island)	17	11	3.3	0.42	How et al. (2009)
T *Sarcophilus harrisii* (Narawntapu)	37	11	3.3	0.47	Jones et al. (2004)
* *Phascolarctus cinereus* (Brisbane Ranges)[b]	24	6	3.3	0.40	Houlden et al. (1996b)
T *Sarcophilus harrisii* (Freycinet)	61	11	3.2	0.42	Jones et al. (2004)
T *Sarcophilus harrisii* (Little Swanport)	36	11	3.2	0.41	Jones et al. (2004)
T *Sarcophilus harrisii* (Pawleena)	40	11	3.2	0.41	Jones et al. (2004)
T* *Petrogale penicillata* (Jenolan 1985)[d]	20	11	3.2	0.50	Eldridge et al. (2004b)
T *Sarcophilus harrisii* (Bicheno)	49	11	3.1	0.41	Jones et al. (2004)
T* *Petrogale l. lateralis* (Mt Caroline West)	19	11	3.1	0.49	Eldridge et al. (2004a)
T* *Perameles bougainville* (Dryandra)[d]	38	6	3.0	0.54	Smith and Hughes (2008)
* *Phascolarctus cinereus* (Stony Rises)[b]	17	6	2.9	0.37	Houlden et al. (1996b)
* *Phascolarctus cinereus* (French Island)[b]	43	6	2.9	0.36	Houlden et al. (1996b)
Antechinus minimus (Kanowna Island)	~90	3	2.7	0.51	Sale et al. (2009)
T* *Parantechinus apicalis* (Boullanger Island)	19	7	2.7	0.44	Mills et al. (2004)
* *Phascolarctus cinereus* (Mt Lofty Ranges)[b]	32	6	2.7	0.34	Seymour et al. (2001)
T *Sarcophilus harrisii* (Marrawah)	39	11	2.7	0.39	Jones et al. (2004)
T* *Petrogale penicillata* (Kawau Is, NZ)[b,d]	18	11	2.6	0.44	Eldridge et al. (2001a)
T* *Potorous gilbertii* (Two Peoples Bay)	25	5	2.6	0.46	Sinclair et al. (2002)
* *Trichosurus vulpecula* (Stewart Is., NZ)[b]	25	5	2.6	0.30	Taylor et al. (2004)
T* *Petrogale l. lateralis* (Nangeen Hill)	19	11	2.5	0.41	Eldridge et al. (2004a)
* *Phascolarctus cinereus* (Kangaroo Island)[b]	46	15	2.4	0.41	Cristescu et al. (2009)
T* *Petrogale penicillata* (Kangaroo Valley)	19	12	2.4	0.37	Paplinska (2006)
T *Perameles bougainville* (Dorre Island)	26	6	2.3	0.32	Smith and Hughes (2008)
T* *Petrogale l. lateralis* (Querekin)[b]	43	11	2.2	0.27	Eldridge et al. (2004a)
* *Trichosurus vulpecula* (Chatham Is., NZ)[b]	51	5	2.2	0.38	Taylor et al. (2004)

Table 22.2 (continued)

Taxon	*N*	No. of loci	*A*	*He*	References
[*]*Trichosurus vulpecula* (Codfish Is., NZ)[b]	34	5	2.2	0.23	Taylor et al. (2004)
[*]*Phascolarctus cinereus* (Kangaroo Island)[b]	25	6	2.0	0.31	Seymour et al. (2001)
[T]*Perameles bougainville* (Bernier Island)	33	6	2.0	0.27	Smith and Hughes (2008)
[T*]*Petrogale penicillata* (Rocky Plains Creek)	8	11	1.9	0.34	Browning et al. (2001)
[T*]*Lasiorhinus krefftii* (Epping Forest)	28	16	1.8	0.27	Taylor et al. (1994)
[T]*Lagorchestes hirsutus* (Bernier Island)	17	8	1.8	0.25	Eldridge et al. (2004a)
[T*]*Perameles bougainville* (Heirisson)[b]	29	6	1.8	0.27	Smith and Hughes (2008)
[*]*Phascolarctus cinereus* (Eyre Peninsula)[b]	20	6	1.7	0.21	Seymour et al. (2001)
[T*]*Parantechinus apicalis* (Whitlock Island)	31	7	1.6	0.20	Mills et al. (2004)
[T]*Dasyurus hallucatus* (Marchinbar Island)	27	5	1.6	0.14	Cardoso et al. (2009)
[T*]*Petrogale l. lateralis* (Barrow Island)	28	10	1.2	0.05	Eldridge et al. (1999)

Data are only presented for studies where $N \geq 15$. Populations are ordered by decreasing values of *A*
[T]Threatened taxon (EPBC Act 1999)
[*]Populations that have experienced a recent demographic bottleneck or a recent substantial range contraction (see Van Dyck and Strahan, 2008 or cited reference)
[#]Samples pooled from multiple localities
[a]Personal communication
[b]Population established via translocation
[c]From museum specimens
[d]Captive population
[e]Extinct population
NZ = New Zealand

Table 22.2), while large dasyurids, bandicoots, wombats, koalas and small–medium macropodoids tend to have low-medium levels of diversity (Table 22.2). Populations with the lowest levels are typically from species with restricted distributions, that are threatened, found on islands, are captive or have been established via translocation (Table 22.2). Statistical analysis confirms these observations, showing that populations of threatened marsupial species have significantly reduced genetic diversity (*A*; t=8.1, d*f*=203, p<0.001: *He*; t=6.1, d*f*=137, p<0.001), compared to non-threatened, while populations that have experienced a recent demographic bottleneck have significantly lower genetic diversity (*A*; t=9.3, d*f*=116, p<0.001: *He*; t=8.2, d*f*=70.0., p<0.001) compared to non-bottlenecked populations. For example, the critically endangered, and recently bottlenecked, northern hairy-nosed wombat (*Lasiorhinus krefftii*) has only 36–41% of the genetic diversity found in the more common, closely related, southern hairy-nosed wombat (*Lasiorhinus latifrons*) (Taylor et al., 1994). In addition, island populations of marsupial species have significantly reduced diversity compared to mainland populations (*A*; t=6.5, d*f*=53, p<0.001: *He*; t=6.2, d*f*=36, p<0.001). For example, the Barrow Island (WA) (see Fig. 22.1) population of the black-footed rock-wallaby (*Petrogale lateralis*) (Fig. 22.2) showed reductions in genetic diversity of 52–92% compared to conspecific mainland populations (Eldridge et al., 2004a). Similarly, captive (*A*; t=2.1, d*f*=15, p=0.05) populations and those established via translocation (*A*; t=4.6,

Fig. 22.1 Australia, showing the location of sites mentioned in the text

Fig. 22.2 Black-footed rock-wallaby (*Petrogale lateralis*). Photograph by EJ Miller

df=61, p<0.001: *He*; t=4.5, df=46, p<0.001) have significantly lower genetic diversity than natural populations. For example, the reintroduced Idalia National Park (Qld) (Fig. 22.1) population of the bridled nailtail wallaby (*Onychogalea fraenata*) has only 51% of the allelic diversity present in the remaining wild population at Taunton (Sigg, 2006).

Although data for many marsupial families remain limited, some emerging trends are evident, with bandicoots, koalas and wombats appearing to be less, and phalangerids more, diverse than the marsupial average (Table 22.3). Within the Dasyuridae there is a significant (*A*; t=4.7, df=34, p<0.001: *He*; t=4.2, df=38, p<0.001) pattern of small dasyurids being more diverse than large (Table 22.3); although small dasyurids from the arid or tropical zone are yet to be examined. Within the Macropodidae, rock-wallaby (*Petrogale*) populations have significantly lower levels of genetic diversity (*A*; t=6.5, df=56, p<0.001: *He*; t=3.8, df=31, p=0.001) than other macropodids. Whether this is a result of the highly patchy nature of rock-wallaby habitat naturally limiting the size of individual populations, or an artefact of most studies being conducted on rock-wallaby species from southern Australia which have all in the last 200 years experienced major range contractions

Table 22.3 Mean levels of microsatellite diversity, *A* (allelic diversity) and *He* (average heterozygosity), amongst and within Australian marsupial families

Family (number of recent species[a])	*A* ± *SE*	*He* ± *SE*	*N* species	*N* populations
Dasyuridae (69)	7.2 ± 0.6	0.65 ± 0.03	11	44
Small (<200 g)	9.3± 0.9	0.74 ± 0.03	6	24
Large (>200 g)	4.6 ± 0.5	0.54 ± 0.04	5	19
Peramelidae/Thylacomyidae (20)	5.2 ± 0.9	0.59 ± 0.08	3	9
Vombatidae (3)	6.0 ± 0.6	0.67 ± 0.04	3	13
Phascolarctidae (1)	4.2 ± 0.6	0.50 ± 0.05	1	12
Phalangeridae (27)	11.3 ± 1.6	0.79 ± 0.04	2	8
Burramyidae (5)	6.5 ± 0.3	0.61 ±0.03	1	11
Petauridae (11)	6.9 ± 2.3	0.68 ± 0.07	2	3
Pseudocheiridae (17)	7.9 ± 1.5	0.73 ± 0.01	1	2
Potoroidae (10)	7.5 ± 0.9	0.76 ± 0.04	3	10
Macropodidae (65)	7.9 ± 0.4	0.71 ± 0.02	16	69
Rock-wallabies	4.7 ± 0.5	0.59 ± 0.04	4	23
Other macropodids	9.1 ± 0.5	0.76 ± 0.02	12	46
Marsupials overall	7.2 ± 0.27	0.67 ± 0.01	43	180

Data for captive and introduced New Zealand populations has been excluded
[a]See Wilson and Reeder (2005)

and population declines, remains unclear (Eldridge et al., 2010). However, the widespread testing of relationships between genetic diversity and body size, species biology, population history, environmental conditions and phylogenetic relationship within marsupials must await a more comprehensive and representative data set.

22.3 Mitochondrial DNA

The advantages and limitations of mitochondrial DNA (mtDNA) for population and evolutionary studies have been extensively discussed (e.g., Moritz et al., 1987; Harrison, 1989; Moritz, 1994b). Being maternally inherited, mtDNA allows the impact of female-mediated population processes to be investigated over both the short-term (haplotype frequencies) and long-term (sequence divergence and phylogeny) (Sunnucks, 2000). Most population genetic studies that have utilised mtDNA have focussed on the hypervariable, non-coding control region or D-loop (Taberlet, 1996). In marsupials the 5′ end of this region (Domain 1) has typically been accessed using a set of conserved marsupial primers (developed by Fumagalli et al., 1997), with individual haplotypes being identified via direct sequencing or SSCP techniques (Sunnucks et al., 2000). Compared to microsatellite studies, relatively few marsupial species (n=43 vs. n=20) or populations (n=209 vs. n=50), have been examined for mtDNA control region variation, with most studies again focusing on populations of threatened or restricted range taxa in eastern or southern Australia rather than widespread or common species (Table 22.4). The relatively small number of population genetic studies utilizing mtDNA may be related

to the greater difficulty and expense of assessing variation through sequencing a DNA fragment compared with the ease of high throughput microsatellite genotyping.

Table 22.4 Haplotype diversity (*h*) (Nei, 1987) for mtDNA control region in 50 populations of 20 marsupial species

Taxon	*N*	No. of haplotypes	*h*	References
[T]*Setonix brachyurus*[#]	21	21	1.00	Sinclair (2001)
Macropus rufus[#]	34	31	0.99	Clegg et al. (1998)
[T]*Macrotis lagotis*[#]	30	20	0.95	Moritz et al. (1997)
Macropus giganteus (South Bathurst)	35	8	0.91	Zenger et al. (2003a)
Isoodon macrourus (North Qld)[#]	55	18	–	Pope et al. (2001)
[T*]*Myrmecobius fasciatus*[#]	22	11	0.90	Fumagalli et al. (1999)
[T]*Gymnobelideus leadbeateri* (Lake Mountain)	71	12	0.87[a]	Hansen et al. (2009)
Antechinus flavipes (Chilton)	115	10	0.83	Lada et al. (2008b)
Antechinus flavipes (Koondrook-Gunbower)	235	13	0.81	Lada et al. (2008a)
[T*]*Parantechinus apicalis*[#]	16	8	–	Mills et al. (2004)
Macropus giganteus (Marulan)	15	7	0.76	Zenger et al. (2003a)
Antechinus flavipes (Millewa-Barmah)	54	9	0.75	Lada et al. (2008b)
Antechinus flavipes (Campbells-Guttrum)	35	8	0.75	Lada et al. (2008a)
Antechinus flavipes (Warby-Ovens)	29	7	0.74	Lada et al. (2008b)
[T]*Petrogale penicillata* (Hurdle Cr)	39	5	0.72	Hazlitt et al. (2006)
Macropus giganteus (Tidbinbilla)	48	7	0.71	Zenger et al. (2003a)
[T*]*Macrotis lagotis* (Astrebla Downs NP)	19	4	0.67[a]	Smith et al. (2009)
Petrogale xanthopus celeris[#]	29	7	0.66	Pope et al. (1996)
[T]*Perameles bougainville* (Dorre Island)	27	3	0.66	Smith and Hughes (2008)
[T*]*Isoodon obesulus* (Sydney)	30	3	0.66	Zenger et al. (2005)
Antechinus flavipes (Rushworth-Reedy)	52	6	0.65	Lada et al. (2008b)
Antechinus agilis (Buccleuch SF)	170	13	0.63	Banks et al. (2005a)
[T]*Bettongia tropica* (Emu Creek)	20	4	0.63	Pope et al. (2000)
Macropus giganteus (Breadalbane)	29	5	0.60	Zenger et al. (2003a)
Phascolarctus cinereus (Mutdapilly)	16	3	0.59	Houlden et al. (1999)
[T*]*Macrotis lagotis* (Charleville)[b]	30	3	0.58[a]	Smith et al. (2009)
[T]*Dasyurus maculatus* (Chichester SF)	16	2	0.50	Firestone et al. (1999)
Phascolarctus cinereus (Iluka)	18	2	0.50	Houlden et al. (1999)
Phascolarctus cinereus (Coonabarabran)	22	4	0.45	Houlden et al. (1999)

Table 22.4 (continued)

Taxon	N	No. of haplotypes	h	References
[T]*Isoodon obesulus* (East Gippsland)	22	4	0.44	Zenger et al. (2005)
[T]*Isoodon obesulus* (Mt Gambier)	22	2	0.40	Zenger et al. (2005)
Dendrolagus lumholtzi[#]	45	3	0.31	Bowyer et al. (2002)
[T*]*Potorous gilbertii* (Two Peoples Bay)	23	3	0.24	Sinclair et al. (2002)
[T]*Perameles bougainville* (Bernier Island)	38	3	0.34	Smith and Hughes (2008)
[*]*Macropus rufogriseus* (NZ)[c]	40	3	–	Le Page et al. (2000)
[T*]*Perameles bougainville* (Dryandra)[b]	37	2	0.28	Smith and Hughes (2008)
[T]*Dasyurus maculatus* (Barrington Guest House)	21	2	0.26	Firestone et al. (1999)
Phascolarctus cinereus (South Gippsland)	19	3	0.20	Houlden et al. (1999)
[T*]*Gymnobelideus leadbeateri* (Yellingbo)	97	2	0.10[a]	Hansen et al. (2009)
[*]*Phascolarctus cinereus* (French Island)[c]	15	1	0.0	Houlden et al. (1999)
Phascolarctus cinereus (Nowendoc)	20	1	0.0	Houlden et al. (1999)
Phascolarctus cinereus (Gold Coast)	19	1	0.0	Houlden et al. (1999)
Isoodon obesulus (Lamb Range)	20	1	0.0	Pope et al. (2001)
[T*]*Perameles bougainville* (Heirisson)[c]	39	1	0.0	Smith and Hughes (2008)
[T]*Bettongia tropica* (Tinaroo)	20	1	0.0	Pope et al. (2000)
[T]*Petrogale penicillata* (Ingles Rd)	28	1	0.0	Eldridge et al. (2004b)
[T]*Petrogale penicillata* (Bowmans Rd)	20	1	0.0	Eldridge et al. (2004b)
[T]*Petrogale penicillata* (Jenolan 1985)[b]	20	1	0.0	Eldridge et al. (2004b)
[T*]*Petrogale penicillata* (Kawau Is., NZ)[c]	18	1	0.0	Eldridge et al. (2001a)
[T]*Petrogale penicillata* (Broke)	17	1	0.0	Eldridge et al. (2004b)

All data were obtained from the 5′ end of the control region by sequencing and/or SSCP techniques (Sunnucks et al., 2000)

Data are only presented for studies where $N \geq 15$. Populations are ordered by decreasing values of h

[T]Threatened taxon (EPBC Act 1999)

[*]Populations that have experienced a recent demographic bottleneck or a recent substantial range contraction (see Van Dyck and Strahan, 2008 or cited reference)

[#]Samples pooled from multiple localities

[a]Personal communication

[b]Captive population

[c]Population established via translocation

NZ = New Zealand

22.3.1 Genetic Diversity at the mtDNA Control Region

Levels of mtDNA control region diversity identified within the 50 marsupial populations examined to date vary considerably (Table 22.4). Haplotype diversity (h), varies across its full range from 0.0 to 1.0 (average ± se = 0.47±0.05); while the number of haplotypes identified ranged from 1 to 31 (average ± se = 5.8±0.9). These values are similar to those reported in populations of eutherians (Fernando et al., 2000; Nyakaana et al., 2002; Arrendal et al., 2004). The limited number of marsupial population studies that have utilized the control region currently limits detailed and conclusive analyses. Analysis is further compromised by the fact that in many studies (including the three reporting the highest values), samples were pooled from multiple localities (Table 22.4) which is likely to inflate the value of diversity parameters. Although trends in the levels of diversity appear broadly similar (Table 22.4) to those observed for the microsatellite data (Table 22.2), only differences between the number of haplotypes for bottlenecked vs. non-bottlenecked populations (t=2.2, df=42, p=0.03), as well as differences between h (t=2.5, df=13.1, p=0.03) for rock-wallabies vs. other macropodids were significant.

