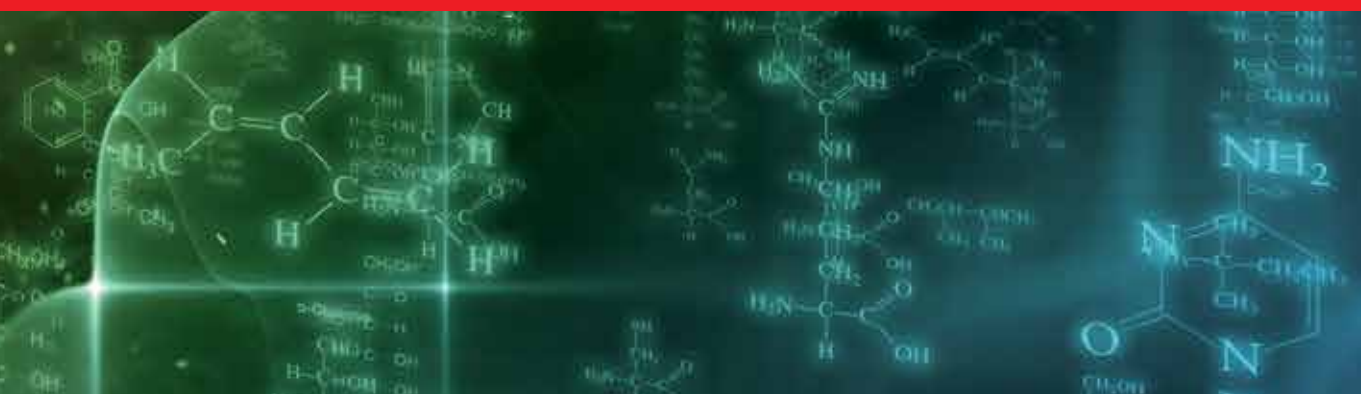# The Molecular Basis of Plant Genetic Diversity

Edited by Mahmut Çalışkan

THE MOLECULAR BASIS OF PLANT GENETIC DIVERSITY

Edited by **Mahmut Çalışkan**

The Molecular Basis of Plant Genetic Diversity
http://dx.doi.org/10.5772/2639
Edited by Mahmut Caliskan

Contributors

Ines Eugenia Garcia De Salamone, Marta Leonor Marulanda, Ana M López, Marcela Uribe, Elsayed Hafez, Snežana Mladenović Drinić, Violeta Andjelkovic, Dragana Ignjatovic-Micic, Ester Wickert, Klaus Konrad Scheuermann, Juliana Vieira Raimondi, Alexander De Andrade, Rubens Marschalek, Elizabeth Ryan, Erica Daniell, María Virginia Moreno, Sebastián Stenglein, Analía Perelló, Peter Poczai, Ildikó Varga, Neil E. Bell, Jaakko Hyvönen, Silvana Petrofeza, Luiz Carlos Berhing Nasser, Murat Ertekin, Łukasz Wolko, Ryszard Slomski, Yaoting Wu, Suping Feng, Weiguo Li, Fei Yu, Jingyi Wang, Yu-Chung Chiang, Pei-Chun Liao, Yi-Shan Chao, Stefano Castiglione, Renato Lupi, Angela Cicatelli, Nasya Borisova Tomlekova, Josphert Ngui Kimatu, Geoffrey M. Muluvi, Bao Liu

Notice

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,000+
Open access books available

116,000+
International authors and editors

120M+
Downloads

151
Countries delivered to

Our authors are among the

Top 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Meet the editor

Mahmut Çalışkan is a Professor of genetics and molecular biology in the Department of Biology at Mustafa Kemal University in Turkey. He obtained his BSc at Middle East Technical University in Ankara, and in the same year was appointed as Research Assistant to the Biology Department of the same University. During his MSc study he was awarded a grant from the Turkish High Education Council to continue his MSc and PhD studies abroad. He attended the University of Leeds Department of Genetics in England for his PhD study, focusing on analysis of germin gene products. After he received his PhD degree, he was appointed as Assistant Professor at Mustafa Kemal University, Hatay. In 2000 he obtained the title of Associate Professor in genetics from the Turkish High Education Council. Following five years of extensive research and lecturing he was appointed as Professor of Molecular Biology in 2006. Prof. Çalışkan was awarded for his postdoctoral research in England with grants from NATO and University. In addition to acting as a reviewer and advisor to numerous journals and grant awarding institutions, he has been involved in the establishment of the Biology Department at Mustafa Kemal University.

Contents

Preface

Ever since human beings adopted an agricultural way of life, they have attempted continuously to improve on what they found in nature. Once they came to understand that plant characteristics are inherited and genetic richness is an advantage against changing environmental conditions, they have developed new strategies to reveal and conserve the genetic diversity in plant populations. Plants are a distinct kingdom of organisms that possess unique properties of reproduction, development, physiology, and adaptation. Plant diversity refers to the variety of plants that exist on the Earth. Plants, in order to survive, have to compete with other plants and organisms in an ecosystem. Over time, they have developed various characteristics to help them survive, which leads to plant diversity. It is essential to have regular assessments of the conservation status of all plant species, in order to prioritize those in need of conservation action and to provide a measure of the success of actions being taken. The improvement of cultivated plants considerably depends on the extent of genetic variability available within the species. Evaluation of different genotypes, varieties, population, etc., for genetic diversity and genetic relatedness, is extremely important before taking up actions for breeding or conservation projects.

Over the last few years there have been dramatic advances in plant molecular genetics and these developments have provided scientists involved in the conservation of plant genetic resources with a range of new techniques for their research. Nowadays techniques are available to analyze variations in the DNA level of plants and animals or other organisms. Differences in gene sequence can be directly observed and described - a degree of precision previously impossible to achieve. Many of the techniques that have been developed have already been used to study the extent and distribution of variation in species gene-pools and to investigate evolutionary and taxonomic relationships among plant populations. With the development of the polymerase chain reaction (PCR) based techniques, in particular, numerous molecular technologies have been, and still are being, developed, which can be used for the detection, characterization, and evaluation of genetic diversity in plant populations. These techniques can be chosen on the basis of how they display genetic differences, the type of data which they produce, the taxonomic levels at which they can be most appropriately applied, and their technical and financial requirements.

The purpose of *The Molecular Basis of Plant Genetic Diversity* is to provide a glimpse into the dynamic process of genetic variation by presenting the thoughts of some of the scientists who are engaged in the development of new tools and ideas used to reveal genetic variation, often from very different perspectives. I would like to express my deepest gratitude to all theauthors who contributed to this book by sharing their valuable works with us. This book should prove useful to students, researchers, and experts in the area of conservation biology, genetic diversity, and molecular biology. The year 2010 was celebrated as the international year of biodiversity by the United Nations and it has been a unique opportunity to realize the vital role that biodiversity plays in sustaining the life on Earth. Let us all wish much success to all projects and initiatives dealing with the conservation of diversity of life because rich genetic resources are a prerequisite for future generations to be able to breed crop varieties and face new challenges.

Prof. Dr. Mahmut Çalışkan
Mustafa Kemal University,
Department of Biology, Hatay
Turkey

Part 1

Genetic Diversity in Plant Populations

Genomics Meets Biodiversity: Advances in Molecular Marker Development and Their Applications in Plant Genetic Diversity Assessment

Péter Poczai[1,2], Ildikó Varga[2], Neil E. Bell[1,3] and Jaakko Hyvönen[1]

[1]*Plant Biology, University of Helsinki, Helsinki*
[2]*Department of Plant Science and Biotechnology, Georgikon Faculty*
University of Pannonia, Keszthely
[3]*Botanical Museum, University of Helsinki, Helsinki*
[1,3]*Finland*
[2]*Hungary*

1. Introduction

Genetic diversity is the fundamental source of biodiversity – the total number of genetic characters contributing to variation within species. In other words it is the measure that quantifies the variation found within a population of a given species. Genetic diversity among individuals reflects the presence of different alleles in the gene pool, and hence different genotypes within populations. Genetic diversity should be distinguished from genetic variability, which describes the tendency of genetic traits found within populations to vary (Laikre et al., 2009). Since the beginning of the 20th century, the study of genetic diversity has been the major focus of core evolutionary biology. The theoretical metrics developed, such as genetic variance and heritability (Fisher, 1930; Wright, 1931), provided the quantitative standards necessary for the evolutionary synthesis. Further research has focused on the origin of genetic diversity, its maintenance and its role in evolution. Simple questions such as "who breeds with whom" initiated studies on the relatedness of populations. These investigations led to the formation of metapopulation theory, where a group of spatially separated populations of the same species interact at some level and form a coherent larger group (Hanski, 1998). The discovery of spatial structure in populations was a key element in the early concepts and models of population ecology, genetics and adaptive evolution (Wright, 1931; Andrewartha & Birch, 1954). How different levels of genetic variance affect the rate of evolutionary change within populations has also been intensively studied. Such changes were originally studied using phenotypic markers: variation among individual plants in traits, such as leaf shape or flower color (Ward et al., 2008). Subsequently the detection of genetic variation has become more sensitive, firstly through the utilization of variation in enzymes (allozymes) and then through PCR-based marker systems allowing direct examination of DNA sequence variation. The precise detection of genetic variation/diversity has greatly enhanced studies of evolution. There is

no doubt that genetic variation influences the fitness of individuals, and that this is reflected in natural selection. In this regard, individual genotypes must vary in ecologically important ways. Does this mean that differences in genetic diversity levels will have predictable ecological consequences? The answer is no, because only one portion of genetic diversity is connected to ecological factors, i.e. adaptation. Ecological adaptation is a significant factor for example, in range expansion of plant species. Plants with different genotypes conferring the highest levels of fitness are expected to survive and reproduce better, shifting the gene pool over time towards higher frequencies of the alleles making up the more successful genotypes (Ward et al., 2008). This can be seen through Fisher's (1930) example: when increase in fitness is allowed, genetic diversity can increase the population growth rate, but only if the population is not regulated by other factors and if it is experiencing directional selection. This is due to the simple fact that individuals with different genetic traits can interact in unpredictable ways. Despite the presence of genetic variation in ecologically important traits, relatively little is known about the range of potential ecological effects of genetic diversity for population dynamics, species interactions and ecosystem processes (Hughes et al., 2008). This has led to the rise of the field molecular ecology, which is integrating the application of molecular population genetics, phylogenetics and genomics to answer ecological questions. These different disciplines are united in an attempt to study genetic-based questions of ecology, e.g. species-area relationships, species delimitation and conservation assessment.

Information about plant genetic diversity is necessary for the development of appropriate strategies in conservation biology as well as in many other applied fields. From a basic evolutionary standpoint, genetic diversity is assumed to be crucial for the evolutionary potential of a species. Research programs that aim to investigate population structure provide evolutionary insights into the demographic patterns of diverse organisms (Milligan et al., 1994). Interest in genetic diversity has also risen in applied fields of biology such as agriculture, where the technique of multiline sowing of varietal mixes within a single field has long been known to increase crop yields (Smithson & Lenne, 1996) and to decrease damage by pests and pathogens (Cantelo & Sanford, 1984). In some crops (e.g. rice) this is applied on a large scale to maximize yield by minimizing damage by pathogens (Zhu et al., 2000). Information on genetic diversity and population structure also assists plant breeding in the selection of parents for crossing, providing a more rational basis for expanding the gene pool, and for identifying materials that harbor genes of value for plant improvement. Furthermore, knowledge of population structure of genetic resources is necessary for the development of strategies for appropriate conservation of genetic diversity. Therefore, in recent decades a primary activity in genetic resource management has been to characterize the structure of diversity within crop species. Increasing attention has also been given to dissecting and understanding diversity in relation to genes underlying important agronomic traits in a number of crops. Molecular phylogenetics and genetic diversity analysis can help to clarify the taxonomic identity and evolutionary relationships of the wild relatives of crop species. These methods can also help prevent misidentification and carefully plan effective germplasm management strategies.

Variability and genetic diversity are important factors in evolution and also in applied sciences because they determine the responses of a given organism to, for example, to environmental stress, natural selection and susceptibility to different diseases. In fact, reactions to different types of stress (biotic or abiotic) receive the most attention in genetic

diversity research. In an anthropocentric world the diversity of our own species also affects the study of other organisms, because human genetic variation determines sensitivity to toxins and different drugs, leading to the development of personalized medicine (Jain 2002). This has initiated the International HapMap project (http://hapmap.ncbi.nlm.nih.gov/).

To accurately study genetic diversity and to answer questions relating to organisms, some basic terminology is required. Thus before discussing genetic diversity assessment in detail, we will define some frequently used concepts and terms:

Allele frequency: the measure of the relative proportion of all copies of a gene, i.e., the frequency of an allele at a particular genetic locus. Allele frequency is used to represent genetic diversity at different scales (e.g. population, species).

Allelic richness: the average number of alleles per locus.

Biodiversity: the degree of variation of life forms within a given ecosystem, biome or entire planet. It is **not** synonymous with **species diversity**.

Coalescent theory: a model in population genetics often referred to as backward **genetic drift**, which attempts to trace alleles of a gene shared by members of a population to a single ancestral copy. Standard coalescence assumes non-overlapping generations, random reproduction, constant population size and random mating (no population structure). However, advanced analytic models allow extension of these parameters.

Evolutionary fitness: the probability that a population or species will not become extinct. Evolutionary fitness has more relevance to evolution and population control than **genetic fitness**, which describes the reproductive success of a genotype.

Founder effect: a special case of a population bottleneck, occurring when a new population is established by a random and small number of individuals from a larger population. The founder population forming the new colony from randomly sampled alleles of the original population may poorly represent the original population.

Gene flow: also called migration, is the transfer of genes from one population to another. This may happen through mobility (e.g. transfer of pollen from one location to another). Gene flow is a very important source of genetic variation, since it enables the free "flow" of genes into a population where they did not previously exist. It can also lead to the combination of gene pools or to the reduction of genetic variation between two populations.

Gene pool: the complete set of unique alleles in a species or population.

Genetic drift: the change in the frequency of an allele in a population due to random events. It is one of the basic mechanisms of evolution affecting the genetic makeup of a population. In contrast to **natural selection** it is an entirely random process.

Genotypic richness: the number of genotypes within a population. It can be measured using molecular markers as the number of haplotypes.

Natural selection: one of the key mechanisms of evolution. Unlike **genetic drift** it is a non-random process, although not constant and varying through time. It initiates the selection of biological traits which consequently become more frequent in a population.

Population bottleneck: is an event in which a significant percentage of the genetic diversity of a population is lost. It occurs when population size is reduced for at least one generation. Through time this could lead to recovery or extinction.

Population fragmentation: a form of population segregation, most often caused by habitat fragmentation. The fragmentation of a population may have significant effects on its genetic diversity through mechanisms including **genetic drift**, reduced **gene flow** and inbreeding depression.

Population structure: heterogeneity in allele frequencies across a population caused by limited gene flow.

Species diversity: an index incorporating the number of species in a given area and their relative abundance.

Subpopulation: a portion of a total population that experiences limited gene flow from other parts of the population, such that its allele frequencies evolve independently to some degree.

2. Why does genetic diversity matter?

The genetic characteristics of a population describe its structure as formed by the interplay of genetic drift, gene flow and natural selection, and perhaps also spatial distribution. In the case of threatened species, the assessment of these characteristics can help in developing accurate strategies for maintaining and preserving species diversity. This may be crucial in the case of endangered species. Ishwaran and Erdelen (2006) correctly note that the diversity of life and living systems are necessary conditions for human development, although genetic diversity assessment projects are costly and time consuming. However, we have many reasons to make efforts to maintain biodiversity. According to Wilcove and Master (2008) biodiversity (i) has an aesthetic and moral justification; moreover (ii) wild species provide products and "services" essential to human welfare; (iii) certain species are indicators of environmental health; (iv) certain species may be keystones of functioning ecosystems; and lastly (v) studies of wild organisms have led to several scientific breakthroughs. Genetic diversity is also important for the applied fields of science and even for economic growth. Plant evolution under domestication has led to increased productivity, but at the same time, domestication has dramatically narrowed the genetic basis of crop species. Due to lack of genetic diversity in crop gene pools, efforts have been made to explore the gene pools of wild species for potential utilization in meeting the future challenges of plant breeding. Thus the main aim of breeding programs nowadays is to trace diversity and to find new traits – particularly genes conferring resistance to diseases and pests – present in wild genetic resources. This is done in order to maintain current levels of agricultural productivity. Good examples can be found in solanaceous plants, including many success stories as well as some cautionary tales. In the Solanaceae, crops such as potato, tomato and pepper have greatly benefited from the use of wild relatives in breeding programs (Albrecht et al., 2010). Virtually all of the disease resistance genes introduced into modern tomato varieties are derived from related wild species (Rick & Chetelat, 1995), demonstrating how a successful breeding program can benefit from the genetic diversity of wild species. But what would have happened without the remarkable improvement in the potato gene pool? This leads to the other side of the story. During the 1840s a lack of genetic

diversity in Irish potato cultivars caused by the monocultural production of basically a single variety led to one of the biggest disasters in human history. How could low genetic diversity cause the deaths of millions of people within a very short space of time? To understand this we need to look closely at the diet of Irish people in the 19th century. Potato was introduced to Ireland in the 17th century and quickly became the staple food of the poor. It was also known as the garden crop of the gentry (Gráda, 2000). Nutritionally, potato was excellent and it was cheap to produce. Most commonly eaten with added milk or butter it formed the basis of a diet with ample protein, carbohydrates, minerals and energy. Around 1700, an average farmer would have eaten one meal with potatoes every day, while by the 1840s this had risen to three meals per day (Whelan, 1997). In an expanding agrarian economy this represented a serious potato dependency. In 1810, a new variety called "Lumper Potato" was introduced to Ireland (Whelan, 1997). This cultivar quickly became widespread, since it required minimum work and tolerated poor soil while maintaining high yields (Whelan, 1997). Lumper Potato dominated when a new disease, late blight, caused by the oomycete *Phytophthora infestans* (Fig.1), started to spread in Ireland. A plant pathogen more closely related to brown algae and diatoms than to fungi (Beakes et al., 2008), *P. infestans* can destroy susceptible crops within 10 to 14 days in favorable weather conditions and in the presence of an inoculum source (Lebecka, 2008). After a few days the whole plant collapses and infected tubers rapidly degrade into a pungent mush. Tubers that appear healthy may also turn into slime at a later stage, during storage. Even today, this pathogen ranks economically as the most important disease of potato and tomato.

Fig. 1. The symptoms of late blight (*Phytophthora infestans*) infection in potato leaves and tubers. Photos kindly provided by Dr. Zsolt Polgár.

Prior to the arrival of *Phytophthora infestans* only two other potato diseases were known in Ireland (Bourke, 1964). The disease was first observed in September 1845 on Lumper potatoes. It spread quickly and destroyed a vast proportion of the yield in that same year. In subsequent years until 1850 the disease recurred each season, destroying yields to such an extent that people did not have sufficient seed potatoes to plant, far less to eat (Fraser, 2003). Between 1845 and 1850 approximately 2 million people lost their lives (Cousens, 1960) and 1–1.5million emigrated (Boyle & Gráda, 1986). Later in the 19th century the disease spread throughout the whole of Europe. Current studies suggest a three-step introduction process for the pathogen, including a first migration from central Mexico to South America several

centuries ago, a migration from South America to the USA around 1841-42, and then a third and final migration to Europe from either South America, the USA or both in 1843-44 on ships carrying guano (Andrivon, 1996).

The epidemic had catastrophic social and economic consequences. Besides death and emigration it resulted in economic losses of 800 million US dollars (equaling 24 billion recalculated for today's market values, O'Rourke, 1994). The Great Famine also had an impact on the regional genetic structure of Ireland, resulting in a rapid decrease in population size in a short period and thus increasing the possibility of genetic drift (Relethford et al., 1997). These experiences tell us that low genetic variation can lead to disaster in ecosystems as well as in society, economy and everyday life. These consequences would not have been so severe if farmers had planted potatoes with more genetic variation.

3. The molecular toolbox: Recent technical advances in genetic diversity assessment

Understanding the forces that influence natural variation within and among populations has been a major objective of evolutionary biology for decades (Csilléry et al., 2010). It can be problematic, as natural populations may have complex demographic histories: size and range changes over time can lead to bottlenecks, fusions and expansions that leave signatures on the genetic composition of the population (Avise, 2004). Detection and analysis of genetic diversity can help us to understand the molecular basis of various biological phenomena in these natural systems. Established molecular databases providing a large volume of information on a range of different markers have potential to help in uncovering the complexity of the demographic and adaptive processes acting in natural populations. The widespread availability of different molecular markers and increased computing power has fostered the development of sophisticated methods and techniques that have begun to fulfill these expectations (Csilléry et al., 2010). Since the entire plant kingdom cannot presently be covered under sequencing projects, molecular markers and their correlation to phenotypes provide us with requisite landmarks for elucidation of genetic variation. In recent years there has been a trend away from arbitrarily amplified dominant (AAD) markers towards gene-targeted functional markers in genetic diversity assessment. Traditional genetic techniques, e.g. RAPDs (random amplified polymorphic DNA), SSRs (simple sequence repeats) and AFLPs (amplified fragment length polymorphism) are now routinely used in ecological, evolutionary, taxonomic, phylogenetic and, most importantly, in genetic diversity studies of plant species. These techniques are well established and their advantages as well as their limitations are well known. In recent years, a new class of advanced techniques has emerged, primarily derived from combination of earlier basic techniques. Genomic databases such as NCBI's GenBank have become primary sources for marker development. Based on the characterization of plant genes and gene families, new methods have been developed to analyze genetic diversity based on genomic database mining. Following this, many recent studies have suggested that polymorphism in functional regions of the genome should be exploited to achieve better estimates of genetic relationships that are relevant for conservation purposes. A wide variety of DNA arrays has been developed to meet these goals. Advanced marker techniques tend to combine adventitious features of several basic techniques, while also

incorporating modifications in methodology to increase sensitivity and resolution in order to detect genetic discontinuity and distinctiveness. These advanced techniques utilize specific classes of genome elements such as (retro) transposons and cytochrome P450 genes and thus reveal genetic variation through increased genome coverage. They differ from each other with respect to important features such as genomic abundance, level of polymorphism detected, locus specify, reproducibility, technical requirements and cost. Depending on requirements, modifications have been made to these techniques, leading to second and third generations of advanced molecular markers. Major advanced techniques will be summarized here.

3.1 Targeting fingerprinting markers (TFMs)

3.1.1 Conserved DNA-Derived Polymorphism (CDDP)

The technique developed by Collard & Mackill (2009a) uses short primers to generate useful genetic markers across functional domains of well-characterized plant genes. It targets short conserved gene sequences present in the plant genome in multiple copies. Primers are specifically designed to anneal to these genes to generate polymorphic banding patterns detected on agarose gels. Theoretically any type of conserved gene region or plant gene family can be tagged using this technique. Collard & Mackill (2009a) described a set of primers that target well characterized plant genes involved in responses to abiotic and biotic stress or plant development. CDDP can easily generate functional markers (FM) related to a given plant phenotype, which is advantageous in many genetic studies. The conserved nature of priming sites in these gene regions makes the technique transferable to a wide variety of species. Since highly conserved DNA regions share the same priming site, but differ in their genomic distribution, variation can be detected as length polymorphism within these regions. The technique is based on single long primer amplifications with a high annealing temperature, which improves reproducibility. CDDP also differs in several ways from a recently developed technique called CoRAP (see 3.1.2.). Firstly, in CDDP a single primer is used in the PCR reaction, while CoRAP uses primer pairs. However, there have also been attempts to combine primers in CDDP reactions to amplify polymorphic regions representing DNA stretches between two identical or very similar conserved primer binding sites (Poczai et al., 2011). Secondly, CoRAP uses expressed sequence tags (ESTs) for specific taxa to detect polymorphism. Finally, the fragments generated by CoRAP are separated on polyacrylamide gels. This may seem to be only a minor technical difference, but the rationale is that CDDP generates fragments in the 200-1,500 bp range, while with CoRAP polymorphism manifests in small size differences in the banding patterns which require polyacrylamide gels to detect. The reproducibility of the technique has proved to be high compared to traditional RAPD (Williams et al., 1990). However, some primers problems can occur, suggesting that primer length and high annealing temperature may not ensure complete reproducibility. This indicates that scoring of banding patterns should be based on replicates and results should be treated cautiously. CDDP markers have not been widely applied in genetic diversity studies, due to their recent development and easy confusion with the other markers mentioned. They have been used to fingerprint rice varieties (Collard & Mackill, 2009a) and to assess the genetic diversity of *Solanum dulcamara* L. germplasm in Europe (Poczai et al., 2011). In these studies CDDP proved to be useful for achieving these goals.

3.1.2 Cytochrome P450 Based Analogues (PBA)

Cytochrome (Cyt) P450 mono-oxygenases are widely found in animals, plants and microorganisms (Shalk et al., 1999). In embryophytes they play important roles in oxidative detoxification and in the biosynthesis of secondary metabolites (Kessmann et al., 1990), while many P450 gene families have been found in various plant species. Sequence diversity of P450 gene-analogues in different plant species has been studied and it has been reported that such analogues can be used as genetic markers for diversity studies in plants, at both functional and genome-wide scales (Somerville & Somerville, 1999). Based on these findings a technique called Cytochrome P450 Based Analog (PBA) markers was developed by Yamanaka et al. (2003). Further, data mining of the genome of the model plant *Arabidopsis thaliana* (L.) Heynh. has resulted in the development of number of primer-sets derived from Cyt P450 genes, which could be used as universal tools for the assessment of genome-wide diversity in diverse plant species lacking relevant genetic markers (Yamanaka et al., 2003). Since Cyt P450 genes are widely distributed within the plant genome they can be universally utilized to create polymorphic fingerprints to characterize genetic diversity within and among populations of a wide verity of plant species. The genomic annotation of *Arabidopsis* revealed that out of the ~ 29,000 genes in the genome, nearly 0.9% (272 genes and 26 pseudogenes) are putative Cyt P450 genes (Riechmann et al., 2001). This indicates that these genes are very diverse, providing an opportunity to utilize them in diversity assessment. In the method developed by Yamanaka et al. (2003) universal primer pairs, designed to anneal to specific conserved exon regions of Cyt P450 genes, are arbitrarily paired. Forward and reverse primers flanking the intron regions are then used to initiate PCR amplification. Based on the random distribution of Cyt P450 genes in the genome, the resulting banding patterns will reflect polymorphism based on the variation found across the targeted (pseudo)genes. The cross-species amplification and transferability of PBAs was reported and verified in 52 different taxa from 28 families (Yamanaka et al., 2003). The generated PBA markers have diverse applications in population biology and have already been used to characterize genetic diversity in variety of species (e.g. Ahmad et al., 2009; Gilani et al., 2009; Panwar et al., 2010).

3.1.3 Intron-Targeting Polymorphism (IT)

The close proximity of introns to exons makes them especially well suited for the detection of length polymorphism in their structure that can be utilized for generating powerful new markers for genetic diversity analysis. This can be effectively achieved using the technique of intron-targeting (IT; Choi et al., 2004; Poczai et al., 2010). The basic principle of IT relies on the fact that intron sequences are generally less conserved than exons and they display polymorphism due to length and/or nucleotide variation in their alleles. Primers designed to anneal in conserved exons to amplify across introns can reveal length polymorphism in the targeted intron. Such primers can be designed for any organism using the available sequences of known genes or by exploiting expressed sequence tag (EST) records from the NCBI database. This method was first applied by Choi et al. (2004) to construct a linkage map of the legume *Medicago truncatula* Gaertn., so it represents a technological transfer from applied biology. However, it was further used to examine genetic diversity in *Solanum nigrum* L. (Poczai et al., 2010) and *S. dulcamara* (Poczai et al., 2011) populations. The results confirm that IT markers are suitable for characterizing genetic diversity. Moreover, the

designed primer is transferable across species of the same genus, or in some cases possibly family. The successful transferability and cross-species amplification capacity of IT markers will depend on the conservation of exon-intron junctions and gene structures across related genomes in the genus or family. If the shared syntenies of the targeted genes as well as their sequence features are relatively conserved, the primers can be transferred easily between higher taxonomic levels. This is valuable for generating functional markers directly related to gene regions and facilitating the discovery of specific markers linked to a given phenotype. It is also possible to tag specific genes related to environmental factors that could have useful applications in molecular ecology in the broader sense. This is because IT uses primers based on allele sequences of functionally characterized genes, and thus specific banding patterns corresponding to plant phenotypes can be identified (Cernák, 2008; Cernák et al., 2008; Gizaw, 2011). However, development of such markers depends on the availability of robust genomic databases holding several target sequences for IT marker development. Functional gene characterization criteria might be limiting factors, since it is not possible to establish gene functions in a molecular ecological sense for all genes. The crucial question is whether useful allelic variation can be identified for all genes of ecological relevance in the targeted organism.

3.1.4 Start Codon Targeted (SCoT) Polymorphism

Molecular markers from the transcribed region of the genome can offer potential for various applications in plant genotyping as they reveal polymorphisms that might be directly related to gene function. A novel marker system called Start Codon Targeted Polymorphism (SCoT) was described by Collard & Mackill (2009b), based on the observation that the short conserved regions of plant genes are surrounded by the ATG translation start codon (Sawant et al., 1999). The technique uses single primers designed to anneal to the flanking regions of the ATG initiation codon on both DNA strands. The generated amplicons are possibly distributed within gene regions that contain genes on both plus and minus DNA strands. The utility of primer pairs in SCoTs was advocated by Gorji et al. (2011). SCoT markers are generally reproducible, and it is suggested that primer length and annealing temperature are not the sole factors determining reproducibility (Gorji et al., 2011). They are dominant markers, however, while a number of co-dominant markers are also generated during amplification, and thus they could be used for genetic diversity analysis (Collard & Mackill, 2009b). This has been validated through study of genetic diversity among rice varieties (Collard & Mackill, 2009b). Further studies have utilized the technique to investigate genetic relationships in a number of plant species (Xiong et al., 2011; Murthy, 2011). These studies show that SCoTs can be used solely, or in combination with other techniques, to assess genetic diversity and to obtain reliable information about population processes and structure across different plant families.

3.1.5 Sequence-Related Amplified Polymorphism (SRAP) and Targeted Region Amplified Polymorphism (TRAP)

With primers targeting short recognition sites a large number of polymorphisms can be generated. In these cases almost any primer can initiate PCR amplification. However, due to the nature of PCR itself, short primer pairs will amplify many sequences distributed throughout the genome. This is a feature of the widely used RAPD technique (Williams et

al., 1990). The main disadvantages of this technique are poor consistency, low reproducibility and multiplexing of output, which limit its use. The sequence-related amplified polymorphism technique (SRAP), developed by Li & Quiros (2001), also uses arbitrary primers to generate a specific banding pattern, originally detected on polyacrylamide gels. Compared to RAPD, the primers used are much longer (17-21 nt). The forward primers are designed to contain GC-rich sequences near the 3'-end, while reverse primers contain AT-rich sequences at the 3'end. This is based on the rationale that protein coding regions tend to contain GC-rich codons, while 3'UTRs frequently consist of AT-stretches (Lin et al., 1999). More specifically, "CCGG" sequences in the core of the forward primer target exon regions. Li & Quiros (2001) found that in their experiments with randomly selected bacterial artificial colonies (BACs), 66% of the produced sequences of exon regions contained the "CCGG" motif. Lin et al. (1999) noted that approximately one-third of the part of the *Arabidopsis* genome found on chromosomes 2 and 4 represents exon regions. With the inclusion of the "CCGG" core in the primer set, sequences containing this element are preferentially amplified. However, because exons are generally conserved and might fail to produce sufficient polymorphism, the reverse primer in SRAP is designed to contain a second core with the aforementioned "AATT" motif. This second core is normally a frequent element in promoters, introns and spacers (Lin et al., 1999). Since these regions are more variable between different individuals, the intrinsic dissimilarity incorporated in the primer sets makes it feasible to generate polymorphic bands based on introns and exons (Li & Quiros, 2001). The arbitrary primers also contain further modifications, as they have 10 bases at the 5'end called filter sequences that have no specific constitution. These are followed by the core sequences (CCGG for forward and AATT for reverse), while at the 3'end three selective nucleotides are added. The PCR profile is also modified to ensure specificity and high stringency. The profile consists of two parts: the early and late cycles. During the early cycles mismatch is tolerated by the use of a lower annealing temperature (35°C), enabling the production of many amplicons. The rationale behind this is that primer annealing to the DNA template depends on the matching level of both sequences, and amplification efficiency is determined by the effectiveness of primer binding capability (Li & Quiros, 2001). The low initial annealing temperature ensures the binding of both primers to sites with partial matches in the target DNA, creating a population of amplicons that contain the priming sites. During the late cycles at higher annealing temperature (50°C), the initially generated amplicons serve as the template rather than the genomic DNA, ensuring high reliability, efficiency and reproducibility due to perfect base pairing of primers with the template. Because mismatches are allowed in the early cycles, the 5'ends of the PCR primers are usually "forced" into the PCR products. This is similar to *in vitro* mutagenesis using PCR primers (Cadwell & Joyce, 1992). The limiting factors in amplification are the 3'end sequences of the primers, which must match perfectly during the PCR cycles (Telenius et al., 1992). Therefore polymorphisms are limited by the 3' ends of the primers, leading to exclusive amplification of alleles matching the 3' ends, while alleles with mutations in these regions will not amplify. With changes in the selective nucleotides at the 3' ends, new polymorphic patterns can be observed. The technique has gained popularity due to several advantages: i) a large number of polymorphic fragments are amplified in each reaction, ii) there is no *a priori* need for information about sequences, iii) primers can be applied to any species, iv) it is cost effective and easy to perform, v) reproducibility is high, and vi) PCR products can be directly sequenced using the original primers without cloning. These

advantages have led to its widespread application for genetic diversity analysis in diverse organisms (e.g. Sun et al., 2006; Feng et al., 2009; Song et al., 2010).

The Targeted Region Amplified Polymorphism technique (TRAP) developed by Hu & Vick (2003) is similar to SRAP, but it incorporates the advantage of the availability of sequence information. It uses information from partial sequences of candidate genes in a fixed primer, while an arbitrary primer is used to amplify the putative candidate gene regions. The PCR cycling conditions are similar to SRAP, with the priming and amplification procedure having the same rationale. Arbitrary primers incorporate the characteristics of SRAP primers, having either AT- or GC-motifs in their sequences, with 5′ end filter sequences and 3′ end selective nucleotides. The major difference lies in the other primer pair, called the fixed primer. This primer is specifically designed from ESTs of the analyzed organism to target candidate gene regions. The generation of fixed primers limits the use of this technique to species where ESTs are known, or requires the generation of new sequence information for primer development. Despite this limitation it has been widely used for multiple purposes in different plant species (e.g. Hu et al., 2005; Alwala et al., 2006; Hu et al., 2007).

3.1.6 Conserved Region Amplification Polymorphism (CoRAP)

The CoRAP technique (Wang et al., 2009) is based on the use of two primers, a fixed and an arbitrary one, to detect polymorphism. The fixed primer is derived from directly targeted ESTs in repository sequence databases (e.g. Genbank), while the arbitrary primer contains a core sequence motif (CACGC) commonly found in plant gene introns. This core sequence ensures the utilization of conserved intron sequences in plant genotyping while the fixed (conserved) primers target coding sequences, together generating highly reproducible and reliable fingerprints. Generally, fixed primers are designed specifically for the analyzed organism, meaning that *a priori* sequence knowledge is required. The technique is similar to TRAP (see 3.1.5.) but differs in the incorporation of sequence motifs in the arbitrary primer. Fixed primers derived from ESTs will have a specific binding site on the exon of the target sequence, while the arbitrary primers will bind to most of the introns during the PCR amplifications. If the distribution of these gene elements allows successful PCR, banding patterns resulting from a specific fingerprint will be amplified. Indels in these regions will certainly generate different distributions of amplified production. The closer the genetic relationship between the two individuals, the more similar the corresponding band patterns of the amplified PCR products will be (Wang et al., 2009). The technique has been used successfully to identify economically important traits in sugarcane (Khan et al., 2011).

3.2 Mobile element based molecular markers

The following techniques can be regarded as a special group of TFMs, since they all utilize mobile elements of the plant genome to generate fingerprints. Mobile element based markers have great potential as tools for investigating aspects of molecular ecology, including population structure, conservation genetics, the genetics of speciation, phylogeography and phylogeny (Ray, 2007). Currently they remain under-utilized tools in molecular ecology. In particular, one group of mobile elements, retrotransposons, provides an excellent basis for the development of markers systems. Retrotransposons replicate by successive transcription, reverse transcription and insertion of the new cDNA copies back

into the genome, very much like retroviruses (Scheifele et al., 2009). The structure and replication strategy of retrotransposons give them several advantages as markers (Kalendar et al., 1999). Firstly they are ubiquitous, present in high copy numbers as highly heterogeneous populations that are widely dispersed on chromosomes and show insertional polymorphism both within and among plant taxa (Kumar et al., 1997). Secondly, active mobile elements produce new insertions in the genome, leading to polymorphism which can be detected and used to temporally order insertion events in a lineage (Kumar & Bennetzen, 1999). Thirdly, many types of mobile elements are widely distributed in the euchromatin domains of chromosomes, making it possible to generate markers linked to a given phenotype (Kenward et al., 1999). Fourthly, an important trait of mobile element markers is that they provide homoplasy-free characters for genetic diversity assessments (Ray et al., 2006). Character states are clearly derived from a common ancestor and they are almost invariably identical by descent, but not identical by state (Ray, 2007). Fifthly, their ancestral state is known and stable, which means that the ancestral state at any amplified locus is the absence of the element, and once the element is present it will almost invariably remain there indefinitely (Shedlock & Okada, 2000). Finally, mobile element based markers are expected to be co-dominant. However, despite the fact that they are extremely useful for population genetics, all mobile element based markers have the same drawback: difficulty of data interpretation and uncertainty about the true nature of the polymorphism. Specifically, the question may arise as to whether differences in banding patterns are due to the absence or presence of retrotransposons, or are caused by some other mechanism, e.g. indels or restriction site loss. Fortunately, advances in analytical methods and a number of successful studies indicate that these drawbacks can be overcome.

3.2.1 Inter-Retrotransposon Amplified Polymorphism (IRAP) and Retrotransposon-Microsatellite Amplified Polymorphism (REMAP)

Both IRAP and REMAP were first described by Kalendar et al. (1999) for generating specific fingerprints, i.e. distinctive fragment patterns in samples. They both target a group of retrotransposons that contain direct long terminal repeats (LTRs) of variable size from 100-5,000 bp (Kumar & Bennetzen, 1999). LTRs do not code for any protein but instead contain the promoters and terminators for transcription. IRAP primers anneal to these regions and amplify DNA segments between two LTR sequences. Either one or two primers specifically designed for LTRs can be used in the same PCR, but the results will be determined by the orientation of these regions. Retrotransposons spread through the generation of new copies, rather than moving directly by excising themselves from their original position and "jumping" to a new locus. These newly generated daughter copies are then integrated back into the genome, which can happen in either orientation (5' to 3' or 3' to 5'). This leads to the generation of differently oriented gene (copy) clusters which may be found in head-to-head, tail-to-tail or head-to-tail arrangements. These can be repeat type daughter sequences (head-to-tail), when the gene copy of the mobile element is directly repeated after the original copy. In other words, the two copies are in the same strand and are transcribed in the same direction. Head-to-head or bidirectional gene (copy) pairs are inverted orientations and describe a genomic locus where two adjacent genes (copies) are divergently transcribed from opposite strands of DNA, with the region between two transcription start sites (TSSs) being shared by the gene and used as a putative bidirectional promoter. In other words, the two adjacent genes are in two different strands separated by a short intergenic distance and

are oriented in a divergent (-+) transcription configuration, meaning that they are transcribed away from each other. The tail-to-tail inverted orientation is the opposite of head-to-head orientation, where the two genes (copies) are located in opposite strands and are transcribed towards each other. For head-to-head and tail-to-tail arrangements, only a single primer is necessary to generate IRAP products. For head-to-tail orientation, both 5′ and 3′ LTR primers must be used to amplify the intervening genomic DNA (Kalendar et al., 1999). The other technique, REMAP, exploits the polymorphisms among regions amplified between an anchored simple sequence repeat (SSR) and an LTR sequence. To achieve this, one specifically designed LTR primer is mixed with another arbitrarily chosen primer directly containing a simple repeat [e.g. $(CA)_n$, $(GA)_n$] plus an additional and randomly chosen anchoring nucleotide in the 5′ or 3′ end [e.g. $C(CA)_n$, $(GA)_nG$]. This technique can be regarded as a modified or extended version of the inter-simple sequences repeat (ISSR) technique, since one primer in a REMAP reaction is an anchored ISSR primer combined with an IRAP primer. IRAP and REMAP have been used individually and in combination to study genetic diversity in several plant genera (e.g. Bretó et al., 2001; Branco et al., 2007; Carvalho et al., 2010), because they produce a reliable and reproducible banding profile.

3.2.2 Retrotransposon-Based Insertion Polymorphisms (RBIP)

It is also possible to tag the presence or absence of a given complete retrotransposon insertion at a specific locus and to use this information as a genetic marker. A PCR-based approach has been developed to detect such individual insertions by Flavell et al. (1998), based on a method originally developed to detect transposition of the PDR1 element in pea (*Pisum sativum* L.). The technique generates co-dominant markers, because the different allelic states caused by the presence or absence of an insertion are scored at a single locus (Kumar & Hirochika, 2001). This is achieved by PCR amplification using primers flanking the exact insertion point of the retrotransposon. Primers are designed for both the 5′ and 3′ end regions flanking the targeted retrotransposon insertion and used to amplify a product if the insertion is absent (Grzebelus, 2006). Further primers are generated from the mobile element, mostly from LTRs, and used together with a primer from the flanking region to detect the presence of the insertion. The results are then scored on agarose gels or by dot hybridization. For agarose detection, flanking primers can be combined with element based primers in a triplex-PCR for genotyping, so that detection of the occupied and unoccupied site is performed in a single tube reaction. The main advantage of this method is that it can be fully automated and applied to many samples in a high throughput system, e.g. the tagged microarray marker (TAM) system (Flavell et al., 2003). This microarray based method enables scoring of thousands of DNAs for RBIP markers on a glass slide, based on fluorescent labeling (Kalendar, 2011). The method does not suffer from the drawbacks that all gel-based markers have, particularly band sharing from one gel to another. However, the insertions of RBIPs can also be detected on agarose gels. Because of the investment needed to sequence flanking regions it is a more costly technique than those based on other mobile element based markers. In addition, it can be technically more complicated. RBIP markers have been utilized in a number of plant taxa for genetic diversity analysis, mostly in crop species where it has been possible to take advantage of sequence information on genomic structure provided by molecular breeding programs (Vershinin et al., 2003; Vitte et al., 2004; Jing et al., 2010).