22.4 Managing Genetic Diversity

A priority of conservation genetics is to understand and address the impact of loss of genetic diversity and therefore evolutionary potential, increased inbreeding and reduced fitness, that are usually a consequence of isolation and small population size, which act to increase the likelihood of population extinction (Saccheri et al., 1998; Frankham et al., 2002; Briskie and Mackingtosh, 2004; Frankham, 2005). The loss of genetic diversity, the raw material of evolution, reduces a population's ability to evolve and adapt to changing environmental conditions (Frankham et al., 2002). Inbreeding (mating between relatives) increases homozygosity throughout the genome and so elevates the expression of deleterious recessive alleles. This typically results in the production of offspring with reduced fitness (ie, inbreeding depression) and so increases an inbred population's vulnerability to extinction (Frankham et al., 2002). It is important then that genetic factors are taken into account when managing marsupial populations, especially those that typically have reduced diversity (threatened species, bottlenecked, island, translocated and captive populations; Tables 22.2 and 22.4) and to avoid management strategies (limited founders, serial bottlenecking, lack of ongoing supplementation) that will further deplete variation. Koala (*Phascolarctos cinereus*) populations in southern Australia provide a good example of the consequences of poor management, with many populations established historically via small numbers of founders and serial bottlenecking now showing significantly reduced diversity, and some, high levels of inbreeding and reduced fitness (Houlden et al., 1996b; Seymour et al., 2001; Cristescu et al., 2009). For example, the Eyre Peninsula (SA) (Fig. 22.1) koala

population has an effective inbreeding coefficient of 0.75, and 24% of males show testicular aplasia (Seymour et al., 2001). Koalas (n=6) where introduced to Eyre Peninsula from Kangaroo Island (SA) in the 1960s. The Kangaroo Island population had itself been introduced in the 1920s using 18 individuals sourced from French Island (Vic), where a population had originally been established in the 1880s using 2–3 mainland animals as founders (Seymour et al., 2001). By contrast, the populations of northern quoll (*Dasyurus hallucatus*) recently introduced to Astell and Pobassoo Islands (NT) (Fig. 22.1) and the captive western barred bandicoot (*Perameles bougainville*) population established at Dryandra (WA) (Fig. 22.1), were established using founders from multiple wild populations and have resulted in new relatively diverse populations (Smith and Hughes, 2008; Cardoso et al., 2009).

The significantly reduced diversity observed in threatened marsupial species is most likely a direct consequence of their persistence at small effective population size for multiple generations (Spielman et al., 2004), but this is also likely to make population recovery more difficult (Evans and Sheldon, 2008) and increase their vulnerability to extinction in the future (Frankham et al., 2002). In some cases however, a lack of genetic diversity may be the cause of their endangerment. For example, the Tasmanian devil (*Sarcophilis harrisii*) provides an alarming and salutary example of a formerly abundant, but ultimately genetically compromised, population being rapidly brought to the edge of extinction by widespread susceptibility to a novel disease (McCallum et al., 2007; Siddle et al., 2007). In the late 1990s the Tasmanian devil was widespread and common in Tasmanian (McCallum, 2008) and regarded as secure (Maxwell et al., 1996). However, the emergence in 1996 of the invariably fatal contagious cancer DFTD (devil facial tumour disease; see Bender, Chapter 23), and its subsequent rapid spread, has decimated devil populations, with extinction in the wild now seeming likely (Jones et al., 2007). Although devils do have moderately low diversity at microsatellite loci (Jones et al., 2004) (Table 22.2), the spread of DFTD appears specifically linked to reduced diversity at MHC loci which prevent the devils' immune system from recognising the DFTD cells as foreign and so rejecting them (Siddle et al., 2007). This example highlights that the genetic quality of individuals can be more important in ensuring long term survival than simply the quantity of individuals, and that numerically numerous but genetically uniform populations are likely to be inherently vulnerable to extinction.

Although reduced diversity is also typical of bottlenecked marsupial populations, there are some notable exceptions (Tables 22.2 and 22.4), where bottlenecked population retain high levels of diversity (e.g., bridled nailtail wallaby, tammar wallaby, numbat (*Myrmecobius fasciatus*)). These exceptions illustrate that not all demographic bottlenecks are sufficiently severe or prolonged to result in a major loss of genetic diversity, as population genetic theory predicts (Frankham et al., 2002). Nevertheless the overall pattern in the marsupial (Table 22.2) and other data (Briskie and Mackingtosh, 2004; Garner et al., 2005) is clear and it would remain prudent to pay close attention to the management of bottlenecked populations, many of which are already of conservation concern.

Also of conservation concern are island populations of Australian marsupials, especially those that currently preserve at least 10 marsupial species most of which

disappeared from the Australian mainland in the last 200 years (Eldridge, 1998; Burbidge, 1999; Abbott, 2000). Increasingly, these island populations are being used to source conservation initiatives including captive breeding, reintroductions and translocations (Christensen and Burrows, 1994; Maxwell et al., 1996; Finlayson and Moseby, 2004; Smith and Hughes, 2008). For example, since 1990 burrowing bettongs (*Bettongia lesueur*) from Bernier and Dorre Islands (WA) (Fig. 22.1) have been translocated to six sites in western and central Australia (Finlayson et al., 2010) in attempts to re-establish populations. In addition, threatened species are still being actively introduced to islands as a management strategy to established "back-up" populations isolated from mainland threatening processes such as introduced predators/competitors, changed fire regimes and poisoning by introduced cane toads (e.g., Gilbert's potoroo (*Potorous gilbertii*), northern quoll, golden bandicoot (*Isoodon auratus*), Van Dyck and Strahan (2008)). Although islands and island populations currently play a significant role in biodiversity conservation in Australia (and internationally, e.g., Lloyd and Powlesland, 1994; Daltry et al., 2001; Boessenkool et al., 2007; Cheke and Hume, 2008), it is becoming clear that the unique island environment (e.g., isolation, limited size, reduced ecological complexity etc.) can also negatively impact the long-term evolutionary potential of resident populations, through the erosion of genetic diversity (as a consequence of finite size and isolation Frankham et al., 2002), inbreeding (Frankham, 1998; Eldridge et al., 2004a), reduced fitness (Eldridge et al., 1999) and adaptations to the island environment (reduced dispersal and fecundity, predator naivety etc. (Grant, 1998)). These factors make most island populations inherently vulnerable to extinction, as history has powerfully demonstrated with island taxa being significantly over-represented in recorded extinctions (Frankham, 1998), and caution against an over reliance on islands and islands population for conservation (Eldridge, 1998).

The predicted pattern of reduced genetic diversity in island populations (Frankham, 1997) is strongly supported by the marsupial data, with reduction up to 90% compared to conspecific mainland populations being observed (Table 22.2). The only exceptions (Table 22.2) were some high density macropodid population inhabiting large (>100,000 ha), temperate, fertile islands (Kangaroo, Flinders, King Islands) where presumably large effective population sizes have enabled moderate to high levels of genetic diversity to be maintained. However for most island populations, erosion of diversity, increased inbreeding and an increased risk of extinction is inevitable and populations recently marooned on islands as a conservation strategy will, in the absence of gene flow, follow a similar trajectory. The same is true for any small isolated population, whether inhabiting an isolated patch of habitat in a fragmented landscape, a fenced enclosure or a captive colony, and makes the regular exchange of individuals or the reestablishment of gene flow a high priority and a major challenge for wildlife managers and conservation biologists.

For those marsupial species naturally found on multiple islands, the potential exists to recreate more genetically diverse and robust populations for use in conservation initiatives. For although individual islands retain little diversity, each island is likely to preserve a different subsets of mainland alleles (e.g., Eldridge et al., 1999; Cardoso et al., 2009; How et al., 2009), due to the random nature of genetic

drift. As a consequence, a population established with founders from multiple island populations will have significantly increased diversity (e.g., Smith and Hughes, 2008), while the addition of unrelated individuals is also likely to increase the fitness of those populations suffering inbreeding depression (Eldridge et al., 1999). as has been reported in a wide variety of species (Saccheri et al., 1996; Westemeier et al., 1998; Madsen et al., 1999; Vila et al., 2003; Pimm et al., 2006).

22.5 Genetics Informing Management

As well as assessing a population's "genetic health", molecular genetics techniques, especially when coupled with detailed ecological, behavioural or geographic data, can also be harnessed as a powerful tool to inform the management of marsupials by increasing knowledge of their often cryptic behaviour ecology, reproduction and evolutionary history. For South American marsupials most recent population genetics studies have explored phylogeography (Steiner and Catzeflis, 2004; Himes et al., 2008), as well as clarifying taxonomy, species identification and species boundaries (Lavergne et al., 1997; Patton and Costa, 2003; Lavergne et al., 2010). For Australian species population genetics studies have addressed a broader array of questions (Table 22.5) and have fundamentally transformed our understanding of many aspects of species biology and greatly enhanced the effective management of populations and entire species.

An aim of many molecular studies has been the use of genetic markers to determine the most appropriate population units and geographic scale for within species management. These are usually based around the concept of management units (MUs) and evolutionary significant units (ESUs) of Moritz (1994a) which seek to ensure that major evolutionary lineages and unique diversity within species is identified and managed appropriately. While for some marsupials the presence or absence of major contemporary or historic barriers to gene flow is predictable (e.g., Houlden et al., 1999; Pope et al., 2000; Spencer et al., 2001; Brown et al., 2006) for many it is not (e.g., Cooper et al., 2000; Neaves et al., 2009) and significant unexpected divergence or lack of divergence is regularly reported. For example, the Tasmanian population of the spot-tailed quoll (*Dasyurus maculatus*) unexpectedly shows significant genetic divergence from the mainland population (Firestone et al., 1999), while the in the eastern grey kangaroo (*Macropus giganteus*) the reverse was reported, with the difference between the Tasmania and mainland populations being trivial, when it was expected to be significant (Zenger et al., 2003a). In south-eastern Australia, significant divergence was surprisingly identified within the brush-tailed rock-wallaby (*Petrogale penicillata*) with a major discontinuity, suggestive of long term isolation, present in mtDNA haplotypes south of Sydney, NSW (Eldridge et al., 2001a). Intriguingly in this same area the distributions of sibling species of *Antechinus* and *Trichosurus* abut (Dickman et al., 1998; Lindenmayer et al., 2002), suggesting that a major (and previously unknown) barrier to gene flow was present in this area millennia ago, resulting in substantial divergence and genetic

Table 22.5 Examples of conservation applications of genetic studies in marsupials

Aspects addressed	References
Establishing identity	
Taxonomy	Eldridge et al. (2001c)
Species boundaries	Patton and Costa (2003)
Species distribution	Johnson et al. (2001), Pope et al. (2001)
Cryptic species	Zenger et al. (2003b)
Scats	McGreevy et al. (2010)
Predators	Banks et al. (2003a), Glen et al. (2010)
Provenance of populations	Eldridge et al. (2001a), Hansen and Taylor (2008)
Increased knowledge of species biology	
Dispersal	Pope et al. (2005), Paplinska et al. (2009)
Recolonisation	Eldridge et al. (2001b)
Mating system	Taylor et al. (2000), Glen et al. (2009)
Social structure	Banks et al. (2002b), Martin et al. (2007b)
Behaviour	Walker et al. (2006), Hazlitt et al. (2010)
Significance of phenotypic variation	Beckman et al. (2007)
Impact of contemporary landscape features	Lada et al. (2008a), Neaves et al. (2009)
Influence of environmental variables	Lada et al. (2007), Walker et al. (2007)
Impact of ancient climatic fluctuations	Pope et al. (2000), Bowyer et al. (2002)
Impact of long-term natural isolation	Eldridge et al. (2004a)
Species long-term trajectory	Zenger et al. (2003a), Hansen et al. (2009)
Detection of hybridisation	Neaves et al. (2010)
Disease susceptibility	Siddle et al. (2007)
Management tools	
Population census	Banks et al. (2003b), Piggott et al. (2006a)
Detect species presence	Alacs et al. (2003)
Construct pedigree	Ivy et al. (2009)
Prioritise populations for management	Eldridge et al. (2004a), Hazlitt et al. (2006)
Determine geographic scale for management	Brown et al. (2006), Neaves et al. (2009)
Monitor reintroduced population	Sigg (2006), Smith et al. (2009)
Increased knowledge of anthropogenic impact	
Forestry operations	Banks et al. (2005b), Taylor et al. (2007)
Habitat fragmentation	Lada et al. (2008b), Walker et al. (2008)
Roads	Martin et al. (2007a)
Inbreeding	Seymour et al. (2001)
Bottlenecks	Taylor et al. (1994), Eldridge et al. (2004b)

differentiation in a range of broadly sympatric species. Additional studies of other south-eastern Australia species are currently underway to better assess the impact and nature of this putative ancient barrier. Despite the advances in knowledge that have been made, these unexpected findings highlight our current fundamental lack of understanding of the interactions between species biology and evolutionary history that have so profoundly shaped the distribution of genetic diversity within many marsupial species and continues to challenge and intrigue wildlife biologists and managers.

22.6 Future Directions

Population genetic studies utilising microsatellites or the mtDNA control region have so far been conducted on less than 15% of marsupial species. However, the potential of these hypervariable markers has been powerfully demonstrated as these studies have begun to transform our understanding of marsupial population biology, ecology, behaviour, reproduction and evolution. They have also been instrumental in formulating and guiding conservation initiatives and management plans for many threatened taxa. It is hoped that in the future, population genetic studies can be expanded to include a larger number of South American species as well as widespread and abundant Australian species, especially those in northern Australia that have been less severely impacted by European settlement. Another priority should be an examination of the unique marsupial fauna of New Guinea and the islands of eastern Indonesia which remains virtually unknown from a population genetics perspective (but see McGreevy, 2009).

The scale and scope of marsupial population genetic studies is becoming increasingly easy through the continuing development of novel genetic markers, methods of high through-put genotypic screening and sequencing and the increasing reliability of non-invasive sampling. In particular, the availability of the DNA sequence for the entire genome of two marsupials (see Papenfuss et al., Chapter 6) will dramatically increase the ease with which additional microsatellite loci, and other genetic markers (e.g., SNPs – single nucleotide polymorphisms), can be identified and developed. The availability of marsupial genome sequences and linkage maps (Deakin, Chapter 5; Samollow, Chapter 4) will also enable genetic markers to be targeted to link to specific genes or genome region of interest (e.g., MHC), allowing the detailed exploration of life history, behavioural and reproductive strategies (e.g., mate choice, kin recognition, sociality, disease/parasite susceptibility etc.). The recently developed X and Y chromosome-specific (MacDonald et al., 2006; 2007) and MHC-linked (Cheng et al., 2009a, b) microsatellite loci are examples of the evolving molecular tools available to researchers and managers. The scientific and conservation community can expect to see continued growth in the number of population and conservation genetics studies of marsupials and a corresponding increase in our detailed knowledge of the biology of these unique mammals.

Glossary

Allele Alternate versions of a locus that differ in DNA sequence.

Allelic diversity (*A*) A standard measure of population genetic diversity: the average number of alleles per locus.

Allozyme Different forms of a blood or tissue enzyme/protein detected by electrophoresis.

Average heterozygosity (*He*) A standard measure of population genetic diversity: the sum of heterozygosities at all loci, divided by the number of loci examined. Range 0–1.

Codominant Where both alleles at a locus are expressed.

Control region or D-loop A non-coding segment of the circular mitochondrial DNA molecular that is the initiation/regulation site for replication. Commonly used in population genetic studies because its DNA sequence is typically highly variable.

Demographic bottleneck A major reduction in population size.

Effective inbreeding coefficient (*Fe*) An indirect estimate of the inbreeding coefficient (*F*) obtained by measuring the loss of genetic diversity over time. Range 0-1.

Evolutionary potential The genetic diversity present within a population that enables it to evolve in response to environmental changes.