3.2.3 Retrotransposon-based Sequence-Specific Amplification Polymorphism (SSAP)

The principle of this method is to convert a subset of retrotransposon based insertion sites into a corresponding set of bands on a gel, by PCR amplification of the junctions between the transposon and the host genome (Grzebelus, 2006). This technique was developed by Waugh et al. (1997) to investigate the location of the BARE-1 retrotransposon in the barley genome. The basic principle is similar to the commonly used amplified fragment length polymorphism technique (AFLP; Vos et al., 1995) – it is an extension of AFLP with generation of mobile element based molecular markers. As in AFLP, the procedure begins with the digestion of genomic DNA using a rare cutting restriction enzyme paired with a frequently cutting one (usually *Mse*I and *Pst*I, or any other restriction enzyme). The sequence of the retrotransposon is known *a priori* and the transposon family, group and type is carefully chosen before starting the reactions. The restricted DNA is then "tagged" by the ligation of short double-stranded adapters with known sequences. The ligation phase is followed by a pre-selective PCR amplification with adaptor–homologous primers. This pre-amplification is useful to reduce the genome complexity by amplifying a subset of restriction fragments when work is performed with organisms having a large genome. This step also promotes higher reproducibility of the downstream SSAP banding pattern. Pre-PCR with selective bases reduces the number of bands amplified, and this step can also be introduced to reduce the restriction fragments further if necessary (Waugh et al., 1997). The next step is selective amplification with a retrotransposon specific primer, paired with either a rare or a frequent site adaptor homologous primer. Primers are usually designed for the LTR region, but could also correspond to an internal part of the mobile element. This generates an amplified fragment for each of the transposon host junctions. The size of the fragments are determined by the distance between the transposon insertion site and the adjacent restriction cut site and differences in insertion sites between genomes are easily visible as different banding patterns (Syed & Flavell, 2006). These can by visualized on polyacrylamide gels, or a transposon specific primer can be fluorescently labeled to ensure detection. SSAP systems commonly utilize retrotransposons with LTRs from the *Ty1-copia* or *Ty3-gypsy* retrotransposon groups. The technique has been successfully used in a number of plant groups for phylogenetic and genetic diversity studies (e.g. Tam et al., 2005; Luo & Chen, 2007; Moisy et al., 2008; Ragupaty et al., 2010). It is important to note that retrotransposon families for tagging must be chosen carefully, because they may differ in the amount of information (variation) that they provide.

4. Beyond molecular marker diversity: Statistical analysis of polymorphism data

The techniques discussed in the previous section produce a number of reliable and unambiguous sets of bands, easily detectable with gel electrophoresis or by high throughput systems with fluorescently labeled primers. Polymorphism data can be scored in presence/absence matrices manually or with the aid of specific software. However, because these techniques are based on the incorporation of genomic elements in the primer sets or else target specific regions in the genome, biases affecting the evaluation process can occur. Although many recently developed targeting methods detect large numbers of polymorphisms, not many studies to date have utilized them, largely due to their unfamiliarity. In many cases the drawbacks are unknown. These mainly affect the analysis

of the banding patterns produced, largely depending on the nature of the methods and whether they generate dominant or co-dominant markers. Here we give a brief description of common analytical approaches that can be applied to statistically evaluate the polymorphism data produced by recently developed targeting methods. Until such time as further studies reveal major weaknesses of targeting fingerprinting techniques, we assume that these drawbacks do not cause serious biases in the data sets produced. Our descriptions are based on those provided by the developers of the discussed techniques as detailed in the original studies cited, or on our own experiences with the application of some of the marker systems (Poczai & Hyvönen, 2011; Poczai et al., 2011; Gorji et al., 2011). There are two features of targeting fingerprinting markers (TFM) that considerably constrain how they can be statistically analyzed. Firstly, polymorphic TFM loci are generally scored for two alleles, the "band-presence" allele and the "band-absence" allele. This makes each locus less informative than a multi-allelic microsatellite locus. However, the large number of TFMs available across the genome, or the specific exploration of multiple regions and their distributions, can counteract this drawback. Secondly, the produced banding patterns for each technique are treated during the evaluation process as dominant markers, and the methods are described accordingly. Thus, it is difficult to distinguish heterozygous individuals from homozygous ones with respect to the band-presence allele, unless exact genotypes can be inferred from pedigree studies (van Haeringen et al., 2002). However, this assumption may be challenged in future and methods appropriately modified if the technical nature of each marker system is properly recognized. To date only a few studies have been reported on TFMs, which may prove to be co-dominant (Poczai et al., 2010), or with mixed banding patterns of dominant and co-dominant markers (Gorji et al., 2011). In the following sections some commonly used basic statistical approaches for measuring genetic diversity will be discussed.

4.1 Band-based approaches

The easiest way to analyze banding patterns and to measure the diversity represented by them is to focus on band presence or absence and to compare it between samples. This method is routinely used at the level of the individual, and produces distances generated from the bands rather than taking into account the diversity of the population. The disadvantage is that it treats the data as polymorphism produced by the markers rather than describing the genetic diversity of the organism. Although the results correlate with genetic diversity, the indices often have no biological meaning in a population–genetic sense.

4.1.1 Measuring polymorphism

A locus is considered polymorphic when the band is present at a frequency of between 5% and 95%. The polymorphism represented in the banding patterns can be extracted simply by observing the total number of polymorphic bands (PB) and then calculating the percentage of polymorphic bands (PP) present in any individual based on this number. The value of a particular band position can also be measured by its similarity to the optimal condition, which is 50% of genotypes containing the band. This "band informativeness" (I_b) can be represented on a scale ranging from 0 to 1 according to the formula: $I_b = 1 - (2 \times |0.5 - p|)$, where p is the portion of genotypes containing the band.

4.1.2 Shannon's information index (*I*)

A coefficient that is often called the Shannon index of phenotypic diversity is widely applied to quantify the degree of band polymorphism provided by banding patterns. It can be calculated as $I = -\sum p_i \log_2 p_i$, where it is assumed that i) the population is large, ii) random mating takes place , iii) the population is isolated from others, and iv) the effects of migration, mutation and selection are not relevant. Due to biallelic coding, some specific features also need to be taken into account. Because of these drawbacks, allele-frequency statistics are commonly supplemented with another estimation that allows for the distortion from HWE (see 4.3). These methods depend on the extraction of allelic frequencies from the data, which can be accomplished using a number of different procedures (see Lynch & Milligan, 1994; Stewart & Excoffier, 1996; Zhivotovsky, 1999; Hill & Weir, 2004; Holsinger et al., 2002). The accurate estimation of frequencies essentially determines the accuracy of the different indices calculated for further quantification of genetic diversity.

4.1.3 Similarity coefficients

Band-based approaches usually utilize similarity or dissimilarity (the inverse of the previous one) coefficients. Many coefficients have been developed, but they all have common properties that they incorporate the pattern of matches and mismatches among their respective band presences and absences. We will just only discuss the three most widely applied indices in this section. Abbreviations will be used for both three as the following: *a* = the number of bands (1s) shared by both individuals; *b* = number of positions where individual *i* has a band, but *j* does not; *c* = number of positions where individual *j* has a band, but *i* does not; and *n* = the total number of bands (0s and 1s). The following measures are commonly used for comparison between individuals: i) **Jaccard similarity coefficient** or Jaccard index $J = a/(a + b + c)$ (Jaccard, 1908); ii) **Simple matching coefficient** or index $SM = (n - b - c)/n$ (Sokal & Michener, 1958); iii) **Sørensen-Dice index** or Nei & Li index $SD = 2a/2a + b + c$ (Dice, 1945; Sørensen, 1948; Nei & Li, 1979). The Jaccard coefficient (*J*) only takes into account the bands present in at least one of the two individuals. It is therefore unaffected by homoplasic absent bands (where the absence of the same band is due to different mutations). The simple-matching index (*SM*) maximizes the amount of information provided by the banding patterns considering all scored loci. Doubling the band absence and presence receive the same importance, which is not the case for frequent band absence homoplasy. The Neil & Li index (*SD*) doubles the weight for bands present in both individuals, thus giving more attention to similarity than dissimilarity.

4.2 Allele frequency based approaches

The methods discussed here measure variability by describing changes in allele frequencies for a particular trait over time. Using the presence/absence dominant data allele frequencies are estimated at each locus. This strategy is more population oriented, than band-based approaches; however, data is treated as dominant for each locus and thus allele frequencies are accessible only with preliminary assumptions made additionally. Moreover, they expect that the population is in Hardy-Weinberg equilibrium (HWE), it is assumed that the i) population is large, ii) random mating takes place , iii) the population is isolated from others, and iv) the effects of migration, mutation and selection are neglected. Due to biallelic coding some specific features should also be taken into account. Because of these drawbacks

allele-frequency statistics are commonly supplemented with other estimation which allows the distortion from HWE (see 4.3). These methods depend on the extraction of allelic frequencies from the data. This can be done according to many procedures (see Lynch & Milligan, 1994; Stewart & Excoffier, 1996; Zhivotovsky, 1999; Hill & Weir, 2004; Holsinger et al., 2002). The accurate estimates of frequencies essentially influence the results of different indices calculated for further measurements of genetic diversity.

4.2.1 Allelic diversity (*A*)

One of the easiest ways to measure genetic diversity is to quantify the number of alleles present. Allelic diversity (*A*) is the average number of alleles per locus and is used to describe genetic diversity. When there is more than one locus, *A* is calculated as the number of alleles averaged across all loci: $A = n_i / n_l$ where n_i is the total number of alleles over all loci and n_l is the number of loci. The observed number of alleles (n_a) can also be given by counting the number of alleles per locus and taking the average of these counts. These values then can be compared with the effective number of alleles (n_e), which is the number of alleles needed to provide the same heterozygosity if all alleles were equally frequent. It is less sensitive to sample size and rare alleles, and calculated as $n_e = 1 / \sum p_i^2$ pi2 ability, so it is expected to provide information about the dispersal ability of the organism and the degree of isolation among populations.

4.2.2 Effective population size (*N*$_e$)

The concept of effective population size (N_e) plays an important role in conservation management. It provides a measure of the rate of genetic drift, the rate of genetic diversity loss and increase of inbreeding within a population. The effective population size by definition is the number of individuals that would give rise to the calculated inbreeding coefficient, loss of heterozygosity, or variance in allele frequency frequency if they behaved in the manner of an idealized population. (Frankham et al., 2002). In other words, it is the size of an idealized population that would lose genetic diversity (or become inbred) at the same rate as the actual population. It is a value which depicts the general core number individuals in a population who contribute offspring to the next generation. The effective population size is usually smaller than the absolute population size. The major factors effecting Ne are i) fluctuating population size, ii) breeding sex ratio, iii) overlap of generations and iv) spatial dispersion. There are many formulas to calculate effective population size that will not be discussed here due to their specific nature. In should be noted that the effective size of a population is an idealized number, since many calculations depend on the genetic parameters used and on the reference generation. Therefore, a single population may have many different effective sizes which are biologically meaningful but distinct from each other.

4.2.3 Heterozygosity (*H*)

Heterozygosity can be regarded as the average portion of loci with two different alleles at a single locus within an individual. It is commonly extended to the whole or a sub-portion of an entire population and differentiated into observed and expected heterozygosity. Expected heterozygosity (H_E) or Nei's gene diversity (*D*) is the expected probability that an individual will be heterozygous at a given locus, in multi-locus systems over the assayed loci. In other words, it is the estimated fraction of all individuals that would be

heterozygous for any randomly chosen locus. It is often calculated based on the square root of the frequency of the null (recessive) allele as follows: $H_E = 1 - \sum_i^n p_i^2$, where p_i is the frequency of the ith allele. Observed heterozygosity (H_O) is the portion of genes that are heterozygous in a population. It is calculated for each locus as the total number of heterozygotes divided by sample size. Typically values for H_E and H_O range from 0 (no heterozygosity) to nearly 1 (a large number of equally frequent alleles). Expected heterozygosity is usually calculated when describing genetic diversity, as it is less sensitive to sample size than observed heterozygosity. If H_O and H_E are similar (they do not differ significantly) mating in the populations is approximately random. If H_O is less then H_E, the population is inbreeding; if H_O exceeds H_E, the population has a mating system avoiding inbreeding.

4.2.4 *F*-statistics

In population genetics the most widely applied measurements besides heterozygosity are *F*-statistics, or fixation indices. These were originally designed to measure the amount of allelic fixation by genetic drift. *F*-statistics are used to describe the structure of the population at different levels. The measure of Wright's index is based on the comparison of frequencies of identical alleles within and between groups. It can describe the inbreeding (coefficient) of an individual relative to the total population (F_{IT}) or the inbreeding (coefficient) of an individual relative to the subpopulation (F_{IS}), and can also express the "fixation" (index) resulting from comparing subpopulations to the total population (F_{ST}). In other words, different *F*-coefficients explain the correlation of genes within individuals over all populations (F_{IT}), the correlation of genes of within individuals within populations (F_{IS}), and the correlation of genes of different individuals in the same population (F_{ST}). For a given species this can be interpreted as the overall inbreeding (F_{IT}), the inbreeding within taxa (F_{IS}) and the coefficient of co-ancestry (F_{ST}), which provides an estimate of interspecific genetic differentiation. The *F*-statistics are related to heterozygosity and genetic drift. Since inbreeding increases the frequency of homozygotes, as a consequence, it decreases the frequency of heterozygotes. In other words, population substructure results in a reduction of heterozygosity. The general result of genetic drift in small populations is a deficiency of heterozygotes and an excess of homozygotes. Moreover, mating between relatives (inbreeding) and genetic drift are related phenomena, which allow us to determine fixation indices when expected and observed heterozygosity are known (see 4.2.4). Thus the three indexes can be calculated as: $F_{IT} = 1 - (H_I/H_T)$; $F_{IS} = 1 - (H_I/H_S)$ and $F_{ST} = 1 - (H_S/H_T)$ where H_I is the average observed heterozygosity within each population, H_S is the average expected heterozygosity of subpopulations assuming random mating within each population and H_T is the expected heterozygosity of the total population assuming random mating within subpopulations and no divergence of allele frequencies among subpopulations. Fixation indices are interrelated as follows: $1 - F_{IT} = (1 - F_{ST})(1 - F_{IS})$ and thus $F_{ST} = (F_{IT} - F_{IS})/(1 - F_{IS})$. They vary from 0 (no isolation) to 1 (complete isolation), but in the case of a bias towards polymorphic loci, lower values (e.g. 0.5) may indicate complete isolation between populations.

4.2.5 Gene flow (*Nm*)

The parameter of gene flow (*Nm*) is the product of the effective size of individual populations (*N*) and the rate of migration among them (*m*). However, it is difficult to

measure the rate of gene flow directly by tracking individuals and also to mathematically infer its value. Many estimates have been developed concerning particular types of gene flow that will not be discussed here. However, commonly used approach is to make an estimation based on the degree of genetic differentiation among populations by calculating F-statistics (detailed in 4.2.5). F_{ST} is related to population size and migration rate, based on a standard island model, which is not applicable to every natural population, but provides an easy way to calculate gene flow as the following: $Nm = (1/F_{ST} – 1)$. Gene flow among fragmented populations is related to dispersal ability, so it is expected to provide information about the dispersal ability of the organism and the degree of isolation among populations.

4.3 Bayesian estimates of genetic diversity

Classical population genetic methods are based on the calculation of various values as discussed in the previous sections. The principles of Bayesian inference rely on the perspective that both model parameters and data are random variables with a joint probability distribution. In general terms, it is a statistical inference method where the evidence is used to revise the ambiguity of parameters and predictions in probability models. In these estimates, the initial belief is termed the *prior*, while the modified belief is the *posterior*. The prior summarizes the initial information about the value of a parameter or hypothesis, before the data is analyzed to produce a probability distribution. The aim of Bayesian calculations is to calculate the *posterior* distribution of the parameters given the data. These functions are useful since no natural population can possibly meet all of the requirements of HWE. However in most cases, allele frequency based methods provide *a priori* knowledge of what dynamics may influence the population. Later, these calculations can be refined making no assumption about HWE and estimating the effective size of populations, levels of heterozygosity, and inbreeding using Bayesian methods. Given that calculation of allele frequencies based on dominant markers requires the assumption of HWE, a more direct estimate of the relative magnitude of population divergence can be obtained using the Bayesian approach. These methods do not assume that genotypes are in Hardy-Weinberg proportion within the population, while taking full advantage of the information provided by molecular markers (Holsinger & Wallace, 2004). Bayesian methods are favored for estimating parameters such as mutation rate (μ) and effective population size (N_e) in demographic models. Bayesian assignment methods are also used to study population differences and to infer parameters in genealogical models concerning coalescent theory, and to detect selection. Other areas of Bayesian computation include inferring changes in population size, analyzing population structure, and even identifying haplotypes through population samples. In many cases, Bayesian methods can address the question of interest more directly than classical approaches. Consequently they are an essential part of genetic diversity estimation and population genetics. Bayesian inference allows the incorporation of complicated models that can be studied, and biologically relevant parameters that can be estimated with the incorporation of prior information. In addition to allele frequency based metrics, the posterior distribution of almost every classically used parameter can be estimated using Bayesian models. In this regard they have Bayesian analogues, e.g. Bayesian fixation index, θ^B (equivalent of Wright's F_{ST}), Bayesian inbreeding, f (the estimate of F_{IS}), etc. A detailed discussion of available Bayesian methods in genetic diversity assessment is beyond the scope of this chapter due to their

complexity, but they have been reviewed for example by Shoemaker et al. (1999) and Beaumont & Rannala (2004).

5. Phylogenetic methods and phylogenetic diversity

Besides genetic diversity estimates it is often important in molecular ecological studies to represent the genetic structure of the population(s) examined. Such information is extremely valuable because it can address many important issues such as estimation of migration, identification of conservation units, or even resolution of biogeographical patterns (Mantel et al., 2005). When geographical locations are available for populations and sampling allows the detection of spatial structure, then tree-based methods should be considered. For many taxa the demarcation of populations is problematic when there is no *a priori* knowledge of population structure (Hollingsworth & Ennos, 2004). In these cases and in general, tree-based methods allow graphical representation of the relatedness of individuals and exploration of information on the (spatial) structure of populations. After identifying population structure the results can be evaluated in the light of biological phenomena, e.g. ancient vicariance, recent dispersal or other biogeographical patterns, genetic differentiation or even migration-drift equilibrium. These methods might be used to describe the history of a population over a long or a short time period depending on the design of the study. Various population processes can be traced and identified, such as range expansion or contraction (Tang et al., 2008), responses to environmental or climatic conditions including climate change (Schmidt & Jensen, 2000; Jump et al., 2006), sources of colonization and origins of founder populations (Gladieux et al., 2011). The simultaneous evaluation of the results of tree-based methods coupled with the distribution of individual markers within and among identified groups provides some information on the distribution of genetic diversity (Bonin et al., 2007). Genetic diversity is assumed to be higher in refugial areas or in meeting zones, while a large number of separate fragments might indicate ancient divergence (Schonswetter & Tribsch, 2005). More recent advances allow the addressing of questions such as "how many populations exist in the study area?" or "is this individual a migrant from another population?" (see Mantel et al., 2005). In the following section some basic tree-based methods will be briefly discussed in relation to the above mentioned goals.

5.1 Phylogenetic analyses

All life on Earth is connected through shared history. Reconstruction of this history, phylogeny, is one of the most difficult challenges in contemporary biology. However it is well worth the effort, due to the considerable benefits that robust phylogenetic hypotheses can provide for diverse fields of biology, both basic and applied. Because even distantly related organisms are surprisingly similar genetically and share the same regulatory genes, it would be counterproductive not to take advantage of the full range of variation produced by this great experiment conducted over billions of years. Robust phylogenetic hypotheses are essential tools for sampling this variation efficiently. In recent decades, phylogenetics has been revolutionized by (i) a re-evaluation of its fundamental theoretical premises, beginning with the seminal publications of Hennig (1950, 1966), (ii) the routine use of molecular sequence level data, (iii) the development of efficient new algorithms, and (iv) the availability of parallel computing resources allowing the use of large datasets (and the study of distant history). The latter two developments especially have enabled studies to be

undertaken (e.g. Goloboff et al., 2009) that would not have been feasible only a few years ago. We are now able to provide robust phylogenetic hypotheses that can be used as reliable starting points for further studies that benefit practically all fields within the life sciences.

Phylogenetic analyses are now routinely performed and are an integral part of systematics. The methods and programs used are diverse, and it is beyond the scope of this treatment to list them comprehensively or to discuss the pros and cons of various fundamentally different methodologies. The primary division in contemporary approaches is in the amount of background knowledge they incorporate. Classical methods are non-statistical, and use parsimony as an optimality criterion. This approach is largely based on the numerous original contributions by Farris et al. (e.g. Kluge & Farris, 1969; Farris, 1970 and Farris et al., 1970), as summarized by Farris (1983). What is assumed *a priori* is only descent with modification. A fundamentally different approach, following Felsenstein (1973) and based on maximum likelihood, incorporates explicit statistical assumptions in the analyses. These model-based methods also include those based on Bayesian argumentation, pioneered by Edwards (1970) and Farris (1973) and lately further developed by e.g. Li et al. (1996) and Mau & Newton (1997). However, both of these approaches are set apart from purely algorithmic distance methods (e.g. Michener & Sokal, 1957) by their discrete nature, i.e. they operate directly on characters rather than on distances calculated from them (Page & Holmes, 1998).

The search for optimal trees, irrespective of the optimality criterion used, is computationally demanding for real data sets and discrete characters (the problem is NP-hard), and thus only heuristic searches are feasible. The advent of molecular characters has exposed another, sometimes equally difficult task in phylogenetic analyses, namely alignment. As succinctly noted by Higgs & Attwood (2005) „...the problems of multiple alignment and tree construction are inextricably linked." The most thorough treatment of this and of the related problems has been provided by Wheeler (e.g. 1996, 1999, 2003a, b).

The methods listed above are further united in that the trees produced are connected graphs without cycles. While macroevolution can be efficiently described using this type of divergent tree, it would be unrealistic to ignore the merging of lineages, i.e. hybridization, as one source of contemporary diversity. Although the methods described above can only indirect hint at hybridization, recently methods that also attempt to consider networks as phylogenetic hypotheses have been developed (e.g. Bryant & Moulton, 2004; Huson & Bryant, 2006). Their development is, however, still in its infancy. It should be noted that finding an optimal network is computationally an extremely demanding task – even for a pair of trees it has been shown to be NP-hard (Bordewich & Semple 2005, 2006). Thus the currently available methods are purely algorithmic, i.e. they do not even attempt to find the optimal solution.

5.2 Phylogenetic diversity

Traditionally there has been a conceptual gulf between research into genetic diversity and research into phylogeny. This partially reflects the differing concerns and working methods of biologists interested in genetic structure at and below the species level, where relationships are often reticulate, and those looking at variation between monophyletic groups, where hierarchical (tree-like) relationships predominate. Because conservationists

and others interested in quantifying diversity have often come from ecological backgrounds, the phylogenetic structure of genetic variation in populations, taxa and ecosystems has frequently been neglected. A knowledge of this structure is essential for understanding what the genetic diversity within a given species (or, more ideally, a least-inclusive taxonomic unit, LITU; Pleijel & Rouse, 2000) is likely to represent as part of the total diversity within any more inclusive group. Thus it is also essential for any informed conservation strategy where the units of interest are discrete populations, species or higher-level taxa, or for optimizing any search strategy for genetic resources within a group of LITUs or higher-level taxa.

Faith (1992) showed how the sum of the branch lengths on a minimum-spanning path between any set of terminals on a cladogram could be used to calculate a metric he termed Phylogenetic Diversity (PD) for that set. As a simple example, take an isolated species with a very long branch that is sister to a clade of 20 species, also with a long branch but with the branches separating the species all very short, perhaps due to a recent, rapid adaptive radiation event. The PD of the first species would simply be its branch length, while the total PD of its sister clade would be the sum of the branch lengths connecting all terminals within that clade (including the single stem branch). While the much more speciose clade of 20 species might have a somewhat higher PD than its isolated sister species, it would only be necessary to sample one or two members of the clade to account for a very high proportion of that total PD, whereas neglecting to sample the single isolated species would miss a large amount of PD. Assuming that variation in the data used to construct the cladogram adequately reflects total genetic variation among the species, then PD effectively measures the proportion of diversity represented by any subset of terminals relative to any other by precisely accounting for shared diversity (represented as synapomorphy on the cladogram). In other words, maximizing PD when selecting a subset of genetic diversity to conserve or to explore for genetic resources is a way of maximizing the amount of hierarchical variation being considered and avoiding repeated sampling of similar or identical diversity derived from shared ancestry.

From a genetic conservation perspective, PD is important not only because species in isolated, monotypic lineages usually represent much more unique genetic diversity than ones in younger, more speciose groups, but also because the very nature of these species means that they may actually be disproportionately threatened by extinction (Purvis et al., 2000). Isaac et al. (2007) devised a measure they termed Evolutionary Distinctiveness (ED), which represents the relative contribution of an individual species to a clade's PD (with the standard PD measure, shared PD is included in the PD of any subset of terminals) . They further showed how extinction risk (derived from Red List categories; IUCN, 2011) could be combined with ED to identify species that are both Evolutionarily Distinct and Globally Endangered ("EDGE species"). Significantly, they also showed that ED scores (and by extension EDGE scores) are robust to clade size, missing species and poor phylogenetic resolution, issues that have been seen as obstacles to the widespread application of PD in practical biodiversity assessment in the past.

Often it is desirable to quantify genetic diversity on a geographic as well as a phylogenetic basis. Rosauer et al. (2009) provide a method of combining PD with weighted endemism measures (Williams et al., 1994; Crisp et al., 2001) to identify geographic areas in which high levels of PD are restricted. Phylogenetic Endemism (PE) effectively corrects for phylogenetic

uniqueness when identifying high value areas based on levels of endemism. Thus when calculated using appropriate genetic data it will identify areas of high (phylo)genetic diversity where that diversity is less likely to occur elsewhere.

Just as standard genetic diversity is a measure of variation in a population or species, phylogenetic diversity measures variation between and amongst clades. It is the shared ancestry of clades (and the corresponding differentially shared genetic diversity) that makes a phylogenetic context essential for useful comparisons of genetic diversity between monophyletic taxa.

6. Acknowledgements

7. References

Ahmad, S.M.; Hoot, S.B.; Qazi, P.H. & Verma, V. (2009) Phylogenetic patterns and genetic diversity of Indian *Tinospora* species based on chloroplast sequence data and cytochrome P450 polymorphisms. *Plant Systematics and Evolution* 281:87–96

Albrecht, E.; Escobar, M. & Chetelat, R.T. (2010) Genetic diversity and population structure in the tomato-like nightshades *Solanum lycopersicoides* and *S. sitens*. *Annals of Botany* 105: 535–554

Alwala, S.; Suman, A.; Arro, J.A.; Vermis, J.C. & Kimbeng, C.A. (2006) Target region amplification polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. *Crop Science* 46:448–455

Andrewartha, H.G. & Birch, L.C. (1954) *The distribution and abundance of animals*. Unniversity of Chicago Press, Chicago

Andrivon, D. (1996) The origin of *Phytophthora infestans* populations present in Europe in the 1840s: a critical review of historical and scientific evidence. *Plant Pathology* 45:1027–1035

Avise, J.C. (2004) *Molecular markers, natural history and evolution*. Sinauer Associates

Beakes, G. W. & Sekimoto, S. (2008) The Evolutionary Phylogeny of Oomycetes—Insights Gained from Studies of Holocarpic Parasites of Algae and Invertebrates, In: *Oomycete Genetics and Genomics: Diversity, Interactions, and Research Tools,* Lamour, K. & Kamoun, S.; pp.1–25, John Wiley & Sons, Inc., Hoboken, NJ, USA

Beumont, M.A. & Rannala, B. (2004) The Bayesian revolution in genetics. *Nature Reviews* 5:251-261

Bonin, A.; Ehrich, D. & Mantel, S. (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology* 16:3737–3758

Bordewich, M. & Semple, C. (2005) On the computational complexity of the rooted subtree prine and regraft distance. *Annals of Combinatorics* 8:409–423

Bordewich, M. & Semple, C. (2006) Computing the minimum number of hybridization events for a consistent evolutionary history. *Discrete Applied Mathematics* 155:914–928

Bourke, P.M.A. (1964) Emergence of potato blight, 1843-1846. *Nature* 203:805–808

Boyle, P.P. & Gráda, C.Ó. (1986) Fertility trends, excess mortality, and the Great Irish Famine. *Demography* 23:543–562

Branco, J.S.C.; Vieira, E.A.; Malone, G.; Kopp, M.M.; Malone, E.; Bernardes, A.; Mistura, C.C.; Carvalho, F.I.F. & Oliveira, C.A. (2007) IRAP and REMAP assessments of genetic similarity in rice. *Journal of Applied Genetics* 48:107–113

Bretó, M.P.; Ruiz, C.; Pina, J.A. & Asíns, M.J. (2001) The diversification of *Citrus clementina* Hort. ex Tan., a vegetatively propagated crop species. *Molecular Phylogenetics and Evolution* 21:285–293

Bryant, D. & Moulton, V. (2004) NeighborNet: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution* 21:255–265

Cadwell, R.C. & Joyce, G.F. (1992) Randomization of genes by PCR mutagenesis. *Genome Research* 2:28–33

Cantelo, W.W. & Sanford, L.L. (1984) Insect population response to mixed and uniform plantings of resistant and susceptible plant material. *Environmental Entomology* 13:1443–1445

Carvalho, A.; Guedes-Pinto, H.; Martins-Lopes, P. & Lima-Birto, J. (2010) Genetic variability of Old Portuguese bread wheat cultivars assayed by IRAP and REMAP markers. *Annals of Applied Biology* 156:337–345.

Cernák, I. (2008) *Development of molecular markers to the potato virus Y (PVY) extreme resistance gene (Ry$_{sto}$) originating from the wild potato species Solanum stoloniferum*. PhD thesis, University of Pannonia, Keszthely, Hungary

Cernák, I.; Decsi, K.; Nagy, S.; Wolf, I.; Polgár, Z.; Gulyás, G.; Hirata, Y. & Taller, J. (2008) Development of a locus-specific marker and localization of the *Ry$_{sto}$* gene based on linkage to a catalase gene on chromosome XII in the tetraploid potato genome. *Breeding Science* 58: 309–314

Choi, H.K.; Kim, D.; Uhm, T.; Limpens, E.; Lim, H.; Mun, J-H.; Kalo, P. et al. (2004) A sequence-based genetic map of *Medicago trunculata* and comparison of marker colinearity with *M. sativa*. *Genetics* 166:1463–1502

Collard, B.C.Y. & Mackill, D.J. (2009a) Conserved DNA-derived polymorphism (CDDP): a simple and novel method for generating DNA markers in plants. *Plant Molecular Biology Reporter* 27:558–562

Collard, B.C.Y. & Mackill, D.J. (2009b) Start Codon Targeted (SCOT) polymorphism: A simple novel DNA marker technique for generating gene-targeted markers in plants. *Plant Molecular Biology* 27:86–93

Cousens, S.H. (1960) Regional death rates in Ireland during the Great Famine, from 1846 to 1851. *Population Studies* 14:55–74

Crisp, M.; Laffan, S; Linder, H & Monro, A. (2001) Endemism in the Australian flora. *Journal of Biogeography* 28:183–198

Csilléry, K.; Blum, M.G.B.; Gaggiotti, O.E. & François, O. (2010) Approximate Bayesian computation (ABC) in practice. *Trends in Ecology and Evolution* 25:410–418

Dice, L.R. (1945) Measures of the amount of ecologic association between species. *Ecology* 26: 297–302.

Edwards, A.W.F. (1970) Estimation of the branch points of a branching diffusion process. *Journal of the Royal Statistical Society Series B* 32: 155–174

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation* 61:1-10

Farris, J.S. (1970) A method for computing Wagner trees. *Systematic Zoology* 19: 83–92

Farris, J.S. (1973) A probability model for inferring evolutionary trees. *Systematic Zoology* 22: 250–256

Farris, J.S. (1983) *The logical basis of phylogenetic analysis.* In: Platnich, N.I. & Funk, V.A. (Eds.) *Advances in Cladistics Vol. 2.* pp. 7-36, Columbia University Press, New York.

Farris, J.S.; Kluge, A. G. & Eckardt, M.J. (1970) A numerical approach to phylogenetic systematics. *Systematic Zoology* 19: 172–189

Felsenstein, J. (1973) Maximum likelihood and minimum-steps methods for estimating evolutionary trees from data on descrete characters. *Systematic Zoology* 22: 240–249

Feng, N.; Sue, Q.; Guo, Q.; Zhao, R. & Guo, M. (2009) Genetic diversity and population structure of *Celosia argentea* and related species revealed by SRAP. *Biochemical Genetics* 47:521–532

Fisher, R.A. (1930) *The genetical theory of natural selection.* Oxford University Press, Oxford

Flavell, A.J.; Bolshakov, V.N.; Booth, A.; Jing, R.; Russell, J.; Ellis, T.H.N. & Isaac, P. (2003) A microarray-based high throughput molecular marker genotyping method: the tagged microarray marker (TAM) approach. *Nucleic Acids Research* 31:e115

Flavell, A.J.; Knox, M.R.; Pearce, S.R. & Ellis, T.H.N. (1998) Retrotransposon-based insertion polymorphisms (RBIP) for high throughput marker analysis. *The Plant Journal* 16:643–650

Frankham, R.; Briscoe, D.A. & Ballou, J.D. (2002) *Introduction to conservation genetics.* Cambridge University Press, UK

Fraser, E.D.G. (2003) Social vulnerability and ecological fragility: building bridges between social and natural sciences using the Irish Potato Famine as a case study. *Conservation Ecology* 7:9

Gilani, S.A.; Kikuchi, A. & Watanabe, K.N. (2009) Genetic variation within and among fragmented populations of endangered medical plant, *Withania coagulans* (Solanaceae) from Pakistan and its implications for conservation. *African Journal of Biotechnology* 8:2948–2958

Gizaw, S.A. (2011) *Molecular marker development and genetic mapping in potato (Solanum tuberosum) genome of their use in breeding.* MSc thesis, University of Pannonia, Keszthely, Hungary

Gladieux, P.; Giraud, T.; Kiss, L.; Genton, B.J.; Jonot, O. & Shykoff, J.A. (2011) Distinct invasion sources of common ragweed (*Ambrosia artemisiifolia*) in Eastern and Western Europe. *Biological Invasions* 13:933–944

Goloboff, P.A.; Catalano, S.A.; Mirande, J.M.; Szumik, C.A.; Arias, J.S.; Källersjö, M. & Farris, J.S. (2009) Phylogenetic analysis of 73 060 taxa corroborates major eukaryotic groups. *Cladistics* 25:211–230

Gorji, A.M.; Poczai, P.; Polgar, Z. & Taller, J. (2011) Efficiency of arbitrarily amplified dominant merkers (SCoT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. *American Journal of Potato Research* 88:226–237

Gráda, C.Ó. (2000) *The Great Irish Famine.* Cambridge University Press, Cambridge, UK

Grzebelus, D. (2006) Transposon insertion polymorphism as a new source of molecular markers. *Journal of Fruit and Ornamental Plant Research* 14:21–29

van Haeringen, W.A.; Den Bieman, M.G.; Lankhorst, A.E.; van Lith, H.A. & van Zutphen, L.F. (2002) Application of AFLP markers for QTL mapping in the rabbit. *Genome* 45:924–921

Hanski, I. (1998) Metapopulation dynamics. *Nature* 396:41–49

Hennig, W. (1950) *Grundzüge einer Theorie der Phylogenetischen Systematik*. Deutscher Zentralverlag, Berlin.

Hennig, W. (1966) *Phylogenetic Systematics*. University of Illinois Press, Urbana.

Higgs, P.G. & Attwood, T.K. (2005) *Bioinformatics and molecular evolution*. Blackwell, London.

Hill, W.G. & Weir BS (2004) Moment estimation of population diversity and genetic distance from data on recessive markers. *Molecular Ecology* 13: 895–908

Holsinger, K.E. & Wallace, L.E. (2004) Bayesian approaches for the analysis of population genetic structure: an example from *Platanthera leucophaea* (Orchidaceae). *Molecular Ecology* 13:887–894

Holsinger, K.E.; Lewis, P.O. & Dey, D.K. (2002) A Bayesian approach to inferring population structure from dominant markers. *Molecular Ecology* 11:1157–1164

Hu, J. & Vick, B.A. (2003) Target region amplification polymorphism: a novel marker technique for plant genotyping. *Plant Molecular Biology Reporter* 21:289–294

Hu, J.; Mou, B. & Vick, B.A. (2007) Genetic diversity of 38 spinach (*Spinacia oleracea* L.) germplasm accessions and 10 commercial hybrids assessed by TRAP markers. *Genetic Resources and Crop Evolution* 54:1667–1674

Hu, J.; Ochoa, O.E.; Truco, M.J. & Vick, B.A. (2005) Application of the TRAP technique to lettuce (*Lactuca sativa* L.) genotyping. *Euphytica* 144:225–235

Hughes, A.R.; Inouye, B.D.; Johnson, M.T.J; Underwood, N. & Vellend, M. (2008) Ecological consequences of genetic diversity. *Ecology Letters* 11:609–623

Huson, D.H. & Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution 23:254–267

Isaac, N.J.B.; Turvey, S.T.; Collen, B.; Waterman, C. & Baillie, J.E.M. (2007) Mammals on the EDGE: conservation priorities based on threat and phylogeny. *PLoS ONE* 2: e296.

IUCN (2011) The IUCN Red List of Threatened Species. http://www.iucnredlist.org.

Jaccard, P. (1908) Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise Des Sciences Naturelles* 44:223–270

Jain, K.K. (2002) Personalized medicine. *Current Opinion in Molecular Therapeutics* 4:548–558

Jing, R.; Vershinin, A.; Grezbyta, J.; Shaw, P.; Smýkal, P.; Marshall, D.; Ambrose, M.J. et al. (2010) The genetic diversity and evolution of field pea (Pisum) studied by high throughput retrotransposon based insertion polymorphism (RBIP) marker analysis. BMC Evolutionary Biology 10:44

Jump, A.S.; Hunt, J.M.; Martínez-Izquirdo, J.A. & Peñuelas, J. (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in Fagus sylvatica. Molecular Ecology 15:3469–3480

Kalendar, R. (2011) The use of retrotransposon-based molecular markers to analyze genetic diversity. *Field and Vegetable Crops Research* 48:261–274

Kenward, K.D.; Bai, D.; Ban, M.R. & Brandle, J.E. (1999) Isolation and characterization of *Tnd*-1, a Retrotransposon marker linked to black root resistance in tobacco. *Theoretic and Applied Genetics* 98:387–395

Genomics Meets Biodiversity: Advances in Molecular Marker
Development and Their Applications in Plant Genetic Diversity Assessment

29

Kessmann, H.; Choudhary, A.D. & Dixon, A. (1990) Stress responses in alfalfa (*Medicago sativa* L.). III. Induction of medicarpin and cytochrome P450 enzyme activities in elicitor-treated cell suspension cultures and protoplasts. *Plant Cell Reports* 9:38–41

Khan, M.S.; Yadav, S.; Srivastava, S.; Swapna, M.; Chandra, A. & Singh, R.K. (2011) Development and utilization of conserved-intron scanning marker in sugarcane. *Australian Journal of Botany* 59:38–45

Kluge, A.G. & Farris, J.S. (1969) Quantitative phyletics and the evolution of Anurans. *Systematic Zoology* 18: 1–32

Kumar, A. & Bennetzen, J.L. (1999) Plant retrotransposons. *Annual Review of Genetics* 33:479–532

Kumar, A. & Hirochika, H. (2001) Applications of retrotransposons as genetic tools in plant biology. *Trends in Plant Science* 6:127–134

Kumar, A.; Pearce, S.R.; McLean, K.; Harrison, G.; Helsop-Harrison, J.S.; Waugh, R. & Flavell, A.J. (1997) The *Ty*1-*copia* group of retrotransposons in plants: genomic organization, evolution, and use as molecular markers. *Genetica* 100:205–217

Laikre, L.; Allendorf, F.W.; Aroner, L.C.; Baker, C.S.; Gregovich, D.P.; Hansen, M.M.; Jackson, J.A. et al. (2009) Neglect of genetic diversity in implementation of the conservation on biological diversity. *Conservation Biology* 24:86–88

Lebecka, R. (2008) Host-pathogen interaction between *Phytophthora infestans* and *Solanum nigrum*, *S. villosum* and *S. scabrum*. *European Journal of Plant Pathology* 120: 233–240

Li, G. & Quiros, C.F. (2001) Sequnce-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica*. *Theoretic and Applied Genetics* 103:455–461

Li, S.; Pearl, D.K. & Doss, H. (2000) Phylogenetic tree construction using Markov chain Monte Carlo. *Journal of American Statistical Association* 95:493–508

Lin, X.; Kaul, S.; Rounsley, S.; Shea, T.P.; Benito, M-I.; Town, C.D.; Fujii, C.Y. et al. (1999) Sequencing and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* 402:761–768

Lou, Q. & Chen, J. (2007) *Ty*1-*copia* retrotransposon-based SSAP marker development and its potential in the genetic study of cucurbits. *Genome* 50:802–810

Lynch, M. & Milligan, B.G. (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology* 3:91–99

Mantel, S.; Gaggiotti, O.E. & Waples, R.S. (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends in Ecology and Evolution* 20:136–142

Mau, B. & Newton, M. (1997) Phylogenetic inference for binary data on dendrograms using Markov chain Monte Carlo. *Journal of Computational and Graphical Statistics* 6: 122–131

Michener, C.D. & Sokal, R.R. (1957) A quantitative approach to a problem in classification. *Evolution* 11:130–162

Milligan, B.G.; Leebens-Mack, J. & Strand, A.E. (1994) Conservation genetics: beyond the maintenance of marker diversity. *Molecular Ecology* 3:423–435

Moisy, C.; Garrison, K.E.; Meredith, C.P. & Pelsy, F. (2008) Characterization of ten novel *Ty*1/*copia*-like retrotransposon families of the grapevine genome. *BMC Genomics* 9:469

Murthy, V. (2011) *Population genetic analysis of common ragweed (Ambrosia artemisiifolia L.) in Europe using DNA-based molecular markers*. MSc thesis, University of Pannonia, Keszthely, Hungary

Nei, M. & Li, W.H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA* 76:5269–5273

O'Rourke, K. (1994) The economic impact of the famine in the short and long run. *American Economic Review* 84:309–313

Page, R.D.M. & Holmes, E.C. (1998) *Molecular evolution*. Blackwell, London.