Founders The individuals used to establish a new population.

Genetic bottleneck A major reduction in population size that results in a significant loss of genetic diversity.

Genetic drift Changes in allele frequency, including the loss of diversity, due to the effects of random sampling from one generation to the next in small populations.

Haplotype diversity (*h*) A standard measure of genetic diversity in mitochondrial DNA studies: the probability that two randomly selected individuals will have different haplotypes. Range 0–1.

Haplotype A haploid genotype; the allelic combination at different loci on a single chromosome.

Heterozygote A diploid individual having two copies of the same allele at a locus.

Homozygote A diploid individual having two different alleles at a locus.

Inbreeding depression A reduction in fitness due to inbreeding.

Inbreeding Mating between related individuals.

Locus A specific gene or location on a chromosome.

Non-coding A sequence of DNA that does not code for a protein.

Non-invasive sampling Obtaining a genetic sample without capture and handling (e.g., using scats, hairs etc.).

Population structure The differences in genetic composition amongst populations within a species as the consequence of restrictions/limitations to gene flow.

Recessive An allele whose phenotype is not expressed when heterozygous with a dominant allele.

Selectively neutral An allele that is not subject to strong natural selection.

Serial bottleneck A sequence of multiple independent demographic bottlenecks.

SSCP (single-strand conformation polymorphism) A method to detect mutations in a DNA sequence through differences in the electrophoretic mobility of denatured DNA, which arises from the different folding pattern of the individual strands.

Supplementation The addition of unrelated individuals into an existing population.

Sympatric Species/populations that have overlapping distributions.

Testicular aplasia A failure in testicular development leading to the absence of one (unilateral) or both (bilateral) mature testes from the scrotum.

References

Abbott I (2000) Improving the conservation of threatened and rare mammal species through translocation to islands: case study Western Australia. Biol Conserv 93:195–201.

Alacs E, Alpers D, de Tores PJ, Dillon MJ, Spencer PBS (2003) Identifying the presence of quokkas (*Setonix brachyurus*) and other macropods using cytochrome *b* analyses from faeces. Wildl Res 30:41–47.

Alpers DL, Taylor AC, Sherwin WB (1998) Genetic structure of populations of the southern hairy-nosed wombat *Lasiorhinus latifrons*. In: Wells RT, Pridmore PA (eds) Wombats. Surrey Beatty and Sons, Sydney, pp. 192–197.

Arrendal J, Walker CW, Sundqvist A-K, Hellborg L, Vila C (2004) Genetic evaluation of an otter program. Conserv Genet 5:79–88.

Avise JC (1994) Molecular Markers, Natural History and Evolution. Chapman and Hall, New York and London.

Banks SC, Piggott MP, Hansen BD, Robinson NA, Taylor AC (2002a) Wombat coprogenetics: enumerating a common wombat population by microsatellite analysis of faecal DNA. Aust J Zool 50:193–204.

Banks SC, Skerratt LF, Taylor AC (2002b) Female dispersal and relatedness structure in common wombats (*Vombatus ursinus*) J Zool Lond 256:389–399.

Banks SC, Horsup A, Wilton AN, Taylor AC (2003a) Genetic marker investigation of the source and impact of predation on a highly endangered mammal. Mol Ecol 12:192–197.

Banks SC, Hoyle SD, Horsup A, Sunnucks P, Taylor AC (2003b) Demographic monitoring of an entire species (the northern hairy-nosed wombat, *Lasiorhinus krefftii*) by genetic analysis of non-invasively collected material. Anim Conserv 6:101–107.

Banks SC, Finlayson GR, Lawson SJ, et al. (2005a) The effects of habitat fragmentation due to forestry plantation establishment on the demography and genetic variation of a marsupial carnivore, *Antechinus agilis*. Biol Conserv 122:581–597.

Banks SC, Ward SJ, Lindenmayer DB, Finlayson GR, Lawson SJ, Taylor AC (2005b) The effects of habitat fragmentation on the social kin structure and mating system of the agile antechinus, *Antechinus agilis*. Mol Ecol 14:1789–1801.

Baverstock PR, Adams M, Archer M (1984) Electrophoretic resolution of species boundaries in the *Sminthopsis murina* complex (Dasyuridae). Aust J Zool 32:823–832.

Beckman J, Banks SC, Sunnucks P, Lill A, Taylor AC (2007) Phylogeography and environmental correlates of a cap on reproduction: teat number in a small marsupial, *Antechinus agilis*. Mol Ecol 16:1069–1083.

Beheregaray LB, Sunnucks P, Alpers D, Banks SC, Taylor AC (2000) A set of microsatellite loci for the hairy-nosed wombats (*Lasiorhinus krefftii* and *L. latifrons*). Conserv Genet 1: 89–92.

Boessenkool S, Taylor SS, Tepolt CK, Komdeur J, Jamieson IG (2007) Large mainland populations of South Island robins retain greater genetic diversity than offshore island refuges. Conserv Genet 8:705–714.

Bowyer JC, Newell GR, Eldridge MDB (2002) Genetic effects of habitat contraction on Lumholtz's tree-kangaroo (*Dendrolagus lumholtzi*) in the Australian Wet Tropics. Conserv Genet 3:61–69.

Briskie JV, Mackingtosh M (2004) Hatching failure increases with severity of population bottlenecks in birds. Proc Natl Acad Sci USA 101:558–561.

Brown M, Kendal TA, Cooksley H, et al. (2004) Polymorphic microsatellite markers for the gliding marsupials *Petaurus australis* and *Petaurus breviceps*. Mol Ecol Notes 4:704–706.

Brown M, Cooksley H, Carthew SM, Cooper SJB (2006) Conservation units and phylogeographic structure of an arboreal marsupial, the yellow-bellied glider (*Petaurus australis*). Aust J Zool 54:305–317.

Browning TL, Taggart DA, Rummery C, Close RL, Eldridge MDB (2001) Multifaceted genetic analysis of the "Critically Endangered" brush-tailed rock-wallaby *Petrogale penicillata* in Victoria, Australia: implications for management. Conserv Genet 2:145–156.

Bruford MW, Cheesman DJ, Coote T, et al. (1996) Microsatellites and their application to conservation genetics. In: Smith TB, Wayne RK (eds) Molecular Genetic Approaches in Conservation. Oxford University Press, New York, pp. 278–297.

Burbidge AA (1999) Conservation values and management of Australian islands for non-volant mammal conservation. Aust Mammal 21:67–74.

Cardoso M, Eldridge MDB, Oakwood M, Rankmore B, Sherwin WB, Firestone KB (2009) Effects of founder events on the genetic variation of translocated island populations: implications for conservation management of the northern quoll. Conserv Genet 10:1719–1733.

Cheke A, Hume J (eds) (2008) Lost Land of the Dodo. An Ecological History of Mauritius, Reunion and Rodrigues. T & A Poyser, London.

Cheng Y, Siddle HV, Beck S, Eldridge MDB, Belov K (2009a) High levels of variation at MHC class II DBB loci in the tammar wallaby (*Macropus eugenii*). Immunogenetics 61: 111–118.

Cheng Y, Wakefield MJ, Siddle HV, et al. (2009b) Isolation and characterization of 10 MHC Class I-associated microsatellite loci in tammar wallaby (*Macropus eugenii*). Mol Ecol Resour 9:346–349.

Christensen P, Burrows N (1994) Project Desert Dreaming: an experimental reintroduction of mammals to the Gibson Desert, Western Australia. In: Serena M (ed) Reintroduction Biology of Australian and New Zealand Fauna. Surrey Beatty and Sons, Sydney, pp. 199–207.

Clegg SM, Hale P, Moritz C (1998) Molecular population genetics of the red kangaroo (*Macropus rufus*): mtDNA variation. Mol Ecol 7:679–686.

Clinchy M, Taylor AC, Zanette LY, Krebs CJ, Jarman PJ (2004) Body size, age and paternity in common brushtail possums (*Trichosurus vulpecula*). Mol Ecol 13:195–202.

Cooper SJB, Adams M, Labrinidis A (2000) Phylogeography of the Australian dunnart *Sminthopsis crassicaudata* (Marsupialia: Dasyuridae). Aust J Zool 48:461–473.

Cristescu R, Cahill V, Sherwin WB, et al. (2009) Inbreeding and testicular abnormalities in a bottlenecked population of (*Phascolarctos cinereus*). Wildl Res 36:299–308.

Daltry JC, Blooxam Q, Cooper G, et al. (2001) Five years of conserving the 'world's rarest snake', the Antiguan racer *Alsophis antiguae*. Oryx 35:119–127.

DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. J Fish Biol 56:461–473.

Dias IMG, Amato G, Carvalho MRS, et al. (2008) Characterization of eight microsatellite loci in the woolly mouse opossum *Micoureus paraguayanus,* isolated from *Micoureus demerare*. Mol Ecol Resour 8:345–347.

Dickman CR, King DH, Adams M, Baverstock PR (1988) Electrophoretic identification of a new species of *Antechinus* (Marsupialia: Dasyuridae) in south-eastern Australia. Aust J Zool 36:455–463.

Dickman CR, Parnaby HE, Crowther MS, King DH (1998) *Antechinus agilis* (Marsupialia: Dasyuridae), a new species from the *A. stuartii* complex in south-eastern Australia. Aust J Zool 46:1–26.

Donaldson FR, Vercoe PE (2008) Cross-family amplification: microsatellites isolated from Macropodidae are polymorphic in Potoroidae. Mol Ecol Resour 8:452–454.

Eldridge MDB (1998) Trouble in paradise? Nat Aust 26:24–31.

Eldridge MDB, King JM, Loupis AK, et al. (1999) Unprecedented low levels of genetic variation and inbreeding depression in an island population of the black-footed rock-wallaby. Conserv Biol 13:531–541.

Eldridge MDB, Browning TL, Close RL (2001a) Provenance of a New Zealand brush-tailed rock-wallaby (*Petrogale penicillata*) population determined by mitochondrial DNA sequence analysis. Mol Ecol 10:2561–2567.

Eldridge MDB, Kinnear JE, Onus ML (2001b) Source population of dispersing rock-wallabies (*Petrogale lateralis*) identified by assignment tests on multilocus genotypic data. Mol Ecol 10:2867–2876.

Eldridge MDB, Wilson AC, Metcalfe CJ, et al. (2001c) Taxonomy of rock-wallabies, *Petrogale* (Marsupialia: Macropodidae). III. Molecular data confirms the species status of the purple-necked rock-wallaby *Petrogale purpureicollis* Le Souef 1924. Aust J Zool 49:323–343.

Eldridge MDB, Kinnear JE, Zenger KR, McKenzie LM, Spencer PBS (2004a) Genetic diversity in remnant mainland and 'pristine' island populations of three endemic Australian macropodids (Marsupialia): *Macropus eugenii, Lagorchestes hirsutus* and *Petrogale lateralis*. Conserv Genet 5:325–338.

Eldridge MDB, Rummery C, Bray C, Zenger K, Browning TL, Close RL (2004b) Genetic consequences of a population crash in brush-tailed rock-wallabies (*Petrogale penicillata*) from Jenolan Caves, south-eastern Australia. Wildl Res 31:229–240.

Eldridge MDB, Piggott MP, Hazlitt SL (2010) Population genetic studies of the Macropodoidea: a review. In: Coulson G, Eldridge MDB (eds) Macropods: The Biology of Kangaroos, Wallabies and Rat-Kangaroos. CSIRO Publishing, Melbourne, pp. 35–51.

Evans SR, Sheldon BC (2008) Interspecific patterns of genetic diversity in birds: correlations with extinction risk. Conserv Biol 22:1016–1025.

Excoffier L, Heckel G (2006) Computer programs for population genetics data analysis. Nat Rev Genet 7:745–748.

Fernando P, Pfrender ME, Encalada SE, Lande R (2000) Mitochondrial DNA variation, phylogeography and population structure of the Asian elephant. Heredity 84:362–372.

Fike JA, Beasley JC, Rhodes OE Jr (2009) Isolation of 21 polymorphic microsatellite markers for the Virginia opossum (*Didelphis virginiana*). Mol Ecol Resour 9:1200–1202.

Finlayson GR, Moseby KE (2004) Managing confined populations: the influence of density on the home range and habitat use of burrowing bettongs (*Bettongia lesueur*). Wildl Res 31:457–463.

Finlayson GR, Finlayson ST, Dickman CR (2010) Returning the rat-kangraoos: translocation attempts in the family Potoroidae (superfamily Macropodoidea) and recommendations for conservation. In: Coulson G, Eldridge MDB (eds) Macropods: The Biology of Kangaroos, Wallabies and Rat-Kangaroos. CSIRO Publishing, Melbourne, pp. 245–262.

Firestone KB (1999) Isolation and characterization of microsatellites from carnivorous marsupials (Dasyuridae: Marsupialia). Mol Ecol 8:1084–1086.

Firestone KB, Elphinstone MS, Sherwin WB, Houlden BA (1999) Phylogeographical population structure of tiger quolls *Dasyurus maculatus* (Dasyuridae: Marsupialia), an endangered carnivorous marsupial. Mol Ecol 8:1613–1625.

Firestone KB, Houlden BA, Sherwin WB, Geffen E (2000) Variability and differentiation of microsatellites in the genus *Dasyurus* and conservation implications for large Australian carnivorous marsupials. Conserv Genet 1:115–133.

Frankham R (1997) Do island populations have less genetic variation than mainland populations? Heredity 78:311–327.

Frankham R (1998) Inbreeding and extinction: island populations. Conserv Biol 12:665–675.

Frankham R (2005) Genetics and extinction. Biol Conserv 126:131–140.

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to Conservation Genetics. Cambridge University Press, Cambridge.

Fumagalli L, Pope LC, Taberlet P, Moritz C (1997) Versatile primers for the amplification of the mitochondrial DNA control region in marsupials. Mol Ecol 6:1199–1201.

Fumagalli L, Moritz C, Taberlet P, Friend JA (1999) Mitochondrial DNA sequence variation within the remnant populations of the endangered (Marsupialia: Myrmecobiidae: *Myrmecobius fasciatus*). Mol Ecol 8:1545–1549.

Garner A, Rachlow JL, Hicks JF (2005) Patterns of genetic diversity and its loss in mammalian populations. Conserv Biol 19:1215–1221.

Glen AS, Berry O, Sutherland DR, Garretson S, Robinson TJ, de Tores PJ (2010) Forensic DNA confirms intraguild killing of a chuditch (*Dasyurus geoffroii*) by a feral cat (*Felis catus*). Conserv Genet 11 (in press).

Glen AS, Cardoso MJ, Dickman CR, Firestone KB (2009) Who's your daddy? Paternity testing reveals promiscuity and multiple paternity in the carnivorous marsupial *Dasyurus maculatus* (Marsupialia: Dasyuridae). Biol J Linn Soc 96:1–7.

Gouin N, Westenberger SJ, Mahaney SM, Lindley P, VandeBerg JL, Samollow PB (2005a) Development, inheritance, and linkage-group assignment of 60 novel microsatellite markers for the gray short-tailed opossum *Monodelphis domestica*. Genome 48:1019–1027.

Gouin N, Westenberger SJ, Mahaney SM, Samollow PB (2005b) Isolation and characterization of polymorphic microsatellite markers in the gray short-tailed opossum (*Monodelphis domestica*). Mol Ecol Notes 5:87–89.

Grant PR (1998) Patterns on islands and microevolution. In: Grant PR (ed) Evolution on Islands. Oxford University Press, Oxford, pp. 1–17.

Guillemin M-L, Lavergne A, Catzeflis FM (2000) Microsatellite markers in the common grey four-eyed opossum (*Philander opossum*: Didelphidae, Marsupialia). Mol Ecol 9:1440–1442.

Hansen BD, Taylor AC (2008) Isolated remnant or recent introduction? Estimating the provenance of Yellingbo Leadbeater's possums by genetic analysis and bottleneck simulation. Mol Ecol 17:4039–4052.

Hansen BD, French J, Handasyde KA, Kendal TA, Taylor AC (2003) A set of microsatellite primers for the striped possum, *Dactylopsila trivirgata* (Petauridae: Marsupialia). Mol Ecol Notes 3:212–214.

Hansen BD, Sunnucks P, Blacket M, Taylor AC (2005) A set of microsatellite markers for an endangered arboreal marsupial Leadbeater's possum. Mol Ecol Notes 5:796–799.

Hansen BD, Harley DKP, Lindenmayer DB, Taylor AC (2009) Population genetic analysis reveals a long-term decline of a threatened endemic Australian marsupial. Mol Ecol 18:3346–3362.

Harrison RG (1989) Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol Evol 4:6–11.

Hawken RJ, Maccarone P, Toder R, Graves JAM, Maddox JF (1999) Isolation and characterisation of marsupial IL5 genes. Immunogenetics 49:942–948.

Hazlitt SL, Goldizen AW, Eldridge MDB (2006) Significant patterns of population genetic structure and limited gene flow in a threatened macropodid marsupial despite continuous habitat in southeast Queensland, Australia. Conserv Genet 7:675–689.