Panwar, P.; Saini, R.K.; Sharma, N.; Yadav, D. & Kumar, A. (2010) Efficiency of RAPD, SSR and Cytochrome P450 gene based marker sin accessing genetic variability amongst finger millet (*Eleusine coracana*) accession. *Molecular Biology Report* 37:4075–4082

Pleijel, F. & Rouse, G.W. (2000) Least-inclusive taxonomic unit: a new taxonomic concept for biology. *Procdings of the Royal Society of London B* 267:627-630

Poczai, P. & Hyvönen, J. (2011) On the origin of *Solanum nigrum*: can networks help? *Molecular Biology Reports* 38:1171–1185

Poczai, P.; Cernák, I.; Gorji, A.M.; Nagy, S.; Taller, J. & Polgár, Z. (2010) Development of intron targeting (IT) markers for potato and cross-species amplification in *Solanum nigrum* (Solanaceae). *American Journal of Botany* 97:e142–e145

Poczai, P.; Varga, I; Bell, N.E. & Hyvönen, J. (2011) Genetic diversity assessment of bittersweet (*Solanum dulcamara*, Solanaceae) germplasm using conserved DNA-derived polymorphism and intron-targeting markers. *Annals of Applied Biology* 159:141–153

Purvis, A.; Agapow, P.M.; Gittleman, J.L. & Mace, G.M. (2000) Nonrandom extinction and the loss of evolutionary history. *Science* 288:328–330

Ragupathy, R.; Banks, T. & Cloutier, S. (2010) Molecular characterization of the Sasanda LTR copia retrotransposon family uncovers their recent amplification in *Triticum aestivum* (L.) genome. *Molecular Genetics and Genomics* 283:255–271

Ray, D.A. (2007) SINEs of progress: mobile element applications to molecular ecology. *Molecular Ecology* 16:19–33

Ray, D.A.; Xing, J.; Salem, A-H. & Batzer, M.A. (2006) SINEs of a nearly perfect character. *Systematic Biology* 55:928–935

Relethford, J.H., Crawford, M.H. & Blangero, J. (1997) Genetic drift and gene flow in post-famine Ireland. *Human Biology* 69:443–465

Rick, C.M. & Chetelat, R.T. (1995) Utilization of related wild species for tomato improvement. *Acta Horticulturae* 412:21–38

Riechmann, J.L.; Heard, J.; Martin, G.; Reuber, L.; Jiang, C-Z.; Keddie, J.; Adam, L et al. (2001) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290:2105–2110

Rosauer, D.; Laffan, S.W.; Crisp, M.D.; Donnellan, S.C. & Cook, L.G. (2009) Phylogenetic endemism: a new approach for identifying geographical concentrations of evolutionary history. *Molecular Ecology* 18:4061–4072

Sawant, S.V.; Singh, P.K.; Gupta, S.K.; Madnala, R. & Tuli, R. (1999) Conserved nucleotide sequences in highly expressed genes in plants. *Journal of Genetics* 78:123–131

Scheifele, L.Z.; Cost, G.J.; Zupancic, M.L.; Caputo, E.M. & Boeke, J.D. (2009) Retrotransposon overdose and genome integrity. *Proceedings of the National Academy of Sciences of the United States of America* 106:13927–13932

Schmidt, K. & Jensen, K. (2000) Genetic structure and AFLP variation of remnant populations in the rare plant *Pedicularis palustris* (Scrophulariaceae) and its relation to population size and reproductive components. *American Journal of Botany* 87:678–689

Schonswetter, P. & Tribsch, A. (2005) Vicariance and dispersal in the alpine perennial *Bupleurum stellatum* L. (Apiaceae). *Taxon* 54:725–732

Shalk, M.; Nedelkina, S.; Schoch, G.; Batard, Y. & Werck-Reichhart, D. (1999) Role of unusual amino-acid residues in the proximal and distal heme regions of a plant P450, CYP73A1. *Biochemistry* 38:6093–6103

Shedlock, A.M. & Okada, N. (2000) SINE insertions: powerful tools for molecular systematics. *BioEssays* 22:148–160

Showmaker, J.S.; Painter, I.S. & Weir, B.S. (1999) Bayesian statistics in genetics: a guide for the uninitiated. *Trends in Genetics* 15: 354-358

Smithson, J.B. & Lenne, J.M. (1996) Varietal mixes: a viable strategy for sustainable productivity in subsistence agriculture. *Annals of Applied Biology* 128:127–158

Sokal, R.R. & Michener, C.D. (1958) A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38:1409–1438.

Somerville, C. & Somerville, S. (1999) Plant functional genomics. *Science* 285:380–383

Song, Z.; Li, X.; Wang, H. & Wang, J. (2010) Genetic diversity and population structure of *Salvia miltiorrhiza* Bge in China revealed by ISSR and SRAP. *Genetica* 138:241–249

Sørensen, T. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskabs Biologiske Skrifter* 5:1–34.

Stewart, C.N. & Excoffier, L. (1996) Assessing population genetic structure and variability with RADP data: application to *Vaccinium macrocarpon* (American Cranberry). *Journal of Evolutionary Biology* 9:153–171

Sun, S-J.; Gao, W.; Lin, S-Q; Zhu, J.; Xie, B-G. & Lin, Z-B. (2006) Analysis of genetic diversity in *Ganoderma* population with a novel molecular marker SRAP. *Applied Microbiology and Biotechnology* 72:537–543

Syed, N.H. & Flavell, A.J. (2006) Sequcne-specific amplification polymorphisms (SSAPs): a multi-locus approach for analyzing transposon insertions. *Nature Protocols* 1:2746–2752

Tam, S.M.; Mhiri, C.; Vogelaar, A.; Kerkveld, M.; Pearce, S.R. & Grandbastien, M.A. (2005) Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. *Theoretic and Applied Genetics* 110:819–831

Tang, S.; Dai, W.; Li, M.; Zhang, Y.; Geng, Y.; Wang, L. & Zhong, Y. (2008) Genetic diversity of relictual and endangered plant *Abies ziyuanensis* (Pinaceae) revealed by AFLP and SSR markers. *Genetica* 133:21–30

Telenius, H.; Carter, N.P.; Beb, C.E.; Nordenskjold, M.; Ponder, B.A. & Tunnacliffe, A. (1992) Degenerate oligonucleote-primered PCR: general amplification of target DNA by single degenerate primer. *Genomics* 13:718–725

Vershinin, A.V.; Allnutt, T.R.; Knox, M.R.; Ambrose, M. J. & Ellis, T.H.N. (2003) Transposable elements reveal the impact of introgression, rather than transposition, in *Pisum* diversity, evolution and domestication. Molecular Biology and Evolution 20:2067–2075

Vitte, C.; Ishii, T.; Lamy, F.; Brar, D. & Panaud, O. (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L). *Molecular Genetics and Genomics* 272:504–511

Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; van de Lee, T.; Hornes, M.; Friters, A. et al. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* 23:4407–4414

Wang, Q.; Zhang, B. & Lu, Q. (2009) Conserved region amplification polymorphism (CoRAP) a novel marker technique for plant genotyping in *Salivia miltiorrhiza*. *Plant Molecular Biology Reporter* 27:139–143

Ward, S.M.; Gaskin, J.F. & Wilson, L.M. (2008) Ecological genetics of plant invasions: what do we know? *Invasive Plant Science and Management* 1:98–109

Wheeler, W.C. (1996) Optimization alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics* 12: 1–9

Wheeler, W.C. (1999) Fixed character states and the optimization of molecular sequence data. *Cladistics* 15: 379–385

Wheeler, W.C. (2003a) Implied alignment. *Cladistics* 19:261–268

Wheeler, W.C. (2003b) Iterative pass optimization. *Cladistics* 19:254–260

Whelan, K. (1997) *The atlas of the Irish rural landscape*. Cork University Press, Cork

Williams, J.G.K.; Kubelik, A.R.; Livak, K.J.; Rafalski, J.A. & Tingey, S.C. (1990) DNA polymorphsims amplified by arbitrary primers are useful as genetic markers. *Nucleaic Acids Research* 18:6531–6535

Williams, P.H.; Humphries, C.J.; Forey, P.L.; Humphries, C.J. & Vane-Wright, R.I. (1994) Biodiversity, taxonomic relatedness, and endemism in conservation, In: *Systematics and Conservation Evaluation,* Forey, P.L., Humphries, C.J. & Vane-Wright, R.I.; p. 438, Oxford University Press, Oxford

Wright, S. (1931) Evolution in Mendelian populations. *Genetics* 16:97–159

Xiong, F.; Zhong, R.; Han, Z.; Jiang, J.; He, L.; Zhuang, W. & Tang, R. (2011) Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachis hypogaea* L.) genotypes. *Molecular Biology Reports* 38:3487–3494

Yamanaka, S.; Suzuki, E.; Tanaka, M.; Takeda, Y.; Watanabe, J.A. & Watanabe, K.N. (2003) Assessment of cytochrome P450 sequences offers a useful tool for determining genetic diversity in higher plant species. *Theoretic and Applied Genetics* 108:1–9

Zhivotovsky, L. (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology* 8:907–913.

Zhu, Y.; Chen, H.; Fan, J.; Wang, Y.; Li, Y.; Chen, J.; Fan, J.; Yang, S.; Hu, L.; Leung, H.; Mew, T.W.; Teng, P.S.; Wang, Z. & Mundt, C.C. (2000) Genetic diversity and disease control in rice. *Nature* 406:718–722

Species Occurrence, Hybridization and Speciation in Postglacial East and South Asia

Liao Pei-Chun[1,3], Yi-Shan Chao[2,3] and Yu-Chung Chiang[2,3]
[1]Department of Biological Science & Technology
National Pingtung University of Science & Technology
[2]Department of Biological Science, National Sun Yat-sen University
[3]Equal contributions
Taiwan, R.O.C.

1. Introduction

Geological disjunction and the changing climate affect the distribution of plants and animals in East Asia. The species of postglacial East and South Asia have gone through the process of speciation. In the Northern Hemisphere, a disjunction of the distribution of closely related plants between eastern Asia and North America has been reported in many previous studies. These plants include *Rhus* (Yi et al. 2004), *Cornus* (Xiang et al. 2005), *Berberis* (Kim et al. 2004), *Chamaecyparis* (Wang et al. 2003), *Panax* (Wen and Zimmer, 1996), and *Pontica* (Milne 2004), *Fraxinus* (Jeandroz et al. 1997), etc. The modern distribution of plant species between eastern Asia and North America reflects migration, speciation, and extinction due to climatic exchange in past glaciating time periods (Milne 2004). Numerous case studies pertaining to eastern Asia and eastern North America have helped to reconstruct the relationships among morphologically similar species to explain the differences in species diversity among areas with similar environmental conditions (Wen et al., 1996, 1998; Wen, 1999, 2000, 2001; Qiu et al., 1995a, b; Soltis 2001). During the glacial and postglacial periods, climate and sea level changes affected the migration and distribution of plants and animals, which caused landbridge formations and a breakdown between islands and continents. The different species showed variant patterns of phytogeography in postglacial East and South Asia. For example, the *Cycas revoluta*, the insular species distributed from Kyushu to South Ryukyu, and *C. taitungensis*, the insular species restricted on South Taiwan, were allopatrically distributed but paraphyly due to ancient ancestral polymorphisms (Chiang et al., 2009). Likewise, *Amentotaxus* is restricted to Taiwan, to the west across southern China, to Assam in the eastern Himalaya and to the south in Vietnam. *A. argotaenia*, *A. formosana*, *A. poilanei* and *A. yunnanensis* are considered complex species and a result of speciation (Ge et al., 2005). In addition, most plant species in the Western Pacific island chain, including the islands from Ilyushin, Japan, and Taiwan to the Philippines, are derived from temperate and tropical Asia, particularly from the mainland of China. Due to the variable topography and the short distance from the continent of Asia, the island forest species are diverse and similar to their relatives that are found on continent of Asia. In addition, the tropical currents and the Kuroshio Current provide an ideal environment for long-distance colonization of seeds from Oceania to the circum-Pacific island arc. In the

past circles of glacial and postglacial time periods, the continental islands such as Taiwan, Japan, and the Philippines became melting pots for organisms, i.e., several taxa were likely to be multiple originated, showing the process of speciation. Here, we review various case studies to discuss the phytogeographic patterns and the process of speciation among related species in East and South Asia.

2. Case study: Diversification rate shift in SE Asian *Lithocarpus*

2.1 Introduction

The stone oaks (*Lithocarpus*, Fagaceae) originated from Southeast Asia (about 300 extant species) and underwent two range-expansion events: a westward dispersal to Europe identified by fossil records (no extant species) and an eastward colonization to western North America (one extant species) (Cannon & Manos, 2003). The diversification rate of *Lithocarpus* was primarily affected by the complicated geohistory, the geographic isolation, and the heterogeneous (micro-) habitats of Southeast Asia. Interspecific hybridization also affected the rate of speciation and the diversity of species of SE Asian stone oaks (Cannon & Manos, 2003). Early researchers had reported the break of reproductive isolation in the related taxa, Fagales (Hardin, 1975). Even later reports regarding the white oak group of *Quercus* indicated that the sympatric *Quercus* species shared identical or similar haplotypes with different species but had distinct haplotypes among different populations of the same species (Whittemore & Schaal, 1991; Petit et al., 1993; Petit et al., 2002).

The largest impact of the Pleistocene glacial time period on the geohistory in Southeast Asia was the emergence and submergence of Sundaland, the historic continent that once connected the Thai-Malay Peninsula to the Borneo and the Philippine Archipelagos. Several reports on the mangroves indicated that the Pleistocene glacial oscillations have affected the historical demography of the mangroves in this region (Liao et al., 2007; Liao et al., 2011). Postglacial range expansion of the *Ceriops tagal* (Rhizophoraceae) around the coasts of the South China Sea reflected the geohistory of the Sundaland submergence (Liao et al., 2009). The glacial-dependent change in population size, especially due to the postglacial expansion, has also been reported for several East and Southeast Asian organisms, such as *Castanopsis carlesii* (Fagaceae) (Cheng et al., 2005), *Machilus* (Lauraceae) (Wu et al., 2006), *Dysosma versipellis* (Berberidaceae) (Fu et al., 2009), and *Cinnamomum kanehirae* (Lauraceae) (Liao et al., 2010), etc. Another impact of the Pleistocene glacials pertained to the connection between the Borneo and Thai-Malay Peninsula (the continent of Asia). The land bridge of the Sundaland accelerated the migration (gene flow) of organisms between these two regions and increased the opportunity of interspecific hybridization, which may have resulted in the formation of new species (speciation by hybridization) (Comes & Kadereit, 1998). The postglacial isolation between Indochina and Borneo also provided opportunities for reproductive isolation and allopatric speciation. Both interspecific hybridization and geographic-dependent reproductive isolation can explain the high biodiversity of Southeast Asia, which includes the regions of Borneo, Java, Thai-Malay and the Indochina Peninsula (Swallow et al., 2005; Blattner et al., 2006; Sheldon et al., 2009).

Cannon and Manos (2003) explored the phylogeographic structure of the stone oaks of Southeast Asia and identified local molecular endemism in Southeast Asia. Significant fragmentation led to independent genetic diversification, and the independent evolutionary history of species in each locality increased the "rate of endemism", which is an indicator of biodiversity. Such events corresponding to the high diversity of stone oaks in Southeast Asia

probably predated the Pleistocene Periods, which, therefore, suggests a relation to the glacial oscillations (Cannon & Manos, 2003). The changes in the phylogeographic patterns of endemic species caused by the catastrophic events may have resulted in a shift in the diversification rate. For example, the 1999 Chi-Chi earthquake in central Taiwan resulted in the extinction of populations of *L. konishii* and a loss of the genetic diversity (Hung et al., 2005). A loss of diversity can result in an increase in the genetic differentiation between extant populations due to the loss of shared haplotypes and the rapid drift of the small population size. In other words, such catastrophic events may play important roles in either extinction or speciation, *i.e.,* an alteration of the diversification rate.

In this case study, the systematic analysis of *Lithocarpus* (Fagaceae) based on the chloroplast *atp*B-*rbc*L spacer was used to explore the diversification rate shift in phylogeographic light. To resolve the question as to whether *Lithocarpus* evolved at heterogeneous speciation/extinction rates (*i.e.,* diversification rate) between distinct clades, the cpDNA was collected and used as the marker for phylogeographic inference.

2.2 Data collection

Partial sequences of the chloroplast *atp*B-*rbc*L spacer were obtained from the GenBank (NCBI) and were directly sequenced and collected in China and Taiwan (Table 1). In total, 110 sequences of *Lithocarpus* were obtained from the NCBI, and 21 sequences were obtained from direct sequencing. The alignment length of the sequence fragment used for systematic analysis was 623 bps. The species sampled in this study were distributed in the diversity center of Southeast Asia, *i.e.,* southern China, Japan, Taiwan, the Indochina Peninsula, the Thai-Malay Peninsula, Indonesia (Sumatra and Java), Borneo, the Philippines, and other areas (e.g., *L. densiflorus* of N. America). *Chrysolepis chrysophylla*, *Quercus suber*, *Castanopsis argyrophylla*, *Castanea crenata*, *Trigonobalanus vertic*, *Colombobalanus excels*, *Formanodendron doicha*, and *Fagus hyatae* were used as outgroups for the phylogenetic tree reconstruction.

Species	Source	Distribution
L. amygdalifolius	FBG	SE China & Taiwan
L. bancanus	NCBI	Malaysia & Borneo
L. beccarianus	NCBI	Borneo
L. bennettii	NCBI	Malaysia & Borneo
L. blumeanus	NCBI	Borneo
L. brevicaudatus	FBG	SE China & Taiwan
L. bullatus	NCBI	Malaysia & Borneo
L. cantleyanus	NCBI	Malaysia & Borneo
L. clementianus	NCBI	Malaysia & Borneo
L. conocarpus	NCBI	Malaysia & Borneo
L. coopertus	NCBI	Philippines, Malaysia & Borneo
L. cornea	FBG	Taiwan (endemic)
L. cyrtocarpus	NCBI	China & Vietnam
L. dasystachyus	NCBI	Borneo
L. dealbatus	NCBI	China
L. densiflorus	NCBI	W America
L. dodoniifolia	FBG	Taiwan (endemic)
L. echinifer	NCBI	Borneo & Malaysia
L. echinophorus	NCBI	Lao, Vietnam & China
L. echinotholus	NCBI	China & Vietnam
L. edulis	NCBI	Japan
L. elegans	NCBI	Borneo
L. encleisocarpus	NCBI	Thailand, Malaysia & Borneo

Species	Source	Distribution
L. ewyckii	NCBI	Malaya, Sumatera & Borneo
L. fenestratus	NCBI	China
L. ferrugineus	NCBI	Borneo
L. formosana	FBG	Taiwan (endemic)
L. glaber	FBG	China, Japan & Taiwan
L. gracilis	NCBI	Malaysia & Borneo
L. grandifloius	NCBI	China, Vietnam, Laos, Tibet & Assam
L. hancei	SCBG	S China
L. hancei var. *ternaticupula* f. *subreticulata*	FBG	Taiwan
L. hancei var. *ternaticupula* f. *ternaticupula*	FBG	Taiwan
L. harlandii	FBG	Taiwan (endemic)
L. hatusimae	NCBI	Borneo
L. havilandii	NCBI	Borneo
L. jacobsii	NCBI	Borneo
L. kalkmanii	NCBI	Borneo
L. kawakamii	FBG	Taiwan (endemic)
L. keningauensis	NCBI	Borneo
L. konishii	FBG	Taiwan (endemic)
L. lampadarius	NCBI	Borneo
L. lepidocarpus	FBG	Taiwan (endemic)
L. lindleyanus	NCBI	Thailand & Myanmar
L. litseifolius	SCBG	S China, Laos, Myanmar & Vietnam
L. lucidus	NCBI	Malaysia & Borneo
L. luteus	NCBI	Borneo
L. meijeri	NCBI	Borneo
L. naiadarum	SCBG	China (endemic to Hainan)
L. nantoensis	FBG	Taiwan (endemic)
L. nieuwenhuisii	NCBI	Borneo & Philippines
L. pachylepis	NCBI	S China & Vietnam
L. pachyphyllus	NCBI	China
L. palungensis	NCBI	Borneo
L. papillifer	NCBI	Malaysia & Borneo
L. porcatus	NCBI	Borneo
L. pseudokunstleri	NCBI	Borneo
L. pulcher	NCBI	Borneo
L. rassa	NCBI	Malaysia & Borneo
L. revolutus	NCBI	Borneo
L. rhombocarpus	SCBG	Taiwan & China
L. rotundatus	NCBI	Borneo, Java & Philippines
L. ruminatus	NCBI	Borneo
L. sericobalanus	NCBI	Borneo
L. shinsuiensis	FBG	Taiwan (endemic)
L. taitoensis	FBG	Taiwan & China
L. truncatus	NCBI	China & Assam
L. turbinatus	NCBI	Borneo
L. variolosus	NCBI	China & Vietnam
L. wallichianus	NCBI	Thailand & Sumatra
L. xylocarpus	NCBI	Lao, Vietnam, Assam & China
L. xylocarpus	NCBI	Assam, Tibet, Myanmar, Laos, Vietnam & China
Castanopsis argyrophylla	NCBI	China, Myanmar, Laos
C. sieboldii	NCBI	Japan & Korea
Chrysolepis chrysophyllus	NCBI	California
Fagus hayatae	FBG	China & Taiwan

Table 1. The sampling for the tree reconstruction and diversification rate shift analysis in this case study. FBG: Fushan Botanical Garden; SCBG: South China Botanical Garden; NCBI: National Center for Biotechnology Information.

2.3 Phylogenetic inference

Before reconstructing the phylogenetic tree, the substitution model was tested using MEGA 5.05 (Tamura et al., 2011), and the general-time-reversible model using a discrete Gamma distribution of the evolutionary rate among sites (GTR+G) was evaluated as the best model according to the corrected Akaike Information Criterion (AICc) (-lnL = 1989.72, AICc = 4549.56). The phylogenetic relationships of *Lithocarpus* were reconstructed by Bayesian inference (BI) using MrBayes ver. 3.1.2 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003), and the species tree was inferred by the Bayesian approach under the Yule's pure-birth model that was implemented using BEAST ver. 1.6.1 (Drummond & Rambaut, 2007). For reconstructing the gene tree (MrBayes), the GTR+G model was used, and two parallel runs of the Markov chain Monte Carlo (MCMC) searches were performed for ten million generations, sampling every 1000 generations for a total of 10,000 trees of each run. The first 10% of the generations were discarded (burn-in). Bayesian posterior probabilities were estimated as the proportion of the trees that contained each node over the number of trees that were sampled during runs. To reconstruct the species tree (BEAST), three independent pre-runs of ten million generations with the length of the MCMC were performed to obtain better parameter priors for the following five independent ten million generations of the MCMC process. Genealogies were sampled every 1000 generations, and the first 10% were discarded as burn-in. All the statistics of the output values of species-tree inference were summarized using TRACER ver. 1.5 (Rambaut & Drummond, 2003), and both the log files and the tree files of the last five runs were combined using LogCombiner ver. 1.6.1 (Drummond & Rambaut, 2007). TreeAnnotator ver. 1.6.1 (Drummond & Rambaut, 2007) and FigTree ver. 1.3.1 (Rambaut, 2008) were used to summarize and display the sampled trees, respectively. Both gene tree (reconstructed by MrBayes) and species tree (reconstructed by BEAST) have similar topologies except that there are more unresolved nodes in the gene tree. The reconstructed BI species tree is displayed in Figure 1.

The phylogenetic tree revealed six major clades, namely clades 1~6 (Figure 1). Three outgroup species, *Chrysolepis chrysophylla*, *Quercus suber*, and *Castanopsis argyrophylla*, are grouped into Clade 1: *L. pachylepi* is grouped with *Castanea crenata*, and *L. rotundatu* is grouped with *Colombobalanus excels* (Figure 1). These phylogenetic mal-positions are consistent with the chaos state of the taxonomy of *Lithocarpus*, and the non-monophyly (*i.e.*, poly- or paraphyly) of *Lithocarpus* also suggests (1) an inappropriate taxonomic treatment for *Lithocarpus*, (2) an incomplete lineage sorting of the cp *atp*B-*rbc*L marker, or (3) a reflection of incomplete sampling. In fact, a consensus of the taxonomy of the genera *Lithocarpus*, *Cyclobalanus*, *Synaedrys*, and *Pasania* was absent: the lumpers preferred combining the above genera as *Lithocarpus sensu lato*, but the splitters suggested that *Lithocarpus sensu stricto* should be separated from the others. In addition, the American species *L. densiflorus* was recognized as a novel genus, *Notholithocarpus*, which was based on recent morphological and phylogenetic evidence (Manos et al., 2008). Notwithstanding the taxonomic controversy of *Lithocaprus*, the occurrence of lineages in the phylogenetic tree is not affected by subjective classification, and it does not influence the following analysis of the diversification-rate shift because it only considers the "appearance" (speciation) and "loss" (extinction) of lineages instead of the taxonomic positions.

2.4 Inferring the diversification rate

The diversification rate shift of the clades of the *Lithocarpus* species were inferred based on a topological-based approach using the SymmeTree program (Moore & Chan, 2005). The asymmetric rate shift of the nodes of the species tree, which was inferred by the Yule's pure-birth model using BEAST, was estimated using the single-tree analysis that searched for the rapid evolving clades. Although Venditti et al. (Venditti et al., 2010) raised an objection that the Yule process in general may not explain the topology and branch-length distributions in real trees, the Yule process of diversification expects a constant rate of speciation with no extinction and a log-linear increase in the appearance of lineages. Therefore the deviation from the log-linear increase of lineage (i.e., rejecting the constant rate of speciation) can be used to infer the diversification rate variation (Rabosky, 2006). The polytomy was randomly resolved for sampling possible dichotomous resolutions using the taxon-size insensitive (TSI) model plus the equal-rate Markov (ERM) random branching model. The TSI-EMR model was selected due to its incomplete and biased sampling of taxa of the *Lithocarpus* species. One million random resolutions with 1000 TSI-ERM resolved trees under one million ERM simulations were performed to obtain the probability of a diversification rate shift of each node. A shift of the diversification rate of six clades of the *Lithocarpus* phylogeny was also tested separately to confirm the constant rate of diversification in these evolutionary lineages. The rate shift was also estimated using the gene tree inferred by MrBayes (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). Comparing the gene tree and the species tree aids in understanding whether the inferred rate variation is caused by the rapid substitution rate of the marker.

The result of the diversification rate shift in the species tree indicated that the inconstant diversification rate of the topology of the whole tree was determined by a significant deviation ($P<0.05$) from the constant-rate hypothesis by indices IC, M_Π^*, M_Π, M_Σ^*, M_Σ, and B1 (Table 2). We further detected the diversification rate shift based on the Bayesian tree using the Δ-shift statistic, which assessed the probability of a rate shift along an internal branch by calculating a function of likelihood difference between the homogeneous model (one diversification rate parameter) and the heterogeneous model (two-rate parameter). Three nodes, A, B, and C, were detected to have asymmetric diversification rates for the derived clades, in which the constant diversification rate was rejected (P = 0.014, 0.001, and 0.013, respectively) by the Δ1 statistic (Figure 1). Three deriving clades of the nodes A, B, and C, which had relatively higher diversification rates than the sister clades, were named clades A_R, B_R, and C_R, respectively (R signifies "radiation clade"). However, the homogeneous model was rejected at two nodes (indicated by the stars in Figure 1) in the gene tree: one node was located within Clade 1, and the other was identical to node B of the species tree (Figure 1). This result might suggest that the inference of evolutionary radiation in clade B_R could be due to a rapid evolutionary (substitution) rate of cp *atp*B-*rbc*L, although more evidence is needed to confirm this speculation.

2.5 Diversification rate shift and biogeographic inference

The formation of the diversity center of *Lithocarpus* in Southeast Asia is a serial process that is composed of fragmentation, dispersal, and hybridization, etc. Although the climate change in Southeast Asia was not as severe as temperate regions, the rise and fall of the sea level severely altered the area of landmass of Southeast Asia, *i.e.*, the emergence and

submergence of the Sundaland during glacials and interglacials, respectively. Such geographic processes might result in the differentiation of the composition of species in various areas. Three of the six clades of *Lithocarpus*, Clades 4~6, are mainly composed of species distributed in Borneo and the nearby regions (Indochina). In contrast, the species distributed throughout continental Asia (China) are primarily found in Clades 1~3. However, the species composition of the Clades 1~3 is not as simple as that for Clades 4~6;
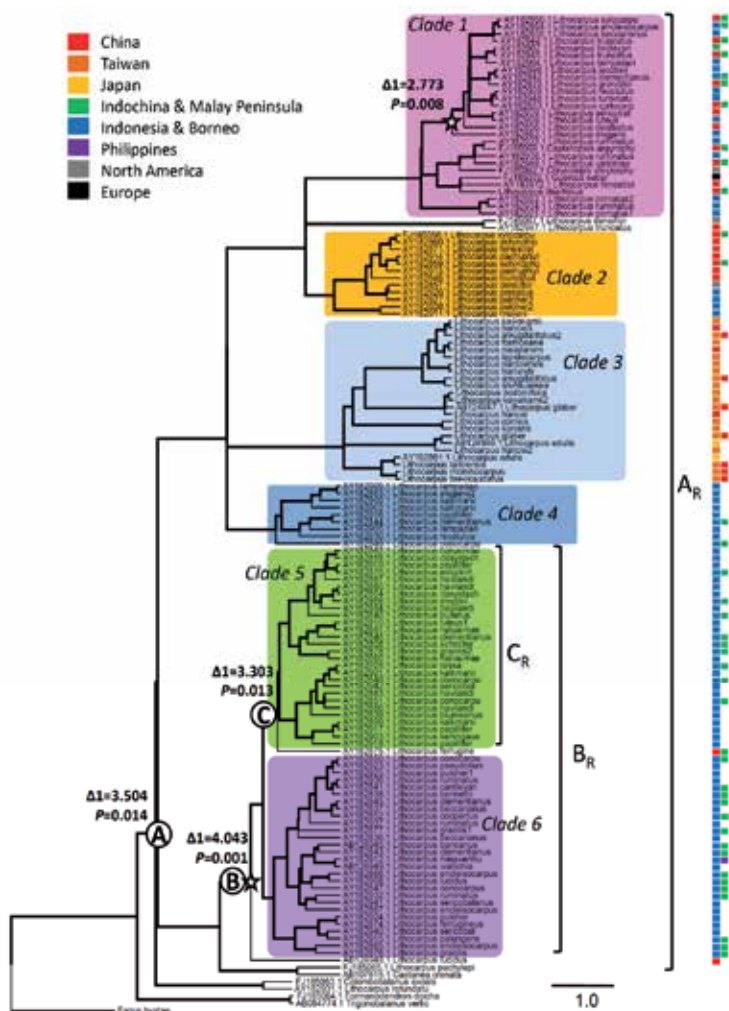
Fig. 1. The Bayesian inference species tree inferred by the Yule's pure-birth model. The Δ statistic (Δ1) and its tail probability are indicated in the nodes of the significant diversification rate shift (A~C). The nodes marked by the stars are significantly shifted in the diversification rate that was estimated in the gene tree. Clades A_R, B_R, and C_R are the derived clades with high diversification rates of nodes A, B, and C, respectively, namely radiation clades. Six major clades of *Lithocarpus* (Clades 1~6) are indicated in colored blocks. The colored boxes to the right indicate the distribution area of the species. The bold lines represent branches with >90% statistical support by posterior probability.

Clade	IC	M_Π*	M_Π	M_Σ*	M_Σ	B1
Whole tree	0.0012	0.0001	0.0012	0.0076	0.0768	0.2193
Clade 1	0.0986	0.1064	0.1690	0.2039	0.3229	0.1771
Clade 2	0.3322	0.3396	0.4043	0.3425	0.4236	0.5100
Clade 3	0.1675	0.2450	0.3973	0.1859	0.3900	0.8635
Clade 4	0.3111	0.2611	0.2873	0.3202	0.2754	0.1937
Clade 5	0.3329	0.2434	0.2357	0.4177	0.3300	0.1384
Clade 6	0.5280	0.4005	0.4886	0.5365	0.6662	0.8031

Table 2. The tail probabilities of the phylogenetic uncertainty of the species tree on inferences of among-lineage diversification rate variation that was assessed using one-million random resolutions with the TSI-EMR algorithm.

the Bornean species and the continental species are phylogenetically interlaced in Clade 1 but distinctly positioned at different subclades in Clade 2, and most of species of Clade 3 are distributed in the continental islands of Taiwan and Japan (Figure 1). Based on the phylogenetic position and geographic distribution of the species of *Lithocarpus*, Borneo is suggested as their origin and as the diversity center, and the Indochinese and Chinese species are probably allopatrically derived due to vicariance or long-distance colonization. Such biogeographic inference is consistent with Cannon & Manos's phylogeographic inference of the network by the separation of a widespread clade and the Bornean clade (Cannon & Manos, 2003). Introgression and bottleneck events of Pleistocene climatic change were therefore used to explain the unapparent genetic structure and the homogeneity of the species composition of the widespread species of *Lithocarpus* that was distributed in Indochina and Borneo (Cannon & Manos, 2003). The Pleistocene glacial/interglacial oscillations largely affected the route of species colonization and their evolution, which was very similar to the European Fagaceae (*Quercus* and *Fagus*) (Comes & Kadereit, 1998). The influence of Quaternary climate change on the Asian tropical forests has been mainly reported on the spatial/demographic expansion and long-distance dispersal of mangroves (Liao et al., 2007; Liao et al., 2009; Liao et al., 2011) or the tropic forest shift at elevational gradients (Colwell & Rangel, 2010). However, the emergence/submergence of the Sundaland also played an important role in the range shift of the SE Asian Fagaceae (*Lithocarpus*), such as in the contraction and expansion of the European *Quercus* and *Fagus* during the Quaternary period.

A serial event of radiative evolution (diversification rate shifts at nodes A, B, and C) was detected in the evolutionary history of *Lithocarpus* (Figure 1). The rate shifts at these three nodes indicate that three catastrophic events caused the burst of appearance of three species: (1) the occurrence of the core *Lithocarpus* (Clade A_R), (2) the rapid evolution of the core Bornean *Lithocarpus* (Clade B_R), and (3) one subclade of B_R with a high speciation rate (Clade C_R). In addition, it has been inferred that the subclades of Clade 1 and the Clade B_R evolve relatively more rapidly than the sister groups in the cp *atp*B-*rbc*L sequences in the gene-tree analysis (the nodes marked as stars in Figure 1). The rate shift in the gene tree fails to exclude the influence of local molecular clocks of selected markers on the inference of the diversification rate shift in the species tree. In other words, we cannot exclude the possibility that the inference of the evolutionary radiation of Clade B_R was caused by the rapid substitution rate of the cp *atp*B-*rbc*L spacer. The use of multilocus markers will hopefully resolve this ambiguity.

In addition to these three rate shift events, the other diversifying events (*i.e.*, the other nodes) derived descendents in the saturating phases of speciation rate and the extinction

rate. None of six major clades (Clades 1~6) had a significant rate shift in their descendents (Table 2). This indicated that the diversification rate shift may not have occurred in the recent past but rather a long time ago. However, by performing a temporal analysis of each clade using lineage-through-time (LTT) plots (with the assistance of Mesquite ver. 2.6 (Maddison & Maddison, 2009)), we observed that the rates of lineage appearance were varied at the early stage of Clade 1, at the late stage of Clade 3, and at the middle stage of Clade 6 (Figure 2). In contrast, the growth rate of the lineage appearance is constant in Clades 2, 4, and 5, which indicated that the appearance (speciation) and loss (extinction) of lineages of these three clades reached balanced phases. Clade 1 was composed of species distributed in diverse areas (China, Indochina, Thai-Malay Peninsulas, Indonesia, and Borneo). Species diversification (which involved in the integration of speciation and extinction) in this clade was probably accomplished through multiple processes, such as repeated colonization and fragmentation. Severe climate change is adverse to the survival and reproduction of organisms but accelerates the speciation rate by natural selection (in strict habitats), isolation (between refugia), and rapid drift (in small populations) (Rieseberg & Wendel, 2004; Rieseberg & Willis, 2007; Pagel et al., 2010). The past climate fluctuations (for example, the Pleistocene glacial/interglacial oscillations) that resulted in the connection and disconnection of Indochina and Borneo (*i.e.*, the emergence/submergence of the Sundaland) may be responsible for the early increase in the diversification rate in Clade 1.

Sea-level changes during the Pleistocene glacials were also responsible for the composition of species around the coasts and the continental islands of East Asia and the genetic diversity of species, including, for example, gobioid fishes in the northeastern Pacific coasts (Fu et al., 2010), lilies in Taiwan and Ryukyus (Hiramatsu et al., 2001), the genus *Kirengeshoma* in East China and South Korea and Japan (Comes et al., 2009), and monkeys in China, Japan, and Taiwan (Wu et al., 2007). The varied diversification rate through the time period of Clade 3 (Figure 2) that primarily encompassed the East Asian continental islands (Taiwan and Japan) was probably a consequence of repeated fragmentation and secondary contacts of populations between the islands (Japan and Taiwan) and the continent (China) during the Pleistocene glacial oscillations. The diversification rate variation of lineages in these continental islands was also involved in the ragged topographies, which provided multiple heterogeneous microhabitats and resulted in niche differentiation of these species. Therefore, the integration of colonization and both adaptive and non-adaptive radiation is responsible for the species endemism (diversity) of these continental islands.

The diversification rate changes in Clade 6 inferred by the LTT analysis are relatively more complicated than those of other clades: the diversification rate increased initially and was followed by a phase of decreased rate, but then ended with a rapid rate recovery (Figure 2). Clade 6 is mostly composed of species from Borneo, the Indochina Peninsula and the Philippines (Figure 1), which has been suggested as the diversity center of this genus (Cannon & Manos, 2003). The factors that caused the change of the speciation rate of Clade 6 are not known yet. Several factors may have separately or jointly affected the diversification rate of this clade in regions such as the diversity center. In actuality, the violation of the constant rate may not necessary involve natural selection or any biotic effects but only stochastic processes (Venditti et al., 2010). Notwithstanding the unknown causal effects of diversification rate variations of the Bornean species, the fluctuated rate reveals an unbalanced speciation-extinction process of *Lithocarpus* in Borneo and the

nearby regions (e.g., Indonesia and Malay Peninsula), which may explains the high diversity of these regions.
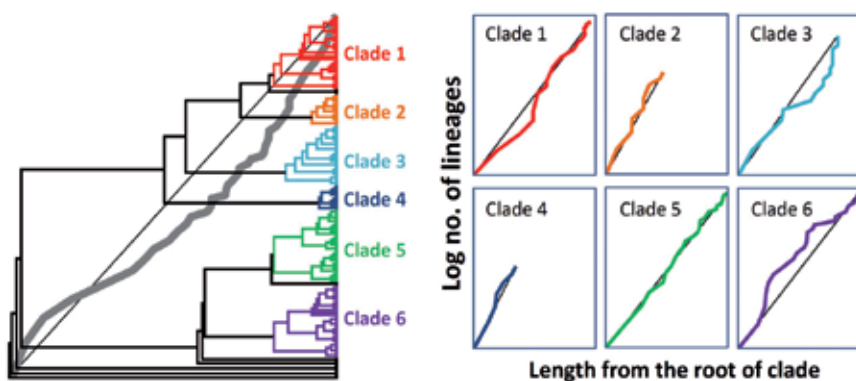
Fig. 2. Plot of the lineage-through-time (LTT) analysis of *Lithocarpus* using the species tree inferred by the Bayesian-based Yule's pure birth model. The left panel is the speciation tree inferred by Yule's pure-birth model, and the gray curve indicates the LTT curve (scaled by the log of the number of lineages). The right panel is the LTT curves of the specific clades (Clade 1~6, indicated in different colors). Straight lines under the LTT curves indicate the averaged LTT in constant growth rate.

2.6 Brief summary

Pleistocene climate changes predominately affected the migration, isolation, and local extinction of species and resulted in diversification rate variation (Comes & Kadereit, 1998; Willis & Niklas, 2004). Major increases in the vascular plant diversification rate occurred at least three times in the past 50 Mya (Willis & Niklas, 2004), during which the climate dramatically changed. The glacial refugia and the postglacial expansion routes of the European and American trees are not unequivocal. Although the glacials affected the climates in the Asian tropics and subtropics less, they changed the SE Asian topography, *i.e.,* in the Sundaland. The phylogenetic and geographic association of *Lithocarpus* (Figure 1) implied (1) its origination in Borneo and its subsequent northward colonization and (2) independent evolution of *Lithocarpus* in the continental islands (Taiwan and Japan). In addition, based on estimates of diversification rate variation, radiation events were detected in the Bornean species by topological analysis (SymmeTree, Figure 1), and an inconstant growing rate of speciation was detected using temporal analysis (Figure 2) in the lineages of the colonization clade (Clade 1), the continental-island clade (Clade 2), and one derived clade in Borneo (Clade 6). The causal event of diversification rate variation is likely related to the emergence and submergence of Sundaland by Pleistocene glacial/interglacial oscillations. However, we cannot exclude other possibilities, such as intrinsic or stochastic factors.

3. Case study: Hybridization and fern speciation

3.1 Introduction

Hybridization is an important mechanism for speciation. Most studies of fern hybrids are noticed by marked morphological variation, which appears to be more than the

morphological variation within a species. The ferns have continuous morphologies between two species. Consequently, the plants with intermediate morphologies are regarded as hybrids, such as some *Pteris* species (Walker, 1958; Wagner & Nauman, 1982; Pangtey et al., 1990). This continuous morphology also results in possible complicated phylogenies, and it is often difficult to identify specific species (Baack & Rieseberg, 2007; Chapman et al., 2007; Rieseberg & Willis, 2007). The earliest reported could be the reticulated evolution of *Asplenium* species in the Appalachian, which was supported by the mispairing of chromosomes (Wagner, 1954). Together with morphological characters, a network relationship constructed using multiple and recurrent hybridizations is revealed.

3.2 Fern speciation under hybridization

Hybridization is a kind of sympatric speciation. After a hybridization event, reproduction isolation is an important mechanism for reinforcement (Turelli et al., 2001). Hybrids are able to maintain sexual function by means of direct genome addition during crossing or genome duplication; the offspring could have successful chromosome pairing during meiosis. For example, two diploid species can cross and produce an allotetraploid, or first have a diploid (homoploid) hybrid which then undergos chromosome doubling to become an allotetraploid, that is, an amphidiploid. For example, *Isoetes sinensis* is an amphidiploid formed from the hybridization of two diploids, *I. taieanensis* and *I. yunguiensis* (Taylor et al., 2004). The amphidiploid species can have sexual function and further hybrid with other species. It has been suggested that *Isoetes sinensis* may backcross with it parental species, *I. taieanensis*, and produce another species, *I. japonica* (Kim et al., 2010).

In another case, hybrids can be ampogamous, such as hybrids that maintain homoploid speciation, or hybrids that have an odd number of chromosome sets (e.g., allotriploids and allopentaploids) (Park & Kato, 2003). For some hybrid ferns, apogamy increases their fitness by providing a mechanism by which failed meiosis due to chromosome mispairing can be avoided (Mehra & Singh, 1957; Braithwaite, 1964; Bretagnolle & Thompson, 1995).

Strictly speaking, it would be untrue to think that fern hybrids should have morphological variation. Apogamy is common in ferns. Apogamous hybrids and their progenies have almost the same morphologies due to the same genetic characters. If an apogamous hybrid species indeed has genetic variation, it may be caused by: (1) the recurrent origins of apogamous races from the same sexual species (Lin et al., 1995; Takamiya et al., 2001), (2) the recurrent hybridization between apogamous races and the same sexual species (Walker, 1962; Suzuki & Iwatsuki, 1990; Lin et al., 1992), (3) somatic mutations (Stebbins, 1950), (4) unequal segregation during meiosis (Lin et al., 1992), or (5) homoeologous pairing at meiosis (Klekowski, 1973).

For apomictic hybrids, the lack of recombination preserves their heterozygosity (Normark et al., 2003). As a result, apogamy creates reproductive barriers that prevent gene flow among closely related taxa, which facilitates sympatric speciation. Excepting some possibilities that can increase the genetic variation of apogamous ferns, apogamy can actually be thought of as reproductive cloning. It is doubtful that the apogamous hybrid is only alive for a short time period and would be soon extinct. Recently, it has been revealed that the asexual *Astrolepis* taxa exit 7 to 47 times younger than sexual parents (Beck et al., 2011). Could this signify that the apogamous ferns have much shorter lives than sexual ferns? This idea requires more research for clarification.

Introgression is led by backcrossing and has been frequently reported in angiosperms. Although the mechanism is discussed in fern hybridization, it is mostly limited in presumption. The limited reports may be because most fern hybrid offspring immediately gain apogamy. Possible introgression cases are found in genus *Asplenium* (Polypodiaceae) and *Serpocaulo* (Aspleniaceae) (Van den heede et al., 2003; Kreier et al., 2008; Hunt et al., 2011). A hybrid population of *Polystichum munitum* and *P. imbricans* may be the only study to provide strong and direct evidence of introgression (Kentner & Mesler, 2000). The hybrids have intermediate morphologies in the hybrid zone, which maintain environmental gradients. It is suggested that the hybrids backcross with parental species, then exhibit clinal patterns of variation.

Although the morphological gradient due to backcrossing is seldom recorded in ferns, a kind of "hybrid swarm" is found in ferns. Limited to a specific geographic area, several parental species cross with one other and produce a morphological gradient, which is formed by a mass of hybrid offspring. In New Zealand, seven tetraploid *Asplenium* species cross with each other to produce 10 octoploid hybrids (Shepherd et al., 2008). One of the parental species, *A. hookerianum*, even devotes it genetic element to seven hybrid species. In this morphological gradient, it seems impossible to understand the phylogenetic relationship without molecular evidence.

3.3 Species concept in fern hybridization

In light of the high frequency of hybridization, the concept of a biological species does not appear appropriate to all fern species. The high frequency of hybridization has been explained by ferns' relative lack of pre-zygotic reproductive isolation mechanisms. Many hybrid species are sexual and cross with other species, especially *Asplenium* species (Trewick et al., 2002; Van den heede et al., 2003; Shepherd et al., 2008). However, as we discussed above, apogamy is common in ferns and is maintained in absolute reproductive isolation. Asexual organisms are not discussed with respect to the biological species concept (Arnold, 1997).

In practice, the species boundaries in ferns are often problematic (Haufler, 2008). Comparatively, "to define a species" is easier. The morphological species concept is more commonly accepted in ferns than are other species concepts. In other words, a fern taxon with identical morphologies could be treated as a species. Most hybrid ferns, such as *Acrorumohra subreflexipinna* (Ogata) H. Ito., *Archangiopteris itoi* Shieh, and *Adiantum meishanianum* F. S. Hsu ex Y. C. Liu & W. L. Chiou (Chang et al., 2009; Liu et al., 2009), were considered as species before hybridization was evident.

Ideal morphological analyses provide the independent clustering of taxa (Thomson, 2000). However, these analyses are not always sufficient, especially for a species complex. Taxa in a species complex seem to exhibit continuous morphologies with one other, and the gaps, for species delimitation, are difficult to identify. In some cases, taxa have clear hybridization patterns but have poorly differentiated morphologies, and it is difficult to determine their taxonomic ranker. Three cryptic species in *Ceratopteris thalictroides* species complex have been isolated based on crossing tests, that is, the biological species concept instead of the morphological species concept (Masuyama et al., 2002). Sequential studies have provided evidence that they are of hybrid origin and have a different number of chromosomes (Adjie

et al., 2007; Masuyama, 2008). Although a detailed diagnosis of the characters of the species is analyzed for the final taxonomic treatment, the continuous variation in morphologies remains elusive (Masuyama, 2008; Masuyama & Watano, 2010).

However, it is still necessary to face the problem: should the hybrid taxon be thought of as a hybrid species or just a hybrid offspring? It is a more difficult issue during the taxonomic treatment of a reticulate phylogeny arising from hybridization. The biological species concept could be applied, although there is high frequency of hybridizations in ferns. In the reticulated relationship of the *Vandenboschia radicans* complex, the fertile amphidiploid taxa are treated as species, whereas the taxa are treated as hybrids (Ebihara et al., 2005; Ebihara et al., 2009).

3.4 Abundant hybridizations in the genus *Pteris* (Pteridaceae)

Genus *Pteris* L., belonging to Pteridaceae, is a cosmopolitan genus with more than 250 species and is diverse in Asia. Many hybridizations of *Pteris* have been reported (Walker, 1962). According to a review of the breeding systems of 99 *Pteris* species, 31 species are apogamous (Chao et al., unpublished). This ratio is higher than the previously estimated value of 10% (Walker, 1984). Apogamous fern species are usually considered potential hybrids. However, this does not mean that the speciation of 31 species from the 99 *Pteris* species are of hybrid origin. As chromosomal mispairing of some autopolyploids could also cause apogamy, such as in autotriploids and autopentaploids, it is difficult to distinguish before a more detailed investigation. A species can have several cytotypes. If we consider the frequency of the reproductive mode combined with cytotypical data, the frequency of apogamous triploids is 21%, and we still have 11% apogamous diploid and 6% apogamous 4X. In total, there is at least a 17% frequency of hybrids and apogamy in the *Pteris* species.

In addition, there are many species complex in *Pteris*. One of the most famous of *Pteris* is the *P. quadriaurita* complex. Species complex may involve multiple mechanisms, including hybridization, polyploidy, morphological plasticity, gene expression, geographical divergence, and reproductive traits (Rosenthal et al., 2002; Guo et al., 2004; Lai et al., 2005; Zhuang & Adams, 2007; Guggisberg et al., 2009). As all of these are closely associated with speciation, species complex could be considered a group of taxa undergoing speciation.

In this study, we supplied more species in the sampling for the genus *Pteris*. We sequenced the *rbc*L gene of cpDNA from 25 *Pteris* species. Relevant sequences were collected from the previous studies and via NCBI BLAST (http://blast.ncbi.nlm.nih.gov/) searches, including 17 *Pteris* species and related species belong to other genera. Maximum likelihood calculations were performed for phylogenetic relationships (Zwickl, 2006).

First, within the maximum likelihood tree (Figure 3), which includes 40 *Pteris* species, *Pteris* is a paraphyletic genus. We provided more samplings than previous studies, and our analysis still supports previous studies (Jefferson et al., 2007; Schuettpelz et al., 2007). Two genera are clustered in the major clade of *Pteris*, *Neurocallis* and *Ochropteris*. Furthermore, there is a strange species, *Pteris vittata*, which is not closely allied to the other *Pteris* species. That could be another issue about taxonomy (Christenhusz et al., 2011). *P. vittata* could arise from multiple hybridizations. The species shows considerable morphological variation and diverse cytotypes, including sexual diploids, triploids, tetraploids, pentaploids, and

hexaploids (Srivastava et al., 2007). Furthermore, the chromosomes of the spore mother cells show varied multivalence, such as 20I+26II+5III, 9I+45II+3III+2IV, 29II+29I, and 29II+87I (Khare & Kaur, 1983; Wang, 1989), which indicates its hybrid origin.

Fig. 3. Chloroplast DNA tree based on the *rbc*L gene and assessed using maximum likelihood (ML) analysis. Arrows indicate two genera included in the clade of genus *Pteris*. Numbers above the branches are ML bootstrap (> 50%).

3.5 Species complex in the genus *Pteris* (Pteridaceae)

Division of the lamina has been considered as key characters for section classification of *Pteris* (Shieh, 1966). These characters, simple pinnate, bipinnatifid, or semi-bipinatifid (irregular bipinnatifid) were mapped on the ML cladogram, using the trace character function and ACCTRAN optimization as implemented in MacClade (Maddison & Maddison, 2000) (Figure 4). It is indicated that the character is not so consistent in a single clade and could be not a synapomorphic character.
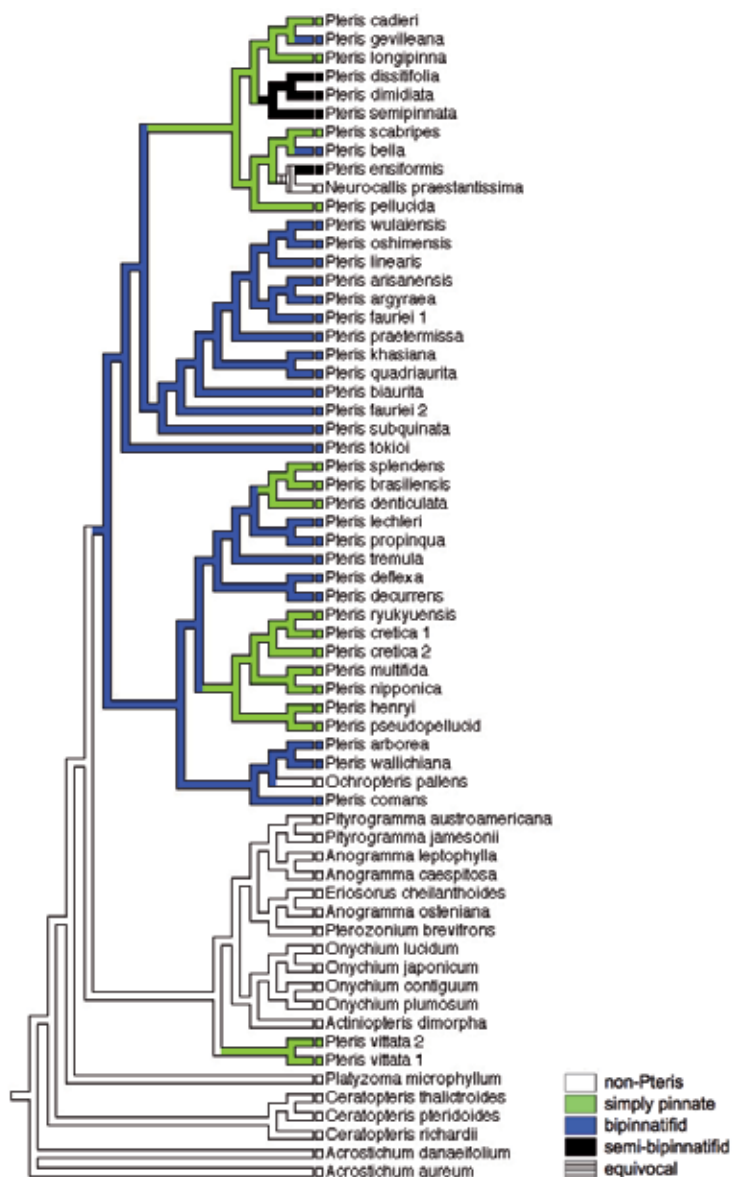

Fig. 4. Division of the lamina are mapped onto the MacClade optimization tree inferred by the maximum likelihood analysis.

Several species complexes can be found in the tree, and which indeed prove the species within the particular species complex having close relationship. They are *P. cadieri* complex, *P. cretica* complex, *P. quadriaurita* complex, and *P. vittata* complex (Figure 3). *Pteris cadieri* complex is distributed throughout eastern Asia, from including Japan, China, Taiwan, Indochina and Malaysia. The species complex exhibits markedly morphological variations, from simply pinnate to bipinnatifid fronds. Previous studies suggested that *P. grevilleana* was parental species of *P. cadieri* (Chao et al., 2010). However, it has been found that the two species are apagamous. Both of they could arise from hybridization.

The clade consisted by species with simple pinnate frond is *Pteris cretica* complex, including 6 species in this tree (Figure 3). *Pteris cretica* is widely distributed in warm-temperate and tropical parts of the Old World (Kramer & McCarthy, 1998). The two "*P. cretica*" have different molecular traits. The first study on *P. cretica* has revealed its apogamy and different cytotypes (Manton, 1950). Although sexual diploid *P. cretica* was reported, apogamous diploids, triploids, and tetraploids suggested possible hybridization (Walker, 1962; Roy et al., 1971; Nakatô, 1975; Huang et al., 2011) and supported by mispairing during meiosis (Verma & Khullar, 1965; Jha & Sinha, 1987). Based on the evidence of allozyme, triploid *P. cretica* may be produced by diploid apogamous *P. cretica* and diploid sexual *P. kidoi* (Suzuki & Iwatsuki, 1990). In *P. cretica* complex, some ferns with intermediate morphologies are suggested hybrid origin. For example, *Pteris namegatae*, recorded in Japan, is thought as a hybrid species between *P. multifida* and *P. ryukyunesis* (Kurata, 1969).

In the earliest reports, it was proposed that *P. quadriaurita* complex is consisted by *P. quadriaurita sensu stricto*, *P. multiarita*, *P. confusa*, and *P. otaria* in Ceylon (Walker, 1954; Walker, 1958). In Dr. Walker's hybridization experiment, the four plants were from hybridizations among tree parental species. Practically, there are more bipinnatifid ferns difficult to distinguish from *P. quadriaurita*. In figure 3, 10 species clustered together with *P. quadriaurita* complex, are so similar and difficult to identify. For this great group of *Pteris quadriaurita* complex, it is unsure if the problem is just taxonomical.

Totally, the difficulty of studying *Pteris* taxonomy and phylogeny could be species complexes, hybridization. Serious hybridizations form reticulate evolution. Different methodologies and synthetical analysis are in need.

3.6 Brief summary

Hybridization is an important mechanism for speciation. It increases the opportunity for hybrids to reproduce and maintain their lineages. Beginning with recurrent hybridizations that were revealed in the genus *Asplenium*, more complicated depictions have been reported in the years of the previous decade, including genus *Equisetum*, *Isoetes*, *Cheilanthes*, and *Vandenboschia*. The genus *Pteris* L., which belongs to Pteridaceae, is a cosmopolitan genus with more than 250 species and is diverse in Asia. It has been found that plants with different morphologies have the same scientific name in different local floras. Meanwhile, species complexes have confused morphologies and are often distributed across different geographical areas. Based on phylogenetic evidence, some taxa have been shown to be of hybrid origins. We intended to reconstruct the evolutionary history of the *Pteris* in Asia and to investigate the role of hybridization in fern speciation.

4. Case study: Divergence of *Toxicodendron radicans* ssp *hispidum* and *orientale* in East Asia

4.1 Introduction

In the past 2 million years, climate changes have often been interpreted as the primary driver of range fragmentation and the speciation of plants and animals (Hewitt 2000). Tree species have responded by migrating to regions where the environmental conditions allowed them to survive (Hewitt 1996). Phylogeography investigates the spatio-temporal dynamics of populations. The field relies on inferences that are made from macrofossils and pollen in sediment profiles and molecular evidence that can reveal aspects of history, such as the location of cryptic refugia (cf. Ge et al., 2005). Few paleoecological studies have identified the pollen and fossils of tree species. The molecular evidence provides an effective method for inferring the number of genetically distinct ancestral lineages, their location during the last glacial maximum, and the postglacial migration routes for tree and plant species (Chiang et al. 2006; Naydenov et al. 2007; Cheng et al. 2008). Altogether, phylogeography and genetic characteristics of populations/species are dictated by the interplay between historical vicariance and recurrent genetic exchanges (cf. Chiang & Schaal 2006). These evolutionary events leave evolutionary footprints in the spatial apportionment of genetic polymorphisms within and among populations across the distributional range of *P. massoniana*, which in turn provides information for recovering their demographic histories (Schaal & Olsen 2000). Of the geological events, historical regular glacial cycles in the Eurasia Continent have had prevalent influences on the survival (extinction) and recolonization of populations/ species. According to geological evidence, during the early Pleistocene period, ice ages occurred at regular intervals of 100,000 years followed by a 20,000-year warm period (Milankovitch cycles) (cf. Bennett 1990).

In the continental island of Taiwan and Southwest China, only one endemic subspecies, *Toxicodendron radicans* ssp. *hispidum*, is present with an extended distribution over all of the Taiwan middle high mountain area and Yunnan, Szechwan, Guizhou, Hunan, Hubai and Fujian. However, in the continental islands of Japan, another endemic subspecies, *T. radicans* ssp. *orientale*, is present with an extended distribution over all of Japan. The close sister for both subspecies is distributed in eastern North America and is divided into several subspecies (Gillis 1975). The total world distribution of the genus *Toxicodendron* includes eastern Asia and North America south to Colombia in South America. Poison ivy is found only in North America and eastern Asia, which is a classical distribution pattern that was reported by a different study and included species and genus levels (Yi et al. 2004; Xiang et al. 2005; Kim et al. 2004; Wang et al. 2003; Wen and Zimmer, 1996; Milne 2004; Jeandroz et al. 1997). Population genetic variation of *T. radicans* using molecular evidence is important for the reconstruction of the distribution patterns of poison ivy between continents and islands and to rebuild the evolutionary paths of these subspecies. *T. radicans* in East Asia is a good model system to use the population genetics approach to explain the effects of the land bridge disconnection after the ice age.

4.2 Data collections

Toxicodendron radicans ssp. *hispidum* is distributed in Southwest China and between the island of Taiwan and the continental mountains of Southwest China, including Yunnan, Szechwan, Guizhou, Hunan, Hubai and Fujian. One-hundred and eighty individuals from

Taiwan and 174 from China were sampled for ssp. *hispidum*. *Toxicodendron radicans* ssp. *orientale* is restricted to Japan, ranging from Hokkaido to Kyushu; a total of 259 samples were collected from 35 locations from Hokkaido to Kyushu. More than 500 individuals were sampled in two subspecies to compare the population differentiation by plate disjunction affected by the landbridge formation and breakdown. Young and healthy leaves were collected from the field, rinsed with tap water and dried in silica gel. Leaf tissue of the above materials were ground to powder in liquid nitrogen and stored in a -70℃ freezer. Genomic DNA was extracted from the powdered tissue according to a CTAB methodology (Doyle and Doyle 1987). For ISSR fingerprinting, a total of 200 primers from UBC were pretested, and 11 primers were chosen for population genetic analyses. PCR amplification was carried out in a total volume of 20 ul reaction using 2 ul of 10X reaction buffer, 2 ul of MgCl2 (25 mM), 2 ul of dNTP mix (8 mM), 10 pmole of ISSR primer, and 2 U of Taq polymerase (Promega, Madison, USA). The reaction was optimized and programmed on a MJ Thermal Cycler (PTC 100). The reaction was performed as follows: one cycle of denaturation at 95℃ for 4 min, 40 cycles of 45 s denaturation at 94℃, 1 min 15 s annealing at 37℃, 1 min 30 s extension at 72℃, and 10 min extension at 72℃. A negative control, in which no template DNA was added, was included in the PCR reaction to test for contamination. The PCR products were separated on 2.0% NuSieve® 3:1 Agarose gels using 1X TAE buffer. The gel was stained with ethidium bromide for 30 min, and DNA fragments were identified using the Bio-Rad Molecular Imager Gel Doc XR+ System for gel documentation. Each sample was run three times to obtain repeatable results. The ISSR bands were scored by dominate markers using 1 (present) or 0 (absent) binary characters. Analysis of molecular variance (AMOVA) and principal coordinate analysis (PCoA) were performed to estimate the hierarchical apportionment of variation using GenAlEx v. 6.1 (Peakall and Smouse 2006).

4.3 Population differentiation based on ISSR molecular markers

Population differentiation based on ISSR dominate markers was estimated by AMOVA and PCoA analyses. Both showed the high differentiation among populations of islands and continents. In total, the analyses showed 36.67% variation among populations and 63.33% variation among individuals within the populations. Based on the AMOVA results, the mean Φ_{ST} was 0.367, and the result was significant (p<0.010). The PCoA result showed three groups separated in Japan, Taiwan and China (Figure 5). Almost 38.42%, 18.76, and 12.86 of the variation is explained by the first, second, and third axes in this test, respectively. Based on the results, populations of spp. *hispidum* and spp. *orientale* were differentiated among the three regions of the islands and the continent. According to a hierarchical AMOVA analysis, significant genetic differentiation was detected among these populations, with values of Φ_{ST} ranging from 0.021 to 0.850. Using three geographic regions, with values of Φ_{ST} ranging from 0.198 to 0.224, significant genetic differentiation was detected that was much higher than the average Φ_{ST} of 0.17 for outcrossing plant species (Nybom and Bartish 2000; Nybom 2004). The mating system and gene flow were probably two of the main factors that determined the levels of genetic variability and the spatial distribution of the genetic polymorphisms in *Toxicodendron radicans* spp. *hispidum* and spp. *orientale*. In general, outcrossing plants maintained the most genetic variation within the populations, whereas predominantly selfing, short-lived species harbored comparatively higher variation among populations (Hamrick and Godt 1989).
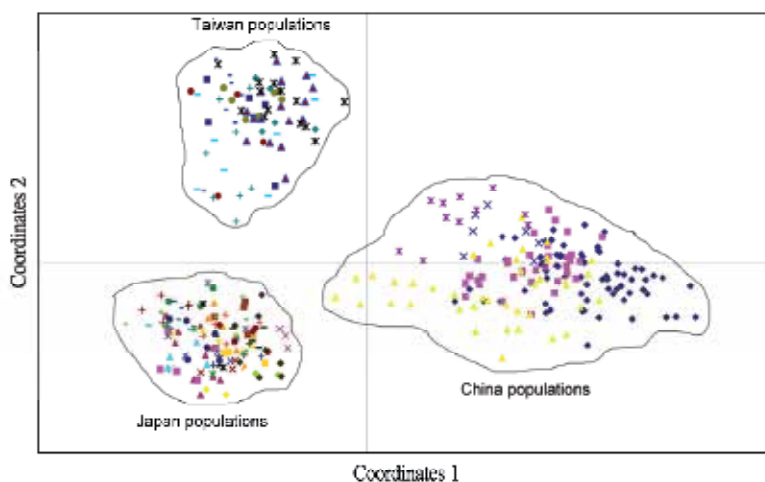
Fig. 5. Distribution of varieties in a two-dimensional PCoA space based on the ISSR data in *Toxicodendron radicans* spp. *hispidum* and spp. *orientale*. According to the ISSR result, the three groups were separated by three different regions, including Japan, Taiwan, and China.

An UPGMA topogram based on the pairwise Nei genetic distance among the populations identified three major clusters that corresponded to the three geographic regions (Figure 6). A close phenetic relationship between the populations of spp. *hispidum* from Taiwan and China was recovered. Populations of Taiwan and populations of China were clustered together in the UPGMA dendrogram. All of the populations of spp. *orientale* from Japan were grouped together within a mono cluster. Geographically, the distribution of spp. *hispidum* in Taiwan is close to that of spp. *orientale* in Japan. Nevertheless, the genetic affinity between spp. *hispidum* of Taiwan and China was identified in the UPGMA topogram, which reflected the different historical migratory route. According to the geological evidence, Japan and Taiwan, now continental islands 150 km east of the Chinese mainland, attained their modern shapes via collisions between the continent and the oceanic tectonic plates only several million years ago (Sibuet and Hsu 1997, 2004). During the Pleistocene glacial expansion, the drops in the sea levels of the East and South China Sea to about 120 m formed the land bridge that emerged from the oceans, which linked Taiwan (Chiang et al. 2001) and Japan to the mainland. Meanwhile, plants and animals migrated from north to south because of the dramatic climatic changes (Comes and Kadereit 1998; Hewitt 1996). The ancestral populations of *T. radicans* were able to migrate across the Bering Straits via a land bridge connection between North America and Asia. After that, the ancestral population separated into two paths: one from Sakhalin to Japan or via the Korean Peninsula to Japan, and one from the north to the south via Siberia which then separated into several paths of scattered refugia in Yunnan, Szechwan, Guizhou, Hunan, Hubai, Fujian and Taiwan (cf Lu et al. 2001; Ge et al. 2005). The genetic drift effected those isolated populations and promoted the process of speciation in various geographic regions.

4.4 Brief summary

The phylogeographic pattern of *Toxicodendron radicans* in East Asia is indicated by the population differentiation and the process of speciation. The spp. *hispidum* were distributed

in the middle high mountain area of Taiwan and the areas of southern China. Their populations were restricted only to refugial regions, which was most likely due to the refugial effect of the Quaternary glaciations. The spp. *orientale* were distributed in lat and hilly areas in Japan and were affected by the bottleneck event in the founder populations, such that they expanded their range and spread to the entire area of Japan. ISSR results
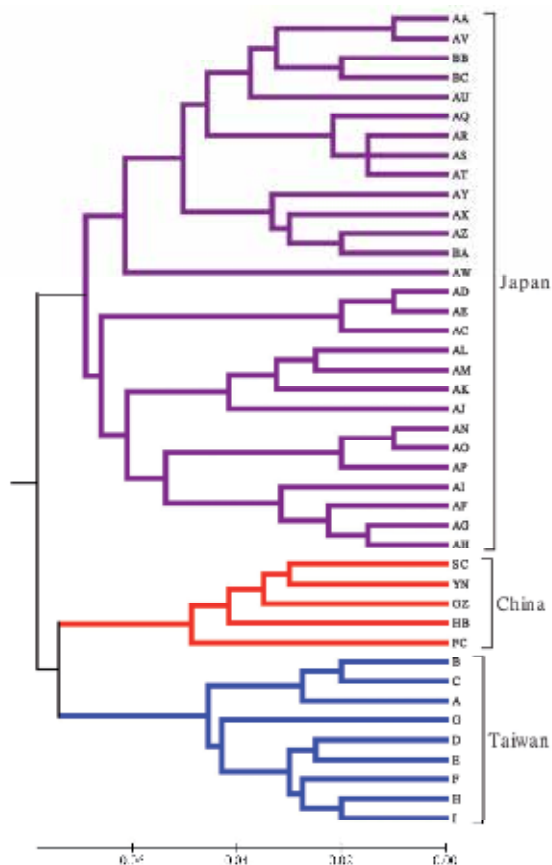
Fig. 6. UPGMA dendrogram of the populations of *Toxicodendron radicans* spp. *hispidum* and spp. *orientale* based on ISSR fingerprints.

showed population differentiation between islands to continents. Different independent lineages with separate clusters are indicative of genetic differentiation among geographic regions and should be considered as evolutionary significant units for plant migration and speciation study.

5. Conclusion

The maintenance of genetic diversity is critical for the long-term survival of species (Frankel and Soulé 1981) because loss of variation may largely limit the adaptability of populations to changing environments. In glacial and postglacial periods, climate and sea level changes affected the distribution of plants and animals. At the same time, isolated populations within

refugial regions might be affected by bottlenecks, which cause genetic variation loss. Meanwhile, the different species showed variant patterns of phylogeography in the postglacial East and South Asia. The process of species occurrence and speciation in the past geologic history was caused by isolation, adaption, and even hybridization in different cases.

6. Acknowledgments

7. References

Adjie, B., Masuyama, S., Ishikawa, H. & Watano, Y. (2007). Independent origins of tetraploid cryptic species in the fern *Ceratopteris thalictroides*. *Journal of Plant Research*, Vol. 120, No. 1, pp. 129-138

Arnold, M.L. (1997). *Natural Hybridization and Evolution*. Oxford Univ. Press, New York

Baack, E.J. & Rieseberg, L.H. (2007). A genomic view of introgression and hybrid speciation. *Current Opinion in Genetics & Development*, Vol. 17, pp. 513-518

Beck, J.B., Windham, M.D. & Pryer, K.M. (2011). Do asexual polyploid lineages lead short evolutionary lives? A case-study from the fern genus *Astrolepis*. *Evolution*, Vol. 65, pp. 3217-3229

Bennett, K.D. (1990). Milankovitch cycles and their effects on species in ecological and evolutionary time. *Paleobiology*, Vol. 16, pp. 11-21.

Blattner, F.R., Banfer, G., Moog, U., Fiala, B., Mohamed, M. & Weising, K. (2006). A chloroplast genealogy of myrmecophytic *Macaranga* species (Euphorbiaceae) in Southeast Asia reveals hybridization, vicariance and long-distance dispersals. *Molecular Ecology*, Vol.15, No.14, pp. 4409-4424

Braithwaite, A.F. (1964). A new type of apogamy in ferns. *New Phytologist*, Vol. 63, pp. 293-305

Bretagnolle, F. & Thompson, J.D. (1995). Tansley Review No. 78. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytologist*, Vol. 129, pp. 1-22

Cannon, C.H. & Manos, P.S. (2003). Phylogeography of the Southeast Asian stone oaks (*Lithocarpus*). *Journal of Biogeography*, Vol.30, No.2, pp. 211-226

Chang, H.M., Chiou, W.L. & Wang, J.C. (2009). Molecular evidence for genetic heterogeneity and the hybrid origin of *Acrorumohra subreflexipinna* from Taiwan. *Amercan Fern Journal*, Vol. 99, pp. 61-77

Chao, Y.S., Liu, H.Y., Chiang, Y.C. & Chiou, W.L. Polyploidy and speciation in the *Pteris* (Pteridaceae). *Journal of Botany*, pp. in review

Chao, Y.S., Liu, H.Y., Huang, Y.M. & Chiou, W.L. (2010). Reproductive traits of *Pteris cadieri* and *P. grevilleana* in Taiwan: Implications for their hybrid origins. *Botanical Studies*, Vol. 51, pp. 209-216

Chapman, M.A., Burke, J.M. & Conti, E. (2007). Genetic divergence and hybrid speciation. *Evolution*, Vol. 61, No. 7, pp. 1773-1780

Chen, K. M., Abbott, R. J., Milne, R. I., Tian, X. M. & Liu, J. Q. (2008). Phylogeography of *Pinus tabulaeformis* Carr. (Pinaceae), a dominant species of coniferous forest in northern China. *Molecular Ecology*, Vol. 17, pp. 4276–4288.

Cheng, Y.P., Hwang, S.Y. & Lin, T.P. (2005). Potential refugia in Taiwan revealed by the phylogeographical study of *Castanopsis carlesii* Hayata (Fagaceae). *Molecular Ecology*, Vol.14, No.7, pp. 2075-2085

Chiang, Y.C., Hung, K.H., Schaal, B.A., Ge, X.J., Hsu, T.W. & Chiang, T.Y. (2006). Contrasting phylogeographical patterns between mainland and island taxa of the *Pinus luchensis* complex. *Molecular Ecology*, Vol. 15, pp. 765–779.

Chiang, T.Y., Schaal, B.A. (2006). Phylogeography of plants in Taiwan and the Ryukyu Archipelago. *Taxon*, Vol. 55, pp. 3–41.

Chiang, T.Y., Hong, K.H. & Peng, C.I. (2001). Experimental hybridization reveals biased inheritance of the internal transcribed spacer of the nuclear ribosomal DNA in *Begonia × taipeiensis*. *Journal of Plant Research*, Vol. 114, pp. 343–351

Christenhusz, M.J.M., Zhang, X.C. & Schneider, H. (2011). A linear sequence of extant families and genera of lycophytes and ferns. *Phytotaxa*, Vol. 19, pp. 7-54

Colwell, R.K. & Rangel, T.F. (2010). A stochastic, evolutionary model for range shifts and richness on tropical elevational gradients under Quaternary glacial cycles. *Philosophical Transactions of the Royal Society B-Biological Sciences*, Vol.365, No.1558, pp. 3695-3707

Comes, H.P. & Kadereit, J.W. (1998). The effect of quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, Vol.3, No.11, pp. 432-438

Comes, H.P., Qiu, Y.X., Sun, Y., Zhang, X.P., Lee, J. & Fu, C.X. (2009). Molecular phylogeography of East Asian Kirengeshoma (Hydrangeaceae) in relation to Quaternary climate change and landbridge configurations. *New Phytologist*, Vol.183, No.2, pp. 480-495

Doyle, J. J. &Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, Vol. 19, pp. 11-15.

Drummond, A.J. & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, Vol.7, pp. 214

Ebihara, A., Ishikawa, H., Matsumoto, S., Lin, S.J., Iwatsuki, K., Takamiya, M., Watano, Y. & Ito, M. (2005). Nuclear DNA, chloroplast DNA, and ploidy analysis clarified biological complexity of the *Vandenboschia radicans* complex (Hymenophyllaceae) in Japan and adjacent areas. *American Journal of Botany*, Vol. 92, No. 9, pp. 1535-1547

Ebihara, A., Matsumoto, S. & Ito, M. (2009). Taxonomy of the reticulate *Vandenboschia radicans* complex (Hymenophyllaceae) in Japan. *Acta Phytotaxonomica et Geobotanica*, Vol. 60, pp. 26-40

Fu, C.X., Qiu, Y.X., Guan, B.C. & Comes, H.P. (2009). Did glacials and/or interglacials promote allopatric incipient speciation in East Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dysosma versipellis*. *Molecular Phylogenetics and Evolution*, Vol.51, No.2, pp. 281-293

Fu, C.Z., Tang, W.X., Ishimatsu, A., Yin, W., Li, G., Chen, H., Wu, Q.H. & Li, B. (2010). Cryptic species and historical biogeography of eel gobies (Gobioidei: Odontamblyopus) along the northwestern Pacific coast. *Zoological Science*, Vol.27, No.1, pp. 8-13

Ge, X. J., Zhou, X. L., Li, Z. C., Hsu, T. W., Schaal, B. A. & Chiang, T. Y. (2005). Genetic structure of the relict *Amentotaxus argotaenia* complex (Taxaceae) based on ISSR fingerprinting. *Journal of Plant Research*, Vol. 118, pp. 415-422.

Gillis, W. T. (1975). Poison-ivy and its kin. *Arnoldia*, Vol. 35, pp. 93-123.

Guggisberg, A., Bretagnolle, F. & Mansion, G. (2009). Allopolyploid origin of the mediterranean endemic, *Centaurium bianoris* (Gentianaceae), inferred by molecular markers. *Systematic Botany*, Vol. 31, No. 2, pp. 368-379

Guo, M., Rupe, M.A., Zinselmeier, C., Habben, J., Bowen, B.A. & Smith, O.S. (2004). Allelic variation of gene expression in maize hybrids. *Plant Cell*, Vol. 16, No. 7, pp. 1707-1716

Hamrick, J.L. & Godt, M. J. W. (1989). Allozyme diversity in plant species. In: Brown, A.H.D., Clegg, M.T., Kahler, A.L. & Weir, B.S. (eds) Plant population genetics, breeding and genetic resources. Sinauer Associates, Sunderland, MA, USA, pp 43–63.

Hardin, J.W. (1975). Hybridization and introgression in *Quercus alba*. *Journal of Arnold Arboretum*, Vol.56, pp. 336-363

Haufler, C.H. (2008). In: *The Biology and Evolution of Ferns and Lycophytes*, T.A. Ranker & C.H. Haufler (eds), 303-331, Cambridge University Press, Cambridge

Hewitt, G. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, Vol. 58, pp. 247-276.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, Vol. 405, pp. 907-913.

Hiramatsu, M., Ii, K., Okubo, H., Huang, K.L. & Huang, C.W. (2001). Biogeography and origin of *Lilium longiflorum* and *L. formosanum* (Liliaceae) endemic to the Ryukyu Archipelago and Taiwan as determined by allozyme diversity. *American Journal of Botany*, Vol.88, No.7, pp. 1230-1239

Huang, Y. M., Hsu, S. Y., Hsieh, T.H., Chou, H.M. & Chiou, W.L. (2011). Three *Pteris* species (Pteridaceae: Pteridophyta) reproduce by apogamy. *Botanical Studies*, Vol. 52, pp. 79-87

Huelsenbeck, J.P. & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, Vol.17, No.8, pp. 754-755

Hung, K.H., Hsu, T.W., Schaal, B.A. & Chiang, T.Y. (2005). Loss of genetic diversity and erroneous phylogeographical inferences in *Lithocarpus konishii* (Fagaceae) of Taiwan caused by the Chi-Chi earthquake: Implications for conservation. *Annals of the Missouri Botanical Garden*, Vol.92, No.1, pp. 52-65

Hunt, H.V., Ansell, S.W., Russell, S.J., Schneider, H. & Vogel, J.C. (2011). Dynamics of polyploid formation and establishment in the allotetraploid rock fern *Asplenium majoricum*. *Annals of Botany*, Vol. 108, No. 1, pp. 143-157

Jeandroz, S., Roy, A. & Bousquet, J. (1997). Phylogeny and phylogeography of the circumpolar genus *Fraxinus* (OLEACEAE) based on internal transcribed spacer sequences of nuclear ribosomal DNA. *Molecular Phylogenetic and Evolution*, Vol. 7, pp. 241–251.

Jefferson, P., Rodrigues, C.D.N., Salatino, A. & Salatino, M.L.F. (2007). Phylogenetic relationships among Pteridaceae, including Brazilian species, inferred from rbcL sequences. *Taxon*, Vol. 56, No. 2, pp. 355-368

Jha, J. & Sinha, B.M. (1987). Cytomorphological variability in apogamous populations of *Pteris cretica* L. *Caryologia*, Vol. 40, pp. 71–78

Kentner, E.K. & Mesler, M.R. (2000). Evidence for natural selection in a fern hybrid zone. *American Journal of Botany*, Vol. 87, pp. 1168-1174

Khare, P. & Kaur, S. (1983). Intraspecific polyploidy in Pteris vittata Linn. *Cytologia (Japan)*, Vol. 48, pp. 21-25

Kim, Y. D., Kim, S. H. & Landrum, L. R. (2004). Taxonomic and phytogeographic implications from ITS phylogeny in Berberis (Berberidaceae). *Journal of Plant Research*, Vol. 117, pp. 175-182.

Kim, C., Shin, H., Chang, Y.T. & Choi, H.K. (2010). Speciation pathway of *Isoetes* (Isoetaceae) in East Asia inferred from molecular phylogenetic relationships. *American Journal of Botany*, Vol. 97, No. 6, pp. 958-969

Klekowski, E.J., Jr. (1973). Sexual and subsexual systems in the homosporous ferns: a new hypothesis. *American Journal of Botany*, Vol. 60, pp. 535-544

Kramer, K.U. & McCarthy, P.M. (1998). *Flora of Australia: Ferns, gymnosperms and allied genera*. Collingwood

Kreier, H.P., Rex, M., Weising, K., Kessler, M., Smith, A. & Schneider, H. (2008). Inferring the diversification of the epiphytic fern genus *Serpocaulo* (Polypodiaceae) in South America using chloroplast sequences and amplified fragment length polymorphisms. *Plant Systematics and Evolution*, Vol. 274, No. 1, pp. 1-16

Kurata, S. (1969). Notes on Japanese Ferns (48). *The Journal of Geobotany*, Vol. 17, pp. 58-61

Lai, Z., Nakazato, T., Salmaso, M., Burke, J.M., Tang, S., Knapp, S.J. & Rieseberg, L.H. (2005). Extensive chromosomal repatterning and the evolution of sterility barriers in hybrid sunflower species. *Genetics*, Vol. 171, No. 1, pp. 291-303

Liao, P.C., Hwang, S.Y., Huang, S., Chiang, Y.C. & Wang, J.C. (2011). Contrasting demographic patterns of *Ceriops tagal* (Rhizophoraceae) populations in the South China Sea. *Australian Journal of Botany*, Vol.59, No.6, pp. 523-532

Liao, P.C., Havanond, S. & Huang, S. (2007). Phylogeography of *Ceriops tagal* (Rhizophoraceae) in Southeast Asia: the land barrier of the Malay Peninsula has caused population differentiation between the Indian Ocean and South China Sea. *Conservation Genetics*, Vol.8, No.1, pp. 89-98

Liao, P.C., Chiang, Y.C., Huang, S. & Wang, J.C. (2009). Gene flow of *Ceriops tagal* (Rhizophoraceae) across the Kra Isthmus in the Thai Malay Peninsula. *Botanical Studies*, Vol.50, No.2, pp. 193-204

Liao, P.C., Kuo, D.C., Lin, C.C., Ho, K.C., Lin, T.P. & Hwang, S.Y. (2010). Historical spatial range expansion and a very recent bottleneck of *Cinnamomum kanehirae* Hay. (Lauraceae) in Taiwan inferred from nuclear genes. *BMC Evolutionary Biology*, Vol.10, pp. 124

Lin, S.J., Kato, M. & Iwatsuki, K. (1992). Diploid and triploid offspring of triploid agamosporous fern *Dryopteris pacifica*. *Botanic Magazine (Tokyo)*, Vol. 105, pp. 443-452

Lin, S.J., Kato, M. & Iwatsuki, K. (1995). Electrophoretic variation of apogamous *Dryopteris varia* group (Drypteridaceae). *Journal of Plant Research*, Vol. 108, pp. 451-456

Liu, Y.C., Huang, Y.M. & Chiou, W.L. (2009). Validation of the name *Adiantum meishanianum* (Pteridaceae), a species endemic to Taiwan. *Novon: A Journal for Botanical Nomenclature*, Vol. 19, No. 1, pp. 59-61

Lu, S.Y., Peng, C.I., Cheng, Y.P., Hong, K.H. & Chiang, T.Y. (2001). Chloroplast DNA phylogeography of *Cunninghamia konishii* (Cupressaceae), an endemic conifer of Taiwan. *Genome*, Vol. 44, pp. 797–807.

Maddison, W.P. & Maddison, D.R. (2009) Mesquite: a modular system for evolutionary analysis. version 2.6. http://mesquiteproject.org

Manos, P.S., Cannon, C.H. & Oh, S. H. (2008). Phylogenetic relationships and taxonomic status of the paleoendemic Fagaceae of Western North America: recognition of a new genus, *Notholithocarpus. Madrono*, Vol.55, No.3, pp. 181-190

Manton, I. (1950). *Problems of Cytology and Evolution in the Pteridophyta*, p.316. Cambridge Univ. Press, New York

Masuyama, S., Yatabe, Y., Murakami, N. & Watano, Y. (2002). Cryptic species in the fern *Ceratopteris thalictroides* (L.) Brongn. (Parkeriaceae). I. Molecular analyses and crossing tests. *Journal of Plant Research*, Vol. 115, No. 2, pp. 87-97

Masuyama, S. (2008). Cryptic species in the fern *Ceratopteris thalictroides* (L.) (Parkeriaceae). III. Referential diagnostic characters of three cryptic species. *Journal of Plant Research*, Vol. 121, No. 3, pp. 279-286

Masuyama, S. & Watano, Y. (2010). Cryptic species in the fern *Ceratopteris thalictroides* (L.) (Parkeriaceae). IV. Taxonomic revision. *Acta Phytotaxonomica et Geobotanica,* Vol. 61, No. 2, pp. 75-86

Mehra, P.N. & Singh, G. (1957). Cytology of Hymenophyllaceae. *Journal of Genetics*, Vol. 55, pp. 379-393

Milne, R. I. (2004). Phylogeny and biogeography of *Rhododendron* subsection Pontica, a group with a tertiary relict distribution. *Molecular Phylogenetic and Evolution*, Vol. 33, pp. 389-401.