Hazlitt SL, Eldridge MDB, Goldizen AW (2010) Strong matrilineal structuring in the brush-tailed rock-wallaby (*Petrogale penicillata*) confirmed by spatial patterns of mtDNA. In: Coulson G, Eldridge MDB (eds) Macropods: The Biology of Kangaroos, Wallabies and Rat-Kangaroos. CSIRO Publishing, Melbourne, pp. 87–95.

Heckenberg KA (1997) DNA studies of *Thylogale*: relationships and potential hybridisation of species and subspecies. Honours thesis, University of Western Sydney.

Himes CT, Gallardo MH, Kenagy GJ (2008) Historical biogeography and post-glacial recolonization of South American temperate rainforest by the relictual marsupial *Dromiciops gliroides*. J Biogeogr 35:1415–1424.

Holleley CE, Dickman CR, Crowther MS, Oldroyd BP (2006) Size breeds success: multiple paternity, multivariate selection and male semelparity in a small marsupial, *Antechinus stuartii*. Mol Ecol 15:3439–3448.

Houlden BA, England PA, Sherwin WB (1996a) Paternity exclusion in koalas using hypervariable microsatellites. J Hered 87:149–152.

Houlden BA, England PR, Taylor AC, Greville WD, Sherwin WB (1996b) Low genetic variability of the koala *Phascolarctos cinereus* in south-eastern Australia following a severe population bottleneck. Mol Ecol 5:269–281.

Houlden BA, Costello BH, Sharkey D, et al. (1999) Phylogeographic differentiation in the mitochondrial control region in the koala, *Phascolarctos cinereus* (Goldfuss 1817). Mol Ecol 8:999–1011.

How RA, Spencer PBS, Schmitt LH (2009) Island populations have high conservation value for northern Australia's top marsupial predator ahead of a threatening process. J Zool Lond 278:206–217.

Ivy JA, Miller A, Lacy RC, DeWoody JA (2009) Methods and prospects for using molecular data in captive breeding programs: an empirical example using parma wallabies (*Macropus parma*). J Hered 100:441–454.

Johnson PM, Eldridge MDB, Kiernan V, Cupitt RJ (2001) A significant range extension of the purple-necked rock-wallaby *Petrogale purpureicollis*. Aust Mammal 23:71–73.

Johnson PM, Lloyd S, Vallance T, Eldridge MDB (2005) First record of quadruplets in the musky rat-kangaroo *Hypsiprymnodon moschatus*. Aust Mammal 27:95–97.

Johnston PG, Sharman GB (1979) Electrophoretic, chromosomal and morphometric studies on the red-necked wallaby *Macropus rufogriseus* (Desmarest). Aust J Zool 27:433–441.

Jones ME, Paetkau D, Geffen E, Moritz C (2003) Microsatellites for the Tasmanian devil (*Sarcophilus laniarius*). Mol Ecol Notes 3:277–279.

Jones ME, Paetkau D, Geffen E, Moritz C (2004) Genetic diversity and population structure of Tasmanian devils, the largest marsupial carnivore. Mol Ecol 13:2197–2209.

Jones ME, Jarman PJ, Lees CM, et al. (2007) Conservation management of Tasmanian devils in the context of an emerging, extinction-threatening disease: devil facial tumour disease. EcoHealth 4:326–337.

Kraaijeveld-Smit FJL, Lindenmayer DB, Taylor AC (2002a) Dispersal patterns and population structure in a small marsupial, *Antechinus agilis*, from two forests analysed using microsatellite markers. Aust J Zool 50:325–338.

Kraaijeveld-Smit FJL, Ward SJ, Temple-Smith PD (2002b) Multiple paternity in a field population of a small carnivorous marsupial, the agile antechinus, *Antechinus agilis*. Behav Ecol Sociobiol 52:84–91.

Kraaijeveld-Smit FJL, Ward SJ, Temple-Smith PD, Paetkau D (2002c) Factors influencing paternity success in *Antechinus agilis:* last-male sperm precedence, timing of mating and genetic compatibility. J Evol Biol 15:325–338.

Kraaijeveld-Smit FJL, Lindenmayer DB, Taylor AC, MacGregor C, Wertheim B (2007) Comparative genetic structure reflects underlying life histories of three sympatric small mammal species in continuous forest of south-eastern Australia. Oikos 116:1819–1830.

Lada H, Mac Nally R, Taylor AC (2007) Genetic reconstruction of the population dynamics of a carnivorous marsupial (*Antechinus flavipes*) in response to floods. Mol Ecol 16:2934–2947.

Lada H, Mac Nally R, Taylor AC (2008a) Distinguishing past from present gene flow along and across a river: the case of the carnivorous marsupial (*Antechinus flavipes*) on southern Australian floodplains. Conserv Genet 9:569–580.

Lada H, Thomson JR, Mac Nally R, Taylor AC (2008b) Impacts of massive landscape change on a carnivorous marsupial in south-eastern Australia: inferences from landscape genetics analysis. J Appl Ecol 45:1732–1741.

Lam M (2000) A major histocompatability (MHC) microsatellite locus in brushtail possums (*Trichosurus vulpecula*). Online J Vet Res 4:139–141.
Lancaster ML, Cooper SJB, Carthew SM, Taylor AC (2009) Microsatellite markers for the common ringtail possum (*Pseudocheirus peregrinus*) and their amplification in other psuedocheirids. Mol Ecol Resour 9:1535–1537.
Lavergne A, Verneau O, Patton JL, Catzeflis FM (1997) Molecular discrimination of two sympatric species of opossum (genus *Didelphis*: Didelphidae) in French Guiana. Mol Ecol 6:889–891.
Lavergne A, Douady C, Catzeflis FM (1999) Isolation and characterization of microsatellite loci in *Didelphis marsupialis* (Marsupialia: Didelphidae). Mol Ecol 8:517–518.
Lavergne A, Verneau O, Patton JL, Catzeflis FM (2010) Delineating geographic boundaries of the woolly mouse opossums, *Micoureus demerarae* and *Micoureus paraguayanus* (Didelphimorphia: Didelphidae). Conserv Genet 11 (in press).
Le Page SL, Livermore RA, Cooper DW, Taylor AC (2000) Genetic analysis of a documented population bottleneck: introduced Bennett's wallabies (*Macropus rufogriseus rufogriseus*) in New Zealand. Mol Ecol 9:753–763.
Lindenmayer DB, Dubach J, Viggers KL (2002) Geographic dimorphism in the mountain brushtail possum (*Trichosurus caninus*): the case for a new species. Aust J Zool 50:369–393.
Lloyd BD, Powlesland RG (1994) The decline of the kakapo *Strigops habroptilus* and attempts at conservation by translocation. Biol Conserv 69:75–85.
Luikart G, Painter J, Crozier RH, Westerman M, Sherwin WB (1997) Characterization of microsatellite loci in the endangered long-footed potoroo *Potorous longipes*. Mol Ecol 6: 497–498.
MacDonald AJ, Sankovic N, Sarre SD, et al. (2006) Y chromosome microsatellite markers identified from the tammar wallaby (*Macropus eugenii*) and their amplification in three other macropod species. Mol Ecol Notes 6:1202–1204.
MacDonald AJ, Sarre SD, Fitzsimmons NN, Marshall Graves JA (2007) Chromosome-specific microsatellites from the tammar wallaby X chromosome and chromosome 2. Mol Ecol Notes 7:1063–1066.
Macqueen P, Goldizen AW, Seddon JM (2009) Response of a southern temperate marsupial, the Tasmanian pademelon (*Thylogale billardierii*), to historical and contemporary forest fragmentation. Mol Ecol 18:3291–3306.
Madsen T, Shine R, Olsson M, Wittzell H (1999) Restoration of an inbred adder population. Nature 402:34–35.
Majoram P, Tavare S (2006) Modern computational approaches for analysing molecular genetic variation data. Nat Rev Genet 7:759–770.
Martin JK, Handasyde KA, Taylor AC (2007a) Linear roadside remnants: their influence on den-use, home range and mating system in bobucks (*Trichosurus cunninghami*). Austral Ecol 32:686–696.
Martin JK, Handasyde KA, Taylor AC, Coulson G (2007b) Long-term pair-bonds without mating fidelity in a mammal. Behaviour 144:1419–1445.
Maxwell S, Burbidge AA, Morris K (eds) (1996) The 1996 Action Plan for Australian Marsupials and Monotremes. Wildlife Australia, Canberra.
McCallum HI (2008) Tasmanian devil facial tumour disease: lessons for conservation biology. Trends Ecol Evol 23:631–637.
McCallum HI, Tompkins DM, Jones ME, et al. (2007) Distribution and impacts of Tasmanian devil facial tumour disease. EcoHealth 4:318–325.
McGreevy Jr, TJ (2009) Genetic diversity in captive and wild Matschie's tree-kangaroo (*Dendrolagus matschiei*) from Huon Peninsula, Papua New Guinea, based on mtDNA control region sequences. Zoo Biol 28:183–196.
McGreevy TJ Jr, Dabek L, Husband TP (2010) A multiplex PCR assay to distinguish among three sympatric marsupial taxa from Huon Peninsula, Papua New Guinea, using mitochondrial control region. Mol Ecol Resour 9:397–400.

Millis AL (2000) Isolation and characterization of microsatellite loci in marsupial gliders (*Petaurus norfolcensis*, *P. breviceps* and *P. gracilis*). Mol Ecol 9:1681–1683.

Mills HR, Spencer PBS (2003) Polymorphic microsatellites identified in an endangered dasyurid marsupial, the dibbler (*Parantechinus apicalis*). Mol Ecol Notes 3:218–220.

Mills HR, Moro D, Spencer PBS (2004) Conservation significance of island versus mainland populations: a case study of dibblers (*Parantechinus apicalis*) in Western Australia. Anim Conserv 7:387–395.

Mitrovski P, Heinze DA, Guthridge K, Weeks AR (2005) Isolation and characterization of microsatellite loci from the Australian endemic mountain pygmy-possum, *Burramys parvus* Broom. Mol Ecol Notes 5:395–397.

Mitrovski P, Heinze DA, Broome L, Hoffmann AA, Weeks AR (2007) High levels of variation despite genetic fragmentation in populations of the endangered mountain pygmy-posum, *Burramys parvus*, in alpine Australia. Mol Ecol 16:75–87.

Moritz C (1994a) Defining "Evolutionary Significant Units" for conservation. Trends Ecol Evol 9:373–375.

Moritz C (1994b) Applications of mitochondrial DNA analysis in conservation: a critical review. Mol Ecol 3:401–411.

Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu Rev Ecol Syst 18:269–292.

Moritz C, Heideman A, Geffen E, McRae P (1997) Genetic population structure of the greater bilby *Macrotis lagotis*, a marsupial in decline. Mol Ecol 6:925–936.

Neaves LE, Zenger KR, Cooper DW, Eldridge MDB (2010) Molecular detection of hybridization between sympatric kangaroo species in south-eastern Australia. Heredity (in press).

Neaves LE, Zenger KR, Prince RIT, Eldridge MDB, Cooper DW (2009) Landscape discontinuities influence gene flow and genetic structure in a large, vagile Australian mammal, *Macropus fuliginosus*. Mol Ecol 18:3363–3378.

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York.

Nyakaana S, Arctander P, Siegismund HR (2002) Population structure of the African savannah elephant inferred from mitochondrial control region sequences and nuclear microsatellite loci. Heredity 89:90–98.

Oliveira EJ, Padua JG, Zucchi MI, Vencovsky R, Vieira MLC (2006) Origin, evolution and genome distribution of microsatellites. Genet Mol Biol 29:294–307.

Paplinska JZ (2005) Seasonal breeding, puberty and population genetics of the swamp wallaby (*Wallabia bicolor*). Ph.D. thesis, University of Melbourne.

Paplinska JZ (2006) Final Report of Faecal DNA Analysis of *Petrogale penicillata* in Kangaroo Valley for NSW National Parks and Wildlife Service. New South Wales National Parks and Wildlife Service, Hurstville.

Paplinska JZ, Eldridge MDB, Cooper DW, Temple-Smith PD, Renfree MB (2009) Use of genetic methods to establish male-biased dispersal in a cryptic mammal, the swamp wallaby (*Wallabia bicolor*). Aust J Zool 57:65–72.

Patton JL, Costa LP (2003) Molecular phylogeography and species limits in rainforest didelphid marsupials of South America. In: Jones M, Dickman C, Archer M (eds) Predators with Pouches the Biology of Carnivorous Marsupials. CSIRO Publishing, Melbourne, pp. 63–81.

Piggott MP, Banks SC, Stone N, Banffy C, Taylor AC (2006a) Estimating population size of endangered brush-tailed rock-wallaby (*Petrogale penicillata*) colonies using faecal DNA. Mol Ecol 15:81–91.

Piggott MP, Banks SC, Taylor AC (2006b) Population structure of brush-tailed rock-wallaby (*Petrogale penicillata*) colonies inferred from analysis of faecal DNA. Mol Ecol 15: 93–105.

Pimm SL, Dollar L, Bass OL Jr (2006) The genetic rescue of the Florida panther. Anim Conserv 9:115–122.

Pope LC, Sharp A, Moritz C (1996) Population structure of the yellow-footed rock-wallaby *Petrogale xanthopus* (Gray, 1854) inferred from mtDNA sequences and microsatellite loci. Mol Ecol 5:629–640.

Pope LC, Estoup A, Moritz C (2000) Phylogeography and population structure of an ectonal marsupial, *Bettongia tropica,* determined from mtDNA and microsatellites. Mol Ecol 9: 2041–2053.

Pope LC, Storch D, Adams M, Moritz C, Gordon G (2001) A phylogeny for the genus *Isoodon* and a range extension for *Isoodon obesulus peninsulae* based on mtDNA control region and morphology. Aust J Zool 49:411–434.

Pope LC, Blair D, Johnson CN (2005) Dispersal and population structure of the rufous bettong, *Aepyprymnus rufescens* (Marsupialia: Potoroidae). Austral Ecol 30:572–580.

Rodrigues FP, Rocha FS, Garcia JE, Garcia JF, De Vivo M, Matioli SR (2006) Isolation and characterization of microsatellite loci in the woolly mouse opossum, *Micoureus paraguayanus* (Marsupialia: Didelphimorphia). Mol Ecol Resour 6:686–688.

Saccheri I, Brakefield PM, Nichols RA (1996) Severe inbreeding depression and rapid fitness rebound in the butterfly *Bicyclus anynana* (Satyridae). Evolution 50:2000–2013.

Saccheri I, Kuussaari M, Kankare M, Vikman P, Fortelius W, Hanski I (1998) Inbreeding and extinction in a butterfly metapopulation. Nature 392:491–494.

Sale MG, Kraaijeveld-Smit FJL, Arnould JPY (2009) Natal dispersal and social organization of the swamp antechinus (*Antechinus minimus*) in a high-density island population. Can J Zool 87:262–272.

Samollow PB, Kammerer CM, Mahaney SM, et al. (2004) First-generation linkage map of the gray, short-tailed opossum, *Monodelphis domestica*, reveals genome-wide reduction in female recombination rates. Genetics 166:307–329.

Samollow PB, Gouin N, Miethke P, et al. (2007) A microsatellite-based, physically anchored linkage map for the gray short-tailed opossum (*Monodelphis domestica*). Chromosome Res 15:269–281.

Seebeck JH, Johnston PG (1980) *Potorous longipes* (Marsupialia: Macropodidae): a new species from eastern Victoria. Aust J Zool 28:119–134.

Seymour AM, Montgomery ME, Costello BH, et al. (2001) High effective inbreeding coefficients correlate with morphological abnormalities in populations of South Australian koalas (*Phascolarctos cinereus*). Anim Conserv 4:211–219.

Sherwin WB, Murray ND (1990) Population and conservation genetics of marsupials. Aust J Zool 37:161–180.

Siddle HV, Kreiss A, Eldridge MDB, et al. (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA 104:16221–16226.

Sigg DP (2006) Reduced genetic diversity and significant genetic differentiation after translocation: comparison of the remnant and translocated populations of bridled nailtail wallabies (*Onychogalea fraenata*). Conserv Genet 7:577–589.

Sigg DP, Goldizen AW, Pople AR (2005) The importance of mating system in translocation programs: reproductive success of released male bridled nailtail wallabies. Biol Conserv 123:289–300.

Sinclair AH (2001) Phylogeographic variation in the quokka, *Setonix brachyurus* (Marsupialia: Macropodidae): implications for management. Anim Conserv 4:325–333.

Sinclair EA, Costello B, Courtenay JM, Crandall KA (2002) Detecting a genetic bottleneck in Gilbert's potoroo, *Potorous gilbertii* (Marsupialia: Potoroidae), inferred from microsatellite and mitochondrial DNA sequence data. Conserv Genet 3:191–196.

Smith S, Hughes J (2008) Microsatellite and mitochondrial DNA variation defines island genetic reservoirs for reintroductions of an endangered Australian marsupial, *Perameles bougainville*. Conserv Genet 9:547–557.