Moore, B.R. & Chan, K.M.A. (2005). SymmeTREE: whole-tree analysis of differential diversification rates. *Bioinformatics*, Vol.21, No.8, pp. 1709-1710

Nakatô, N. (1975). A cytological study on an intermediate form between *Pteris multifida* and *P. cretica*. *Journal of Japanese Botany*, Vol. 50, pp. 110-125

Nybom, H. (2004). Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* Vol. 13, pp.1143–1155.

Nybom, H. & Bartish, I.V. (2000). Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, Vol. 3, No. 2, pp. 93–114.

Naydenov, K., Senneville, S., Beaulieu, J., Tremblay, F. & Bousquet, J. (2007). Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biolohy*, Vol. 7:233.

Normark, B.B., Judson, O.P. & Moran, N.A. (2003). Genomic signatures of ancient asexual lineages. *Biological Journal of the Linnean Society*, Vol. 79, pp. 69-84

Pagel, M., Venditti, C. & Meade, A. (2010). Phylogenies reveal new interpretation of speciation and the Red Queen. *Nature*, Vol.463, No.7279, pp. 349-352

Pangtey, Y., Samant, S. & Verma, S. (1990). A new natural hybrid in the genus *Pteris* (Pteridaceae: Pteridophyta) from the Kumaun Himalaya. *Fern Gazette*, Vol. 13, pp. 357-360

Park, C. H. & Kato, M. (2003). Apomixis in the interspecific triploid hybrid fern *Cornopteris christenseniana* (Woodsiaceae). *Journal of Plant Research*, Vol. 116, pp. 93-103

Peakall, R. & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, Vol. 6, pp. 288-295.

Petit, R.J., Kremer, A. & Wagner, D.B. (1993). Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theoretical and Applied Genetics*, Vol.87, No.1-2, pp. 122-128

Petit, R.J., Csaikl, U.M., Bordacs, S., Burg, K., Coart, E., Cottrell, J., van Dam, B., Deans, J.D., Dumolin-Lapegue, S., Fineschi, S., Finkeldey, R., Gillies, A., Glaz, I., Goicoechea, P.G., Jensen, J.S., Konig, A.O., Lowe, A.J., Madsen, S.F., Matyas, G., Munro, R.C., Olalde, M., Pemonge, M.H., Popescu, F., Slade, D., Tabbener, H., Taurchini, D., de Vries, S.G.M., Ziegenhagen, B. & Kremer, A. (2002). Chloroplast DNA variation in European white oaks - Phylogeography and patterns of diversity based on data from over 2600 populations. *Forest Ecology and Management*, Vol.156, No.1-3, pp. 5-26

Rabosky, D.L. (2006). Likelihood methods for detecting temporal shifts in diversification rates. *Evolution*, Vol.60, No.6, pp. 1152-1164

Rambaut, A. & Drummond, A.J. (2003) Tracer ver. 1.4.1, available from http://evolve.zoo.ox.ac.uk/software/ (Institute of Evolutionary Biology, University of Edinburgh, Edinburgh)

Rambaut, A. (2008) FigTree ver. 1.3.1, available from http://tree.bio.ed.ac.uk/software/FigTree/ (Institute of Evolutionary Biology, University of Edinburgh, Edinburgh)

Rieseberg, L.H. & Wendel, J. (2004). Plant speciation - rise of the poor cousins. *New Phytologist*, Vol.161, No.1, pp. 3-8

Rieseberg, L.H. & Willis, J.H. (2007). Plant speciation. *Science*, Vol.317, No.5840, pp. 910-914

Ronquist, F. & Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, Vol.19, No.12, pp. 1572-1574

Rosenthal, D.M., Schwarzbach, A.E., Donovan, L.A., Raymond, O. & Rieseberg, L.H. (2002). Phenotypic differentiation between three ancient hybrid taxa and their parental species. *International Journal of Plant Science*, Vol. 163, pp. 387-398

Roy, R.P., Sinha, B.M.B. & Sakya, A.R. (1971). Cytology of some ferns of Kathmandu valley. *Fern Gazette*, Vol. 10, pp. 193-199

Schaal, B.A. & Olsen, K.M. (2000). Gene genealogies and population variation in plants. *Proceedings of National Academy of Sciences of the United States of America*, Vol. 97, pp. 7024–7029.

Schuettpelz, E., Schneider, H., Huiet, L., Windham, M.D. & Pryer, K.M. (2007). A molecular phylogeny of the fern family Pteridaceae: assessing overall relationships and the affinities of previously unsampled genera. *Molecular Phylogenetics and Evolution*, Vol. 44, pp. 1172–1185

Sibuet, J.C. & Hsu, S.K. (1997). Geodynamics of the Taiwan arc–arc collision. *Tectonophysics*, Vol. 274, pp. 221–251.

Sibuet, J.C. & Hsu, S.K. (2004). How was Taiwan created? *Tectonophysics* Vol. 279, pp. 159–181.

Sheldon, F.H., Lohman, D.J., Lim, H.C., Zou, F., Goodman, S.M., Prawiradilaga, D.M., Winker, K., Braile, T.M. & Moyle, R.G. (2009). Phylogeography of the magpie-robin species complex (Aves: Turdidae: *Copsychus*) reveals a Philippine species, an

interesting isolating barrier and unusual dispersal patterns in the Indian Ocean and Southeast Asia. *Journal of Biogeography*, Vol.36, No.6, pp. 1070-1083

Shepherd, L.D., Perriea, L.R. & Brownsey, P.J. (2008). Low-copy nuclear DNA sequences reveal a predominance of allopolyploids in a New Zealand *Asplenium* fern complex. *Molecular Biology and Evolution*, Vol. 49, pp. 240-248

Shieh, W.C. (1966). A synopsis of the fern genus *Pteris* in Japan, Ryukyu, and Taiwan. *Botanic Magazine (Tokyo)*, Vol. 79, pp. 283-292

Srivastava, J., Ranade, S.A. & Khare, P.B. (2007). Distribution and threat status of the cytotypes of *Pteris vittata* L. (Pteridaceae) species complex in India. *Current Science*, Vol. 93, pp. 81-85

Stebbins, G.L. (1950). *Variation and Evolution in Plants*. Columbia Univ. Press, New York

Suzuki, T. & Iwatsuki, K. (1990). Genetic variation in apogamous fern *Pteris cretica* L. in Japan. *Heredity*, Vol. 65, pp. 221-227

Swallow, J.G., Wallace, L.E., Christianson, S.J., Johns, P.M. & Wilkinson, G.S. (2005). Genetic divergence does not predict change in ornament expression among populations of stalk-eyed flies. *Molecular Ecology*, Vol.14, No.12, pp. 3787-3800

Takamiya, M., Ohta, N., Yatabe, Y. & Murakami, N. (2001). Cytological, morphological, genetic and molecular phylogenetic studies on intraspecific differentiations within *Diplazium doederleinii* (Woodsiaceae: Pteridophyta). *International Journal of Plant Sciences*, Vol. 162, pp. 625-636

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, Vol.28, No.10, pp. 2731-2739

Taylor, W.C., Lekschas, A.R., Wang, Q.F., Liu, X., Napier, N.S. & Hoot, S.B. (2004). Phylogenetic relationships of *Isoëtes* (Isoëtaceae) in China as revealed by nucleotide sequences of the nuclear ribosomal ITS region and the second intron of a *LEAFY* homolog. *American Fern Journal*, Vol. 94, pp. 196-205

Thomson, J.A. (2000). Morphological and genomic diversity in the genus *Pteridium* (Dennstaedtiaceae). *Annuals of Botany*, Vol. 85 (Supplement B), pp. 77-99

Trewick, S.A., Morgan-Richards, M., Russell, S.J., Henderson, S., Rumsey, F.J., Pintér, I., Barrett, J.A., Gibby, M. & Voge, J.C. (2002). Polyploidy, phylogeography and Pleistocene refugia of the rockfern *Asplenium ceterach*: evidence from chloroplast DNA. *Molecular Ecology*, Vol. 11, pp. 2003-2012

Turelli, M., Barton, N.H. & Coyne, J.A. (2001). Theory and speciation. *Trends in Ecology and Evolution*, Vol. 16, No. 7, pp. 330-343

Van den heede, C.J., Viane, R.L.L. & Chase, M.W. (2003). Phylogenetic analysis of *Asplenium* subgenus *Ceterach* (Pteridophyta: Aspleniaceae) based on plastid and nuclear ribosomal ITS DNA sequences. *American Journal of Botany*, Vol. 90, No. 3, pp. 481-495

Venditti, C., Meade, A. & Pagel, M. (2010). Phylogenies reveal new interpretation of speciation and the Red Queen. *Nature*, Vol.463, No.7279, pp. 349-352

Verma, S.C. & Khullar, S.P. (1965). Cytogenetics of the Western Himalayan *Pteris cretica* complex. *Annals of Botany*, Vol. 29, No. 4, pp. 673-681

Wagner, W.H.J. (1954). Reticulate evolution in the Appalachian Aspleniums. *Evolution*, Vol. 8, pp. 103-118

Wagner, W.H.J. & Nauman, C.E. (1982). *Pteris X delchampsii*, a spontaneous fern hybrid from southern Florida. *American Fern Journal*, Vol. 72, pp. 97-102

Walker, T.G. (1954). The *Pteris quadriaurita* complex in Ceylon. *Kew Bulletin*, Vol. 14, pp. 321-332

Walker, T.G. (1958). Hybridization in some species of *Pteris* L. *Evolution*, Vol. 12, pp. 82-92.

Walker, T.G. (1962). Cytology and evolution in the fern genus *Pteris* L. *Evolution*, Vol. 16, pp. 17-43.

Walker, T.G. (1984). Chromosomes and evolution in pteridophytes. In: *Chromosomes in Evolution of Eukaryotic Groups*, Vol. 2, A.K. Sharma & A. Sharma, pp. 101–141, CRC Press, Boca Raton, FL

Wang, Z.R. (1989). A preliminary study on cytology of Chinese *Pteris*. *Acta Phytotaxonomica Sinica*, Vol. 27, pp. 421-438

Wang, W. P., Hwang, C.Y., Lin, T. P. & Hwang, S. Y. (2003). Historical biogeography and phylogenetic relationships of the genus *Chamaecyparis* (Cupressaceae) inferred from chloroplast DNA polymorphism. *Plant Systematics and Evolution* Vol. 241, pp. 13–28.

Wen, J. & Zimmer, E. A. (1996). Phylogeny abd biogeography of *Panax* L. (the Ginseng Genus, Araliaceae): Inferences from ITS sequences of nuclear ribosomal DNA. *Molecular Phylogenetic and Evolution*, Vol. 6, pp. 167-177.

Whittemore, A.T. & Schaal, B.A. (1991). Interspecific Gene Flow in Sympatric Oaks. *Proceedings of the National Academy of Sciences of the United States of America*, Vol.88, No.6, pp. 2540-2544

Willis, K.J. & Niklas, K.J. (2004). The role of Quaternary environmental change in plant macroevolution: the exception or the rule? *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol.359, No.1442, pp. 159-172; discussion 172

Wu, H.Y., Chu, J.H. & Lin, Y.S. (2007). Evolution and dispersal of three closely related macaque species, *Macaca mulatta*, *M. cyclopis*, and *M. fuscata*, in the eastern Asia. *Molecular Phylogenetics and Evolution*, Vol.43, No.2, pp. 418-429

Wu, S.H., Hwang, C.Y., Lin, T.P., Chung, J.D., Cheng, Y.P. & Hwang, S.Y. (2006). Contrasting phylogeographical patterns of two closely related species, *Machilus thunbergii* and *Machilus kusanoi* (Lauraceae), in Taiwan. *Journal of Biogeography*, Vol.33, No.5, pp. 936-947

Xiang, Q. Y., Manchester, S. R., Thomas, D. T., Zhang, W. & Fan, C. (2005). Phylogeny, biogeography, and molecular dating of cornelian cherries (*Cornus*, Cornaceae): tracking Tertiary plant migration. *Evolution*, Vol. 59, pp. 1685-1700.

Yi, T., Miller, A. J. & Wen, J. (2004). Phylogenetic and biogeographic diversification of *Rhus* (Anacardiaceae) in the Northern Hemisphere. *Molecular Phylogenetic and Evolution*, Vol. 33, pp. 861-879.

Zhuang, Y. & Adams, K.L. (2007). Extensive allelic variation in gene expression in *Populus* F1 hybrids. *Genetics*, Vol. 177, No. 4, pp. 1987-1996

Zwickl, D. J., (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin, USA.

Genetic Diversity, Habitat Fragmentation and Epigenetic Variations

Josphert N. Kimatu[1], Geoffrey M. Muluvi[1] and Bao Liu[2]
[1]Department of Biological Sciences, South Eastern University College
(A Constituent College of Nairobi University), Kitui
[2]Plant Molecular Epigenetics Laboratory Northeast Normal University, Changchun
[1]Kenya
[2]China

1. Introduction

There is an increase in the levels in which tropical rainforest and temperate grasslands have become anthropogenically fragmented in recent years. However, an understanding of the genetic and epigenetic mechanisms coupled with conservation biology and biodiversity studies is crucial in explaining habitat fragmentation implications in stable ecosystems (Fahrig, 2003; Henle et al., 2004). The utilization of genetic variability and adaptability can be used as a stepping stone in identifying germplasms that can be used in the search of new ecotypes with novel genes which can be incorporated in crop, fodder and cover tree improvement programs in other similar reclaimable ecological zones. Studies have shown that there is a considerable genetic variability within or between natural populations. This variability provides a genomic flexibility that can be used as a raw material for plant adaptation as pioneer species. There is need to seek for higher genetic variability so as to increase the capacity of an organism to adapt to the ever changing environmental conditions (Ellstrand and Elam, 1993). Low genetic variability has been associated with the inability to cope with abiotic and biotic stresses (Valen, 1965). It is now possible to compare and analyze the pattern and spectrum of genetic variations within or between using molecular genetic tools (Zhang et al., 2009, Madan et al., 2002; Reisch et al., 2005). The amplified fragment length polymorphism (AFLP) is one of the most precise, cost effective polymerase chain reaction (PCR) based ecological tool that has been used for molecular analysis of such population genetic diversity analysis (Vos et al., 1995; Bensch, and Akesson, 2005). It has been shown to be comparatively highly informative, reliable and efficient tool in other grass investigations (Fjellheim and Rognli, 2005). Such tools can be used to determine the correlation between the genetic differentiation and geographic distance among different populations. However, the level of genetic differentiation can be affected by more than one ecological factor in a geographical region, especially environmental heterogeneity (Liu et al., 2004). Other factors could be natural mutation, artificial selection, and combined ecological factors which are instrumental in the differentiation into ecotypes (Liu et al., 2002). Therefore, variations in gene expressions which are caused by abiotic and biotic stresses in the environment trigger epigenetic mechanisms like cytosine DNA methylation which have

recently been used to measure the potential of genome flexibility and intrinsic adaptable evolution mechanisms in dynamic environments. The methylation sensitive amplified polymorphism (MSAP) protocol has been used to combine the genetic diversity studies with epigenetic variation in evaluating the effects of habitat fragmentation especially in plants (Zhang et al., 2009).

2. Genetic differentiation and gene flow

Recent studies have also pinpointed that self-incompatibility can contributes to genetic differentiation especially when coupled with environment habitat selection pressure and low levels of gene flow (Gong et al., 2007). Furthermore, other studies by Liang et al., 2007, showed that wind-mediated dispersal of pollen and seed transport via migrating birds influences gene flow systems and habitat revegetation and could be correlated to artificial selection, habitat selection pressure, environmental heterogeneity, genetic diversity and the eventual differentiation. The occurrence and spectra of gene flow through either seed and/or pollen dispersal was found to possibly be imposing a significant impact on intra- and inter-population genetic variations. The gene flow frequencies between the population pairs were later correlated with the geographical distances (Zhang et al., 2009).

Cluster analysis of the *L. chinensis* also showed that physical barriers like building, mountains and rivers could be instrumental in habitat fragmentation which led to the differentiation and formation of ecotypes due to limitation of gene flow and probable inbreeding alongside an increase in epigenetic variations. Investigations using AFLP in some natural populations of a grass species (*L. chinensis)* identified two easily recognizable phenotypes which had adapted to contrasting ecological habitats due to habitat fragmentation. Some of the phenotypes were in eroded and highly stressful alkaline and saline meadows of pH 8.5 to 9.5 while another was in normal edaphic conditions of pH of less than 8.0 (Zhang et al., 2009, Zhou and Yang, 2003). These genetic and epigenetic variations could have been triggered by the abiotic or biotic stresses. Habitat fragmentation can lead to genetic homogenization among populations; this is because various physical and biological distances usually play a role in determining the level of gene flow and diversity between ecotypes.

3. Inter-clonal genetic diversity

Studies have shown that when levels of gene flow in a population are low, then genetic drift could become the predominant factor in shaping the population genetic structure. This usually results in a within-population homogeneity and causes at the same time an among population differentiation (Reisch et al., 2005). However, with regard to gene flow, for a given plant species, it is most conceivable that causal links with geographic distances, but in some cases it has been shown that there could be little relatedness to geographical distances. This scenario could be envisioned to be caused by multiple other factors which might have imposed their influences on the efficiency or effectiveness of gene flow among populations. Such habitat fragmentation and other physical barriers like mountains, human activities, shielding vegetations and accumulation of valleys could influence the flow of wind and movement of dispersal agents. Further more, recent studies by Li et al., 2009, showed a remarkable observation in their cluster analysis indicating that the soil condition and the proximal distance between the populations have a significant influence in the genetic

diversity analysis. Hence, it is now conceivable that the occurrence and spectra of genetic variations both within a plant population and among populations can be caused by three distinct as well as interlaced factors. These are gene flow through seed dispersal and/or pollen dissemination, genetic drift and habitat induction leading to inter-clonal diversity.

4. Habitat inducement and epigenetic influences in genetic diversity

There is always a possibility of other biologically influenced dispersal factors to influence the enrichment of habitat fragmented clones. These could interfere with wind and pollinator movements (Li et al., 2009) causing homogenization of clones. This possibility was evidenced by an exceptionally low level of gene flow between two populations of *L. chinensis* (Zhang et al., 2009). However, it was clear that physical and/or biological isolation were identified as the most significant factors in causing differentiation of these populations. It was interesting that although habitat heterogeneity like soil pH might not be correlated with genetic variation intrapopulation of the same ecotype, a dendrogram of between these populations constructed by the UPGMA method (Sneath and Sokal, 1973), based on a similarity matrix of AFLP data and calculated according to the Jaccard index showed a close link between the clones grown in the same abiotic conditions (Zhang et al., 2009). Therefore, the possibility of habitat inducement can not be ruled out in causing heterogeneity in stressful abiotic conditions of varied salinity/alkaline which have been known to cause heritable epigenetic variations in plants (Zhang et al., 2010, Kimatu, 2010). Recent studies by Kimatu, 2011, suggested that narrowed genetic base due to inbreeding could contribute to morphological variations and eventually lead to pathogenic attacks to plants. Such eventuality can more be amplified by habitat fragmentation and possible epigenetic variations leading to greater genetic diversity.

5. The impact of anthropogenic destruction of natural habitats

Large continuous natural landscapes are now anthropogenically being destroyed alongside a rapid habitat fragmentation which is a threat to the available biodiversity especially in the tropical forests and temperate grasslands (Young & Clarke 2000, Soulé and Orians 2001). Hence, there has been an increased interest in the assessment of the genetic consequences of habitat fragmentation (Young et al., 1996; Lowe et al., 2005; Ouborg et al., 2006; Honnay & Jacquemyn 2007). Recent studies have established that in contrast to naturally fragmented habitats populations, human impacted fragmented habitats, may suffer from genetic variability erosion as a consequence of isolation and reduced gene flow alongside an increased interpopulation divergence (Zhang et al., 2009, Young, Boyle & Brown 1996). However, many studies have not been done to overrule absolutely that in naturally fragmented landscapes the same phenomenon may not necessarily be occurring. However, some studies have recently shown that not all fragmentation episodes necessarily result in genetic erosion of plant populations (Collevatti et al., 2001; Lowe et al., 2005; Kettle et al., 2007). This isolation can be due to the partitioning of genetic variation among populations and most likely is influenced by the breeding systems and on the reproductive mode of the populations. An increase in clonal propagation may decrease genetic variation within populations although among populations, it might increase genetic differentiation. Adaptations which are triggered by abiotic and biotic stresses in a habitat may increase genetic separation involving epigenetic variations which aid in the formation of epialleles (Liu, 2010).

5.1 Plants in fragmented habitats

Anthropogenic destruction and fragmentation of the formerly continuous natural landscape is increasing and has become now a threat to biodiversity maintenance in ecosystems (Young, Boyle & Brown 1996). The consequences can be the extinction of species and the loss of genetic diversity and the loss of genetic diversity. What can be the causes of these fragmentations? Some could be caused by abiotic causes like nutrient-poor soils, pH differences, toxicity of some minerals etc. These factors are also known to trigger changes in gene expressions causing epigenetic variations. The smaller the population the more likely it is vulnerable to demographic and environmental stochasticity and genetic erosion (Lande 1988; Menges 2000; Schemske et al., 1994; Young et al., 1996). There has been the phenomenon of formation of 'habitat island' which can be found in natural habitats when they become isolated due to abiotic or biotic fragmentation factors. This is most likely when they are surrounded by non-suitable habitats which could hinder gene flow and genetic variation enrichment (Bossuyt, Honnay & Hermy 2003; Pither, Shore & Kellman 2003; Vucetich et al., 2001). Examples of such habitats can be harsh conditions like infertile soils, occasional high solar radiations, seasonal desiccating winds, steep environments or other biotic impacts like competition from invasive or predatory species. This can decrease population genetic diversity and decrease differentiation among populations. This consequently may increase inbreeding, limited genetic flow due to reduced mate availability (Young et al., 1996). The level of heterozygosity may decrease while is possible that in the midst of such genetic drift some advantageous genes may be lost, especially if the drift is selective (Lacy 1987). Studies by Barrett & Kohn 1991 and Ellstrand & Elam 1993 showed that in the short term the genetic diversity may reduce plant performance and lower the viability of the seeds from the population. While in the longer term such drifting may result in a limitation for further adaptive evolution mechanism capabilities (Falconer & Mackay 1996).

The differential expression of genes in plants which is regulated by epigenetic mechanisms like cytosine DNA methylation is crucial for phenotypic plasticity in the matching of certain genotypes to variable environments. Studies have shown that although plant growth and reproduction are determined by the genotype and interactions with the abiotic and biotic environment, an increase in environmental variability favours greater levels of plasticity (Schlichting 1986; Sultan 1987). Recent studies by Sumner et al., 2004, who genotyped 411 individuals at nine microsatellite DNA loci found fewer alleles per locus in prickly forest skinks from small rain forest fragments and a lower ratio of allele number to allele size range in forest fragments than in continuous forest. They also found that heterozygosity and variance in allele size did not differ in continuous forests as expected. However, the relationship between genetic (F_{ST}) and geographical distance was significantly stronger for continuous forest sites than for fragments. This is consistent with the disruption of gene flow due to habitat fragmentation. They noted that it is not easy to detect the genetic consequences of recent habitat fragmentation especially if the population size is small and the dispersal rate is not high (Figure 1). This may be influenced by the species dispersal potential, behavioural response and epigenetic flexibility coupled with the ability to adapt to habitat fragmentation limitations (Weins 1997). Therefore, the loss of genetic diversity reduces a population's ability to respond to biotic and abiotic environmental changes hence greatly reducing the chances of evolution and increasing the extinction probabilities (Nason et al. 1997; Booy et al. 2000).
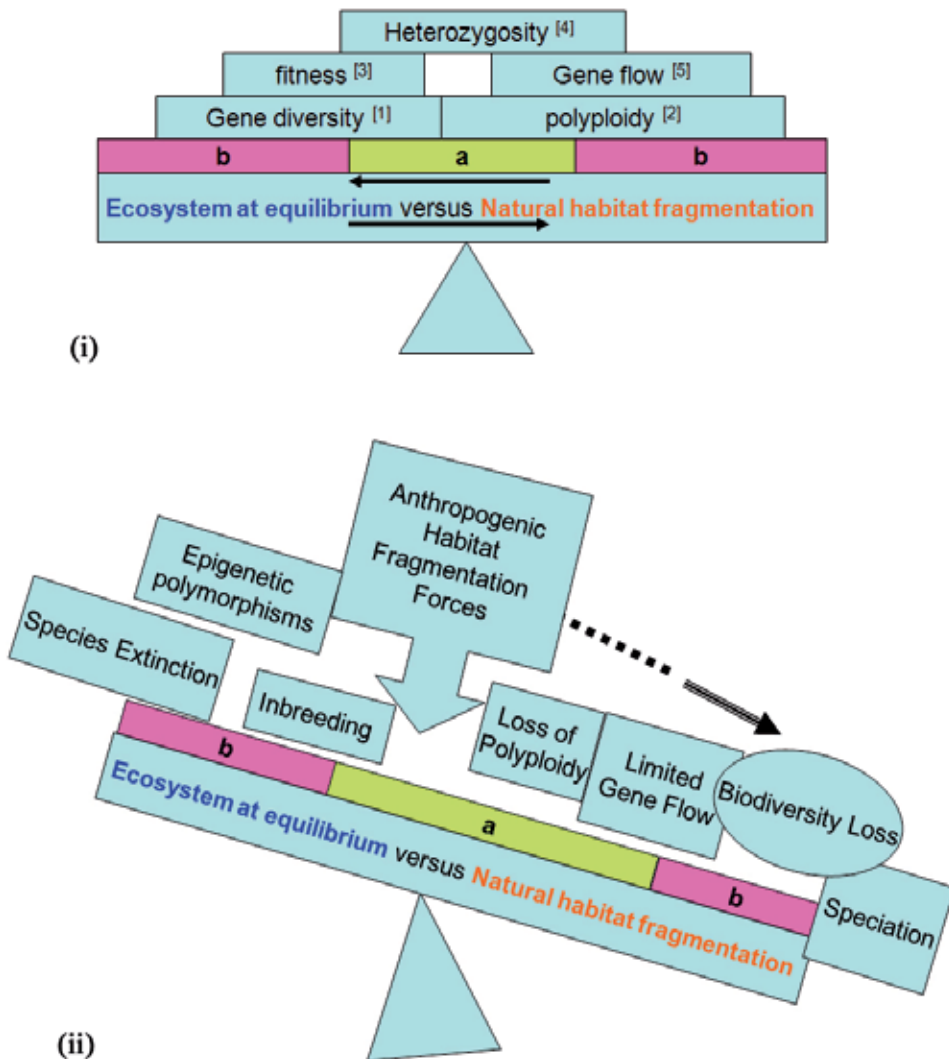
Fig. 1. Natural ecosystems are usually in equilibrium although natural habitat fragmentations are still occurring. This dynamism maintains high biodiversity, gene flow, frequent polyploidy and hybridization events, high fitness etc in the populations (i). Species without narrow geographic distributions (a) are more stable that those with narrow geographical distributions (b). The impact of human activities in natural ecosystems causes habitat fragmentations which cause instability and introduces inbreeding, species extinction and biotic and abiotic triggered epigenetic variations and ecotypes (ii). The references are: [1] and [4]; Lande, 1999; [2], Zhang et al., 2009; [3], Allphin et al., 1998 and [5], Templeton et al., 1990.

5.2 Does habitat fragmentation necessarily affect genetic diversity?

Habitat fragmentation have been found to alters the structure, distribution, and functioning of natural ecosystems (Saunders et al., 1991). This has made many researchers to make the general conclusion that habitat fragmentation definitely affects the genetic diversity. Hence,

many habitat fragmentation studies have been working with the hypothesis that it causes erosion of genetic variability and an increase and interpopulation genetic divergence in plant populations. This could be due to the increase in random genetic drift, inbreeding, and reductions in gene flow (e.g. Young et al., 1996; Sork et al., 1999; Lowe et al., 2005). Furthermore, the genetic diversity and variation of populations is expected to be reduced due to genetic bottlenecks which tend to lower the proportion of polymorphic loci and a reduction in the number of alleles per locus within habitat fragments (Nei et al., 1975; Ellstrand & Elam 1993; Young et al., 1996).

Recent studies by Aguilar et al., 2008, confirmed that indeed there are overall large and negative effects of habitat fragmentation on genetic diversity and outcrossing rates. However, they found that there were no effects on the inbreeding coefficients (F_{IS}). The highest proportion of variation could be explained by the type of mating system, age of the fragment and the rarity status of plant species. They suggested that habitat fragmentation tends to shift mating patterns towards an increased rate of selfing. This can cause an increase in inbreeding and consequently an accumulation of deleterious recessive alleles. This lowers the fecundity of individuals and may in plants, increase seed/seedling mortality, reducing the growth rate of individuals and the eventual diminishing of species or even species extinction (Young et al., 1996). Moreover, the size of the effective population within a fragment and the pattern of genetic variability of the original populations previous to the fragmentation have been shown to be instrumental in determining the immediate effects of habitat fragmentation and the genetic diversity of the organisms in question (Nason et al., 1997; Hamrick, 2004).

Other studied factors which determine the effects of habitat fragmentation are: some particular life-history traits of plants, as is expected to find stronger negative effects on the adult generation of short-lived species compared to long-lived species (Young et al., 1996); vegetative reproduction may also buffer the genetic effects of habitat fragmentation (Honnay & Bossuyt 2005); the mating system has also been implicated in how the genetic diversity can be influenced as cross pollinated generally show a higher genetic variation compared to self pollinating plants (Loveless & Hamrick 1984; Hamrick & Godt 1989); Finally, polyploids are more resistant to the effects on genetic diversity due to fragmentation; (Bever & Felber 1992; Moody et al., 1993). Interestingly, there are now reports which implicate epigenetic variations and there involvement in the adaptation mechanisms of polyploids. For example, it was found that high incidence of alteration in cytosine DNA methylation occurred in allopolyploid plants even in the absence of other structural genomic changes. This suggested that there is an existence of "pure" epigenetic variation which is associated with hybridization and allopolyploidy (Liu et al., 2010). Moreover, it is demonstrated that a recently (less than 150 years) formed natural allopolyploid species, *Spartina anglica*, manifested genome-wide alterations in cytosine DNA methylation which reached a frequency of 35%, and was accompanied by little genetic changes, and which was strikingly hypothesized to be associated with their extraordinary invasiveness in new habitats as a weedy plant (Salmon et al., 2005). Hence, fine-tuned epigenetic states in DNA methylation and histone modifications play critical roles in orchestrating gene expressional trajectories across development and in maintaining genome stability (Liu et al., 2010). Therefore, it is not difficult to imagine that the polyploidy-induced epigenetic alterations will have significant impacts in gene expression, genome integrity and phenotypes even in new abiotic and biotic stresses especially due to habitat fragmentation.

6. Habitat fragmentation and ecology

Because, out crossing plants will be more affected by habitat fragmentation, it is conceivable that, ecological processes like pollination and seed dispersal mechanisms of plants, can have a high impact on genetic connectivity in newly fragmented habitats (Nathan & Muller-Landau 2000; Tewksbury et al., 2002; Hamrick 2004). Hence, the distribution, abundance, composition, and behaviour of pollinator and dispersal organisms will affect the potential of the genetic diversity of the plants in the new habitats (Nason et al., 1997). This is because the attributes of pollinators and seed dispersers are also affected by habitat fragmentation (Aizen & Feinsinger 2003; Griscom et al., 2007). Studies have shown that animal-pollinated and animal seed-dispersed plant species can have decreased genetic connectivity due to habitat fragmentation compared to plants which do not require animals but rely on abiotic factors (Nathan & Muller-Landau 2000 and García et al., 2007).

Studies on tree species have shown that there are their high within-population genetic diversity (0.177) and low genetic diversity among-population differentiation (0.148) for nuclear genes. But, in contrast, annual plants show the opposite (i.e. 0.154 and 0.101 respectively). This suggests a more differentiation for nuclear genes but much less diversity than in trees (Hamrick et al., 1992 and Hamrick and Godt, 1996). Most researchers explain this using the pollen flow hence the gene flow is much higher in trees than in annual plants. However, trees which have been introduced to recolonize temperate and tropical areas have experienced many foundational events, which have led to reduced within-population diversity and increase differentiation due to the narrowing of the genetic base (Kimatu, 2011). There is still need to research on how the pollen flow influences genetic diversity, especially epigenetically. This has its clue on the fact that one key indicator of the pollen flow showed epigenetic gene imprinting in that the paternally inherited chloroplasts showed much less differentiation than maternally inherited mitochondria in pines (Latta and Mitton 1997). Studies have shown that active demethylation, which might be probably acting synergistically with passive loss of methylation, of promoters by the 5-methyl cytosine DNA glycosylase or DEMETER (DME) is required for the uni-parental expression of imprinting genes in endosperm, which is essential for seed viability (Reviewed by Zhang et al., 2010) and most likely lead to habitat influenced genome fragmentation.

7. Molecular diversity and habitat fragmentation

Most studies will easily hypothesize that plant populations in isolated habitats are to a large extend differentiated. However, does population viability in a habitat decrease with a decrease in molecular diversity or is molecular diversity of a population significantly different between contrasting habitats? It could readily be suggested that gene flow among populations is highly dependent on the distance. One general conclusion which is likely to be drawn is that plants from naturally fragmented habitats are more likely to be affected in their genetic diversity level compared to the genetic isolation which is caused by more recent and human induced habitat fragmentation.

How then can we reliably detect the genetic effects of habitat fragmentation which can have genetic and ecologically induced genomic changes? Recent studies using the PCR based amplified fragment length polymorphism (AFLP) showed a higher ability to find habitat fragmentation effects on genetic parameters compared to allozymes. This is because DNA-based genetic markers such as microsatellites have a higher mutation rate, which consequently

causes higher levels of variation and may have higher resolution to detect changes in inter- and intrapopulation genetic variation compared to allozymes. The AFLP (Vos et al., 1995) is a DNA fingerprinting technique that is ideal in resolving genetic diversity among individuals, populations and species (Muller and Wolfenbarger, 1999). It is highly reproducible and polymorphic and can be used to survey overall genetic differences in the nuclear genome in a single assay without any prior sequence knowledge (Vos et al., 1995; Jones et al., 1997). Recently, it has been applied to investigate genetic relationship among species, closely related cultivars and even clones of plants (Mizianty et al., 2006; Yoon et al., 2007; Li et al., 2008; Karimi et al., 2009). Hence, Methylation Sensitive Amplified Polymorphism (MASP) which is a modified version of AFLP may be most suitable in detecting those affected as it combines the genetic diversity evaluation with the epigenetic variation (Zhang et al., 2010).

8. Conclusion

There is an increase in biodiversity study in the last half a decade, due to the opportunities which it offers research scientist and the portfolio of varying characters to the crop and animal breeders. However, the correspondences between genes, alleles and environments on the one hand, and phenotypes on the other have been found to be increasingly complex (Rockman and Kruglyak, 2006). The recent global increases in habitat fragmentations due to natural or human factors have created smaller populations, hereby referred to as species 'islands' from the same population. This process of habitat fragmentation can gradually transform a large population into several smaller islands causing genetic bottlenecks. This transformation can have serious losses in genetic diversity within a population and can result in species extinction. This recent human induced habitat fragmentation may be too recent to be detected genetically. However, variations in gene expressions which can epigenetically be determined can provide important tool in this evaluation. Epigenetic studies have shown that there are variations among individuals in the degree of cytosine DNA methylations of a gene, termed as epialleles. These produce novel phenotypes that are heritable across generations. These epigenetic alleles or epialleles differ at a particular gene sequences. Plants with different epialleles can exhibit distinct phenotypes which are usually mitotically and also sometimes meiotically stable (Kakutani, 2002, Kakutani, et al., 1999). Furthermore, polyploidization has been identified a major evolutionary process where hybridization and chromosome doubling induce enormous genetic and epigenetic variations. Studies of how the specific sequences are restructured are still largely unknown (Liu, 2010). Hence, studies generally seem to point that the number of genes in an organism may not necessarily be used to show the functional complexity of a genome because other factors like abiotic, biotic, genome proteins and epigenetic mechanism are usually acting in harmony to create a unique genetic diversity which can assist in functioning, adaptation and formation of ecotypes.

9. Acknowledgment

10. References

Aguilar R.; Quesada M.; Ashworth L.; Herreriasdiego Y. & Lobo J. (2008). Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology,* 17, 5177–5188.

Allphin, L., Windham, MD. and Harper, KT. (1998). Genetic diversity and gene flow in the endangered dwarf poppy, *Arctomecon humilis* (Papaveraceae). *Am. J. Bot.*, 85: 1251–1261.

Aizen, MA. & Feinsinger P. (2003). Bees not to be? Responses of insect pollinator faunas and flower pollination to habitat fragmentation. In: *Disruptions and Variability: The Dynamics of Climate, Human Disturbance and Ecosystems in the Americas* (eds Bradshaw GA, Marquet PA, Mooney HA), pp. 111–129. Springer-Verlag, Berlin,Germany.

Barrett, SCH & Kohn, JR. (1991). Genetic and evolutionary consequences of small population size in plants: Implication for conservation. *Genetics and conservation of rare plants,* (eds D.A. Falk & K.E. Holsinger), pp. 3-30. Oxford University Press, New York, Oxford.

Bever, JD. & Felber F. (1992). The theoretical population genetics of autopolyploidy. In: *Oxford Surveys in Evolutionary Biology,* (eds Antonovics J, Futuyma D), pp. 185–217. Oxford University Press, Oxford, UK.

Booy, G.; Hendriks RJJ.; Smulders MJM.; Van Groenendael JM. & Vosman B. (2000). Genetic diversity and the survival of populations. *Plant Biology*, 2, 379–395.

Bossuyt, B.; Honnay, O. & Hermy, M. (2003). An island biogeographical view of the successional pathway in wet dune slacks. *Journal of Vegetation Science,* 14, 781-788.

Ellstrand, N. & Elam, D. (1993). Population genetic consequences of small population size: Implications for plant conservation. *Annual Review of Ecology and Systematics,* 24, 217-242.

Fahrig, L., (2003). Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, *Evolution and Systematics,* 34, 487–515.

Falconer, DS. & Mackay, TFC. (1996) Introduction to quantitative genetics. Longman, Harlow, Essex. Geber, M.A. (1990) The cost of meristem limitation in *Polygonum arenastrum* - Negative genetic correlations between fecundity and growth. *Evolution,* 44, 799-819.

Fjellheim, S. & Rognli, OA. (2005). Genetic diversity within and among Nordic meadow fescue (*Festuca pratensis* Huds.) cultivars determined on the basis of AFLP markers. *Crop Sci.,* 45: 2081-2086.

García C.; Jordano P. & Godoy JA (2007). Contemporary pollen and seed dispersal in a *Prunus mahaleb* population: patterns in distance and direction. *Molecular Ecology*, 16, 1947–1955.

Gong, L.; Song, XX; Li, M.; Guo, WL;, Hu, LJ.; Tian, Q.; Yang, YF.; Zhang, YF.; Zhong, XF.; Wang, DL.& Liu, B. (2007). Extent and pattern of genetic differentiation within and between phenotypic populations of *Leymus chinensis* (Poaceae) revealed by AFLP analysis. *Can. J. Bot.*, 85: 813-821.

Griscom PH.; Kalko EKV. & Ashton MS (2007). Frugivory by small vertebrates within a deforested, dry tropical region of Central America. *Biotropica*, 39, 278–282.

Hamrick JL. (2004). Response of forest trees to global environmental changes. *Forest Ecology and Management*, 197, 323–335.

Hamrick JL. & Godt MJW. (1989). Allozyme diversity in plants. In: *Population Genetics, Breeding and Genetic Resources* (eds Brown ADH, Clegg MT, Kahler AL, Weir BS), pp. 43–63. Sinauer & Associates, Sunderland, Massachusetts.

Hamrick, J L.; and Godt MJ W. (1996). Effects of life history traits on genetic diversity in plants. Philos. Trans. R. Soc. Lond. Ser. *B Biol. Sci.,* 351: 1291–1298.

Hamrick, J L, Godt M J W & Sherman-Broyles S L. (1992). Factors influencing levels of genetic diversity in woody plant species. *New For.*, 6: 95–124.

Henle, K.; Lindenmayer, DB.; Margules, CR.; Saunders, DA. & Wissel, C. (2004). Species survival in fragmented landscapes: where are we now? *Biodiversity and Conservation.*, 13, 1–8.

Honnay O.& Bossuyt B. (2005). Prolonged clonal growth: escape route or route to extinction? *Oikos*, 108, 427–432.

Honnay O. & Jacquemyn H. (2007). Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology*, 21, 823–831.

Jones CJ.; Edwards KJ.; Castagllone S.; Winfleld MO.; Sala F.; van de Wiel C.; Bredemeljer G.; Vosman B.; Matthes M.; Daly A.; Brettschneider R.; Bettini P.; Buiatti M.; Maestri E.; Malcevschi A.; Marmiroli N.; Aert R.; Volckaert G.; Rueda J.; Linacero R.; Vazquez A. & Karp A (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Mol. Breed,*. 3: 381-390.

Kakutani, T., (2002). Epi-alleles in plants: inheritance of epigenetic information over generations. *Plant Cell Physiol,*. 43, 1106–1111.

Kakutani, T.; Munakata, K.; Richards, E.J.; & Hirochika, H. (1999). Meiotically and mitotically stable inheritance of DNA hypomethylation induced by ddm1 mutation of Arabidopsis thaliana. Genetics 151, 831–838.

Kalisz, S. & Purugganan MD. (2004). Epialleles via DNA methylation: consequences for plant evolution. *Trends in Ecology & Evolution*, 19, 309–314.

Karimi HR.; Kafkas S.; Zamani Z.; Ebadi A. & Fatahi Moghadam MR (2009). Genetic relationships among Pistacia species using AFLP markers. *Pl. Syst. Evol.,* 279: 21-28.

Kimatu, JN. & Liu, B. (2010). Epigenetic polymorphisms could contribute to the genomic conflicts and gene flow barriers resulting to plant hybrid necrosis. *African Journal of Biotechnology,* Vol. 9(48), pp. 8125-8133, 29, (November, 2010).

Kimatu, JN. (2011). Effects of narrowing genetic base and abiotic stress on leaf-spotting in *Grevillea robusta. Journal of Tropical Forest Science*, 23(2):117-124.

Lacy, RC. (1987). Loss of genetic diversity from managed populations: Interacting effects of drift, mutation, immigration, selection and population subdivision. *Conservation Biology,* 1, 143-158.

Lande, R. (1999). Extinction risks from anthropogenic, ecological, and genetic factors. In *Genetics and Extinction of Species* (L.A. Landweber and A.P. Dobson, eds), pp. 1–22. Princeton, NJ: Princeton University Press.

Lande, R. (1988). Genetics and demography in biological conservation. *Science,* 241, 215-244.

Landolt, E. (1967). Gebirgs- und Tieflandsippen von Blütenpflanzen im Bereich der Schweizer Alpen. *Botanische Jahrbücher,* 86, 463-480.

Latta, RG. & Mitton JB. (1997). A comparison of population differentiation across four classes of gene markers in limber pine (*Pinus flexiblis* James). *Genetics*, 146: 1153–1163.