Smith S, McRae P, Hughes J (2009) Faecal DNA analysis enables genetic monitoring of the species recovery program for an arid-dwelling marsupial. Aust J Zool 57:139–148.

Spencer PBS (1996) Coping with a naturally fragmented environment: a genetic and ecological study of the allied rock-wallaby, *Petrogale assimilis*. James Cook University of North Queensland.

Spencer PBS, Bryant KA (2000) Characterization of highly polymorphic microsatellite markers in the marsupial honey possum (*Tarsipes rostrastus*). Mol Ecol 9:492–494.
Spencer PBS, Marsh HD (1997) Microsatellite DNA fingerprinting confirms dizygotic twinning and paternity in the Allied Rock-wallaby, *Petrogale assimilis* (Marsupialia: Macropodidae). Aust Mammal 19:279–280.
Spencer PBS, Odorico DM, Jones SJ, Marsh HD, Miller DJ (1995) Highly variable microsatellites in isolated colonies of the rock-wallaby (*Petrogale assimilis*). Mol Ecol 4:523–525.
Spencer PBS, Rhind SG, Eldridge MDB (2001) Phylogeographic structure within *Phascogale* (Marsupialia: Dasyuridae) based on partial cytochrome *b* sequence. Aust J Zool 49:369–377.
Spencer PBS, Fletcher TP, Woolley PA (2003) Microsatellite markers from the Julia Creek dunnart, *Sminthopsis douglasi* (Marsupialia: Dasyuridae). Mol Ecol Notes 3:570–571.
Spencer PBS, Cardoso M, How RA, Williams J, Bunce M, Schmitt LH (2007) Cross-species amplification at microsatellite loci in Australian quolls including the description of five new markers from the chuditch (*Dasyurus geoffroii*). Mol Ecol Notes 7:1100–1103.
Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. Proc Natl Acad Sci USA 101:15261–15264.
Steiner C, Catzeflis FM (2004) Genetic variation and geographical structure of five, mouse-sized opposums (Marsupialia, Didelphoidae) throughout the Guiana region. J Biogeogr 31: 959–973.
Stow AJ, Minarovic N, Eymann J, Cooper DW, Webley LS (2006) Genetic structure infer generally high philopatry and male-biased diapersal of brushtail possums (*Trichosurus vulpecula*) in urban Australia. Wildl Res 33:409–415.
Sunnucks P (2000) Efficient genetic markers for population biology. Trends Ecol Evol 15:199–203.
Sunnucks P, Wilson AC, Beheregaray LB, Zenger K, French J, Taylor AC (2000) SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. Mol Ecol 9:1699–1710.
Taberlet P (1996) The use of mitochondrial DNA control region sequencing in conservation genetics. In: Smith TB, Wayne RK (eds) Molecular Genetic Approaches in Conservation. Oxford University Press, New York, pp. 125–142.
Taylor AC, Cooper DW (1998a) Microsatellite markers for the Phalangerid marsupial, the common brushtail possum (*Trichosurus vulpecula*). Mol Ecol 7:1780–1782.
Taylor AC, Cooper DW (1998b) A set of tammar wallaby (*Macropus eugenii*) microsatellites tested for genetic linkage. Mol Ecol 7:925–926.
Taylor AC, Cooper DW (1999) Microsatellites identify introduced New Zealand tammar wallabies (*Macropus eugenii*) as an "extinct" taxon. Anim Conserv 2:41–49.
Taylor AC, Sherwin WB, Wayne RK (1994) Genetic variation of microsatellite loci in a bottle-necked species: the northern hairy-nosed wombat *Lasiorhinus krefftii*. Mol Ecol 3:277–290.
Taylor AC, Sunnucks P, Cooper DW (1999) Retention of reproductive barriers and ecological differences between two introduced sympatric *Macropus* spp. in New Zealand. Anim Conserv 2:195–202.
Taylor AC, Cowan PE, Fricke BL, Cooper DW (2000) Genetic analysis of the mating system of the common brushtail possum (*Trichosurus vulpecula*) in New Zealand farmland. Mol Ecol 9:869–879.
Taylor AC, Kraaijeveld K, Lindenmayer DB (2002) Microsatellites for the greater glider, *Petauroides volans*. Mol Ecol Notes 2:57–59.
Taylor AC, Cowan PE, Fricke BL, et al. (2004) High microsatellite diversity and differential structuring among populations of the introduced common brushtail possum, *Trichosurus vulpecula*, in New Zealand. Genet Res 83:101–111.
Taylor AC, Tyndale-Biscoe CH, Lindenmayer DB (2007) Unexpected persistence on habitat islands: genetic signatures reveal dispersal of a eucalypt-dependent marsupial through a hostile pine matrix. Mol Ecol 16:2655–2666.
Van Dyck S, Strahan R (eds) (2008) The Mammals of Australia, 3rd edn. Reed New Holland, Sydney.

Vila C, Sundqvist A-K, Flagstad O, et al. (2003) Rescue of a severely bottlenecked wolf (*Canis lupus*) population by a single immigrant. Proc R Soc Lond B 270:91–97.
Walker FM (2004) Sociobiology inferred from relatedness structure via remotely-collected DNA in Southern Hairy-nosed Wombats (*Lasiorhinus latifrons*). Ph.D. thesis, Monash University, Clayton.
Walker FM, Sunnucks P, Taylor AC (2006) Genotyping of "captured" hairs reveals burrow-use and ranging behavior of southern hairy-nosed wombats. J Mammal 87:690–699.
Walker FM, Taylor AC, Sunnucks P (2007) Does soil type drive social organization in southern hairy-nosed wombats. Mol Ecol 16:199–208.
Walker FM, Sunnucks P, Taylor AC (2008) Evidence for habitat fragmentation altering within-population processes in wombats. Mol Ecol 17:1674–1684.
Walker FM, Horsup A, Taylor AC (2009) Leader of the pack: faecal pellet deposition order impacts PCR amplification in wombats. Mol Ecol Resour 9:720–724.
Westemeier RL, Brawn JD, Simpson SA, et al. (1998) Tracking the long-term decline and recovery of an isolated population. Science 282:1695–1698.
Wilson DE, Reeder DM (eds) (2005) Mammal Species of the World: A Taxonomic and Geographic Reference, 3rd edn. Johns Hopkins University Press, Baltimore.
Wilson K, de Tores P, Spencer PBS (2009) Isolation and characterisation of polymorphic microsatellite markers in the western ringtail possum, *Pseudocheirus occidentalis*. Conservation Genetics Resources 1:123–125.
Zenger KR, Cooper DW (2001a) Characterization of 14 macropod microsatellite genetic markers. Anim Genet 32:160–167.
Zenger KR, Cooper DW (2001b) A set of highly polymorphic microsatellite markers developed for the eastern grey kangaroo (*Macropus giganteus*). Mol Ecol Notes 1:98–100.
Zenger KR, Johnston PG (2001) Isolation and characterisation of microsatellite loci in the southern brown bandicoot (*Isoodon obesulus*), and their applicability to other marsupial species. Mol Ecol Notes 1:149–151.
Zenger KR, McKenzie LM, Cooper DW (2002) The first comprehensive genetic linkage map for a marsupial: the tammar wallaby (*Macropus eugenii*). Genetics 162:321–330.
Zenger KR, Eldridge MDB, Cooper DW (2003a) Intraspecific variation, sex-biased dispersal and phylogeography of the eastern grey kangaroo (*Macropus giganteus*). Heredity 92:153–162.
Zenger KR, Eldridge MDB, Johnston PG (2003b) A rapid PCR-RFLP diagnostic test for distinguishing sympatric bandicoot species (Marsupialia: Peramelidae) in south-eastern Australia. Conserv Genet 4:655–657.
Zenger KR, Eldridge MDB, Pope LC, Cooper DW (2003c) Characterisation and cross-species utility of microsatellite markers within kangaroos, wallabies and rat kangaroo (Macropodoidae: Marsupialia). Aust J Zool 51:587–596.
Zenger KR, Eldridge MDB, Johnston PG (2005) Phylogenetics, population structure and genetic diversity of the endangered southern brown bandicoot (*Isoodon obesulus*) in south-eastern Australia. Conserv Genet 6:193–204.

Chapter 23
Devil Facial Tumour Disease (DFTD): Using Genetics and Genomics to Investigate Infectious Disease in an Endangered Marsupial

Hannah S. Bender

Abstract The Tasmanian devil (*Sarcophilus harrisii*), so named for its blood-curdling, nocturnal shrieks and snarls, is the largest of the carnivorous marsupials. Although once widely persecuted, concerted efforts are now being made to save the devil from extinction following the emergence of a fatal transmissible malignancy known as devil facial tumour disease (DFTD). DFTD is unusual in that the infectious agent is the cancer cell itself. This chapter discusses the aetiology and pathogenesis of DFTD as well as the profound impact the spread of DFTD has had on the devil's conservation status. Strategies for managing DFTD and conserving the devil will be explored and the contribution of new sequencing technology to the field of conservation genetics and genomics will be examined with regard to the Tasmanian devil and DFTD.

Keywords Tasmanian devil · Devil facial tumour disease · Wildlife disease · Conservation genetics · Genomics

23.1 The Tasmanian Devil and the Emergence of DFTD

The Tasmanian devil (*Sarcophilus harrisii*), is the largest of the Dasyuridae, a speciose and morphologically diverse family of carnivorous marsupials that are represented in most terrestrial habitats in Australia and New Guinea (Crowther and Blacket, 2003). *S. harrisii* diverged in the mid Miocene, approximately 12.4 million years ago, and originally thrived on the Australian mainland (Krajewski et al., 2000). Tasmanian populations were separated from their mainland conspecifics after the formation of the Bass Strait 14,000 years ago (Lambeck and Chappell, 2001). Now the devil is confined to the island state; their extinction on mainland Australia

H.S. Bender (✉)
Department of Evolution, Ecology and Genetics, Research School of Biology, The Australian National University, Canberra, ACT 0200, Australia
e-mail: hannah.bender@anu.edu.au

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2_23, © Springer Science+Business Media B.V. 2010

approximately 400 years ago was likely associated with the introduction and establishment of the dingo (Jones et al., 2003). Since the extinction of the devil's larger, sympatric relative, the thylacine (*Thylacinus cynocephalus*), the Tasmanian devil has held the mantle of World's Largest Remaining Marsupial Carnivore.

Tasmanian devils are non-territorial, nocturnal, hunter-scavengers. Adults typically inhabit wide ranges throughout open sclerophyll forest and coastal scrub and are found at low densities in alpine areas, extensively cleared agricultural land and dense forests (Guiler, 1970; Jones et al., 2004). Although occasionally predatory, devils are largely specialised scavengers and opportunistic feeders (Owen and Pemberton, 2005). Devils are a relatively short-lived species, living for 5–6 years in the wild. Mature devils are sexually dimorphic; males weigh up to 13 kg and females are considerably smaller at around 9 kg. Female devils are monoestrus with mating usually occurring in March. Four pouch young can be accommodated in the female's rear-facing pouch and weaning success is high in the wild. Joeys leave the pouch at about 105 days, are independent at 9 months and reach sexual maturity themselves at 2 years of age (Guiler, 1970).

There is anecdotal evidence that devil populations have suffered from several fluctuations in the 200 years since European settlement (Guiler, 1970). No reliable data exist to verify the extent or cause of these population fluxes, however infectious disease, the extinction of the thylacine and anthropogenic factors, such as habitat destruction and culling, have all been implicated (Jones et al., 2004; Bradshaw and Brook, 2005). Certainly both the devil and the thylacine were singled out as "stock-destroying vermin" and were the subject of intense persecution by graziers for over 100 years (Owen and Pemberton, 2005). Molecular marker profiling of devils across Tasmania has revealed poor population structuring and a distinct lack of genetic diversity at microsatellite and MHC loci, consistent with periods of low population density (Jones et al., 2004; Siddle et al., 2007). Prior to the emergence of DFTD, however, devil populations were robust and widespread across Tasmania with the exception of the southwest where population density was and remains low (Jones et al., 2004).

In 1996 a devil with a large facial lesion was photographed at Mt William in the state's northeast (Hawkins et al., 2006). Although striking, the tumour was not thought to be particularly unusual as devils seem to be particularly susceptible to neoplasia (Griner, 1979; Canfield et al., 1990). In subsequent years, however, an increasing number of devils from northeast Tasmania presented to the Department of Primary Industries and Water's Animal Health Laboratories with grossly similar facial tumours. As tumour cases appeared further afield it became apparent that the developing tumour epidemic was due to an infectious process.

23.2 DFTD Pathogenesis

DFTD is an aggressive, debilitating malignancy. Tumorigenesis is characterised by the proliferation of round to spindle-shaped cells in the dermis of the head and neck or the submucosal tissue of the oral cavity. Small nodules rapidly

Fig. 23.1 Gross pathology – DFTD lesions initially appear as small, well-circumscribed nodules arising in the dermis of the head and neck, or the submucosa of the oral cavity (**a**) and rapidly become locally invasive, often preventing affected animals from feeding (**b**). Metastasis to thoracic viscera is common. Image **c** depicts a shower of small, nodular metastases on the serosal surface of the lungs

progress to become large, multicentric masses with ulcerating, exudative surfaces (Fig. 23.1a, b). DFTD is locally invasive and highly metastatic with regional lymph node involvement and metastasis to thoracic and abdominal viscera occurring in 65% of cases (Loh et al., 2006a) (Fig. 23.1c). Disease progression is rapid and the mortality rate is 100%. Affected animals may die as little as 3 months after the first appearance of lesions (Hawkins et al., 2006). Mortality results when large primary tumours obstruct feeding, leading to starvation, or from complications associated with metastasis (Pyecroft et al., 2007). The typical histological appearance of DFTD, characterised by multinodular aggregates of undifferentiated round cells, prompted early suggestions that DFTD was a lymphosarcoma, possibly of viral aetiology (Ladds et al., 2003). Immunostaining subsequently revealed a neuroectodermal origin for the tumour and it was noted that DFTD shares gross pathological and immunophenotypic features with Merkel cell carcinoma of humans (Loh et al., 2006b). Sophisticated sequencing approaches to investigating DFTD recently elucidated the cell lineage of the tumour which was shown to express genes characteristic of Schwann cells or Schwann cell precursors (Murchison et al., 2010).

23.3 DFTD Cytogenetics: Evidence for Cellular Transmission

The genetic basis of neoplasia has been recognised for over 100 years, with cytogenetic observations providing the first insights into the clinical significance of chromosomal anomalies. In 1914 Theodor Boveri first suggested that an excessive number of chromosomes might be a cause of malignancy. Aneuploidy is now known to be one of the most common features of cancer, found in approximately 90% of solid tumours and 75% of haematological tumours in humans (Weaver and Cleveland, 2006). It remains unclear whether aneuploidy, when present, is a driver of carcinogenesis or whether it is simply an incidental finding, a background

against which other processes initiate and steer carcinogenesis. Evidence from naturally occurring tumours suggests that aneuploidy represents one of several types of genetic instability which renders cells susceptible to the accelerated acquisition of mutations, eventually leading to cancer development (Lengauer et al., 1997). The underlying cause of aneuploidy is also unclear although several causative gene mutations have been determined, many involving the mitotic spindle checkpoint (Rajagopalan and Lengauer, 2004). In addition to aberrations in chromosome number, cancer cells are often characterised by structural chromosome defects such as translocations, inversions, duplications or deletions. Recurrent chromosome rearrangements may even be pathognomonic for specific cancer types and in some cases cytogenetic techniques are crucial prognostic tools (Mitelman et al., 1997). One of the best characterised chromosome rearrangements is the Philadelphia (Ph) chromosome, formed by the reciprocal translocation of the q arms of chromosomes 9 and 22, found in 90% of cases of human chronic myelogenous leukaemia (CML) (Nowell and Hungerford, 1960; Deininger and Druker, 2003). The Ph chromosome was the first recurring chromosome rearrangement to be associated with a specific cancer type. Since this breakthrough additional haematological malignancies, as well as some solid tumours, have been shown to be caused by recurring chromosome rearrangements (Mitelman, 2000).

Tumour development is now known to be a dynamic process akin to Darwinian evolution. Cancer cells progressively acquire somatic mutations and those that afford a growth advantage are selected for and propagated in waves of clonal expansion (Nowell, 1976). This is evident in the heterogeneity of tumour cell populations, each neoplastic clone reflecting a new step in the progression from benign to malignant. Favourable mutations confer on cancer cells a proliferative advantage and are manifest as the physiological traits that define malignancy; namely acquired growth signal autonomy, insensitivity to anti-growth signals, evasion of apoptosis, limitless replicative lifespan, angiogenesis and metastatic capability (Hanahan and Weinberg, 2000).