Li, XX.; Ding XY.; Chu BH.; Zhou Q.; Ge D. & Gu S (2008). Genetic diversity analysis and conservation of the endangered Chinese endemic herb *Dendrobium officina*le Kimura *et* Migo (Orchidaceae) based on AFLP. *Genetica*, 133: 159-166.

Li, M.; Gong, L.; Tian, Q.; Hu, L.; Guo, W.; Kimatu, J.N.; Wang, D. & Liu, B. (2009). Clonal genetic diversity and populational genetic differentiation in *Phragmites australis* distributed in the Songnen Prairie in northeast China as revealed by amplified fragment length polymorphism and sequence-specific amplification polymorphism molecular markers. *Annals of Applied Biology,* 154:1, 43-55

Liang, Y.; Diao, Y.; Liu, G. & Liu, J. (2007). AFLP variations within and among natural populations of *Leymus chinensis* in the northeast of China. *Acta Pratacul Turae Sinica*, 16: 124-134.

Liu, B.; Xu, C.; Zhao, N.; Qi, B.; Kimatu, JN.; Pang, J. & Han, F. (2010). Rapid genomic changes in Polyploid wheat and related species: implications for genome evolution and genetic improvement. Review. *J. Genet. Genomics* , 36 (2009) 519−528.

Liu, HF.; Gao, YB.; Ruan, WB.; Chen, L.; Li, CL.; Zhao, NX. & Wang, D. (2004). Genetic differentiation within and between *Leymus chinens* is population s from different zones of Mid-Eastern Inner Mongolia steppe. *Acta Ecologica Sinica.*, 24: 2157-2164. [in Chinese with English abstract].

Liu, J.; Zhu, ZQ.; Liu, GS.; Qi, DM.& Li, FF. (2002). AFLP variation analysis on the germplasm resources of *Leymus chinensis*. *Acta Bot. Sin,.* 44: 845-851.

Loveless, MD. & Hamrick, JL (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, 15, 60–90.

Lowe, AJ.; Boshier D, Ward M, Bacles CFE, Navarro C (2005). Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity*, 95, 255–273.

Madan, S.; Avi, GG.& David, W. (2002). Population genetic structure and the conservation of isolated populations of *Acacia raddiana* in the Negev Desert. *Biological Conservation*, 108: 119-127.

Menges, ES. (2000). Population viability analyses in plants: challenges and opportunities. *Trends in Ecology & Evolution,* 15, 51-56.

Mizianty, M.; Bieniek, W.; Czech, A,; Strzałka, W. & Szklarczyk, M. (2006). Variability and structure of natural populations of *Elymus caninus* (L.) L. and their possible relationship with *Hordelymus europaeus* (L.) Jess. ex Harz as revealed by AFLP analysis. *Pl. Syst. Evol.* 256: 193-200.

Moody, M.E.; Muellert, LD. & Soltis DE. (1993). Genetic variation and random drift in autotetraploid populations. *Genetics*, 154, 649– 657.

Nason, JD.; Aldrich, PR.; Hamrick, JL. (1997). Dispersal and the dynamics of genetic structure in fragmented tropical tree populations. In: *Tropical Forest Remnants: Ecology, Management and Conservation of Fragmented Communities* (eds Laurance WF, Bierregaard RO), pp. 304–320. University of Chicago Press,Chicago, Illinois.

Nathan, R. & Muller-Landau, HC. (2000). Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology & Evolution*, 15, 278–285.

Nei, M.; Maruyama, T. & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, 29, 1–10.

Pither, R., Shore, JS. & Kellman, M. (2003). Genetic diversity of the tropical tree *Terminalia amazonia* (Combretaceae) in naturally fragmented populations. *Heredity,* 91, 307-313.

Reisch, C., Anke, A., Rohl, M. (2005). Molecular variation within and between ten population of *Primula farinosa* (Primulaceae) along an altitudinal gradient in the northern Alps. Basic Appl. Ecol. 6: 35-45.

Rockman, MV., & Kruglyak, L. (2006). Genetics of global gene expression. *Nature Reviews Genetics,* 7 (11):862-72.

Salmon, A., Ainouche, ML., & Wendel, JF. (2005). Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). Mol. Ecol. 14: 1163−1175.

Saunders DA, Hobbs RJ, Margules CR (1991). Biological consequences of ecosystem fragmentation: a review. *Conservation Biology*, 5, 18–32.

Schemske, DW.; Husband, BC.; Ruckelshaus, MH.; Goodwillie, C.; Parker, IM & Bishop, J.G. (1994). Evaluating approaches to the conservation of rare and endangered plants. *Ecology* 75, 584-606.

Schlichting, C. (1986). The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics,* 17, 667-693.

Sneath, PHA. & Sokal, RR. (1973). Numerical taxonomy. W.H. Freeman, San Francisco, Calif., USA.

Sork, VL.; Nason, J.; Campbell, DR.; Fernandez, JF. (1999). Landscape approaches to historical and contemporary gene flow in plants.*Trends in Ecology & Evolution*, 14, 219–224.

Soulé, ME.; & Orians, GH. (2001). *Conservation biology - Reserach priorities fot the next decade*. Island Press, Washington.

Sultan, SE. (1987). Evolutionary implications of phenotypic plasticity in plants. *Evolutionary Biology,* 21, 127- 178.

Sumner, J.; Jessop T.; Paetkau D. & Moritz C. (2004). Limited effect of anthropogenic habitat fragmentation on molecular diversity in a rain forest skink, *gnypetoscincus queenslandiae. Molecular Ecology,* 13 , 259–269.

Templeton, AR, Shaw, K., Routman, E. and Davis, SK. (1990). The genetic consequences of habitat fragmentation. *Ann. Missouri. Bot. Gard.*, 77: 13–27.

Tewksbury, JJ.; Levey, DJ.; Haddad, NM.; Sargent, S.; Orrock JL.; Weldon A.; Danielson BJ.; Brinkerhoff J.; Damschen EI. & Townsend P. (2002). Corridors affect plants, animals, and their interactions in fragmented landscapes. *Proceedings of the National Academy of Sciences, USA*, 99, 12923–12926.

Valen, LV. (1965). Morphological Variation and Width of Ecological Niche. *Am. Nat.,* 99: 377-390.

Vos, P.; Hogers, R.; Bleeker, M.; Reijans, M.; van de Lee, T.; Hornes, M.; Frijters, A.; Pot, J.; Peleman, J.; Kuiper, M. & Zabeau, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nuc. Acids Res.*, 23: 4407- 4414.

Vucetich, LM.; Vucetich, JA.; Joshi, CP.; Waite, TA. & Peterson, RO. (2001). Genetic (RAPD) diversity in Peromyscus maniculatus populations in a naturally fragmented landscape. *Molecular Ecology,* 10, 35-40.

Weins, J. (1997). Metapopulation dynamics and landscape ecology. In: *Metapopulation Biology. Ecology, Genetics and Evolution* (eds Hanski IA, Gilpin ME), pp. 32–60. Academic Press, San Diego.

Yoon, MS.; Lee, J.; Kim, CY. & Baek, HJ (2007). Genetic relationships among cultivated and wild Vigna angularis (Willd.) Ohwi et Ohashi and relatives from Korea based on AFLP markers. *Genet. Resour. Crop Evol.*, 54: 875-883.

Young, AC. & Clarke GM. (2000). *Genetics, Demography and Viability of Fragmented Populations*. Cambridge University Press, Cambridge.

Young, AG.; Boyle, T. & Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution*, 11, 413–418.

Young, AG. & Clarke, GM. (2000). *Genetics, Demography and Viability of Fragmented Populations*. Cambridge University Press, Cambridge, UK.

Zhang, JF.; Kimatu, JN.; Guo, WL. & Liu, B. (2009). Habitat fragmentation causes rapid genetic differentiation and homogenization in natural plant populations – A case study in *Leymus chinensis. African Journal of Biotechnology,* Vol. 8 (15), pp. 3440-3447.

Zhang, M.; Kimatu, JN.; Xu, K. & Liu, B. (2010). DNA Cytosine Methylation in Plant Development. Review. *J. Genet. Genomic,s* 37 (2010) 1-12.

Zhou, C. & Yang, Y. (2003). Water characteristics on salt-alkali resistance of two divergent types in experimental *Leymus chinensis* populations in the Songnen Plain of China. *Acta Pratacul. Sin.*, 12: 65-68. [In Chinese with English abstract].

Plant Self-Incompatibility: Or, Self-Induction of Population Genetic Diversity

Łukasz Wolko and Ryszard Słomski

Department of Biochemistry and Biotechnology, Poznań University of Life Sciences
Poland

1. Introduction

The purpose of pollination is fertilization and seed production. Charles Darwin initiated studies on the phenomenon of plants "…which are completely sterile with their own pollen, but are fertile with that of any other individual of the same species" (Darwin, 1878). This inability for self-pollination is defined as self-incompatibility (SI). Self-incompatible plants have genetic systems that prevent self-fertilization through the recognition and rejection of pollen expressing the same allelic specificity as that expressed in the pistils. The mechanisms of SI, by promoting the outcrossing, would have an important influence on diversification and speciation of angiosperms because effective genetic mechanism preventing inbreeding allowed flowering plants to adapt to different environmental conditions and to accelerate the pace of evolution. Currently, SI systems are distributed widely throughout angiosperm lineages and are present in approximately 19 orders, 71 families and 250 genera comprising approximately 60% of angiosperm species (Allen & Hiscock, 2008). However, self-incompatibility systems are greatly heterogeneous and it is difficult to establish their evolutionary relationships or to determine the ancestral SI state.

2. Self-incompatibility systems

Two classes of SI can be distinguished: heteromorphic and homomorphic. Heteromorphic systems are characterized by morphological differences between genotypes, particularly variations in the length of the style (heterostyly). In contrast, in homomorphic systems, the S-genotype cannot be distinguished morphologically and the incompatibility response relies only no physiological mechanisms (Castric & Verkmans, 2004).

Most of the homomorphic SI systems are controlled by a single locus (S-locus) (Hiscock & McInnis, 2003). The S-locus should consist of, at least, two tightly linked (non-recombining) and highly polymorphic genes. The first of these genes coding pollen identity and the second is expressed in the pistil as a factor of incompatible pollen recognition. In addition to these primary male and female determinants, many other genes, not linked to the S-locus, are involved in SI functions. These additional factors play some role in pollen tube rejection after the first step of incompatibility recognition (McClure & Franklin-Tong, 2006; Kubo et al., 2010).

In SI plants, pistil tissues are able to recognize and reject pollen of the same S-genotype (incompatible pollen), or to accept growth of genetically different pollen tubes. Thus, the SI

system protects plants from pollination with pollen carrying the same S-haplotype (McClure & Franklin, 2006).

Two main systems of homomorphic SI comprise: gametophytic self-incompatibility (GSI) and sporophytic self-incompatibility (SSI). In GSI, the incompatibility phenotype is determined by the pollen haploid genome. The S-locus determines S-specificity of pollen recognition and rejection, so the pollen is rejected by the GSI system when the S-haplotype is the same as one of two S-alleles of the diploid pistil. This indicates that the S-locus products readily expressed in the pistil and pollen interact with each other and determine the compatibility or incompatibility of the emerging pollen tube.

In sporophytic SI, the pollen S-phenotype is determined by the diploid genome of the parental plant. Therefore, in GSI systems half-compatibility is shown between individuals that share one S-allele, while in SSI systems crosses are always fully compatible or fully incompatible (Allen & Hiscock, 2008).

The SSI system has been reported in six families: *Asteraceae, Berulaceae, Brassicaceae, Caryopgyllaceae, Convolvulaceae* and *Polemoniaceae*. The molecular mechanism of the SSI has only been well characterized in *Brassicaceae* (see below) (Hiscock & McInnis, 2003).

Studies of GSI species at the molecular level have identified two completely different SI mechanisms. One GSI mechanism, which is found in the *Solanaceae, Rosaceae* and *Scrophulariaceae*, has S-RNase as the pistil S-component and an F-box protein as the pollen S-component (see below). The second mechanism has been identified only in *Papaver* (poppy), where the interaction between male and female determinants transmits a cellular signal into the pollen tube, resulting in an influx of calcium cations. This influx interferes with the intracellular concentration gradient of calcium ions which exists inside the pollen tube, essential for its elongation (McClure & Franklin-Tong, 2006; Zhang & Xue, 2008).

2.1 SSI mechanism of *Brassicaceae*

In the *Brassica* species, two highly polymorphic self recognition proteins encoded in the S-locus have been identified: the S-locus receptor kinase (SRK), which is displayed at the surface of stigma epidermal cells and the S-locus cysteine-rich protein (SCR or SP11), which is expressed in the anther tapetum (i.e. sporophytically) and localized in the pollen coat (Schopfer et al., 1999; Suzuki et al., 2000; Takasaki et al., 2000). The SRK gene consists of seven exons, with the first exon encoding the signal peptide and extracellular domain, the second exon encoding the trans membrane domain and the remaining exons encoding the cytoplasmic kinase domain (Stein et al., 1996). All SCRs are small proteins of approximately 50 amino acids consisting of two exons: one encodes a signal peptide and the second one - the mature SCR protein (Schopfer et al., 1999; Takayama 2001).

The interaction between the SRK and SCR proteins results in autophosphorylation of the intracellular kinase domain of the SRK and a signal is transmitted into the papilla cell of the stigma (Takayama et al., 2001). Another protein essential for the SI response is MLPK, a serine-threonine kinase, which is anchored to the plasma membrane from its intracellular side. The downstream cellular and molecular events, leading eventually to pollen inhibition, are poorly described (Murase et al., 2004).

2.2 S-RNase-based GSI

The origin of the so-called S-RNase-based mechanism has been placed early in the eudicot evolution and is likely the most widespread SI mechanism operating among extant flowering plants (Richman et al., 1997; Allen & Hiscock, 2008).

The S-RNase is a N-glycoprotein, consisting of five conservative regions and one or two hyper-variable (HV) sequence regions. The N-terminal signal peptide is short and absent in the mature protein, suggesting its secretory nature. The allelic polymorphism is clustered in the HV regions which are responsible for S-specificity. In *Solanaceae*, there are two such regions (HVA and HVB), while only one is in *Rosaceae* (RHV) (McClure & Franklin-Tong, 2006; Zhang & Xue, 2008). The sequence alignment of different S-RNases revealed large inter-specific diversity. The level of the S-locus polymorphism is similar to the histo-compatibility system of *Vertebrateae* (Newbigin, 1996).

The S-RNases are expressed in the pistil and secreted to extracellular space. Then the S-RNases penetrate pollen tubes where they interact with pollen proteins. Recent discoveries identified S-locus F-box (SFB or SLF) proteins as pollen components of the S-locus in all three families that utilize the S-RNase type of SI (Entani et al., 2003; Ikeda et al., 2004; Ushijiama et al., 2003, 2004; Sassa et al., 2007, 2010; De Franceschi et al., 2011). These proteins are components of the SCF (Skp1-Cullin1-F-box) complex (Qiao et al., 2004a, b).

The "protein degradation" model has explained the mechanism of S-haplotype-specific rejection of pollen tubes by S-RNase. The SLF proteins recognize non-self S-RNases and mediate their degradation by the ubiquitin-25S-proteasom system (Hua & Kao, 2006; Hua et al., 2007). According to this hypothesis, each SLF allelic product mediates degradation of all S-RNases except its own S-RNase. In incompatible pollen tubes, S-RNases exert cytotoxicity specifically by RNA degradation (Takayama & Isogai, 2005; Hua & Kao, 2006; Hua et al., 2007).

3. Evolutionary characteristics of SI

Besides studies on the molecular basis of self-recognition and rejection systems, high polymorphic S-locus might also provide a unique evolutionary and population history information. The evolutionary properties of self-incompatibility systems have been revealed through a fascinating interplay between empirical advances and theoretical developments since Wright (1939) suggested that the main evolution force driving the population genetics of SI species is negative frequency-dependent selection (Castric & Vekemans, 2004).

3.1 Frequency-dependent selection

The self-incompatibility mechanisms put the pressure on individuals inside populations. All of them should be heterozygous at the S-locus and at least three S-alleles are required in the population because, if the number of alleles falls below three, the fertilization is blocked and the population is sentenced to extinction.

All S-alleles in the population are in the state of equilibrium with a similar frequency. If, for some reason, the allele frequency is disturbed, the population should establish a new equilibrium state during a few generiations, because of the negative frequency-dependent selection. This balancing selection consists in favoring rare S-haplotypes by increased

opportunity for successful mating in the population (Newbigin, 1996; Richman & Kohn, 2000; Charlesworth, 2006).

Theoretically, the new S-allele occurring in the population due to mutation, should successfully fertilize all the plants, so its frequency will increase until it reaches a new equilibrium state. The likelihood of successful fertilization is inversely proportional to the frequency of this allele in the population and the size of the particular S-haplotypes in the population can be quantified through the proportion N_e/n_e, where N_e is the effective size of the whole population and n_e is the effective number of alleles in the population (Vekemans & Slatkin, 1994). Accordingly, S-locus has to be much more polymorphic than loci from a non-negative frequency-dependent selection and S-haplotypes are long-lived and resistant to the loss due to genetic drift. (Richman & Kohn, 2000).

Observations from natural populations of species possessing GSI systems show that equal S-phenotypes frequencies are often, but not always, found. Meanwhile, in SSI populations, dominance relations make it more difficult to set expectations for the frequencies of different alleles, because these will depend on the dominance relations between all pairs of alleles (Richman & Kohn, 2000; Bechsgaard et al., 2004). Theoretically, in equal frequencies of incompatibility types, the recessive alleles are expected to reach higher frequencies than dominant alleles, but they have shorter persistence times in the population (Schierup et al., 1997; Uyenoyama, 2000, Bechsgaard et al., 2004).

3.2 Trans-specific evolution

One of the fundamental questions about the volatility of S-alleles is the rate and the role of mutations. Based on the research on a large number of alleles observed, the mutation rate can be very high. However, the experimental studies have not confirmed such a high frequency of mutations in real populations (Newbigin, 1996). The relatively low mutation rate, coupled with the observation of an unexpectedly high polymorphism of the S-locus, lead to the conclusion that the S-allelic polymorphisms should be inherited from ancestors and the process of their formation and differentiation should begin before the currently existing species formation.

The negative selection which increases frequency of rare alleles also preserves polymorphism over a very long period of time. It has been estimated that genetic variability can be preserved at the S-locus for tens of millions of years, periods longer than the lifetime of a species. It means that S-haplotypes present in the related species have originated from a common ancestor (Ioerger, 1990; Takahata, 1990; Vekemans & Slatkin, 1994; Richman & Kohn, 2000; Wolko et al., 2010).

Due to the S-locus ancestral origin, the alleles from one species can be more internally divergent than alleles from different species but which originated from a common evolutional lineage. In the sequence-alignment analysis, the S-alleles from different species and genera can be clustered together indicating that the allelic lineages descended from a common ancestor (Richman et al., 1995; Sasala et al., 1996; Usijima et al., 1998; De Franceschi, 2011). In the *Solanaceae*, it has been estimated that majority of the observed S-locus polymorphisms arose before speciation from the common ancestor. This ancestor is believed to have occurred 35-45 million years ago (Ioerger et al., 1990; Richman at al., 1995, 1996a, 1996b; Igic et al., 2006; Paape et al., 2008; Kohn, 2008).

3.3 Long-term population history recorded in S-locus polymorphism

The evolutional history has been recorded in S-alleles DNA sequences and can be demonstrated by the conservatism of the sequence that takes place between many different species or even between plant types, if the allele lineage is particularly old. The consequence of this is that the alleles occurring after the time of speciation must be descendents from one of the existing alleles. Therefore, the taxonomic spread of new alleles is much more limited than of the older alleles. Polymorphism that persists for tens of millions of years can be used to study the history of S-locus lineages, populations, species and genera.

The analysis of genetic variability of the S-locus in the population should include two main aspects. The first is the number of alleles in the population. This number grows in the population due to negative selection favoring rare alleles. However, an unlimited increase in the number of alleles driven by selection is inhibited by forces of genetic drift. At the equilibrium state, it is expected that the number of alleles in the population will increase proportionally to the population size, because the drift effect in large populations is weaker. Due to the fact that the number of alleles in a population reflects random changes in its size, this parameter exhibits a relatively recent event (the fluctuations in abundance) in the history of taxa.

The second aspect is the age of alleles which can be estimated by the sequence alignment of alleles from common species and genera. It is expected that this parameter would be much less variable than the number of alleles, because the selection mechanisms should prevent the disappearance of alleles. Some mutations in the gene sequence change allelic specificity, but differences of the entire accumulated sequence reflect the history of alleles over millions of years during the creation of species and varieties (Richman at al., 1995, 1996a, 1996b).

Estimation of the S-allele evolution rate can provide information about changes in the population size that took place in the distant past. The sequences of polymorphic alleles of a single species would reflect the history of the species. For example, if sequences have a large number of new alleles, it means that, at some point in the past, most of the older lineage alleles were lost, probably due to a sudden reduction in the population size ("bottleneck"). Therefore, the pedigree analysis of S-polymorphisms provides an opportunity for paleo-population genetic studies (Takahata, 1990).

For example, in the total number of 93 S-allele sequences derived from the species of *Physalis* and *Witheringia,* all have fallen within only three lineages that predate the most recent common ancestor of these two genera (Richman at al., 1996a; Richman & Kohn 1999, 2000; Lu, 2001; Stone & Pierce, 2005). Therefore, the bottleneck of the S-locus clearly occurred prior to the most recent common ancestor of *Physalis* and *Witheringia.* Paape et al. (2008) identified *Physalis* and *Witheringia* S-allele lineage homologes in *Iochromina.* But *Iochromina* S-locus seems much more polymorphic and several lineages not represented in *Physalis* and *Witheringia* have been identified there. This means that the restriction in the S-locus number must have occurred after the most recent common ancestor of the group containing the *Iochrominae* and it has been estimated that the bottleneck must have occurred between 14 and 18 mya (Kohn, 2008; Paape et al., 2008).

However, the extreme type of bottleneck event is needed to explain the fact that only three ancient lineages are represented among the S-alleles of *Physalis* and *Witheringia*. A founder event involving perches as few as two individuals would seem the only likely way for such

a restriction. But also the complete lack of further gene flow between other populations would be necessary because the subsequent gene flow would almost certainly introduce additional S-alleles (Kohn, 2008).

3.4 Sheltered load-constrains S-alleles diversification

Richman et al. (1996b), comparing the S-allele lineage diversity in *Physalis crassifolia* and *Solanum carolinease*, found that allele sequences from natural populations revealed a puzzling lack of correlation between the allele number and the number of trans-specific lineages in the genealogical analysis. *P. crassifolia* possesses 34 alleles residing in two lineages, while *S. carolinense* - 13 alleles scattered over multiple lineages within the *Solanaceae*. The lack of correlation between allele and lineage numbers is surprising because, as observed previously, differences in historical population size should lead to a positive correlation between the allele and lineage number (Richman et al, 1996a,b, Uyenoyama, 1997). In addition, genealogical studies exhibit a rapid diversification of the S-alleles near the base level of the genealogy with a subsequent deceleration of the diversification rate (Richman et al., 1996b).

The lack of correlation between the allele number and the divergence could be caused by differences in the fixation of mutations (Uyenoyama, 1997). The load of deleterious recessive mutations linked to S-locus accumulates due to two factors. First, the S-locus, by definition, is always expressed in heterozygous conditions, as pollen is rejected if it bears either allele held by the maternal plant. Second, the S-locus is in a non-recombining region of the genome (Uyenoyama, 1997; Coleman & Kao, 1992).

Therefore, the deleterious recessive mutations that accumulate in the region surrounding the S-locus are even more sheltered from selection than unlinked deleterious recessives. As these mutations are sheltered by being kept in heterozygous state, the lineage-specific load is called 'sheltered load' (Uyenoyama, 1997; Stone, 2004; Jorge et al., 2009). The sheltered load could be expressed when a mutation leads to the formation of a new S-allele. The new allele is compatible with any other in the population leading to the formation of embryos containing both the new allele and its progenitor. Therefore, the regions surrounding the S-locus will be combined in homozygotes and the sheltered load has a chance to be expressed. If this load is sufficiently deleterious, selection will eliminate one of the two alleles in the population (Uyenoyama, 2003; Stone, 2004; Mena-Alı́ et al., 2009). Thus, the fixation of new mutations and the diversification of the lineage should be blocked. Richman and Kohn (1999) found evidences that more divergent alleles were preserved when ecological factors reduced the effective population size and the number of S-alleles maintained.

Stone (2004) reported that crosses between distinct plants of *S. carolinense* sharing the same alleles at the S-locus led to dramatically high seed abortions, nearly equal to that found upon selfing. An excess of heterozygotes in the surviving progeny supports the hypothesis that these high abortion rates are caused by sheltered loads. Mena-Alı́ et al. (2009) found evidence of variation in the magnitude of deleterious load among S-alleles in *S. carolinense*. These results suggest that sheltered loads might slow the fixation of weak (partially compatible) S-alleles in the population, thus adding to the maintenance of a mixed mating system rather than leading to the fixation of selfing alleles.

4. Origination of self-incompatibility specificities

While the analysis of allele frequencies in populations can shed light on the fate of S-haplotypes, the major question is how new self-incompatibility specificities are generated and how male and female factors co-evolve to maintain their interaction?

Since pollen and pistil factors interact in the SI mechanism, mutations changing specificity of e.g. pollen, could result in a haplotype that is self-compatible. But the nonfunctional *S*-haplotypes are unlikely to be maintained along in population. They should be in the majority eliminated by inbreeding depression and selection against self-fertilization or, if they persist in a population, they should drive to extinction the functional haplotypes from which they arise (Chookajorn et al., 2004). Therefore, to form a new specificity, the mutation in both the female and the male components of the system should be required (Charlesworth, 2000).

4.1 Formation of new S-specificities

Attempts to explain how new specificities arise in two components of the SI system are faced with intriguing and difficult problems. One hypothesis proposes that the diversification process may devolop by means of intermediates and accumulation of polymorphism in the one gene that disrupts SI, followed by a compensatory mutation in the second gene that restores SI (Charlesworth, 2000; Uyenoyama & Newbigin, 2000). Alternatively, the variability in both pollen and pistil components can be tolerated thanks to flexibility of the physical interaction of the factors. The mutual recognition can exist until, by chance, a new mutation produces protein exhibiting stronger affinity with some varieties of the corresponding S-allele. Then the natural selection drives the strengthening of the interactions (Chookajorn et al., 2004, Kemp & Doughty, 2007; Naithani et al., 2007).

Since sequence polymorphism is necessary for the transitional stages of new S-allele formation under this model, variety of functional S-alleles in natural populations should provide opportunities to study this process. In *Brassica oleracea*, polymorphism within-specificity and developmental stages of new specificities were described (Miege et al., 2001, Sato et al. 2006), whereas in *Maloideae* (*Rosaceae*) populations the observed within-specificity polymorphism was low (Uyenoyama et al., 2001; Raspe & Kohn, 2007).

The rate of the formation of new specificities may depend on population conditions. It is known from studies on *Solanaceae* that the diversification rate of S-alleles increases following bottlenecks because of greater selection in the population below equilibrium of allele number (Richman, 2000; Paape et al., 2008; Miller et al., 2008). In the stable equilibrium state of populations, the negative frequency-dependent selection tends to preserve extreme polymorphism among alleles, but should have the opposite effect on the polymorphism within alleles, reducing it below levels of standard loci (Kohn, 2008).

One way of answering questions about S-haplotype creation could be the identification of amino-acid residues determining the specificity on the receptor and ligand. Slow progress has been achieved for the *Brassicaceae* SSI system, largely based on statistical analysis of sequences of the S-locus (Chookajorn et al., 2004; Kemp & Doughty, 2007; Naithani et al., 2007).

4.2 Coevolution hypothesis

The expectations for S-component genes are: tight linkage, high nucleotide polymorphism, long evolution history and evidence for balancing selection. The male and female factors should have long coevolution history because recombination events result in self-compatibility. As expected, in the SSI system of *Brassicaeae*, SRK and SCR genes show high levels of synonymous and non-synonymous nucleotide polymorphism and the hypothesis of strict coevolution of linked pollen and pistil genes could not be rejected (Sato et al., 2002).

However, in S-RNase SI systems, several aspects of current findings differ very markedly from what would be expected if F-box and S-RNaze genes had the same evolutionary history (Kohn 2008). Surprisingly, sequences of particular F-box genes are more closely related to other F-box genes in this species, than to F-box genes of the S-locus lineage in other relative species. This contrasts sharply with S-RNase genes, where sequences show trans-species relationships (Sassa et al., 1996; Ushijama et al., 1998).

Additionally, levels of polymorphism of the F-box genes are surprisingly low in some taxa, while in others they seem to conform to expectations. In *Prunus*, SFBs show similar to S-RNase levels of polymorphism and evidence of positive selection (Ikeda et al., 2004). The distances among *Prunus* S-RNase alleles and corresponding SLF alleles are correlated; however, genealogies of these genes do not strictly correspond, perhaps due to rare recombination events (Nunes et al., 2006).

In apple and pear (*Rosaceae*), multiple F-box genes within the S-locus region have been described. These genes were named S-locus F-Box Brothers (SFBB) (Sassa et al., 2007; De Franceschi et al., 2011). In European pear, for example, the S-locus contains no less than six SFBB members and exhibits much lower sequence diversity than their associated S-RNases and they show little or no evidence for positive selection. However, two of the genes show evidence of coevolution with S-RNases (De Franceschi et al., 2011). In *Antirrhinum* (*Plantaginaceae*), levels of divergence are at least an order of magnitude lower at the SLF locus than at the S-RNase locus (Wheeler & Newbigin, 2007; Newbigin et al., 2008). This fact implies that the histories of the F-box genes and their S-RNases are markedly different. This surprising finding is far from the prevailing theory of the S-locus gene coevolution.

4.3 Collaborative non-self recognition hypothesis

Just recently, Kubo et al. (2010) shed new light on pollen specificity component problem, proposing a collaborative non-self recognition model for the S-RNase-based SI. In this mechanism, pollen S-specificity is conditioned by multiple types of SLF genes and every single F-box gene is able to recognize and mediate degradation of only a small subset of non-self S-RNases. Within an S-haplotype, products of the entire suite of SLF proteins collaborate in recognition and detoxification of different S-RNases (Kubo et al., 2010). In the previous model with a single F-box gene, the loss of SLF function would be lethal because such mutant would be incapable of detoxification of any S-RNases and would be incompatible with any haplotypes. According to collaborative recognition model, the loss of function or mutation of a single SLF would block pollination only of pistils carrying the S-RNases recognized by the type of SLF affected by the mutation. Meanwhile, the mutation of the S-RNase gene causes escape detoxification from the existing repertoire of SLF proteins and blocks pollination by any pollen haplotypes (Kubo et la., 2010; Indriolo & Goring, 2010).

This model may suggest an evolution mechanism similar to pathogen-plant competition. It was previously proposed that SI systems evolved from pathogen defense mechanisms (Hodgkin et al., 1988; Elleman & Dickinson, 1999) and that S-RNases that mediate GSI are related to RNases involved in defense against pathogens (Kao & McCubbin, 1996; Galiana et al., 1997; Alen & Hiscock, 2008).

So, the new allelic specificities may be generated by the process of competition between S-RNases acting as plant-defense-like mechanism and pollen attempting to generate resistance-like response. The competition between stile and pollen components inducing polymorphism are also observed in real plant/pathogen interactions (Brunet & Mundt, 2000a, b). The first step of diversification would be the mutation in the S-RNase gene changing specificity and allowing new S-RNase to escape detoxification with the existing SLFs. It would cause the formation of a female-sterile and male-fertile phenotype. This new S-haplotype would spread across the population until it achieved the state of frequency-dependent selection equilibrium. The increase of frequency of the new S-RNase allele in the population induces the selection pressure to generate pollen resistance. Generation of the new SLF specificity which allows pollen to inactivate the new S-RNase and to pollinate pistils would be advantageous by supplementing potential mating partners. This mechanism should be practically tested on wild populations by studying the temporary existence of hypothesized female sterility plants.

5. Loss of SI

Mutations that promote self-fertility occur in both sporophytic and gametophytic SI systems and disturb S-alleles themselves as well as genes that modify S-locus expression and genes involved in the downstream rejection pathway (Tsukamoto et al., 1999, 2003a, b; Mena-Alí & Stephenson, 2007; Tao et al., 2007). The current models of SI maintain that the products of several unlinked genes are involved in the formation of multi-protein complex that cooperates in SI response (McClure & Franklin-Tong, 2006). Because of the complexity of the SI response, there are many opportunities for mutations in genes involved in the rejection process to have a qualitative or quantitative effect on self-fertility (Good-Avila et al., 2008). Therefore, in many natural populations, SI is rather a quantitative trait due to the segregation of unlinked genes that modify the strength of SI (Levin, 1996; Good-Avila & Stephenson, 2002).

The populations of the SI species with generic polymorphism for self-fertility have been found as populations in transition to SC. However, recently models with population stable state of equilibrium have been examined with mixed mating systems where genetic polymorphism for self-fertility may modify the SI and permit selfing in the same conditions (Vallejo-Marín & Uyenoyama, 2004; Porcher & Lande 2005).

The transition from SI to self-compatibility (SC) is viewed as one of the most frequent processes in plant evolution (Takebayashi & Morrell, 2001, Igic et al., 2006). There are two advantages of self-pollination: reproductive assurance because selfed progeny is better than no progeny at all; lower energy cost to produce selfed seed. Consequently, the frequency of mutation that promotes self-fertilization should increase in a population unless this process is counterbalanced by an opposing selective pressure of inbreeding depression (reduction of the selfed progeny fitness as a result of the increase in homozygosity that exposes deleterious recessive alleles) (Good-Avila et al., 2008).

Although mutations enhancing self-fertility are common in populations of SI species, most of them are eliminated by genetic drift or by purifying selection of inbreeding depression. On the other hand, these mutations are expected to increase in conditions of low numbers of S-alleles, low inbreeding depression (weak sheltered load) and limit of cross-pollen seed availability. In these conditions, mutations may become fixed resulting in the loss of SI.

When the SI mechanism is lost, polymorphism at the S-locus is switched to neutral selection and is expected to collapse in a few population (Kohn 2008). So, once polymorphism is lost, the SI system cannot be regained because, with fewer than three different alleles, all individuals are mutually incompatible. In addition, once the SI is lost, mutations in other genes of self-incompatibility are expected to accumulate (Stone, 2002; Igic et al., 2006, 2008).

If the SI is frequently and irreversibly lost during evolution, is self-incompatibility during this process vanishing in angiosperms? If it is not, than SI species must have a higher diversification rate (defined as the speciation rate minus the extinction rate) than SC taxa (Igic et al., 2004, 2008). The loss of SI mechanisms and transition to selfing should be a rather common evolutionary process. However, this pathway of evolution could often be a short-term solution and one-way ticket to the evolutionary dead-end (Takebayashi & Morrell, 2001; Kohn, 2008). Although selfing taxa usually show even greater levels of genetic diversification and would speciate faster than autcrossing taxa, the SI increases and preserves variation that might reduce the rate of total extinctions of taxa (Glémin et al., 2006; Kohn, 2008).

6. Conclusions

In this chapter, we tried to discuss issues concerning evolutional aspects of plant self-incompatibility. By mechanisms of frequency-dependent selection, self-incompatibility genes manage the population genetic diversity. It is a special case of Dawkins's "selfish gene", which, apart from natural selection, induces their spread in the population. This interesting phenomenon deserves a detailed analysis and verification of theoretical assumptions on the level of natural populations.

7. Acknowledgment

8. References

Alen, M.A. & Hiscock S.J. (2008) Enolution and phylogeny of self-incompatibility systems in angiosperms, In: *Self-Incompatibility in Flowering Plants – Evolution, Diversity and Mechanisms.* Franklin-Tong VE (Ed.), ISBN 978-3-540-68485-5Springer-Verlab, Berlin, Germany

Bechsgaard J., Bataillon T. & Schierup M.H. (2004) Uneven segregation of sporophytic selfincompatibility S-lleles in *Arabidopsis lyrata. J Evol Biol,* Vol. 17, No. 3, (May 2004), pp. 554-561, ISSN 1010-061X

Brunet, J. & Mundt, C.C. (2000a). Disease, frequency-dependent selection, and genetic polymorphisms: experiments with stripe rust and wheat. *Evolution*, Vol. 54, No.2, (Apr 2000), pp. 406–415, ISSN 1936-6426

Brunet, J. & Mundt, C.C. (2000b). Effects of competition on resistance gene polymorphism in a plant/pathogen system. *Heredity,* Vol 85, No 4, (Oct 2000), pp. 393–400, ISSN 0018-067X

Castric, V & Vekemans, X (2004) Plant self-incompatibility in natural populations: A critical assessment of recent theoretical and empirical advances. *Mol Ecol* Vol. 13, No. 10, (Oct 2004), pp. 2873–2889, ISSN 0962-1083

Charlesworth,D. (2000) How can two-gene models of self incompatibility generate new specificities? *Plant Cell,* Vol. 12, No. 3, (Mar 2000), pp. 309–310, ISSN 1835-2693

Chookajorn, T., Kachroo, A., Ripoll, D.R., Clark, A.G. & Nasrallah, J.B. (2004) Specificity determinants and diversification of the *Brassica* self-incompatibility pollen ligand. *Proc Natl Acad Sci USA*. Vol. 101, No.4, (Jan 2004), pp.911–917, ISSN 0027-8424

Charlesworth, D. (2006) Balancing selection and its effects on nearby genome regions. *PLoS Genet* Vol. 2, No. 4, (Apr 2006), pp. :379–384, ISSN 1553-7390

Coleman, C.E. & Kao T-H. (1992). The flanking regions of two *Petunia inflata* S alleles are heterogeneous and contain repetitive sequences. *Plant Mol. Biol.,* Vol. 18 No. 4 (Feb 1992), pp. 725–737, ISSN 0167-4412

Darwin, C.R.(1878) *The effects of cross and self-fertilisation in the vegetable kingdom*, 2nd edn. John Murry (Ed). London, UK

De Franceschi, P., Pierantoni, L., Dondini, L., Grandi, M., Sansavini, S. & Sanzol J. (2011) Evaluation of candidate F-box genes for the pollen S of gametophytic self-incompatibility in the *Pyrinae* (*Rosaceae*) on the basis of their phylogenomic context. Tree Genet. Genom. Vol. 7, No. 4 (Feb 2011), pp. 663–683, ISSN 1614-2942

Elleman, C.J. & Dickinson, H.G. (1999) Commonalities between pollen/stigma and host/pathogen interactions: Calcium accumulation during stigmatic penetration by *Brassica* oleracea pollen tubes. *Sex Plant Reprod* Vol. 12, No. 3, (Jun 1999), pp. 194–202, ISSN 0934-0882

Entani, T., Iwano, M., Shiba, H., Che, F.S., Isogai, A. & Takayama, S. (2003) Comparative analysis of the self-incompatibility S-locus region of *Prunus* mume: identification of a pollen-expressed F-box gene with allelic diversity. *Gene Cell,* Vol. 8, No. 3, (Mar 2003), pp. 203–213. ISSN 1356-959

Galiana, E., Bonnet, P., Conrod, S., Keller, H., Panabières, F., Ponchet, M., Poupet, A. & Ricci P. (1997) RNase activity prevents the growth of a fungal pathogen in tobacco leaves and increases upon induction of systemic acquired resistance with elicitin. *Plant Physiol*. Vol. 115, No. 4 (Dec 1997), pp. 1557-1567. ISSN 0032-0889

Glémin, S., Bazin, E. & Charlesworth D. (2006) Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proc R Soc Biol Sci B*, Vol. 273, No 1604, (Jun 2006), pp. 3011–3019. eISSN 1471-2954

Good-Avila, S.V. & Stephenson, A. (2002) The inheritance of modifiers conferring self-fertility in *Campanula rapunculoides* L. (*Campanulaceae*). *Evolution* Vol. 56, No. 2, (Feb 2002), pp. 263–272, ISSN 0014-3820

Good-Avila, S.V., Mena-Ali,J.I. & Stephenson, S.G. (2008) Genetic and environmental causes and evolutionary consequences of variations in self-fertility in self incompatible species. In: *Self-Incompatibility in Flowering Plants – Evolution, Diversity and*

Mechanisms. Franklin-Tong VE (Ed.), ISBN 978-3-540-68485-5Springer-Verlab, Berlin, Germany

Hiscock, S.J. & McInnis, S.M. (2003) The diversity of self-incompatibility systems in flowering plants. *Plant Biol*.Vol 5, No. 1, (Jan 2003), pp. 23-32. ISSN 1435-8603

Hodgkin, T., Lyon, G.D. & Dickinson HG (1988) Recognition in Flowering Plants: A Comparison of the *Brassica* Self-Incompatibility system and plant pathogen interactions. *New Phytol*, Vol. 110, No. 4 (Dec 1988), pp. 557–569, ISSN 0028646X

Hua, Z. & Kao, T.-h. (2006) Identification and characterization of components of a putative *Petunia* S-Locus F-box-containing E3 ligase complex involved in S-RNase-based self-incompatibility. *Plant Cell* Vol. 18, No.10, (Oct 2006), pp. 2531–2553, ISSN 1835-2693

Hua, Z.H., Meng, X.Y., & Kao, T.H. (2007) Comparison of *Petunia inflata* S-locus F-box protein (Pi SLF) with Pi SLF-like proteins reveals its unique function in S-RNase-based self-incompatibility. *Plant Cell, Vol.* 19, No. 11 (Nov 2007), pp. 3593–3609, ISSN 1835-2693

Igic, B., Bohs, L. & Kohn, J.R. (2004) Historical inferences from the self-incompatibility locus. *New Phytol*, Vol. 161, No. 1, (Jan., 2004), pp. 97–105, ISSN 0028646X

Igic, B., Bohs, L. & Kohn, J.R. (2006) Ancient polymorphism reveals unidirectional breeding system shifts. *Proc Natl Acad Sci USA,* Vol. 103, No. 5, (Jan 2006), pp.1359–1363, ISSN 0027-8424

Igic, B., Lande, R. & Kohn, J.R. (2008) Loss of self-incompatibility and its evolutionary consequences. *Int J Plant Sci* Vol. 169, No. 1, (Jan 2008) pp. 93–104, ISSN 1058-5893

Ikeda, K., Igic, B., Ushijima, K., Yamane, H., Hauck, N.R., Nakano, R., Sassa, H., Iezzoni, A.F., Kohn, J.R. & Tao, R. (2004) Primary structural features of the S-haplotype-specific F-box protein, SFB, in *Prunus*. *Sex Plant Reprod* Vol. 16, No. 5, (Dec 2004), pp. 235–243. ISSN 0934-0882

Indriolo, E. & Goring, D.R. (2010) Plant science. Pollen gets more complex. *Science*, Vol. 330, No. 6005, (Nov 2010), pp. 767-768, ISSN 0036-8075

Ioerger, T.R., Clark, A.G. & Kao, T.H. (1990) Polymorphism at the self-incompatibility locus in *Solanaceae* predates speciation. *Proc Natl Acad Sci USA,* Vol. 87, No. 24, (Dec 1990), pp.9732–9735, ISSN 0027-8424