The discipline of cancer cytogenetics has been crucial to our understanding of DFTD transmission and development. In early investigations of DFTD transmission, electron microscopy failed to detect evidence of viral involvement in DFTD carcinogenesis (Loh et al., 2006b). Not until cytogenetic investigations of DFTD were undertaken was it suggested that DFTD tumour cells were themselves the causative agent (Pearse and Swift, 2006). Diploid *S. harrisii* cells contain 12 autosomes and a pair of sex chromosomes; XX in female devils and XY in males (Fig. 23.2a). G-banding of metaphase spreads revealed DFTD cells to be aneuploid and marked by extensive chromosome rearrangements. Moreover, these chromosome rearrangements were identical in every animal studied. The DFTD karyotype is distinguished by the loss of both copies of chromosome 2, one copy of chromosome 6 and both sex chromosomes, and the addition of four unidentified marker chromosomes (Fig. 23.2b). This signature karyotype is found in all tumours regardless of the sex of the affected animal or the stage of DFTD progression. It is highly improbable that the same complex tumour karyotype independently arises from a recurring rearrangement in every affected animal. Even if it were possible, the

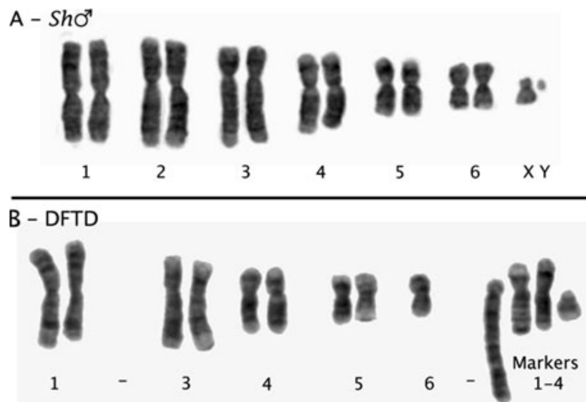


Fig. 23.2 Tumour cytogenetics – (**a**) The diploid Tasmanian devil karyotype consists of 14 chromosomes, including two sex chromosomes (X and Y in this male devil). (**b**) The DFTD karyotype (strain 1) is aneuploid and distinguished by stable chromosome rearrangements. Several chromosomes are rearranged, identifiable only as four marker chromosomes

stepwise accumulation of progressive chromosome aberrations should be evident within individual tumours and between animals. However, no intermediate stages of karyotypic evolution have been observed; all affected animals studied so far have a homogeneous tumour cell population that is more or less identical to the original G-banded tumour. Additional subtle chromosome rearrangements have been recognised in some animals and also appear to be stable and conserved (Department of Primary Industries, 2009). The clinical significance of these karyotypic "strains" is not known, nor whether they represent selection for mutations that confer favourable tumour traits. Nevertheless, the basic clonality of DFTD strongly suggests that tumours are not endogenously derived but are transplanted between animals, much like a tissue allograft. Pearse and Swift's serendipitous discovery of one animal with a pericentric inversion of chromosome 5 confirmed that DFTD could not be host derived, as the inversion was present only in the devil's lymphocyte preparations and not his tumour (Pearse and Swift, 2006). Genotyping of host and tumour tissue at microsatellite and MHC loci has verified that DFTD is a clonal cell line that is genetically different from the animals it is grafted to (Siddle et al., 2007). The infectious nature of DFTD has been confirmed by transmission trials, in which DFTD pathogenesis was reproduced after subcutaneous inoculation of naïve animals with culture-derived tumour cells (Pyecroft et al., 2007). The manner in which cellular transmission occurs is yet to be definitively identified, however, the devil's tendency to bite during mating and communal feeding (Hamede et al., 2008), and the friable consistency of the facial tumours suggests that DFTD is transplanted by biting (Pearse and Swift, 2006).

Although it is tempting to speculate about the possible causes of DFTD, it is unlikely that we will ever know whether the initiating event was exogenously induced or a spontaneously arising somatic mutation. By the time DFTD was

discovered in 1996, the tumour was already a transmissible cell line; a thriving clone that out-competed its precursors to become the sole cell type in a highly malignant tumour.

23.4 Immunogenetics of Transmissible Tumours

Naturally occurring cellular transmission of an infectious tumour has been reported in only one other disease. Canine transmissible venereal tumour (CTVT) is a sexually transmitted histiocytic tumour of domestic dogs (Cohen, 1985; Mozos et al., 1996). The tumour has a global distribution with a higher prevalence in countries with large populations of free-roaming dogs. Transplantation of tumour cells occurs during coitus when primary tumours of the external genitalia make direct contact with vaginal or penile mucosa. Metastasis is uncommon (<5%) and tumours are responsive to vincristine chemotherapy. There are variable reports of naturally occurring remission in immunocompetent dogs; these animals develop a robust antibody response that is protective upon subsequent re-exposure (Das and Das, 2000). Like DFTD, CTVT is characterised by a conserved, highly rearranged karyotype that is found in tumours worldwide (Fujinaga et al., 1989). In addition to its characteristic cytogenetic aberrations, CTVT cells can be distinguished from host cells by the insertion of a LINE (long interspersed nuclear element) transposable element upstream of the c-myc oncogene (Katzir et al., 1987). CTVT is a much older cell line than DFTD, having arisen up to 15,000 years ago in a wolf or early breed of dog (Rebbeck et al., 2009). The common ancestor of existing CTVT clones existed long after the tumour's origin, until about 200 years ago, indicating that today's cell line is a particularly successful clone that has overtaken the original cell line. Indeed, CTVT has evolved an effective means of escaping the host immune response, modifying cell surface expression of major histocompatibility complex (MHC) molecules in order to invade by stealth. The MHC locus is comprised of a highly polymorphic set of genes that encode two classes of antigen-presenting molecules found on the surface of cells. Class I and II MHC molecules bind the targets of antibody and cytotoxic T cell responses and are important in the adaptive immune response against pathogens, tumours and foreign grafts (Parham et al., 1989). Class I molecules are expressed on the surface of all cells and are responsible for recognition of endogenous antigens, whereas class II molecules recognise exogenous antigens and are only found on the surface of antigen presenting cells. Structurally, MHC molecules consist of conserved lengths of amino acids that stabilise the protein, and a rapidly evolving, highly polymorphic peptide binding region (PBR) that enables recognition of a diverse range of antigens (Bjorkman et al., 1987; Hughes and Nei, 1988). Modulation of MHC expression is commonly employed by tumour cells to overcome host immunity and achieve uninhibited growth (Garcia-Lora et al., 2003). CTVT cells downregulate MHC (or DLA; dog leukocyte antigen) class I and II antigen expression during transplantation and progression so that tumour growth can occur in the absence of a host immune response (Murgia et al., 2006; Hsiao

et al., 2008). Those tumours that spontaneously regress do so because MHC expression is reinstated, provoking infiltration by lymphocytes and macrophages, which aid in graft rejection through cytokine secretion (Cohen, 1985; Hsiao et al., 2008).

How then has DFTD managed to out-manoeuvre the devil's immune system in order to become a thriving somatic cell parasite? Unlike CTVT, DFTD does not appear to manipulate its MHC expression to escape the host immune response, as it has been shown to express both class I and II MHC genes (Siddle et al., 2007). Nevertheless, DFTD transplants do not activate host T cells as inflammatory cell infiltration of tumours is limited (Loh et al., 2006a). This aspect of DFTD prompted a number of studies of the basic functionality of the Tasmanian devil immune system.

Devils have a full complement of lymphoid organs and a typical range of circulating white blood cells (Woods et al., 2007). They are able to produce a robust antibody response and their neutrophils effectively phagocytose and digest bacteria (Kreiss et al., 2008; Kreiss et al., 2009). Lymphocytes isolated from the peripheral blood of healthy devils proliferate when stimulated by T-cell mitogens (PHA and Con A), however, pooling of lymphocytes from different devils fails to induce a lymphocyte response (Siddle et al., 2007; Woods et al., 2007). The devil, therefore, is immunologically competent, except when challenged with allogeneic tissue. In an outbred population, polymorphism of MHC class I genes enables recognition and rejection of tissue transplants from unrelated individuals (Hughes and Nei, 1988). Conversely, poor MHC diversity impedes recognition of non-self and has been shown to facilitate allograft acceptance in inbred wild populations of cheetah (*Acinonyx jubatus*) (Yuhki and O'Brien 1990) and pocket gopher (*Thomomys bottae*) (Sanjayan et al., 1996). Genotyping of MHC genes from 21 Tasmanian devils demonstrated extremely low diversity, particularly at class I loci (Siddle et al., 2007). Low MHC polymorphism and failure of mixed lymphocyte responses strongly suggest that DFTD is transmitted due to histocompatibility between tumour and host devils. CTVT may have similarly emerged when the domestication process caused a genetic bottleneck, facilitating allograft acceptance due to host-graft compatibility. In the following 15,000 years the canine tumour has evolved strategies that enable it to affect outbred hosts (Murgia et al., 2006). Another transmissible tumour, which has arisen in inbred laboratory populations of the Syrian hamster is also thought to be due to a paucity of MHC diversity (Banfield et al., 1965; McGuire et al., 1985). DFTD, however, is the first transmissible tumour to threaten a wild species with extinction and provides a conclusive link between MHC polymorphism and population health.

23.5 DFTD Epidemiology and Impact

Wildlife extinction due solely to infectious disease is considered a rare event. In fact, species extirpation due to disease outbreaks has been conclusively shown in very few vertebrate species: *Taudactylus acutirostris*, an Australian frog (Schloegel et al.,

2006); and the Christmas Island rats, *Rattus macleari* and *Rattus nativitatis* (Wyatt et al., 2008). The causative agents in these cases, respectively, *Batrachochytrium dendrobatidis* and *Trypanosoma lewisi*, are both able to affect more than one species or persist in an unaffected reservoir host, an important factor in their success as pathogens. According to standard epidemiological models, it is unlikely for a host-specific pathogen to cause the extinction of a wild species as transmission rates tend to decrease as host density decreases (McCallum and Dobson, 1995). Yet, the emergence of DFTD, a single-host pathogen, has profoundly affected the conservation status of the Tasmanian devil. This is due in large part to the manner in which the disease is transmitted. DFTD transmission is frequency-dependent, meaning the rate of transmission depends on the frequency of host contact, not on host density (McCallum, 2008; Lachish et al., 2009). This important epidemiological detail was determined from mark recapture data in northeast Tasmania. In the Mt William area where DFTD was first reported, tumour prevalence remains as high as 33% despite devil population declines of 90% (Lachish et al., 2007). This model of disease transmission is consistent with our understanding of devil behaviour and DFTD dissemination. The tumour is transmitted by direct contact, most likely via bite wounds. Mating interactions are associated with a higher incidence of penetrating facial bites (Hamede et al., 2008) and DFTD may therefore be considered a sexually transmitted disease, most of which, in humans, are transmitted in a frequency-dependent manner.

One of the major implications of frequency-dependent transmission is that DFTD is propagated even when devil populations are low. Density-dependent diseases (infectious disease such as influenza viruses are modelled in this manner) require host population density to remain above a certain threshold for disease maintenance. When host density decreases to a critical level disease transmission is no longer possible. It is therefore considered improbable for density-dependent disease to cause extinction, except in the presence of other contributing factors (Hilker et al., 2009). By contrast, DFTD is unlikely to peter out even when devil densities are extremely low. It is this aspect of DFTD transmission that has had such a strong impact on devil numbers. In 1996, when the first case of DFTD was observed, the devil was at a low risk of extinction (Maxwell et al., 1996). Since then, the state-wide devil population is thought to have diminished by more than 60%. The impact of DFTD has been modelled from data collected on the Freycinet Peninsula on Tasmania's east coast where devil populations were closely monitored before and after DFTD's appearance there in 2001. The arrival of DFTD on the peninsula coincided with a profound decline in devil numbers and caused acute changes to the life-history and age structure of the population. Because transmission of DFTD usually occurs during mating interactions, the devils most at risk are those of breeding age. Eight years after DFTD first swept through Freycinet National Park, the devil population is almost completely comprised of animals younger than 3 years of age, with adult devils accounting for a mere 20% of the population (Lachish et al., 2009). There has been an increase in precocial breeding throughout the peninsula, with females producing their first litter at 1 year of age instead of 2 (Jones et al., 2008). The compensatory value of this reproductive response is limited however; in affected

populations the prevalence of disease among 2 year olds exceeds 50% and many females will not survive to wean their only litter (McCallum, 2008). In little more than 10 years, DFTD has swept westwards throughout Tasmania and now threatens the last remaining disease-free populations in Tasmania's far northwest (Fig. 23.3). Given the rapid spread of DFTD, it is likely that in 5–10 years time the disease will occur across the entire range of the devil, with wild Tasmanian devil extinction projected to occur within 25–35 years (McCallum et al., 2007). Given this alarming prognosis, the Tasmanian devil was listed as an endangered species by the International Union for the Conservation of Nature and Natural Resources (IUCN) in May 2008 (IUCN, 2009).

An additional complicating factor in the emergence of DFTD relates to the most recent introduction of foxes to Tasmania (Saunders et al., 2006). It is thought that the existence of devils in Tasmania has prevented foxes from becoming

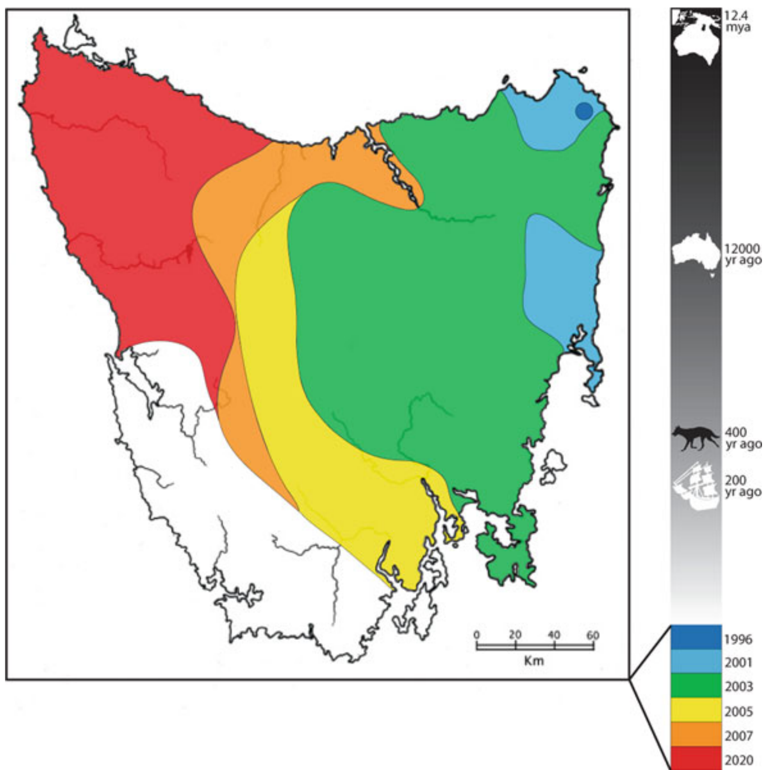


Fig. 23.3 Spread of DFTD throughout Tasmania, adapted from (McCallum, 2008) – the logarithmic scale on the right of the map depicts a timeline from Tasmanian devil divergence to the formation of the Bass Strait and the devil's extirpation from mainland Australia following the introduction of the dingo. The final segment on the scale correlates with the spread of DFTD since 1996 when the disease was first reported. It is predicted that DFTD will be present throughout the devil's range by 2020

established following previous incursions over the last 100 years. The loss of an apex predator can precipitate widespread ecosystem disturbances with effects on mesopredator abundance being of particular concern (Ritchie and Johnson, 2009). The wild extinction of Tasmanian devils would leave their predatory niche vulnerable to appropriation by foxes or feral cats. Foxes have been implicated in the widespread extinction of small mammals on the Australian mainland, a catastrophe that Tasmania's biodiversity has hitherto been free from (Kinnear et al., 2002).

23.6 Disease Management

The emergence of DFTD has highlighted the importance of early, decisive responses in the management of wildlife disease outbreaks. Humane destruction of infected animals may have limited transmission sufficiently to achieve disease eradication when the tumour first emerged, however, DFTD was not immediately suspected to be an infectious disease and culling of infected animals remained contentious even after overwhelming evidence indicated that the tumours were transmissible. As a single-host pathogen transmitted by direct contact in a population contained on a relatively small island, DFTD constituted a comparably manageable pathogen (Cousins and Roberts, 2001). Nevertheless, field monitoring was prioritised and by the time it was agreed that intervention was urgently required, DFTD was too widespread to eradicate by culling. Studies on the Forestier Peninsula on the east coast of Tasmania indicate that some disease suppression can be achieved by systematically removing infected animals from the population (Jones et al., 2007). Although widespread euthanasia of infected devils is not a pragmatic management solution at this stage, strategic culling may still be employed in high priority devil populations, such as those near the advancing disease front, or in infected populations thought to be genetically valuable.

Management options are now focused on conserving "insurance" stocks of disease-free animals for re-release in the event of wild extinction. The insurance strategy aims to conserve as much genetic diversity as possible in a DFTD-free population that will be physically and behaviourally fit for eventual release (DPIW and ARAZPA, 2007). The strategy recommends sourcing founder devils from DFTD-free wild populations according to a strict protocol that minimises the risk of introducing an infected devil into the insurance population. Founder devils are quarantined to verify their DFTD-free status before being transferred to mainland zoos for intensively managed breeding. Based on conservative modelling, a founder base of 150 individuals is predicted to conserve 95% of wild source genetic diversity (Jones et al., 2007). The successful rehabilitation of species such as the black-footed ferret have demonstrated that wild species may be rescued from extinction by breeding from much smaller founding populations (Grenier et al., 2009).

If devils were to become extinct in the wild, the ecosystem-wide impacts of their absence could be catastrophic. Even the transient loss of devils from their predatory niche may facilitate mesopredator release with dire consequences for Tasmania's

biodiversity. The maintenance of free-ranging devils in fenced-off, disease-free exclosures or off-shore Tasmanian islands has been discussed as a management strategy to ensure that insurance devils remain an "ecologically functional" species (Jones et al., 2007). Neither strategy has as yet been employed; a fencing strategy is likely to be expensive and difficult to engineer, and the release of devils onto an island ecosystem that has not previously supported them remains contentious.

Vaccine production remains a long-term solution for DFTD control. Vaccination is a cornerstone of disease prevention in veterinary and human medicine, and is playing an increasingly important role in wildlife disease control (Plumb et al., 2007). The age of targeted cancer therapy was heralded by the development of imatinib, a tyrosine kinase inhibitor that acts directly on the fusion gene product created by the Ph chromosome rearrangement in CML patients (Deininger and Druker, 2003). Like CML, DFTD is associated with a pathognomonic karyotype, however, with little known about the clinical significance of DFTD-associated chromosome rearrangements, no targets have been identified for treatment or prevention. One of the difficulties in developing cancer vaccines lies in the antigenicity of tumour cells. Unlike virus-associated cancers, spontaneously arising tumour cells may be only weakly immunogenic and unless a tumour-specific antigen can be identified there is a risk of inducing autoimmunity (Rice et al., 2008). At this stage DFTD has not been characterised sufficiently to identify tumour-specific therapeutic targets. Should this challenge be negotiated, the stability of the DFTD genome may lend itself to vaccine development. Chromosomally unstable tumours present a therapeutic challenge as CIN provides an adaptive mechanism by which tumours are able to evolve drug resistance (Rajagopalan et al., 2003). Cytogenetic observations of DFTD indicate that the karyotype is stable, suggesting that a targeted therapy might be successful in treating disease. The failure of DFTD to elicit any immune response in the devil, however, may limit attempts to induce immunity through vaccination (Woods et al., 2007). Additionally, the development of a novel vaccine that is safe and efficacious is a lengthy and expensive process, and the successful delivery of vaccines to free-ranging wildlife presents yet another challenge (Haydon et al., 2006). At this stage it is unlikely that a vaccine will be produced in time to prevent DFTD from spreading throughout the range of devil. Nevertheless, even 10 years from now, a DFTD vaccine would be a valuable management resource for demographically strategic use in habitat corridors and along the frontline of the disease, as well as protection of insurance devils prior to release. Further research is required in order to determine whether vaccination remains a viable option in the management of DFTD.

One possible, though improbable, outcome of DFTD emergence is natural selection for resistance to infection (McCallum, 2008). It is not known whether the faint genetic variation between east and west coast devils correlates with sufficient MHC diversity to enable antigen recognition and antibody production in western Tasmanian devils. This seems unlikely as no resistant devils have been observed in any of the widespread mark-recapture programs. The rate at which DFTD continues to sweep through Tasmania leaves little time for the evolution of resistance and the likelihood of adaptive variation occurring in an endangered species with poor genetic diversity is low (Kohn et al., 2006).

23.7 A Role for Genomics in Tasmanian Devil Conservation

Advances in the development of genetic markers have been of great value to conservation programs and are now essential tools in the management of endangered populations. Traditionally these markers have been used to gauge genetic variation and population structure, estimate effective population size, resolve phylogenetic and phylogeographic queries and guide captive breeding programs (Primmer, 2009). Rapid developments in sophisticated new genome sequencing technology and the explosion of genomic resources available to non-model organisms have already had a significant impact on the management of some endangered species. The endangered California condor (*Gymnogyps californianus*), largest of the North American land birds, suffered an acute population bottleneck associated with severe habitat loss and lead poisoning. Genetic variation within the tiny founder population was assessed from restriction fragment length polymorphism (RFLP), mitochondrial DNA and microsatellite studies and was used to advise a successful captive breeding program. A high prevalence of autosomal recessive chondrodystrophy complicated conservation efforts by causing an unacceptable incidence of embryonic mortality. In the absence of an assay to identify carrier birds it was impossible to eliminate the condition without compromising genetic diversity in the breeding population. Cross-species fluorescent in situ hybridisation (FISH) of condor metaphase spreads with chicken chromosome paints indicated strong homology between the two species (Raudsepp et al., 2002). Subsequently, a comparative chicken-condor map was generated by screening a condor BAC library using overgo probes obtained from the zebra finch and chicken genome projects. In total 93 loci were mapped, including several candidates for inherited chondrodystrophy (Romanov et al., 2006). A candidate gene approach to mining the BAC library is being complemented by single nucleotide polymorphism (SNP) screening of the condor transcriptome, generated by 454 sequencing of fibroblast cDNA libraries (Romanov et al., 2009). The California condor conservation program has demonstrated how sophisticated genomic resources, and in particular new sequencing technology, are enabling rapid advances in threatened species conservation. Even those species where the resources are not available for large scale sequencing projects may eventually be "genome-enabled"; nearly 1,000 IUCN red-listed mammals are within the same order or group of a sequenced species (Kohn et al., 2006). Therefore the genome of the domestic cat (*Felis catus*) will not only contribute to the health and well-being of companion animals but will also benefit the management of endangered wild felids (O'Brien et al., 2008).

Like the condor, the Tasmanian devil will benefit from the application of new genomic resources to disease management. Genetic and genomic approaches have already been instrumental in demonstrating the aetiology and pathogenesis of DFTD. The disease was thought to be virally or environmentally induced until cytogenetic and genotyping studies convincingly showed that DFTD was a clonal cell line. The genomic toolkit is gradually being expanded to include genome sequence and gene expression data that will contribute greatly to management of insurance populations and laboratory investigations of DFTD carcinogenesis.

Next generation sequencing technologies have revolutionised cancer genomics. Shotgun sequencing provides an unbiased approach to rapidly discovering pathologically significant mutations. Whereas candidate gene approaches have proved laborious and protracted, tumour genome sequencing has dramatically increased the catalogue of cancer-associated somatic mutations, even for those malignancies that have been rigorously studied for decades (Ding et al., 2008; Ley et al., 2008). Deep sequencing of the DFTD transcriptome sequence revealed a pattern of gene expression consistent with that of myelinating cells (Murchison et al., 2010). It was subsequently determined that the tumour expresses the Schwann cell-specific myelin protein periaxin (PRX), suggesting that DFTD arose from a Schwann cell or a Schwann cell precursor. Using new sequencing technology to refine DFTD cell classification beyond its broad immunostaining characteristics has therefore generated a new, specific diagnostic marker (PRX) and will provide crucial insights into the molecular pathways involved in carcinogenesis, possibly revealing therapeutic targets. The catalogue of expressed genes yielded by the devil transcriptome provides a framework with which to investigate gene expression across a large number of tumours at different stages of progression and from various locations in Tasmania. The significance of the various cytogenetic strains of DFTD may be clarified and surveillance for tumour evolution that is not cytogenetically evident will be possible. Sequencing projects evidently make critical contributions to endangered species management programs and should continue to be a priority in the conservation of the Tasmanian devil.

23.8 Implications of DFTD for Conservation Biology and Future Directions

Emerging diseases of wildlife are increasingly underlining the relationship between environmental, human and wildlife health. The emergence of zoonotic diseases such as severe acute respiratory syndrome (SARS) and Hendra virus have highlighted the importance of conservation medicine to global health (Daszak and Karesh). The resources pouring into wildlife disease are largely directed at those zoonotic pathogens that pose a spill-over risk to the public. Although cancer is an important cause of mortality in wild species, resources tend to be limited for investigating such diseases (McAloose and Newton, 2009). Regardless, considerable progress has been made in DFTD research. An integrated approach has elucidated the immunogenetic and epidemiological background against which such an unusual pathogen has been able to emerge and take hold. Continued efforts towards understanding DFTD tumorigenesis and conserving genetic diversity may yet rescue this iconic species from extinction.

Three particularly important lessons have emerged from the outbreak of DFTD. Firstly, DFTD has illustrated the vulnerability of inbred wild populations to the emergence of new pathogens and the importance of maintaining genetic diversity at functionally important loci such as the MHC. Genetic variation is crucial

to population fitness in the event of emerging infectious disease (O'Brien and Evermann, 1988). The profound impact of DFTD on Tasmanian devil numbers provides an alarming example of the potential consequences of inbreeding and loss of genetic diversity in wild populations. Secondly, DFTD has challenged our understanding of the dynamics of infectious disease in free-living wildlife. It is clear now that, counter to prevailing theory, single-host pathogens can pose a major extinction risk when disease transmission is frequency-dependent. Finally and perhaps most importantly, the DFTD outbreak underscores the importance of rapid responses to wildlife disease, even when very little is known about the causative pathogen or mode of transmission. In the case of DFTD, considerable progress has been made in a short amount of time, however, the outlook for the devil is tenuous at best. New high-throughput sequencing platforms are proving essential to conservation genomics. In the case of the devil, a well-stocked genomic toolkit will equip conservationists with the means to judiciously manage insurance populations and will provide a gene-level insight into DFTD pathogenesis, bringing us closer to the development of preventative or curative therapy.

References

Banfield WG, Woke PA, Mackay CM, Cooper HL (1965) Mosquito transmission of a reticulum cell sarcoma of hamsters. Science 148:1239–1240.

Bjorkman PJ, Saper MA, Samraoui B, et al. (1987) Structure of the human class I histocompatibility antigen, HLA-A2. Nature 329:506–512.

Bradshaw CJA, Brook BW (2005) Disease and the devil: density-dependent epidemiological processes explain historical population fluctuations in the Tasmanian devil. Ecogeography 28:181–190.

Canfield P, Hartley W, Reddacliff G (1990) Spontaneous proliferations in Australian marsupials – a survey and review. 2. Dasyurids and bandicoots. J Comp Pathol 103:149–158.

Cohen D (1985). The canine transmissible venereal tumor: a unique result of tumor progression. Adv Cancer Res 43:75–112.

Cousins DV, Roberts JL (2001) Australia's campaign to eradicate bovine tuberculosis: the battle for freedom and beyond. Tuberculosis 81:5–15.

Crowther MS, Blacket MJ (2003) Biogeography and speciation in the dasyuridae: why are there so many kinds of dasyurids? In: Jones M, Dickman C, Archer M (eds) Predators with Pouches: The Biology of Carnivorous Marsupials. CSIRO, Collingwood.

Das U, Das AK (2000) Review of canine transmissible venereal sarcoma. Vet Res Commun 24:545–556.

Deininger MW, Druker BJ (2003) Specific targeted therapy of chronic myelogenous leukemia with imatinib. Pharmacol Rev 55:401–423.

Department of Primary Industries and Water (2009) New Strains of DFTD Emerging. Retrieved 21 November 2009, from http://www.tassiedevil.com.au

Ding L, Getz G, Wheeler DA, et al. (2008) Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455:1069–1075.

DPIW and ARAZPA (2007) Save the Tasmanian Devil Program – Insurance Population Strategy. Retrieved 30th August 2009, from http://www.dpiw.tas.gov.au/inter.nsf/Attachments/EKOE-75M8ET/$FILE/STTD_IPS_v1_290707.pdf

Fujinaga T, Yamashita M, Yoshida MC, et al. (1989) Chromosome analysis of canine transmissible sarcoma cells. Zentralbl Veterinarmed A 36:481–489.

Garcia-Lora A, Algarra I, Collado A, Garrido F (2003) Tumour immunology, vaccination and escape strategies. Eur J Immunogenet 30:177–183.
Grenier MB, McDonald DB, Buskirk SW (2009) Rapid population growth of a critically endangered carnivore. Science 317:779.
Griner LA (1979) Neoplasms in Tasmanian devils (*Sarcophilus harrisii*). J Natl Cancer Inst 62:589–595.
Guiler ER (1970) Observations of the Tasmanian devil, *Sarcophilus harrisii* (Marsupialia: Dasyuridae) I. Numbers, home range, movements and food in two populations. Aust J Zool 18:49–62.
Hamede RK, McCallum H, Jones M (2008) Seasonal, demographic and density-related patterns of contact between Tasmanian devils (*Sarcophilus harrisii*): implications for transmission of devil facial tumour disease. Austral Ecol 33:614–622.
Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70.
Hawkins CE, Baars C, Hesterman H, et al. (2006) Emerging disease and population decline of an island endemics, the Tasmanian devil *Sarcophilus harrisii*. Biol Conserv 131:307–324.
Haydon DT, Randall DA, Matthews L, et al. (2006) Low-coverage vaccination strategies for the conservation of endangered species. Nature 443:692–695.
Hilker FM, Langlais M, Malchow H (2009) The Allee effect and infectious diseases: extinction, multistability, and the (dis-)appearance of oscillations. Am Nat 173:72–88.
Hsiao YW, Liao KW, Chung TF, et al. (2008) Interactions of host IL-6 and IFN-gamma and cancer-derived TGF-beta1 on MHC molecule expression during tumor spontaneous regression. Cancer Immunol Immunother 57:1091–1104.
Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335:167–170.
IUCN (2009) IUCN Red List of Threatened Species. Version 2009.1. Retrieved 25 August 2009, from www.iucnredlist.org
Jones ME, Cockburn A, Hamede R, et al. (2008) Life-history change in disease-ravaged Tasmanian devil populations. Proc Natl Acad Sci USA 105:10023–10027.
Jones ME, Jarman PJ, Lees CM, et al. (2007) Conservation management of Tasmanian devils in the context of an emerging, extinction-threatening disease: devil facial tumor disease. EcoHealth 4:326–337.
Jones M, Oakwood M, Belcher CA, et al. (2003). Carnivore concerns: problems, issues and solutions for conserving Australia's marsupial carnivores. In: Jones M, Dickman C, Archer M (eds) Predators with Pouches: The Biology of Carnivorous Marsupials. CSIRO, Collingwood.
Jones ME, Paetkau D, Geffen E, Moritz C (2004) Genetic diversity and population structure of Tasmanian devils, the largest marsupial carnivore. Mol Ecol 13:2197–2209.
Katzir N, Arman E, Cohen D, Givol D, Rechavi G (1987) Common origin of transmissible venereal tumors (TVT) in dogs. Oncogene 1:445–448.
Kinnear JE, Sumner NR, Onus ML (2002) The red fox in Australia – an exotic predator turned biocontrol agent. Biol Conserv 108:335–359.
Kohn MH, Murphy WJ, Ostrander EA, Wayne RK (2006) Genomics and conservation genetics. Trends Ecol Evol 21:629–637.
Krajewski C, Wroe S, Westerman M (2000) Molecular evidence for the pattern and timing of cladogenesis in dasyurid marsupials. Zool J Linn Soc 130:375–404.
Kreiss A, Fox N, Bergfeld J, et al. (2008) Assessment of cellular immune responses of healthy and diseased Tasmanian devils (*Sarcophilus harrisii*). Dev Comp Immunol 32:544–553.
Kreiss A, Wells B, Woods GM (2009) The humoral immune response of the Tasmanian devil (*Sarcophilus harrisii*) against horse red blood cells. Vet Immunol Immunopathol 130:135–137.
Lachish S, Jones M, McCallum H (2007) The impact of disease on the survival and population growth rate of the Tasmanian devil. J Anim Ecol 76:926–936.
Lachish S, McCallum H, Jones M (2009) Demography, disease and the devil: life-history changes in a disease-affected population of Tasmanian devils (*Sarcophilus harrisii*). J Anim Ecol 78:427–436.

Ladds P, Loh R, Jones M, Tucker P (2003). Probable Lymphosarcoma in the Tasmanian devil (*Sarcophilus harrisii*). Proceedings of the Australian Society for Veterinary Pathology Annual Conference, Elizabeth Macarthur Agricultural Institute, Menangle, New South Wales, Australia.

Lambeck K, Chappell J (2001) Sea level change through the last glacial cycle. Science 292: 679–686.

Lengauer C, Kinzler KW, Vogelstein B (1997) Genetic instability in colorectal cancers. Nature 386:623–627.

Ley TJ, Mardis ER, Ding L, et al. (2008) DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature 456:66–72.

Loh R, Bergfeld J, Hayes D, et al. (2006a) The pathology of devil facial tumor disease (DFTD) in Tasmanian devils (*Sarcophilus harrisii*). Vet Pathol 43:890–895.