Kao, T.H. & McCubbin, A.G. (1996) How flowering plants discriminate between self and non-self pollen to prevent inbreeding. *Proc Natl Acad Sci USA* Vol. 93, No. 22, (Oct 1996), pp. 12059–12065. ISSN 0027-8424

Kemp, B.P. & Doughty, J. (2007) S cysteine-rich (SCR) binding domain analysis of the *Brassica* selfincompatibility S-locus receptor kinase. *New Phytol*, Vol. 175, No. 4, (Jun 2007), pp. 619–629, ISSN 0028646X

Kohn, J.R. (2008) What genealogies of S-alelles tell us. In: *Self-Incompatibility in Flowering Plants – Evolution, Diversity and Mechanisms*. Franklin-Tong VE (Ed.), ISBN 978-3-540-68485-5Springer-Verlab, Berlin, Germany

Kubo, K., Entani, T., Takara, A., Wang, N., Fields, A.M., Hua, Z., Toyoda, M., Kawashima, S., Ando, T., Isogai, A., Kao, T.H. & Takayama, S. (2010) Collaborative non-self recognition system in S-RNase-based self-incompatibility. *Science*. Vol. 330, No. 6005, (Nov 2010), pp 796-799, ISSN 0036-8075

Lawrence, M.J. (2000) Population genetics of the homomorphic self-incompatibility polymorphisms in flowering plants. *Ann Bot*, Vol. 85 No. suppl. A (Jan 2000), pp. 221–226, ISSN 0305-7364

Levin, D.A. (1996) The evolutionary significance of pseudo self-fertility. *Am Nat*, Vol. 148, No. 2 (Aug 1996), pp. 321–332, ISSN 00030147

Lu, Y. (2001) Roles of lineage sorting and phylogenetic relationship in the genetic diversity at the self-incompatibility locus of *Solanaceae*. *Heredity* Roles of lineage sorting and phylogenetic relationship in the genetic diversity at the self-incompatibility locus of *Solanaceae. Heredity,* Vol. 86, No.2 (Feb 2001), pp. 195–205, ISSN 0018-067X

McClure, B. & Franklin-Tong, V. (2006) Gametophytic self-incompatibility: understanding the cellular mechanisms involved in "self" pollen tube inhibition. *Planta* Vol. 224, No. 22, (Jun 2006), pp. 233-245, ISSN 0032-0935

Mena-Alí, J.I. & Stephenson, A.G. (2007) Segregation analyses of partial self-incompatibility in self and cross progeny of *Solanum carolinense* reveal a leaky S-allele. *Genetics,Vol.* 177, No. 1, (Sept 2007), pp. 501–510, ISSN 0016-6731

Mena-Alí, J.I., Keser, L.H. & Stephenson, A.G. (2009) The effect of sheltered load on reproduction in *Solanum carolinense*, a species with variable self-incompatibility. Sex Plant Reprod Vol. 22, No. 2, (Jun 2009), pp. 63-71, ISSN 0934-0882

Miege, C., Ruffio-Chable, V., Schierup, M.H., Cabrillac, D., Dumas, C., Gaude, T. & Cock, J.M. (2001) Intrahaplotype polymorphism at the *Brassica* S-locus. *Genetics* Vol. 159 No. 2. (Oct 2001), pp. 811–822, ISSN 0016-6731

Miller, J.S., Levin, R.A. & Feliciano, N.M. (2008) A tale of two continents: Baker's rule and the maintenance of self-incompatibility in *Lycium* (*Solanaceae). Evolution,* Vol. 62, No. 5 (May 2008), pp.1052–1065, ISSN 0014-3820

Murase, K., Shiba, H., Iwano, M., Che, F.S., Watanabe, M., Isogai, A., & Takayama S. (2004). A membrane-anchored protein kinase involved in *Brassica* self-incompatibility signaling. *Science* Vol. 303, No. 5663, (Mar 2004), pp.1516–1519, ISSN 0036-8075

Naithani, S., Chookajorn, T., Ripoll, D.R. & Nasrallah, J.B. (2007) Structural modules for receptor dimerization in the S-locus receptor kinase extracellular domain. *Proc Natl Acad Sci USA* Vol. 104, No. 29, (Jul 2007), pp. 12211–12216, ISSN 0027-8424

Newbigin, E. (1996) The evolution of self-incompatibility: a molecular voyeur's perspective. *Sex Plant Reprod*, Vol. 9, No. 6, (Aug 1996) pp. 357–361. ISSN 0934-0882

Newbigin, E., Paape, T. & Kohn, J.R. (2008) RNase-Based Self-Incompatibility: Puzzled by Pollen S^w . *Plant Cell*, Vol. 20, No. 9, (Sept 2008), pp. 2286–2292. ISSN 1835-2693

Nunes, M.D.S., Santos, R.A.M., Ferreira, S.M., Viera, J. & Viera, C.P. (2006) Variability patterns and positively selected sites at the gametophytic self incompatibility pollen SFB gene in a wild self-incompatible *Prunus spinosa* (*Rosaceae*) population. *New Phytol* Vol. 172, No. 3, (Jul 2006), pp. 577–587, ISSN 0028-646X

Paape, T., Igic, B., Smith, S.D., Olmstead, R., Bohs, L., Kohn, J.R. (2008) A 15-million-year-old genetic bottleneck. *Mol Biol Evol,*Vol. 25, No. 4, (Jan 2008), pp. 655–663, ISSN 0737-4038

Porcher, E. & Lande, R. (2005) Loss of gametophytic self-incompatibility with evolution of inbreeding depression. *Evolution*, Vol. 59, No. 1, (Jan 2005), pp. 46–60, ISSN 0014-3820

Qiao, H., Wang, H., Zhao, L., Zhou, J., Huang, J., Zhang, Y. & Xue, Y. (2004a). The F-box protein AhSLF-S2 physically interacts with S-RNases that may be inhibited by the

ubiquitin/26S proteasome pathway of protein degradation during compatible pollination in *Antirrhinum*. *Plant Cell* Vol. 16, No. 3, (Mar 2004), pp. 582–595, ISSN 1835-2693

Qiao, H., Wang, F., Zhao, L., Zhou, J., Lai, Z., Zhang, Y., Robbins, T.P. & Xue, Y. (2004b). The F-box protein AhSLF-S2 controls the pollen function of S-RNase-based self-incompatibility. *Plant Cell,* Vol. 16, No. 3, (Mar 2004), pp. 2307–2322, ISSN 1835-2693

Raspé, O. & Kohn, J.R. (2007) Population structure at the S-locus of *Sorbus aucuparia*. *Mol Ecol*, Vol 16, No. 6, (March 2007), pp.1315–1325, ISSN 0962-1083

Richman, A.D., Kao, Th., Schaeffer, S.W. & Uyenoyama, M.K. (1995) S-allele sequence diversity in natural populations of *Solanum carolinense* (*Horsenettle*). *Heredity,* Vol. 75, No. 4, (Oct 1995), pp. 405-415, ISSN 0018-067X

Richman, A.D., Uyenoyama, M.K. & Kohn, J.R. (1996a) S-allele diversity in a natural population of ground cherry *Physalis crassifolia* (*Solanaceae*) assessed by RT-PCR. *Heredity,* Vol 76, No. 5 (May 1996), pp. 497-505, ISSN 0018-067X

Richman, A.D., Uyenoyama, M.K. & Kohn, J.R. (1996b) Contrasting patterns of allelic diversity and gene genealogy at the self-incompatibility locus in two species of *Solanaceae*. *Science*, Vol. 273, No. 5279 (Aug 1996), pp. 1212-1216, ISSN 0036-8075

Richman, A.D., Broothaerts, W. & Kohn, J.R. (1997) Self-incompatibility RNases from three plant families: homology or convergence? *Am J Bot*, Vol 84, No. 7, (Jul 1997), pp. 912–917, ISSN 1537-2197

Richman, A.D. & Kohn, J.R. (1999) Self-incompatibility alleles from *Physalis*: Implications for historical inference from balanced genetic polymorphisms. *Proc Natl Acad Sci USA*, Vol. 96, No 1, (Jan 1999), pp. 168-172, ISSN 0027-8424

Richman, A.D. (2000) Evolution of balanced genetic polymorphism. *Mol Ecol* , Vol. 9, No. 12, (Dec 2000), pp. 1953–1963, ISSN 0962-1083

Richman, A.D. & Kohn J.R. (2000) Evolutonary genetics of self-incompatibility in the *Solanaceae*. *Plant Mol Biol,* Vol. 42, No. 1, (Jan 2000), pp. 169-179, ISSN 0167-4412

Sassa, H., Nishio, T., Kowyama, Y., Hirano, H., Koba, T. & Ikehashi, H. (1996) Self-incompatibility (S) alleles of the *Rosaceae* encode members of a distinct class of the T2/S ribonuclease superfamily. *Mol Gen Genet,* Vol. 250, No. 1-2 (Aug 1996), pp. 547–557, ISSN 0026-8925

Sassa, H., Kakui, H., Miyamoto, M., Suzuki, Y., Hanada, T., Ushijima, K., Kusaba, M., Hirano, H. & Koba, T. (2007) S-locus F-box brothers: multiple and pollen-specific F-box genes with S-haplotypespecific polymorphisms in apple and Japanese pear. *Genetics*, Vol. 175, No. 4, (Apr 2007), pp.1869-1881, ISSN 0016-6731

Sato, K., Nishio, T., Kimura, R., Kusaba, M., Suzuki, T., Hatakeyama, K., Ockendon, D.J., Satta, Y. (2002) Coevolution of the S-locus genes SRK, SLG and SP11/SCR in *Brassica oleracea* and *B. rapa*. *Genetics,* Vol. 162, No. 2, (Oct 2002), pp. 931-940, ISSN 0016-6731

Sato, Y., Sato, K. & Nishio, T. (2006) Interspecific pairs of class II S-haplotypes having different recognition specificities between *Brassica oleraceae* and *Brassica rapa*. *Plant Cell Physiol,* Vol. 47, No. 3, (Mar 2006), pp, 340-345. ISSN 0032-0781

Schierup, M.H., Vekemans, X. & Christiansen, F.B. (1997) Evolutionary dynamics of sporophytic self-incompatibility alleles in plants. *Genetics,* Vol. 147, No. 2, (Oct 1997), pp. 835–846, ISSN 0016-6731

Schopfer, C.R., Nasrallah, M.E. & Nasrallah, J.B. (1999) The male determinant of self-incompatibility in *Brassica*. *Science*, Vol. 286, No. 5445, (Nov 1999), pp. 1697–700, ISSN 0036-8075

Stein, J., Dixit, R., Nasrallah, M.E. & Nasrallah, JB (1996) SRK, the stigma-specific S-locus receptor kinase of *Brassica*, is targeted to the plasma membrane in transgenic tobacco. *Plant Cell*, Vol. 8, No. 3, (Mar 1996), pp. 429–445, ISSN 1835-2693

Stone, J.L. (2002) Molecular mechanisms underlying the breakdown of gametophytic selfincompatibility. *Quart Rev Biol*, Vol. 77, No. 1, (Mar 2002), pp. 17–32, 0033-5770

Stone, J.L. (2004) Sheltered load associated with S-alleles in *Solanum carolinense*. *Heredity*, Vol. 92, No. 4, (Apr 2004), pp. 335-42, ISSN 0018-067X

Stone, J.L. & Pierce, S.E. (2005) Rapid recent radiation of S-RNase lineages in *Witheringia solanacea* (*Solanaceae*). *Heredity*, Vol. 94, No. 5, (May 2005), pp. 547-555, ISSN 0018-067X

Suzuki, T., Kusaba, M., Matsushita, M., Okazaki, K. & Nishio, T. (2000) Characterization of *Brassica* S-haplotypes lacking S-locus glycoprotein. *FEBS Lett*, Vol. 482, No. 1-2, (Sept 2000), pp. 102-108, ISSN 0014-5793

Takahata, N. (1990) A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. *Proc Natl Acad Sci USA*, Vol. 87, No. 7, (Apr 1990), pp. 2419-2423, ISSN 0027-8424

Takasaki, T., Hatakeyama, K., Suzuki, G., Watanabe, M., Isogai, A. & Hinata, K. (2000) The S receptor kinase determines self-incompatibility in *Brassica* stigma. *Nature*, Vol. 403, No. 6772, (Feb 2000), pp. 913–916. ISSN 0028-0836

Takayama, S.H., Shimosato, H., Shiba, M., Funato, F-E., Che, M., Watanabe, M., Iwano, M. & Isogai, A. (2001). Direct ligand–receptor complex interaction controls *Brassica* self-incompatibility. *Nature*, Vol. 413, No. 6855, (Oct 2001), pp. 534-538, ISSN 0028-0836

Takayama, S. & Isogai, A. (2005) Self-incompatibility in plants. *Annu Rev Plant Biol*, Vol. 56, (Jun 2005), pp. 467-489, ISSN 1543-5008

Takebayashi, N. & Morrell, P.L. (2001) Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *Am J Bot*, Vol. 88, No. 7, (Jul 2001), pp. 1143-1150, ISSN 0002-9122

Tao, R., Watari, A., Hanada, T., Habu, T., Yaegaki, H., Yamaguchi, M. & Yamane, H. (2007) Self-compatible peach (*Prunus persica*) has mutant versions of the S-haplotypes found in self-incompatible *Prunus* species. *Plant Mol Biol*, Vol. 63, No. 1, (Jan 2007), pp. 109-123, ISSN 0167-4412

Tsukamoto, T., Ando, T., Kokubun, H., Watanabe, H., Masada, M., Zhu, X., Marchesi, E. & Kao, T-H. (1999) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* (*Solanaceae*) in Uruguay containing both self-incompatible and self-compatible plants. *Sex Plant Reprod*, Vol. 12. No. 1, (Jan 1999), pp. 6-13, ISSN 0934-0882

Tsukamoto, T., Ando, T., Takahashi, K., Omori, T., Watanabe, H., Kokubun, H., Marchesi, E. & Kao, T-H. (2003a) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by loss of pollen function. *Plant Physiol Vol*, 131, No. 4, (Apr 2003), pp. 1903-1912, ISSN 0032-0889

Tsukamoto, T., Ando, T., Kokubun, H., Watanabe, H., Sato, T., Masada, M., Marchesi, E. & Kao, T-H. (2003b) Breakdown of self-incompatibility in a natural population of *Petunia axillaris* caused by a modifier locus that suppresses the expression of an S-

RNase gene. *Sex Plant Reprod* Vol. 15, No. 5, (May 2003), pp. 255-263, ISSN 0934-0882

Ushijima, K., Sassa, H., Tao, R., Yamane, H., Dandekar, A.M., Gradziel, T.M. & Hirano H. (1998) Cloning and characterization of cDNAs encoding S-RNases in almond (*Prunus dulcis*): primary structural features and sequence diversity of the S-RNases in *Rosaceae. Mol Gen Genet*, Vol. 260, No. 2-3, (Nov 1998), pp.261-268, ISSN 0026-8925

Ushijima, K., Sassa, H., Dandekar, A.M., Gradziel, T.M., Tao, R. & Hirano, H. (2003) Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen expressed F-box gene with haplotype-specific polymorphism. *Plant Cell*, Vol. 15, No. 3, (Mar 2003), pp. 771–781, ISSN 1835-2693

Ushijima, K., Yamane, H, Watari A, Kakehi E, Ikeda K, Hauck NR, Iezzoni A, Tao R (2004) The S-haplotype-specific F-box protein gene, SFB, is defective in self-compatible haplotypes of *Prunus avium* and *P. mume. Plant J*, Vol. 39, No. 4, (Aug 2004), pp. 573–586. ISSN 0960-7412

Uyenoyama, M.K., (1997). Genealogical structure among alleles regulating self-incompatibility in natural populations of flowering plants. *Genetics*, Vol. 147, No. 4, (Aug 2004), pp. 1389–1400, ISSN 0016-6731

Uyenoyama, M.K. (2000) Evolutionary dynamics of self-incompatibility alleles in *Brassica. Genetics,* Vol. 156, No. 1, (Sep 2000), pp. 351-359, ISSN 0016-6731

Uyenoyama, M.K. & Newbigin, E. (2000) Evolutionary dynamics of dual-specificity self-incompatibility alleles. Plant Cell Vol. 12, No. 3, (Mar 2000), pp. 310-312, ISSN 1835-2693

Uyenoyama, M.K., Zhang, Y. & Newbigin, E. (2001) On the origin of selfincompatibility haplotypes: Transition through self-compatible intermediates. Genetics Vol. 157, No. 4, (Apr 2001), pp. 1805–1817, ISSN 0016-6731

Uyenoyama, M.K. (2003). Genealogy-dependent variation in viability among self-incompatibility genotypes. *Theor Popul Biol,* Vol. 63, No. 4, (Jun 2003), pp. 281–293 ISSN 0040-5809

Vallejo-Marín, M. & Uyenoyama, M.K (2004) On the evolutionary costs of self-incompatibility: Incomplete reproductive compensation due to pollen limitation. *Evolution* Vol. 58, No. 9, (Sep 2004), pp. 1924–1935, ISSN 0014-3820

Vekemans, X. & Slatkin, M. (1994) Gene and allelic genealogies at a gametophytic self-incompatibility locus. *Genetics,* Vol. 137, No. 4, (Aug 1994), pp. 1157–1165, ISSN 0016-6731

Wheeler, D. & Newbigin, E. (2007) Expression of 10 S-class *SLF*-like sequences in *Nicotiana alata* pollen and its implications for understanding the pollen factor of the *S*-locus. *Genetics*, Vol. 177, No. 4, (Dec 2007), pp. 1–10. ISSN 0016-6731

Wolko, Ł., Antkowiak, W., Sips M. & Słomski, R. (2010) Self-incompatibility alleles in Polish wild pear (*Pyrus pyraster* (L.) Burgsd.): a preliminary analysis. *J Appl Gene,* Vol. 51, No. 1 (Jan 2010), pp 33-35, ISSN 1234-1983

Wright, S. (1939) The distribution of self-sterility alleles in population. Genetics Vol. 24, No. 4, (Jun 1939), pp. 538-552, ISSN 0016-6731

Zhang, Y. & Xue, Y. (2008) Molecular biology of S-RNase-based self-incompatibility. In: *Self-Incompatibility in Flowering Plants – Evolution, Diversity and Mechanisms*. Franklin-Tong VE (Ed.), ISBN 978-3-540-68485-5Springer-Verlab, Berlin, Germany

Genetic Diversity of Maize Landraces as Sources of Favorable Traits

Snezana Mladenovic Drinic, Violeta Andjelkovic
and Dragana Ignjatovic Micic
Maize Research Insitute, Belgrade
Serbia

1. Introduction

The Maize Research Institute Genebank conserves a sizeable collection of 2217 landraces and 3297 introduced accessions from 40 countries. Landraces are considered to be the most significant genotypes, since they represent the original biological material created by the process of natural selection and adapted to local growing conditions. The landrace collection is designated as national treasure and all the activities regarding the management, characterization and evaluation are given the highest priority in the Genebank. Moreover, accessions maintained in the collection are samples of the original populations which, through propagation, have lost a part of variability due to genetic drift and inbreeding. Genetic diversity of these landraces has never been comprehensively studied. Assessment of genetic diversity is an essential step helping to understand the relationship among and within the landraces, as well as their genetic structure. It is also valuable for setting priorities for genetic conservation, parental selection and as a source of favorable traits.

The search for superior genotypes regarding yielding ability, disease and pest resistance, stress tolerance or better nutritional quality is very hard, competitive and expensive. This is why breeders tend to concentrate to adapted and improved materials, avoiding wild parents, landraces and exotics, available in germplasm banks which would require long time, high financial support besides the difficulty to identify potentially useful genes. Evidently, there is a gap between available genetic resources and breeding program activities. While germplasm banks try to preserve as much as possible genetic variability to be used by breeders, breeding programs do not explore efficiently the available diversity, relying almost exclusively on their working collection. One of the consequences of the massive use of uniform commercial varieties in maize production is loss of its genetic variability. Only about 5% of maize germplasm is in commercial use.

Unlike high yielding varieties the landraces are endowed with tremendous genetic variability as they were not subjected to selection over a long period of time. This aids in the adaptation of landraces to wide agro-ecological niches making them important as genetic resources for useful traits.

With their reserve of ancestral genes, maize landraces are the real sources of genetic diversity and variability for maize improvement programs. Landraces can possess specific

characteristics: tolerance to abiotic and biotic stress, tolerance to herbicides, low antinutritional components content as phytate or a large grains content of proteins, oil or starch.

The objective of the chapter is to present the importance and usage of maize landraces genetic diversity as a source of favorable traits (resistance to abiotic and biotic stress, grain quality and other specific traits) for breeding purposes. Besides identification of target traits, activities designed to identify desirable characteristics encompass estimation of their genetic diversity using morphological and molecular markers.

2. Establishment of genetic diversity of maize landraces

Maize was introduced into Europe at the end of the 15th century from Central America and due to differences in climate needed to be adapted to cooler conditions. Dissemination from Spain, associated transmission over the old continent from different parts of America, and subsequent maize exploitation during the next five centuries contributed to the establishment of native maize genetic diversity in Europe. Alleles specific to European populations also emerged during adaptation to the local climate and environment (Rebourg et al., 2003).

The first introduction of maize from the west Indies, via southern Spain, Portugal and Mediterranean to the territory of the former Yugoslavia, (XVI century), consisted of the flint races, which were mostly maintained as exotic plants and garden speciemen (Radovic et al., 2000). The next introduction of new flint genotypes from Mexico increased the genetic diversity of the existing germplasm through crosses with the formely introduced flint types. Early and medium early flint types with shorter ear type, remained along Adriatic cost and its neighbouring continental area.

The new flint race from North America were brought in the thrid introduction wave (XVII century) and created local flints of unique genetic bases, better adapted to cooler conditions. Dent types from North America, especially Corn Belt Dent, spread rapidly after introduction at the end of 19th century and resulted in creation of orginal genotypes specific for European Corn Belt. Landraces specific for former Yugoslavia Corn Belt, involving Vojvodina, Slavonija and river valleys, were developed from crosses of flint and dent maize genotypes.

The first organized collecting of landraces from theritory of former Yugoslavia started in the 1960s. Nowdays, Maize Research Institute genebank maintains the collection of 2217 landraces collected from ex-Yugoslavian territories. Typical Corn Belt dents and denty flints of late maturity were collected in Yugoslavia Corn Belt region. Along the Adriatic cost flints of medium late maturity as well early maturity flints were predominant. At higher altitudes flints and denty flints were collected.

The most detailed study of Yugoslavian landraces was conducted by Pavlicic and Trifunovic (1966), when all the collected populations were aligned into 18 agro-ecological groups using the method of "natural classification" proposed by Anderson & Cutler (1942). The 18 agroecological groups are: Montenegrin flints, Bosnian early dents, Kosmet flinty dents, Macedonian flints, Eight rowed corn type of North eastern America, Derived flints, Mediterranean flint, Small-kernelled flints, Eight rowed soft dents, Romanian flints, Large-eared flints, White flinty dents Moravac, Dent type of USA corn belt dents, Derived dents, Dent type of southern east USA, Serbian dents, Flinty dents and Denty flints (Figure 1).

The utilization of genetic diversity of landraces in breeding programs started with development of the first inbred lines in Yugoslavia form Ruma Golden Dent, Vukovar Dent, Šidski, Beljski and Novi Sad Golden Dent populations. Crosses of self pollinated landraces with American inbred lines, which expressed a high level of heterosis, were the base for development of line-hybrids. First Yugoslav maize hybrids were created from inbred lines of local origin and foreign inbred lines genetically divergent from local germplasm. In modern maize breeding program landraces are used for creation of broad base synthetic populations, development of core collections and as potential sources of favorable traits.

Fig. 1. Agroecological groups of maize landraces from ex-Yugoslav territories

A core collection consists of a limited set of genotypes derived from the existing collection and has to represent the genetic diversity of that collection. The Maize Genebank adopted two approaches to develop core collections. Representatives of each eco-geographical group (20%) were chosen based on morphological and biological traits for development of eco-core collection. Core subsets are used as donors for traits typical for each ago-ecological group. On the other hand, the main criterion for selecting landraces for the elite-core was the combining ability. Combining abilities of 900 medium and late maturing landraces, from regions with temperate growing conditions, were tested with four commercial inbred lines, Mo17, B73, V395/31 and LD230 (Radovic et al., 2000). Four core subsets of high combining abilities with specific inbred lines were obtained for further use in breeding programs.

3. Genetic diversity assessment

Significant maize genetic diversity decrease in the last few decades is a consequence of development of modern hybrids and agricultural systems. The use of a limited number of elite lines and synthetics heightens the risk of genetic uniformity in commercial maize production fields (Hallauer et al., 1988). Thus, maize breeders became more aware of the need for both maintaining genetic diversity among hybrid varieties and improving the management of genetic resources through the conservation of landraces (Goodman, 1994).

Landraces are the cultivated maize material with the highest genetic variation as well as with the best adaptation to the natural and anthropological environment were they have evolved (Maxted et al., 1997). They contain locally adapted alleles and represent an irreplaceable bank of highly co-adapted genotypes (Quals et al., 1997). In this sense, genetic variability is a valuable natural source of beneficial traits providing plant adaptation to different stresses and undergoing climatic changes.

Different methods can be used to assess genetic variability in plant species, such as pedigree data, morphological and molecular markers. Data obtained by landrace description are further statistically processed. Multivariate analyses (e.g., cluster analysis, principal component analysis and discriminate analysis) for measuring the degree of divergence among populations have been useful in different fields of research, allowing obtaining a summary of the most relevant characteristics. Populations (landraces) can be grouped together based on informative data and be used directly in a breeding program.

A broad approach using phenotypic and molecular markers is required to analyze diversity and to support conservation, management, and development of plant genetic resources (Hammer et al., 1999). Both kinds of markers have their advantages and drawbacks and their combined utilization is recommended to increase the resolving power of genetic diversity analyses (Singh et al., 1991).

Considering the origins of the populations and the fact that trade and communications have always been very intensive on a relatively small ex-Yugoslav territory it can be assumed that genetic relationships between the 2217 landraces from Maize Research Institute gene bank could be inconsistent with the existing classifications. Moreover, accessions that are maintained in the collection are samples of the original populations which, through propagation, have lost a part of variability due to genetic drift and inbreeding. A new phenotypic analysis, as well as molecular analysis, is being performed on these landraces in order to assess their genetic variability and verify their relationships.

3.1 Phenotypic analysis

Phenotypic markers have been of great value in studies of maize landraces (Galarreta & Alvarez, 2001; Lucchin et al., 2003; Ortiz et al., 2008). The most commonly used descriptors are related to plant architecture traits, tassel traits, ear and kernel characteristics. Earliness is also frequently a part of the landrace description. Relationships between traits are statistically analyzed using graphs, correlation coefficient estimations and principal component analysis (PCA). Genetic similarity can be calculated from the morphologic data by unweighted pairgroup method (UPGMA) cluster analysis based on Euclidian distance coefficients. Cluster analysis is used to reveal the association between landraces.

A sample of 54 maize landraces was analyzed with three populations per agro-ecological group. Besides ex-Yugoslavian landraces, six introduced populations were included, used as

a check since they are expected to be genetically distant from local populations. These populations originate from France, Georgia and China.

All populations were sown at Zemun Polje in 2008, in two different sowing densities – 44640 and 64935 plants per ha. The experimental design was RCBD (*Randomized Complete Block Design*) with two replicas, four rows per replica and 20 plants per row. For morphological trait measurements two medium rows per population were used. We measured for each plot 15 morphological traits (plant height, ear height, leaf number, husk leaf length, tassel length, central spike/tassel length, number of tassel primary branches, branched part/tassel length, ear length, ear row number, ear kernel/row number, ear diameter, kernel length, kernel width and kernel thickness) taken from 20 competitive plants (40 plants per population) and from two kernels per ear.

Based on the average values of 15 morphological traits, as well as ASI (anthesis - silking interval) and their standard deviations cluster analysis was performed using square Euclidean distance and Ward complete linkage method. All the statistical analyses were performed using program NTSYSpc 2.1.

Cluster analysis based on phenotypic traits highlights distance between populations originating from ex-Yugoslav territories including two populations from France (clusters 1, 2 and cluster 3) and introduced populations from Georgia and China (cluster 4). Dendrogram can be divided into four clusters. The first cluster includes only two populations, one from Montenegrin flints and the other from Mediterranean flint agro-ecological groups. Cluster II encompasses only local flint populations. The third cluster can further be divided in two sub-clusters (Figure 2). Sub-cluster IIIA includes mostly local flint and intermediary (flinty-dents and denty-flints) populations, while sub-cluster IIIB includes mostly local dent and intermediary populations. Both French populations, with flint type of kernels, fall into cluster III. Introduced populations from Georgia and China form a distinctive cluster IV, which also includes two local populations (one dent and one flinty dent).

Phenotypic analysis revealed large heterogeneity among local populations indicating that populations are highly adapted to specific environmental conditions and uses. As such, they could be a valuable source of genetic variability. These results are in agreement with the findings that maize is the most diverse crop known, containing extensive diversity at both the phenotypic and molecular levels (Buckler et al., 2006). More ever, studies on European maize germplasm have already pointed out the presence of a great variability in morphological traits and large molecular diversity in traditional populations (Gauthier et al., 2002).

3.2 Molecular marker analysis

Development of molecular marker techniques, which enable estimation of genetic diversity directly at the DNA sequence level, are expected to supplement and refine the morphological-based classification. Molecular markers provide a direct measure of genetic diversity and go beyond indirect diversity measures, based on morphological traits or geographical origin.

RFLP (restriction fragment length polymorphisms), as well as PCR (polymerase chain reaction) based markers like RAPD (random amplified polymorphic DNA), AFLP (amplified fragment length polymorphisms) and SSR (simple sequence repeats) have been successfully used in maize genetic diversity studies. RFLP have been extensively used in assessing genetic variability in maize (Dubreil & Charchosset, 1998; Gauthier et al., 2002; Rebourg et al., 2001; Rebourg et al., 2003) due to their high discriminative power. However,

RFLP technique is labor intensive and time consuming, needs large quantities of DNA and is difficult to automate. On the other hand, RAPD which overcome some of the RFLP drawbacks have shown some problems with reproducibility of amplification and scoring of error data (Demeke et al., 1997; Karp et al., 1997).
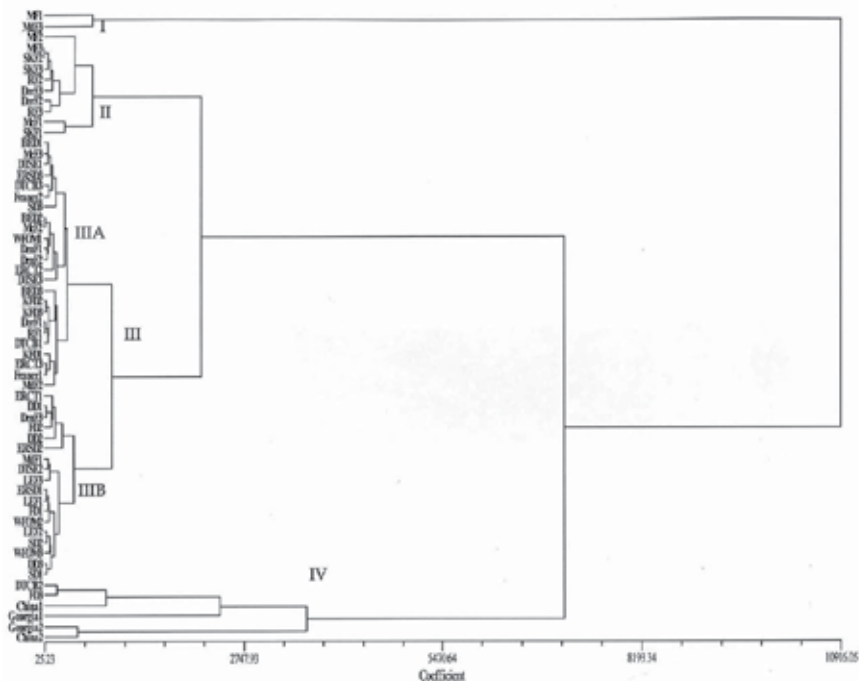
Abbreviations for the 18 agroecological groups are: MF - Montenegrin flints, BED -Bosnian early dents, KFD - Kosmet flinty dents, McF - Macedonian flints, ERCT - Eight rowed corn type of North eastern America, DerF - Derived flints, MdF - Mediterranean flint, SKF - Small-kernelled flints, ERSD - Eight rowed soft dents, RF - Romanian flints, LEF - Large-eared flints, WFDM - White flinty dents Moravac, DTCB - Dent type of USA corn belt dents, DD - Derived dents, DTSE - Dent type of southern east USA, SD - Serbian dents, FD - Flinty dents and DenF - Denty flints.

Fig. 2. Cluster analysis of the analyzed populations based on phenotypic traits using square Euclidean distance and Ward method

SSR and AFLP, the most informative among DNA-based markers, seem to be a better choice for genetic diversity studies (Smith et al., 1997; Pejic et al., 1998; Lubberstedt et al., 2000), since they overcome the problem of reproducibility and error data of RAPD markers, are generally much simpler to apply than RFLP, can be automated and yet provide results correlated with those from RFLP analysis (Pejic et al., 1998). They have been successfully utilized in maize genetic diversity studies (Lubberstedt et al., 2000; Barcaccia et al., 2003; Dubreil et al., 2006; van Eten et al., 2008; Eschholz et al., 2010; Hartings et al., 2010).

Although often used for the same scientific research AFLP and SSR present major distinctions. AFLP markers are genomic fragments detected after selective PCR amplification, while SSR markers consist of tandemly repeated units of short (1-6 bp) nucleotide motifs that show extensive length variation between individuals (Jarne & Lagoda, 1996). AFLP markers are dominant and biallelic, while SSR markers are co-

dominant and multiallelic. Pejic et al. (1998) suggested that SSR appear to be suited for the analysis of outcrossing heterozygous individuals and AFLP are advantageous when genome coverage is a major issue due to the presence of linkage disequilibrium, such as in inbred lines and breeding material. The choice among the markers to use depends on the material to be analyzed, the aim of the experiment and the facilities available.

Two studies with SSR and AFLP markers were conducted at MRI on the landraces from different agro-ecological groups. In the first study 21 landraces belonging to seven different agro ecological groups have been subjected to SSR analysis with aim to develop genetic fingerprint for characterization, identification and classification of the landraces, as well as for estimation of their genetic diversity (Ignjatovic-Micic et al., 2008). In the second study eighteen landraces from six different agro-ecological groups were assayed with AFLP and SSR markers (Ignjatovic-Micic et al., 2007). The main objectives of the work were to compare the level of information provided by SSR and AFLP and the genetic similarities (GS) obtained with the two marker systems. In both studies a DNA-pooling strategy that proved to be effective in RFLP analysis of maize populations (Dubreuil et al., 1999) was employed in order to assess if it could be applied with the same success for population analysis with SSR and AFLP markers.

SSR analysis was performed with 25 primer pairs covering all 10 maize chromosomes. The amplification reaction was carried out in 25µl reaction volume containing 1x enzyme buffer, 2.4 mM MgCl2, 200 µM dNTPs, 0.5 µM primers, 1 x BSA, 1U *Taq* polymerase and 200 ng of DNA. The amplification profiles followed were: an initial denaturation at 95°C/5min, followed by 15 cycles each of denaturation at 95°C/30sec, annealing at 63.5°C/1min (-0.5°C/cycle) and extension at 72°C/1min; another 22 cycles of 95°C/30sec, 56°C/1min and 72°C/1min were performed.

AFLP analysis was conducted using three primer combinations: EACA/M-CAG, E-AGA/M-CAG, E-ACA/M-CAT and E-AAA/CAG. *EcoRI* was used as a rare cutter and *MseI* as a frequent cutter. Pre-amplification was done by the following program: 1. 94°C/60sec, 2. 94°C/30sec, 3. 55°/30sec, 4. 72°C/60sec (30 cycles of steps 2-4). Amplification program consisted of: 1. 94°C/60sec, 2. 94°C/30sec, 3. 62°C/30sec, 4. 72°C/60sec (40 cycles of steps 2-4). Both SSR and AFLP amplified fragments were separated on 6% denatured polyacrilamide gels and visualized by silver staining (Bassam et al., 1991). The gels were scanned and band profiles for each primer were scored visually.

The SSR probes used detected single loci and each detected band was assumed to be an allele. Allele frequency was scored as percentage of individual bands within the sample using peak height estimated by scanning. Genetic distances (GD) between populations were evaluated by the *Rogers' Distance* (Rogers, 1972). For the AFLP primer combinations bands were binary scored (1 and 0) with each band being considered a locus and the genetic similarities among tested populations were calculated by the *Jaccard coefficient* (Jaccard, 1908). Cluster analysis was performed with unweighted pair-group method (UPGMA). Statistical analyses were done by NTSYSpc2.1 program package.

Twenty five SSR markers generated a total of 224 polymorphic bands, while three AFLP primers combinations generated 188 bands with 79.25 being polymorphic. The average polymorphism content was higher for SSR and higher genetic similarity values were estimated for AFLPs. In our study ranking of the agro ecological groups (order from the highest to the lowest average GS) was congruent for both marker systems.

The cluster analysis did not group the populations precisely with expectations based upon their agro-ecological groups. Better agreement with expected results was accomplished with AFLP analysis (Figure 3). In the work of Reif et al. (2005) it was suggested that SSR analysis of bulked populations cannot detect alleles with a frequency below 0.2, so that a large number of alleles with low frequencies remains undetected. They also concluded that differences between populations are emphasized due to a concentration on alleles with high frequencies.
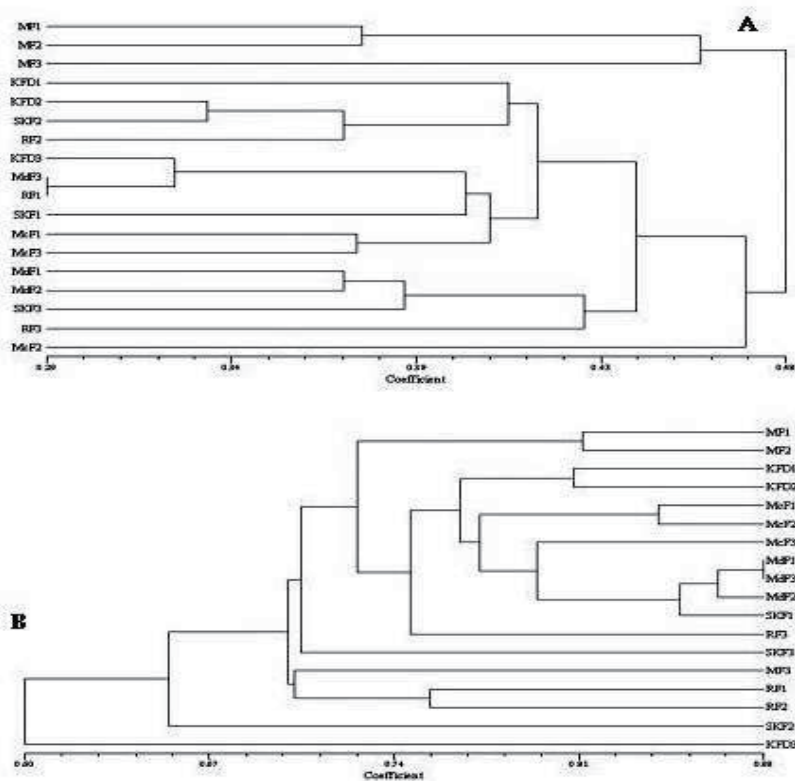
Fig. 3. Dendrograms of 18 maize populations (bulked population samples) constructed using UPGMA cluster analysis of A) Roger's distance values obtained by SSR analysis and B) Jaccard similarity values obtained by AFLP analysis. Abbreviations are the same as in Fig. 2.

The results in our work which indicate that SSRs are not a good choice for bulk analysis of populations could be explained by several reasons, besides the low sensitivity of bulk approach in detecting low frequency alleles. Use of di-nucleotide repeat SSR (excluded in the work of Reif et al., 2005) which tend to be more polymorphic than SSR with longer repeat motifs (Smith et al., 1997) could have caused higher scoring errors. The silver staining of the gels is probably not the right choice for bulk analysis due to its low sensitivity. It would be much better to use fluorescence as was done in the work of Reif et al. (2005), in which it was shown that the allele frequencies of a population could be estimated with high accuracy from the band intensities of SSR profiles of bulks. Populations cannot detect alleles with a frequency below 0.2, so that a large number of alleles with low frequencies remain

undetected. They also concluded that differences between populations are emphasized due to a concentration on alleles with high frequencies. However, in a simulated SSR analysis of bulks they showed that the correlation between genetic distance values of individuals and bulks is high, but it is still to be checked for final conclusions.

Assessing molecular marker diversity among heterogeneous populations using a DNA bulking strategy was shown to be successful in different plant species (Yu & Pauls, 1993; Sweeny & Dannenberger, 1997; Dubreuil et al., 1999; Mian et al., 2002). The question is what marker system to use in order to obtain the most accurate results with maximum efficiency. In our work we compared SSRs and AFLPs, the markers suggested as the best choice in genetic diversity studies. Based on results of this comparison it could be concluded that AFLPs are more reliable than SSRs in assessing diversity of maize populations when using a DNA bulking strategy.

4. Genetic diversity of landraces as source for cytoplasmic male sterility (CMS)

Absence of pollen and plant inability to produce functional pollen grains is known as male sterility. Male sterility can be determined by nuclear (genic male sterility) or cytoplasmic (cytoplasmic genetic male sterility – **CGMS**) genes. CGMS (onwards referred to as CMS) is successfully used in commercial production of hybrid seed, avoiding the drawbacks of hand or mechanical emasculation (Kaul, 1988). Identification of CMS types is important because commercial production of maize hybrid seed today stands upon utilization of C and S cytoplasmic types, as self-pollination in plants can severely jeopardize its production. Detasseling (removal of flowers from maize plants by hand) can be replaced by introducing male sterility in maize. This reduces possibility of self-pollination and also saves seed producers millions of dollars per year in labor costs.

Three main types of CMS were identified in maize: CMS-T, CMS-S and CMS-C. Male sterile cytoplasm is distinguished by specific nuclear genes (*Rf* genes) that restore pollen fertility. These genes, restorers of fertility, suppress the male-sterile effect of the cytoplasm, allowing the production of viable pollen.

The use of uniform materials can be a threat in crop production and genetic vulnerability must be a constant concern in plant breeding for all species. An example is the extensive use of CMS-T in 1950-ies and 1960-ies, which showed to be extremely susceptible to *Helminthosporium maydis* race T. Severe yield losses occurred in the United States in 1970-ies and the next year in other regions of the world, due to the *Helminthosporium maydis* pandemic. For new CMS sources landrace genotypes worldwide were searched and one source was found within Maize Research Institute gene bank collection.