Loh R, Hayes D, Mahjoor A, et al. (2006b) The immunohistochemical characterization of devil facial tumor disease (DFTD) in the Tasmanian devil (*Sarcophilus harrisii*). Vet Pathol 43: 896–903.

Maxwell S, Burbridge AA, Morris K (1996). The 1996 Action Plan for Australian Marsupials and Monotremes. Environment Australia, Canberra.

McAloose D, Newton AL (2009) Wildlife cancer: a conservation perspective. Nat Rev Cancer 9:517–526.

McCallum H (2008) Tasmanian devil facial tumour disease: lessons for conservation biology. Trends Ecol Evol 23:631–637.

McCallum H, Dobson A (1995) Detecting disease and parasite threats to endangered species and ecosystems. Trends Ecol Evol 10:190–194.

McCallum H, Tompkins DM, Jones M, et al. (2007) Distribution and impacts of Tasmanian devil facial tumor disease. EcoHealth 4:318–325.

McGuire KL, Duncan WR, Tucker PW (1985) Syrian hamster DNA shows limited polymorphism at class I-like loci. Immunogenetics 22:257–268.

Mitelman F (2000) Recurrent chromosome aberrations in cancer. Mutat Res 462:247–253.

Mitelman F, Johansson B, Mandahl N, Mertens F (1997) Clinical significance of cytogenetic findings in solid tumors. Cancer Genet Cytogenet 95:1–8.

Mozos E, Mendez A, Gomez-Villamandos JC, Martin De Las Mulas J, Perez J (1996) Immunohistochemical characterization of canine transmissible venereal tumor. Vet Pathol 3:257–263.

Murchison E, Tovar C, Hsu A, et al. (2010) The Tasmanian devil transcriptome reveals Schwann cell origins of a clonally transmissible cancer. Science 327:84–87.

Murgia C, Pritchard JK, Kim SY, Fassati A, Weiss RA (2006) Clonal origin and evolution of a transmissible cancer. Cell 126:477–487.

Nowell PC (1976) The clonal evolution of tumor cell populations. Science 194:23–28.

Nowell PC, Hungerford DA (1960) A minute chromosome in human chronic granulocytic leukemia. Science 132:1497.

Owen D, Pemberton D (2005). Tasmanian Devil: A Unique and Threatened Animal. Allen & Unwin, Crows Nest.

O'Brien SJ, Evermann JF (1988) Interactive influences of infectious disease and genetic diversity in natural populations. Trends Ecol Evol 3:254–259.

O'Brien SJ, Johnson W, Driscoll C, et al. (2008) State of cat genomics. Trends Genet 24:268–279.

Parham P, Lawlor DA, Lomen CE, Ennis PD (1989) Diversity and diversification of HLA-A,B,C alleles. J Immunol 142:3937–3950.

Pearse AM, Swift K (2006) Allograft theory: transmission of devil facial-tumour disease. Nature 439:549.

Plumb G, Babiuk L, Mazet J, et al. (2007) Vaccination in conservation medicine. Rev Sci Tech 26:229–241.

Primmer CR (2009) From conservation genetics to conservation genomics. Ann NY Acad Sci 1162:357–368.

Pyecroft SB, Pearse A-M, Loh R, et al. (2007) Towards a case definition for devil facial tumour disease: what is it? EcoHealth 4:336–351.

Rajagopalan H, Lengauer C (2004) Aneuploidy and cancer. Nature 432:338–341.

Rajagopalan H, Nowak MA, Vogelstein B, Lengauer C (2003) The significance of unstable chromosomes in colorectal cancer. Nat Rev Cancer 3:695–701.

Raudsepp T, Houck ML, O'Brien PC, et al. (2002) Cytogenetic analysis of California condor (*Gymnogyps californianus*) chromosomes: comparison with chicken (*Gallus gallus*) macrochromosomes. Cytogenet Genome Res 98:54–60.

Rebbeck CA, Thomas R, Breen M, Leroi AM, Burt A (2009) Origins and evolution of a transmissible cancer. Evolution 63:2340–2349

Rice J, Ottensmeier CH, Stevenson FK (2008) DNA vaccines: precision tools for activating effective immunity against cancer. Nat Rev Cancer 8:108–120.

Ritchie EG, Johnson CN (2009) Predator interactions, mesopredator release and biodiversity conservation. Ecol Lett 12:982–998

Romanov MN, Koriabine M, Nefedov M, de Jong PJ, Ryder OA (2006) Construction of a California condor BAC library and first-generation chicken-condor comparative physical map as an endangered species conservation genomics resource. Genomics 88:711–718.

Romanov MN, Tuttle EM, Houck ML, et al. (2009) The value of avian genomics to the conservation of wildlife. BMC Genomics 10(Suppl 2):S10.

Sanjayan MA, Crooks K, Zegers G, Foran D (1996) Genetic variation and the immune response in natural populations of pocket gophers. Conserv Biol 10:1519–1527.

Saunders G, Lane C, Harris S, Dickman C (2006). Foxes in Tasmania. A Report on an Incursion of an Invasive Species, Invasive Animals Cooperative Research Centre, Canberra.

Schloegel L, Hero J-M, Berger L, et al. (2006) The decline of the sharp-snouted day frog (*Taudactylus acutirostris*): the first documented case of extinction by infection in a free-ranging wildlife species? EcoHealth 3:35–40.

Siddle HV, Kreiss A, Eldridge MD, et al. (2007) Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial. Proc Natl Acad Sci USA 104:16221–16226.

Weaver BA, Cleveland DW (2006) Does aneuploidy cause cancer? Curr Opin Cell Biol 18: 658–667.

Woods GM, Kreiss A, Belov K, et al. (2007) The immune response of the Tasmanian devil (*Sarcophilus harrisii*) and devil facial tumour disease. EcoHealth 4:338–345.

Wyatt KB, Campos PF, Gilbert MT, et al. (2008) Historical mammal extinction on Christmas Island (Indian Ocean) correlates with introduced infectious disease. PLoS One 3:e3602.

Yuhki N, O'Brien SJ (1990) DNA variation of the mammalian major histocompatibility complex reflects genomic diversity and population history. Proc Natl Acad Sci USA 87:836–840.

Index

A
Antimicrobial, 130, 336, 381–382, 392, 394–396

B
Bioinformatics, 118, 121–122, 124, 126, 129, 131, 133, 336

C
Captive breeding, 1–2, 5–7, 9, 14, 18, 20–21, 24, 26–27, 102, 483, 510
Centromere, 33–34, 37–38, 44, 48–49, 55–66, 68, 91–92, 109, 155, 161, 189, 194, 196, 349
Chromosome, 2, 19, 33–34, 37–51, 55–61, 65–68, 76–79, 81–83, 85, 87–95, 101–113, 118, 133, 135, 141–143, 147–148, 151–167, 173–183, 187–200, 207–222, 229–230, 234, 237–238, 241–244, 246, 259–267, 269–271, 273–274, 300–302, 306, 347, 349, 360–361, 366, 369–370, 374, 386, 392, 394, 404–405, 409–410, 412, 415, 418–422, 426, 428–429, 439–441, 445, 450, 457, 486, 501–503, 509–510
Chromosome evolution, 33, 37–38, 40–44, 46–47, 49–50, 56–57, 143, 148, 187, 200, 211, 226
Chromosome painting, 33–34, 37, 39, 41–45, 47, 49–51, 57, 66, 102, 107–109, 152–154, 159–160, 213
Colour vision, 401, 403–407, 409–412, 446
Conservation, 2, 5–9, 14, 18, 26, 33, 40, 56, 60, 62, 81, 109, 138–139, 141, 157, 159, 162, 174, 188, 200, 217, 219, 239, 263, 335, 339, 343, 358, 366, 374, 385, 395, 427, 450, 457, 461–462, 465–467, 469, 471, 473, 475, 477, 479, 481–483, 485–486, 499, 506–507, 510–512
Cytogenetic, 33, 101–102, 111–112, 160–162, 200, 230, 262, 501–504, 509–511
Cytokine, 295, 336, 347, 381–391, 394, 396–397, 505

D
Dense plate, 148, 177–178, 187, 192, 194, 196, 214
Devil facial tumour disease-DFTD, 2, 8, 104, 112, 122, 335, 351, 390, 482, 491, 501, 503, 505, 507, 509, 511
Diapause, 10–13, 25, 282, 285, 289–290, 296–297
Dosage compensation, 156–157, 158, 164, 167, 187, 200, 230, 259–262, 266–267, 270–271, 274

E
Embryonic development, 13, 282, 285, 287, 289, 293, 295, 300
Epigenetic, 56, 63, 66, 68, 95–96, 101, 136, 141, 157, 165, 199, 229–230, 233–234, 243, 245–247, 259–262, 265, 274, 439, 452
Exaptation, 118, 133, 139–144

G
Gametes, 77–78, 154, 174, 190, 285, 295
G-banding, 77–78, 154, 174, 190, 285, 295
Gene family, 129, 154, 179, 181, 299, 345, 394, 401–402, 435, 437, 439, 441–445, 447, 449–452
Gene regulation, 133, 135, 137, 139, 141, 143, 260
Genetic diversity, 2, 5, 9, 14, 18–22, 24–26, 81, 335, 339, 349, 351–352, 461–463, 477–478, 481–483, 485–486, 500, 508–512
Genome evolution, 34, 79, 101, 104, 106–107, 113, 133, 139, 143, 152

J.E. Deakin et al. (eds.), *Marsupial Genetics and Genomics*,
DOI 10.1007/978-90-481-9023-2, © Springer Science+Business Media B.V. 2010

Genomic imprinting, 51, 79, 101, 106, 112, 229–230, 233–237, 239–241, 243, 245–250, 439, 452
Genomics, 1, 5–6, 33, 37, 51, 68, 101, 103–104, 117–118, 121, 130, 133–134, 138–139, 151, 200, 207, 244, 251, 322, 335–336, 340, 357, 359, 361, 363, 365, 367, 369, 371, 373, 375, 458–499, 510–512
Germ cells, 218, 245, 285, 288–289, 295, 303, 305, 307, 365, 443
α-Globin, 415–416, 418–423, 425–429
β-Globin, 401–402, 415–423, 425–429

H

Haemoglobin, 401, 415–418, 420–422, 427–429

I

Immune response, 19, 335–336, 339–340, 351–352, 357, 368, 375, 381, 383–384, 389–392, 397, 458, 504, 509
Immunity, 128, 327, 357, 359, 375, 383–384, 394, 397, 458, 504, 509
Immunoglobulin, 129–130, 335, 357, 359, 361, 363, 365–367, 369, 371, 373–375, 383, 386, 391
Inbreeding, 5, 9, 14, 18–25, 481, 483, 485, 512

K

Karyotype, 34, 37–51, 58–59, 61–62, 101–102, 106, 113, 152–153, 189, 190, 502–504, 509
KERV, 33, 49, 55, 61–66, 110

L

Linkage mapping, 2, 34, 75–77, 84, 102, 135

M

Major Histocompatibility Complex-MHC, 2, 335, 339, 341, 343, 345, 347, 349, 351, 353, 358, 383, 462, 504
Mammary, 215, 301, 317–319, 321, 323, 325, 327–331, 369, 394
Mating system, 5, 9, 14–18, 24–26
Meiosis, 76–78, 89–90, 143, 148, 153–156, 166, 173–179, 181, 187, 189–191, 193, 195–200, 210, 212–214, 245, 264, 273, 295–296, 307
Meiotic sex chromosome inactivation-MSCI, 143, 147, 166, 173, 176, 178, 187, 192, 197, 244, 264
Mental retardation, 173, 181, 219–220, 222, 231
Microsatellite, 19, 24, 77, 82, 85, 87, 351, 458, 461, 463–464, 468, 478, 481–482, 486, 503, 510
Milk, 123, 127, 250, 281–282, 286, 293, 317–328, 330–331, 367, 395
Mitochondrial DNA, 19, 461, 478, 510
Monotreme, 121, 134, 145, 159–160, 167, 174, 190, 210, 230, 233, 240, 242–243, 259, 260, 270, 330, 343, 403, 411, 418, 422–425, 428, 442

N

Natural killer cell, 381–382, 391
Next generation sequencing, 104, 121–123, 125, 130, 230, 511

O

Olfaction, 402, 435–436, 443–444, 446
Olfactory receptor, 402–423, 429, 435–439, 441, 443, 445, 447, 449, 451, 453
Opossum, 1, –367, 5–6, 12–13, 34, 37–38, 40, 46, 50, 75, 77, 81, 84–96, 101–102, 104–113, 117–118, 121, 123, 125–130, 133–139, 141–144, 147, 154–163, 165, 174, 176–177, 179, 218, 220, 222, 239–240, 267–269, 273, 307–308, 330, 335–336, 339–350, 352, 359, 360–361, 363, 365, 369–370, 372, 374, 381–382, 386–394, 406–407, 409–412, 426, 435, 440, 442, 447–452, 463–464
Opsin, 401, 403–404, 406, 408–409, 412

P

Parental conflict, 229–230, 233, 248–249
Physical map, 1, 101, 107, 109–110, 112–113, 162
Placenta, 10–11, 112, 229–230, 235, 237–239, 246–250, 272, 281, 286, 290–292, 331, 443
Platypus, 90, 134, 151, 156, 160–161, 174, 190, 210, 240–243, 246, 270–271, 330, 344, 375, 403, 407, 410–412, 421–422, 424–425, 428, 440, 442, 448, 450, 452
Population genetics, 19, 461–462, 484, 486

R

Recombination, 34, 68, 75–86, 88–92, 95, 110, 133, 135–137, 139, 143, 147, 152, 155, 175–176, 179–180, 187–188, 190–191, 196–197, 207–209, 211–213, 221, 260–261, 274, 358–359, 365, 368, 372–375, 445
Retroelement, 55, 63, 66, 139

S
Sex chromosomes, 34, 46, 58–60, 103, 147–148, 152–156, 159–161, 165, 167, 174, 177, 179, 187–200, 208–217, 220, 223, 260–261, 271, 273–274, 302, 502–503
Sex determination, 79, 152, 167, 175, 190, 207–208, 211, 214–215, 217, 220, 222, 282, 285, 299–300, 302, 304, 306–308
Sex and reproduction related-SRR, 173, 176
Sexual differentiation, 215–216, 285, 291, 299, 301–308
Small RNA, 55, 63–64, 66, 139
SRY, 104, 147–148, 152–153, 167, 174–175, 188, 190, 207, 210–211, 215–218, 221–223, 260–261, 282, 300, 302–304, 306–308
Synapsis, 187, 190–191, 197, 199, 214

T
Tammar wallaby, 1, 5–8, 13, 25–26, 34, 42–43, 48–50, 57, 64, 68, 75, 77, 81–83, 101–104, 107–110, 117–118, 122–123, 126–128, 134, 142, 147–148, 154–156, 159, 161–163, 174, 176, 179, 216–221, 237, 239, 246, 267–269, 281–282, 287–291, 297, 299–300, 304–305, 307, 317–318, 320, 322–324, 326, 328, 335, 339, 341–352, 363, 365, 381–382, 385–389, 394, 406–407, 410–412, 416, 420, 422, 426, 429, 435, 440, 447, 450–452, 463, 482
Tasmanian devil, 1–2, 7, 9, 15, 104, 112, 118, 122–123, 126, 335–336, 341–346, 351–352, 375, 390, 457–458, 482, 499–500, 503, 505–507, 510–512
T cell receptor, 106, 357, 359, 361, 363, 365, 367, 369, 371, 373, 375, 383, 441
Telomere, 55, 67–68, 160, 196, 349
Transcriptome, 121–123, 125, 127–128, 137, 319, 321, 458, 510–511
Transposable elements, 66, 133, 135, 138–144, 426, 429
Transposition, 66, 105, 139–141, 213, 261, 375, 415, 424–429

W
Wildlife disease, 499, 508–509, 511–512
Wildlife health, 339, 351–353, 511
Wildlife management, 462

X
X chromosome, 2, 34, 48, 57–58, 79, 81–82, 85, 89–90, 101–102, 104, 106, 107, 109, 111–112, 118, 133, 141–143, 147–148, 151–167, 173–176, 179–183, 188, 192, 198–200, 210–212, 215–218, 221–223, 229–230, 234, 243–244, 259–267, 269–271, 273–274, 298–302, 306, 404–405, 409–410, 412, 426, 439
X chromosome inactivation-XCI, 2, 79, 101, 106, 109, 112, 118, 133, 141–143, 162, 164, 166–167, 198–200, 229–230, 244, 259–260, 262, 270, 272, 274, 405
XIST, 106, 109, 118, 141, 143, 151, 162–167, 182, 197–199, 244, 263–266, 269, 271–272, 274

Y
Y chromosome, 19, 104, 123, 124, 147–148, 152–153, 155–156, 163, 165–167, 174–180, 187–191, 193, 199–200, 207, 209–223, 230, 243, 260–261, 271, 274, 300, 302, 306, 421, 426, 457, 486