Tester lines containing nuclear *Rf* genes are traditionally used for identification and classification of CMS types. However, test-crossing procedure is time consuming, labor intensive and not precise. Development and appliance of molecular methods revealed that mutations responsible for CMS are located in mitochondrial DNA (mtDNA) in many plant species (Shnable & Wise, 1998). It was confirmed that chimeric genes – DNA parts with open reading frames (ORF) that comprise sequences derived from different genes, are responsible for cytoplasmic male sterility. Himeric T-*urf13* gene was detected in mtDNA CMS-T (Dewey et al., 1986) chimeric *atp6-C* gene in mtDNA CMS-C (Dewey et al.,

1991) and a repeated DNA region "R" containing two ORFs in mtDNA CMS-S (Zabala et al., 1997).

The unique characteristics of himeric mtDNA regions were used for developing several molecular methods. RFLP analysis is time consuming to be applied routinely in breeding programs, while PCR-based markers are known to be rapid and definitive identifiers of the cytoplasm (Sato, 1998; Nakajima et al., 1999). Liu et al. (2002) developed a multiplex PCR assay combining three primer pairs specific for each type of CMS in a single reaction. This assay reveals 398, 440 and 799 bp specific DNA fragments identifying C, T and S cytoplasm, respectively. No DNA fragment is amplified from the N-type cytoplasm in this assay because an mtDNA fragment unique for the normal type does not exist. It was shown to be a quick and a reliable method, convenient for analysis of a huge number of samples.

The multiplex PCR assay was successfully used for CMS identification of maize accessions from Maize Research Institute (MRI) Genebank collection. Over 100 sources (2% of the total accession numbers) of cytoplasmic male sterility (CMS) were found in field trials, distinguishable by their overwhelming frequency of male sterile plants in segregating test progenies. As the types of CMS involved were unknown, a screening experiment was performed using a multiplex PCR with specific primers for S, C and T cytoplasm (Liu et al., 2002).

For PCR analysis each source of CMS was tested in triplicate, i.e. three separate samples per genotype were analyzed. DNA extraction was done from ten seeds per sample with CTAB buffer by the modified method of (Saghai-Maroof et al., 1984). The PCR reaction mix contained 1× PCR buffer, 1.5 mM MgCl2, 0.8 mM dNTP, 50 pmol of CMSSF, CMSSR, CMSTF, CMSTR, CMSCF and CMSCR primers, 1U of *Taq* polymerase and 50 ng of DNA. The PCR amplification consisted of 40 cycles (steps 2, 3 and 4), as follows: 1) initial denaturation 94°C/2 min, 2) denaturation 94°C/1 min, 3) annealing 55°C/1 min, 4) elongation 72°C/1.30 min and 5) final elongation 72°C/5 min. The amplified fragments were separated on 1.5% agarose gels, stained with ethidium bromide and photographed. An illustration of the multiplex PCR is given in Figure 4.

The multiplex PCR method used in our research, which is recommended for large scale screening, gave conclusive results for most of the analyzed accessions. A small number of accessions could not be identified for cytoplasmic male sterility type as, for example, different PCR products were obtained. The structure of maize sterile cytoplasm types within a gene bank collection is given in Figure 5. The predominant CMS type within the analyzed accessions was S cytoplasm. Clear presence of some type of CMS was found in 4.56% of all the former Yugoslav landraces, which is a significant percentage and S type makes 84% of all CMS found. S type was also predominant (44%) within introduced populations. This significantly lesser presence of CMS-S within introduced populations could be the consequence of a smaller sample size analyzed. On the other hand, introduced populations and lines are of different provenience, coming from different parts of the world and it was shown that, for example, among populations from Argentina the dominant type was CMS-C and from Italy CMS-S (Vancetovic et al., 2010).

The multiplex PCR approach enabled a simple, fast and reliable large scale screening of maize cytoplasm within MRI gene bank accessions, reducing time for cytoplasm characterizations from several years to a few weeks. Indirectly, this also helps in pre-

breeding procedures when presence of specific sterile cytoplasm is required, with the final goal of commercial production of maize hybrids.
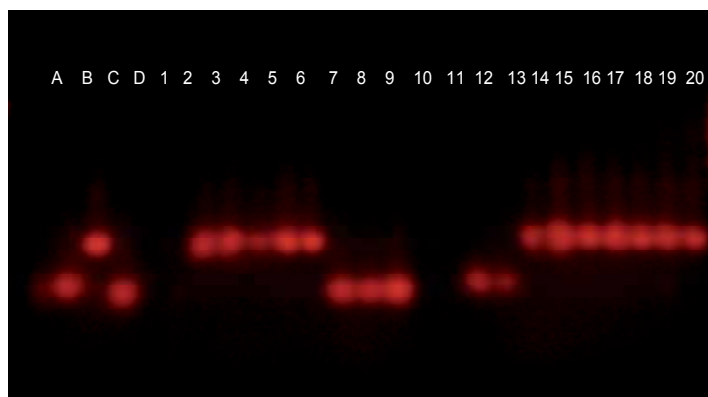
Lanes A – Cms-T (440bp), B – Cms-S (799bp), C – Cms-C (398bp), D – N type of cytoplasm; 1 – no band, 2 to 6 - Cms-S, 7 to 9 - Cms-C, 10 to 11 – no band, 12 to 13 - Cms-T, 14 to 20 - Cms-S.

Fig. 4. PCR amplification of CMS types with specific primers for C, T and S cytoplasm
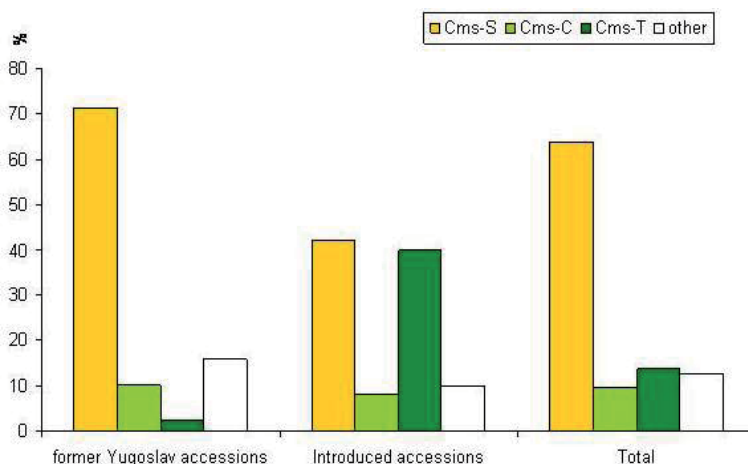
Fig. 5. The structure of maize sterile cytoplasm types within a gene bank collection

5. Genetic diversity as a source of drought tolerance

Drought can be defined as the absence of adequate moisture necessary for normal plant growth and to complete the life cycle (Zhu, 2002). It is a "single most common cause of severe food shortage" (FAO) in developing countries and predicted global warming in XXI century will increase drought impact on crop production. Estimations of Intergovernmental Panel on Climate Change (IPCC, www.ipcc.org) are that between 1990 and 2100, globall increase of temperature will be 1.4 to 5.8 °C. Although the effect of this warming will be regionally distributed, it is assumed that increase in temperature, reduction in rainfalls, together with increase of incidents of insects and pests will reduce crops grain yield, particularly in tropical and subtropical area (Ribaut, 2006).

In Southern Europe, drought followed by heat wave, reduces maize production for 20% (EC, 2004). The similar effect was in Serbia in 2007, when July the most critical period for grain yield formation was the hottest in last 50 years.

Drought severely affects maize grain yields by restricting season length and through unpredictable stress than can occur at any time during the vegetation season. Drought occurring at flowering has the greatest impact on yield losses compared to all the other developmental stages (Grant et al., 1989). Water deficit lasting one or two days during pollination may cause more than 20% reduction in yield (Hall et al., 1981).

However, drought tolerance is the most difficult task for maize breeders to be solved. It is polygenic and complex trait with high genotype x environment interaction. Moreover, this stress occurs randomly in timing and severity, making identification of drought tolerant genotypes more difficult. In agriculture, drought tolerance is generally evaluated at the end of growing season using grain yield (Dhanda et al., 2002). Maize, like other crops, can achieve drought tolerance through different approaches, involving escape, avoidance and tolerance of stress that influenced growth, development and grain yield formation (Cooper et al., 2004). Possible choice could be selection for earliness (drought avoidance or escape) but in years with adequate water supply it could bring lower yields. On the other hand, by given adequate rainfall, yield is usually correlated with late maturity, in maize, sorghum and sunflower (Edmeades et al., 1989). The most widely used strategy is to select for yield under optimal conditions, and latter on field testing under different moisture conditions. Assumptions of the approach are that genes for drought tolerance are present in high yielding material and the selection under optimal conditions can increase performance under sub-optimal conditions (Bolaños & Edmeades, 1996). However, in the absence of stress, drought tolerance in genotypes provides yield losses (Carena et al., 2009). Thus, many different strategies and traits have been proposed for using in maize breeding programs for improved drought tolerance and best improvement is achieved in long term processes.

Environmental conditions that caused drought and expression of genetic variation for drought tolerance are very important for successful breeding. The use of appropriate locations where drought stress can be controlled and applied at designated time and intensity (MSE- manage stress environment) has made significant improvement in maize breeding (Barker et al., 2005).

Moreover, because of difficulties in selecting for yield *per se* under drought conditions, it is recommended to use secondary traits to complement phenotypic selection. Measurement of secondary traits as well as grain yield is necessary in objective identification of drought tolerant genotypes. Secondary traits are "plant characteristics other than grain yield that give additional information about how yield will change under drought" (Fisher et al., 2003). Useful secondary trait is the trait that used together with yield makes expected progress from selection, greater than the progress made using grain yield alone. The use of secondary traits with grain yield, rather than selection for grain yield alone, has been shown to increase selection efficiency by about 20% in maize grown under stress induced by low nitrogen (Bänziger & Lafitte, 1997). Many physiological and morphological secondary traits related to survival or increased production in water limited environments have been proposed (Richards, 1996). Some of them have never been applied in breeding programs, because of their complexity or impossibility of routine application (Ribaut et al., 2006).

Ideally, secondary traits should be genetically associated with grain yield under drought, carry no yield penalty under favorable conditions, highly heritable, cheap and rapid to measure, stable over time and be able to be observed at or before flowering so that undesirable parents are not crossed (Edmeades et al., 1999).

A characteristic of maize under drought stress is a delay in silking resulting in an increase in the anthesis-silking interval (ASI), incomplete fertilization and decreased kernel development (Hall et al., 1984). The recommended secondary traits, routinely used by CIMMYT, for use in maize breeding programs are: flowering date (anthesis-silking interval, ASI), ears per plant (bareness), leaf rolling, tassel size and stay green (Monneveux et al., 2008).

Considering tremendous genetic diversity and importance of drought for maize production, long term project for improved drought tolerance has been started. All maize accessions, 6371, from Maize Research Institute Zemun Polje genebank were imposed to controlled water stress in Egypt, at the Sids Agricultural Research station (150 km south of Cairo). Among all accessions, 2217 landraces from former Yugoslavia were included into the experiment (Vancetovic et al., 2008).

Based on existing data, the material was divided into five groups, according to the duration of the vegetation season: extra early, early, medium, medium late and late. Groups were sown separately and have been irrigated until the appearance of the first tassels. After that, no irrigation was applied. The genotypes were estimated for ASI, staygreen (on the 1-5 scale, where 5 was the highest rate), total appearance of the plant (on the 1-8 scale, where 8 was the highest rate), barrenness, seed set and grain filing. A total of 167 landraces was selected according to reported criteria for further testing in Egypt and on two locations in temperate climate (Macedonia/Skopje and Serbia/Zemun Polje). According to the average of these traits for the three studied locations in two years, 14 most drought tolerant landraces were selected.

Besides field performance *per se*, development of good maize hybrids requires establishment of the best combination among heterotic groups (Barata & Carena, 2006). Therefore, chosen landraces were crossed to inbred lines from three heterotic groups (BSSS, Lancaster and independent source) in order to classify them into heterotic groups. The testcrosses for combining ability were studied in field trials at three locations in two years. The criterion for the selection of good testcrosses was the yield not statistically different from the check or performance index (based on grain yield and moisture at harvest) that was over 100% in comparison to the check (Andjelkovic et al., 2010).

5.1 Drought at seedling stage

Plant responses to environmental stresses include many adaptive mechanisms, changes in physiological and biochemical processes. Adaptation is associated with metabolic adjustments that lead to the accumulation of several organic solutes like sugars, betaines and proline (Yancey et al., 1982). It enables plants to maintain high rates of photosynthesis and can be advantageous in environments with large variation in water availability (Morgan, 1980). The experiments on maize genotypes (Moussa & Abdel-Aziz, 2008) pointed out that stress tolerance mechanism exists in seedlings stage and that early root and shoot development may provide useful information to enable the prediction of maize performance

under field conditions (Ruta, 2010). Usually, in laboratory conditions drought stress is investigated by withholding water. But, very often, different osmolytes (manitol, sorbitol, inorganic salts or polyethylene glycol-PEG) can be used to mimic water stress. Polymers of PEG with a molecular weight above 6000 Da do not penetrate cell wall and therefore mimic dry soil more closely (Verslues et al., 2006).

Maize landraces chosen within our project were analyzed in response to PEG treatment. Seeds were germinated for three days on moistened filter paper and then transferred into plastic pots containing Knopp solution with modified nitrogen content. For the next four days, plants were grown on ¼ strength nutrient solution supplemented with different forms of nitrogen. Nitrogen was supplied in the form of KNO_3, $Ca(NO_3)_2$ and $(NH_4)_2SO_4$. The initial pH of the solution was adjusted to 5.6. Plants were kept in a growth chamber under a 12-h photoperiod at 22/18°C, with the irradiance of $40Wm^{-2}$ and relative humidity of 70%. For the terminal 48 h of the growing period, one half of the plants (treatment) were grown on the fresh aerated nutrient solution supplemented with 4% polyethylene glycol (PEG, Mr 10 000), parallel to control plants grown on the nutrient solution without any addition. Root and shoot length, fresh and dry weight and prolin content were measured (Kravic et al., 2010).

Roots, responsible for water uptake, are the first plant organs sensing water stress. But roots and shoots usually respond differently to water deficit: in contrary to the shoot, relative or absolute size of the root might increase, depending on the root type and duration of stress (van der Weele et al., 2000). In our experiments drought stress at seedling stage affected shoot growth to a greater extent than root growth (Table 1).

In the experimental study of drought tolerant and sensitive maize seedlings (Ruta, 2010) the plant traits most severely affected by water shortage were shoot growth, (−75%) and leaf area (−68%), followed by root growth (−51%). It should be noted that at the same developmental stage, water stressed plants were two days older than the well-watered plants. This slower growth at water stress was much more pronounced at the shoot level than at the root level. Limited water availability retarded shoot development and promoted root development, consistent with results from plants stressed in solid substrates (Sharp & Davies, 1989; Weerathaworn et al., 1992a; 1992b). Genes for root development and proliferation during the seedling stage are also active at later growth stages, i.e. investigations of the genetic control of root traits at the seedling stage provide valuable information about its behavior at later stages of the plant development under the field conditions.

Although drought stress reduces plant water potential, it affects root and leaf growth differently (Frensch, 1997). It shifts the weight based shoot to root ratios in favor of the roots and it is a well known adaptive mechanism of preferring root over shoot development. The larger root-to-shoot ratio under severe stress is a further indicator that a moderate water stress had been achieved as it promotes a relatively better functioning of the roots (El Nadi et al., 1969; Sharp & Davies, 1979; Westgate & Boyer, 1985). Tendency of plants to increase the root to shoot ratio by allocating assimilates to the roots is typical response of plants in the field. In this experiment, and some others (Suharjo, 2007), plants were grown in solution, with mild (4% PEG) osmotic stress and root was not stimulated to grow deeper to extract more water (as in the soil). For this reason, root to shoot ratio does not differ too much compared to control conditions (Figure 6).

Landraces	Root lenght		% reduction	Shoot lenght		% reduction
	tretment	control		tretment	control	
L1	18.4	20.8	11.5	7.8	8.8	11.4
L2	9.7	11.1	12.6	6.5	6.6	1.5
L3	18.2	22.3	18.4	8.2	9.4	12.8
L4	17.1	20.0	**14.5**	7.4	9.2	**19.5**
L5	22.5	23.7	**5.1**	10.2	11.9	**14.3**
L6	17.0	18.0	5.5	8.3	8.5	2.4
L7	17.3	19.5	11.2	7.7	8.3	7.2
L8	17.1	17.2	**0.6**	8.4	9.4	**10.6**
L9	17.1	18.0	**5.0**	8.0	8.8	**9.1**
L10	14.7	16.3	9.8	7.5	8.2	8.5
L11	13.8	16.6	**16.8**	5.7	7.0	**18.6**
L12	16.4	17.7	**7.3**	9.3	10.2	**8.8**
L13	18.8	20.3	**7.4**	8.1	9.9	**18.2**
L14	20.4	22.4	**8.9**	8.4	10.1	**16.8**

Table 1. Root and shoot length and percent (%) of reduction under control and stress condition (PEG)

In the genetic study of quantitative trait loci (QTLs) in maize seedling some QTLs were specific to water stress, such as QTLs for root dry weight, shoot dry weight, and leaf area-to-root length ratio (Ruta, 2010). The QTLs for shoot-to-root ratio, were collocated and overlapped with QTLs for ear number in the field under intermediate and severe stresses (Ribaut et al., 1996) and for the anthesis-silking interval were collocated with QTLs for the numbers of crown roots and seminal roots. An early increase in growth of main roots seemed to be linked to a later decrease in ASI, while an early increase in lateral root length distribution on the main axile root may extend ASI. Considering that conventional breeding may take about ten years from crossing to the final release of new maize hybrids, it is important to find rapid screening technique under laboratory conditions. QTLs controlling the balance between early root and shoot development may provide useful information to enable the prediction of maize performance under field conditions.

6. Genetic diversity of landraces for grain quality

Characterization of genetic diversity of maize landraces aids efficient exploring of the allelic variation for genetic improvement of economically desirable traits, such as grain quality. Maize is a relevant food source, so the quantification of the nutritionally important grain constituents is important for the best exploitation of different genotypes. In this context, the landraces represent a good source of genetic variability and may help to identify the most suitable materials for the development of more nutritious foods. Many studies have documented genetic and phenotypic variability for grain composition traits in maize (Pollak & Scott, 2005; Reynolds et al., 2005; Berardo et al., 2009; Has et al., 2010, Mittelmann et al., 2011).

To determine genetic diversity for grain quality traits a set of 54 landraces was grown in a randomized complete block design (RCBD) with two replications at one location. Grain
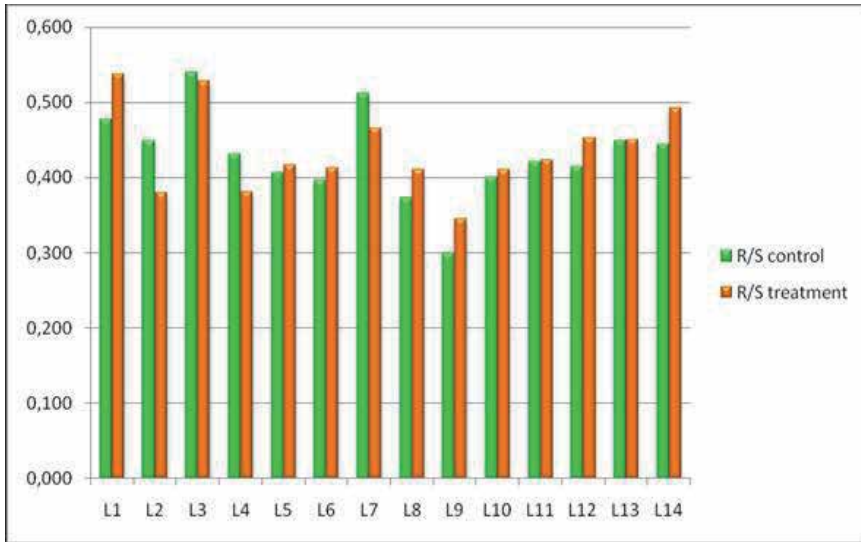
Fig. 6. Root to shoot ratio (dry weight) in maize landraces (L1-L14) at seedling stage

protein, oil and starch content were measured using Infratec 1241 Grain analyzer, (Foss Tecator, Sweden) and expressed in percents (%). Phytate and Pi were determined by the method of Latta & Eskin (1980), modified by Sredojević & Dragicevic (2009).

The genetic variability for protein, oil and starch content was recorded in the analyzed landraces (Andjelkovic et al., 2011). The protein content ranged from a low level of 10.10% (a population from Derivated dents) to a high level of 14.85% (a population from Small-kernelled flints) and averaged 12.34%. The highest average protein content of 14.05% had population from Montenegrin flints and the lowest of 10.75%, one population from Derived dents (Figure 7). Landraces belonging to flint types have higher protein content than dent type landraces.

Populations from Montenegrin flints and Macedonian flints had the highest average and the second highest average oil content, 4.8% and 4.75%, respectively (Figure 8). The lowest average oil content had populations from Large-eared flints, 3.63%. Landraces belonging to flint types had higher oil content than landraces of dents type.

The starch content of maize landraces ranged from 66.35% to 71.50% and averaged 68.61% (Figure 9). The highest average starch content had populations from Derived flints (70.43%), and the lowest populations from Montenegrin flints (67.53%). Among landraces some interesting forms with high level of starch content were identified: Dent flints pop. 709 (71.10%), Derived flints pop. 190 (71.50%) and pop. 186 (70.95%), Dent type of USA Corn Belt Dents pop. 268 (70.85%).

It could be concluded that many tested landraces show an exceptional kernel quality based on the oil, protein and starch content, and this makes them suitable to be used for futher breeding improvment programs.
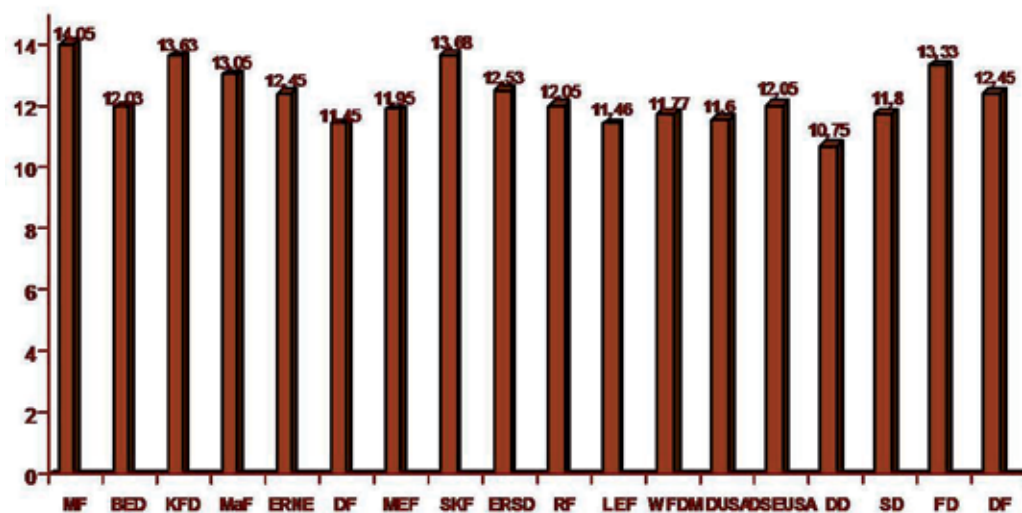
Abbreviations: MF-Montenegrin flints; BED-Bosnian early dent; KFD-Kosmet flinty dents;MaF-Macedonian flints; ERME-Eight rawed corn type of NE America; DF-Derived flints; MEF-Mediterranean flint, SKF-Small kernel flint; ERSD-Eight rowed soft dents; RF-Romanian flints; LEF-large-earted flints, WFCM- White flinty dents Moravac; DUSA-Dent type of USA corn belt dents; DSEUSA-Dent type of southern east USA; DD-Derived dents; SD-Serbian dents; FD-Flinty dents; DF-Denty flints.

Fig. 7. Protein content of landraces of agro ecological groups
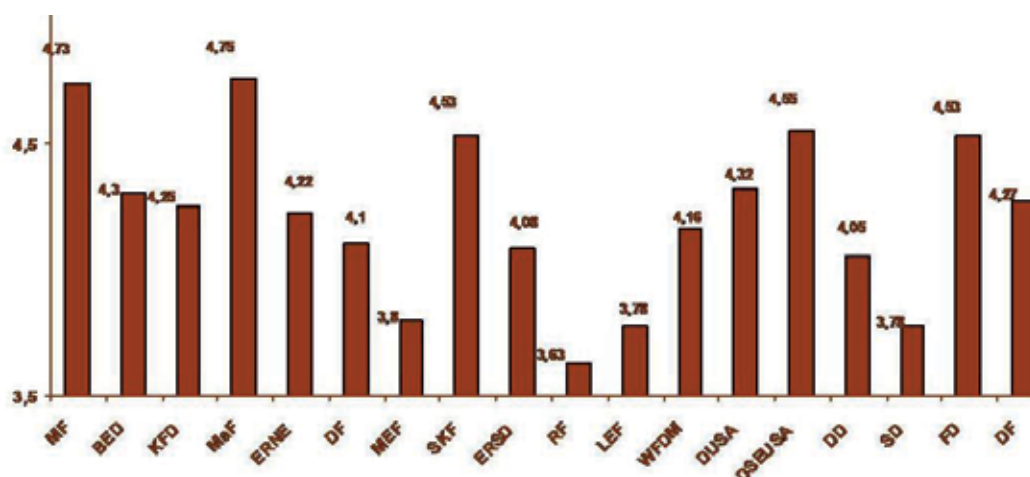
Fig. 8. Oil content of agroecological groups of landraces, abbreviations are same as on Fig. 7
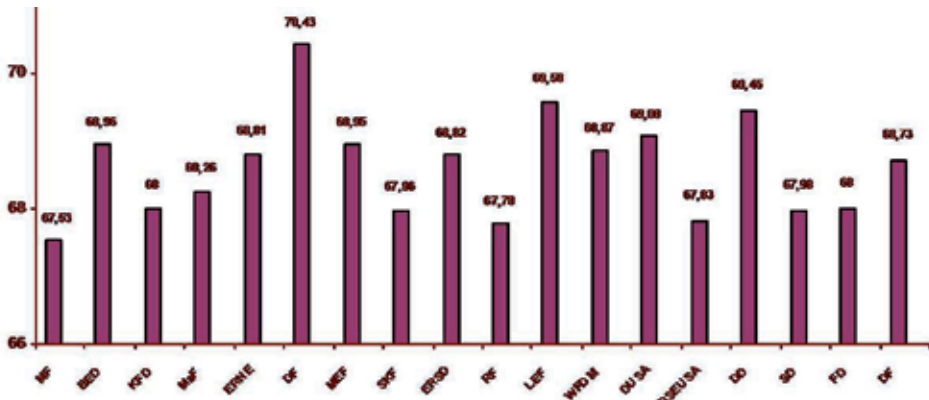
Fig. 9. Starch content of agroecological groups of landraces, abbreviations are same as on Fig.7

Many of the problems associated with P in maize grain are not due to concentration of the total P per se, but rather to the fact that most of the P is bound in phytate. Therefore, it would be desirable to increase the amount of available P and reduce the amount of phytate in maize grain. Approaches to resolve problems of the bad nutritive quality of grain phytate include engineering of crops to express high levels of *phytase* enzyme in seeds (Brinch-Pedersen et al, 2002), or through recurrent selection that uses the indigenous quantitative genetic variation (Raboy, 2009). Genetic variability in phytate contents of 54 landraces was observed, with values in the range from 1,147 to 4,13 and the average of 2.91 g kg-1 (Drinic et al, 2009). Twenty six population had high, twenty intermediate and eight low phytate content. Populations beloging to Derived flints had the lowest (1.95 g kg-1) while populations from Flinty dents had the highest (3.43 g kg-1) phytate content ,(Figure 10).
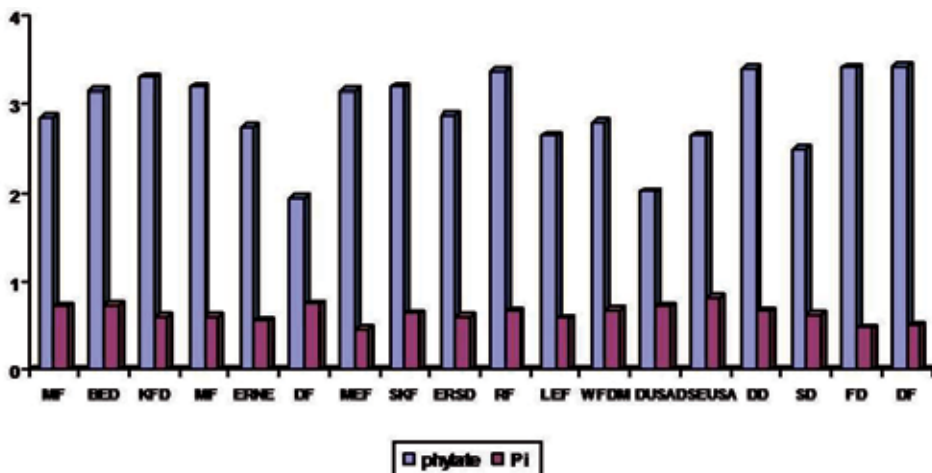
Fig. 10. Phytate and Pi contents of agroecological groups of landraces, abbreviations are same as on Fig.7

Populations from Mediterranean flints had the lowest (0.46 g kg-1) and populations from Dent type of southern east USA the highest (0.82 gkg-1) Pi content. One of the populations, i.e. population 216 was determined to have the lowest phytate concentration of 1.14 g kg-1,

a Pi concentration 30% greater than Pi mean and will be used for further breeding genotypes with low phytate content and good agronomic traits.

7. Conclusion

Landraces were developed through a complex adaptation process of different orginal genotypes to diverse climatic and soil conditions and therefore are the most accessible part of maize biodiversity. Genetic diversity of landraces have important role in maize breeding and could have a significant impact on the improvment of plants as a valuable source of useful traits. A considerable diversity among landraces from MRI genbank has been observed with morphological and molecular markers. The landraces are autochthonous source of potentially useful traits and alleles for improvement of the existing modern varieties. Fourteen most drought tolerant landraces were selected for further breeding programme aimed to broaden the genetic base of the elite breeding material. Presence of a wide genetic diversity of landraces based on their oil, protein, starch and phytate content, makes them more than suitable to be used for further maize breeding improvement programs.

8. References

Anderson, E.&Cutler, C. (1942). Races of Zea mays L.1.Their recognition and classification. *Ann.Missouri Bot.Gard*.29, pp.69-89.

Andjelkovic, V.; Kravic, N.; Vancetovic, J.; Dragicevic V. & Drinic Mladenovic S. (2011). Maize landraces as a natural source of beneficial traits. Proceeding of the 4th International scientific conference, Vukovar, Croatia, 1-3.06.2011, pp. 37-45, ISBN 978-953-7693-01-5

Andjelković, V.; Božinović, S.; Pavlov, M. & Vančetović, J. (2010). Combining abilities for grain yield of the most drought tolerant maize accessions from the MRI gene bank. In: *Book of Proceedings of 45th Croatian & 5th International Symposium on Agriculture*, February 15-19, 2010, Opatija, Croatia, pp. 368-371.

Banziger, M. & Lafitte, H.(1997). Breeding tropical maize for low N environments: II. The values of secondary traits for improving selection gains under low N. *Crop Science*, Vol.37, pp.1110- 1117, ISSN 0931-2250

Barata, C. & Carena, M. (2006). Classification of North Dakota maize inbred lines into heterotic groups based on molecular and testcross data. *Euphytica*, Vol. 151, pp. 339-249. ISSN 0014-2336

Barcaccia,G.; Lucchin, M. & Parrini, P. (2003). Characterization of a flint maize (*Zea mays* var. *indurata*) Italian landrace, II. Genetic diversity and relatedness assessed by SSR and Inter-SSR molecular markers. *Genetic Resources and Crop Evolution,* Vol.50, pp. 253–271, ISSN 0925-9864

Barker, T.; Campos, H.; Cooper, M.; Donlan, D.; Edmeades, G.; Habben, J.; Schussler, J.; Wright, D. & Zinselmeier, C. (2005). Improving drought tolerance in maize. *Plant Breeding Reviews*, Vol. 25. Pp. 173-253. ISSN 0730-2207

Bassam, B.J.; Caetano -Anolles, G.& Gresshoff, P.M. (1991). Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry*, Vol. 196, pp. 80-83, ISSN 0003-2697

Berardo, N.; Mazzinelli, G.; Valoti, P.; Lagana, P. & Redaelli, R. (2009). Characterization of maize germplasm for the chemical composition of the grain. *J. Agric. Food Chem.*, 57, pp.2378-2384, ISSN 0021-8561.

Bolaños, J. & Edmeades, G. (1996). The importance of the anthesissilking interval in breeding for drought tolerance in tropical maize. *Field Crops Research*, Vol, 48, pp.65–80. ISSN 0378-4290

Brinch-Pedersen, H.; Sørensen, L.D & Holm P.B. (2002). Engineering crop plants: getting a handle on phosphate. *Trends Plant Sci*. 7, pp.118–125. ISSN 1360-1385

Buckler, E.S.; Gaut, B.S. & McMullen, M.D. (2006). Molecular and functional diversity of maize. *Current Opinion in Plant Biology*, Vol. 9, pp.172-176, ISSN 1369-5266

Carena; M., Bergman, G.; Riveland, N.; Eriksmoen, E. & Halvorson, M. (2009). Breeding maize for higher yield and quality under drought stress. *Maydica,*Vol. 54, pp.287-296. ISSN 0025-6153

Cooper, M.; Podlich, D. & Smith; O. (2004). Complex traits and gene-to-phenotype models. In: *Proceedings of the 4th International Crop Science Congress*, pp. 67, Brisbane, Australia.

Demeke, T.; Sasikumar, B.; Hucl, P. & Chibbar, R.N. (1997). Random amplified polymorphic DNA (RAPD) in cereal improvement. *Maydica*, Vol.42, pp. 133-142, ISSN 00256153

Dewey, R.E.; Levings, C.S.III; & Timothy, D.H. (1986). Novel recombinations in the maize mitochondrial genome produce a unique transcription unit in the Texas male sterile cytoplasm. *Cell*, Vol. 44, pp.439–449, ISSN 0092-8674

Dewey, R.E.; Timothy, D.H. & Levings, C.S.III. (1991). Chimeric mitochondrial genes expressed in the C male sterile cytoplasm of maize, *Current Genetics*, Vol. 20, pp. 475–482, ISSN 0172-8083

Dhanda, S. & Sethi, S. (2002). Tolerance to drought stress among selected Indian wheat cultivars. *Journal of Agricultural Science,* Vol. 139, pp. 319-326. ISSN 0021-8596

Drinic Mladenovic, S.; Ristic, D.; Sredojevic S.; Dragicevic V.; Ignjatovic Micic, D. & Delic N. (2009). Genetic variation of phytate and ionorganic phosphorus in maize population. *Genetika*, Vol. 41, No. 1, pp. 107 -115, ISSN 0534-0012.

Dubreil, P. & Charcosset, A. (1998). Genetic diversity within and among maize populations: a comparison between isozyme and nuclear RFLP loci. *Theoretical and Applied Genetics,* Vol. 96, pp. 577-587, ISSN 0040-5752

Dubreuil, P.; Rebourg, C.; Merlino, M. & Charcosset, A. (1999). The DNA-pooled sampling strategy for estimating the RFLP diversity of maize populations. *Plant Molecular Biology Reporter*, Vol. 17, pp. 123-138, ISSN 0735-9640

Dubreuil, P.; Warburton, M.; Chastanet, M.; Hoisington, D. & Charcosset, A. (2006). More on the introduction of temperate maize into Europe: large scale bulk SSR genotyping and new historical elements. *Maydica*, Vol. 51, pp. 281-291, ISSN 00256153

Edmeades, G.; Bolaños, J.; Lafitte, R.; Rajaram, S.; Pfeiffer, W. & Fischer, A. (1989). Traditional approaches to breeding for drought resistance in cereals, In: *Drought resistance in cereals,* F.W.G. Baker, (Ed.), 27–52, ICSU and CABI, ISBN 0 85198 641 2, Wallingford, UK.

El-Nadi, A. (1969). Efficiency of water use by irrigated wheat in Sudan. *Journal of Agricultural Science* (Cambridge). Vol. 73, pp. 2267-2266. ISSN 0021-8596

Eschholz, T. W.; Stamp, P.; Peter, R.; Leipner, J. & Hun d, A. (2010). Genetic structure and history of Swiss maize (Zea mays L. ssp. mays) landraces. *Genetic Resources and Crop Evolution,* Vol.57, pp. 71–84, ISSN 0925-9864

Fischer, K.; Lafitte, R.; Fukai, S.; Atlin, G. & Hardy, B. (2003). Breeding rice for drought prone environments. Los Baños (Philippines): International Rice Research Institute. 98 p. ISBN 971-22-0189-9

Frensch, J. (1997). Primary responses of root and leaf elongation to water deficits in the atmosphere and soil solution. *Journal of Experimental Botany*, Vol. 48, pp. 985–999, ISSN 0022-0957

Galarreta de Ruiz, J.I. & Alvarez, A. (2001). Morphological classification of maize landraces from northern Spain. *Genetic Resources and Crop Evolution,* Vol. 48, pp. 391–400, ISSN 0925-9864

Gauthier, P.; Gouesnard, B.; Dallard, J., Redaelli, R.; Rebourg, C.; Charcosset, A. & Boyat, A. (2002). RFLP diversity and relationships among traditional European maize populations. *Theoretical and Applied Genetics,* Vol.105, pp. 91-99, ISSN 0040-5752

Goodman, M.M. (1994). Maize. In: *Evolution of Crop Plants*, Simmonds N.W. (ed.), ISBN-10: 0470233729, Longman Group, Ltd., Essex, UK,

Grant, F.; Jackson, S.; Kiniry, R. & Arkin, F. (1989). Water deficit timing effects on yield components in maize. *Agronomy Journal,* Vol. 81, pp.61–65. ISSN 0002-1962

Hall, A.; Chimenti, N.;Trapani, F. & Cohen de Hunau, R. (1984). Yield in water-stressed maize genotypes: association with traits measured in seedling and in flowering plants. *Field Crops Research,* Vol. 9, pp. 41-46. ISSN 0378-4290

Hall, A.; Lemcroft, J. & Trapani, N. (1981). Water stress before and during flowering in maize and its effects on yield, its components and their determinants. *Maydica,*Vol. 26, pp.19-38. ISSN 0025-6153

Hallauer, A.R.; Russell, W.A. & Lamkey, K.R. (1988). Corn breeding. In: *Corn and Corn Improvemen,* Sprague G.F. and Dudley J.W. (eds), 463–564, ISBN 0-89054-029-2, Am. Soc. Agron., Inc., Madison, Wisconsin, USA,

Hammer, K.; Diederichsen, A. & Spahillari, M. (1999). Basic studies toward strategies for conservation of plant genetic resources. In: Serwinski J, Feberova´ I (eds) Proceedings of technical meeting on the methodology of the FAO World Information and Early Warning System on Plant Genetic Resources, Prague

Hartings, H; Berardo, N.; Mazzinelli, G.F.; Valoti, P.; Verderio, A. & Motto, M. (2008). Assessment of genetic diversity and relationships among maize(Zea mays L.) Italian landraces by morphological traits and AFLP profiling. *Theoretical and Applied Genetics,* Vol. 117, pp. 831–842, ISSN 0040-5752

Has, V.; Has, I.; Pamfil, D.& Copandean A. (2010). Characterization of „turda" maize germplasm for the chemical composition of the grain. *Romanian Agricultural Research*, 27, pp.59-67, ISSN 1222 - 4227

Ignjatović-Micić, D.; Mladenović Drinić, S.; Nikolić, A. & Lazić-Jančić, V. (2007). Comparison of AFLP and SSR markers for genetic diversity studies in maize populations. *Maydica,* Vol. 52, pp. 399, Vol. -406, ISSN 00256153

Ignjatovic-Micic, D.; Mladenovic Drinic, S.; Nikolic, A. & Lazic-Jancic, V. (2008). SSR analysis for genetic structure and diversity determination of maize local populations from former Yugoslavia territories. *Russian Journal of Genetics,* Vol. 44, No. 11, pp. 1317–1324, ISSN 1022-7954

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. Bull. Soc. Vaud. Nat. 44: 223-270.

Jarne, P. & Lagoda, P.J.L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*, Vol. 11, pp. 424-429, ISSN 0169-5347

Karp, A.; Edwards, K.; Bruford, M.; Vosman, B.; Morgante, M.; Seberg, O.; Kremer, A.; Boursot, P.; Arctander, P.; Tautz, D. & Hewitt, G. (1997). Newer molecular technologies for biodiversity evaluation: opportunities and challenges. *Nature Biotechnology*, Vol.15, pp. 625-628, ISSN 1087-0156

Kaul,M.L.H. (1988). *Male Sterility in Higher Plants.* Springer-Verlag Berlin, Germany.ISBN 0387179526

Kravić, N.; Vuletić, M. & Anđelković, V. (2010). Maize roots responses to osmotic stress. *Journal of Scientific Agricultural Research*, Vol.71 (255), pp. 57-65. ISSN 0354-5695

Latta, M. & Eskin M. (1980). A simple and rapid colorimetric method for phytine determination. *J.Agric. Food Chem.28*, pp.308-1311, ISSN 0021-8561.

Liu, Z.; Peter, S.O.; Long, M.; Weingartner, U.; Stamp, P. & Kaeser, O. (2002). A PCR assay for rapid discrimination of sterile cytoplasm types in maize, *Crop Science*, Vol. 42, pp. 566–569, ISSN 1835-2693

Lubberstedt, T.; Melchinger, A.E.; Dussle, C.; Vuylsteke, M. & Kuiper, M. (2000). Relationship among early European maize inbreds. IV.Genetic diversity revealed with AFLP markers and comparison with RFLP, RAPD, and pedigree data. *Crop Science*, Vol. 40, pp. 783–791, ISSN 1835-2693

Lucchin, M.; Barcaccia, G. & Parrini, P. (2003). Characterization of a flint maize (*Zea mays* L. convar. *mays*) Italian landrace: I. Morpho-phenological and agronomic traits. *Genetic Resources and Crop Evolution,* Vol. 50, pp. 315–327, ISSN 0925-9864

Maxted, N.; Ford-Lloyd, B. & Hawkes, J.G. (1997). Complementary conservation strategies. In: *Plant Genetic Conservation: The in situ Approach*, Maxted N., Ford-Lloyd B. and Hawkes J.G. (eds), 15–40, ISBN-10 0412637308, Chapman & Hall, London.

Mian, M.A.; Hopkins, A.A. & Zwonitzer, J.C. (2002). Determination of genetic diversity in tall fescue with AFLP markers. *Crop Science*, Vol. 42, pp. 944-950, ISSN 1835-2693

Mittelmann, A.; Filho, J; Nass,L; Lima,G.; Hara-Klein, C & da Silva R. (2011).Quantitative variation for grain quality in brazilian maize populations. *Sci. Agric. (Piracicaba, Braz*.), Vol.68, n.1, pp.50-56, ISSN 0103-9016

Monneveux, P.; Sanchez, C. & Tiessen, A. (2008). Future progress in drought tolerance in maize needs new secondary traits and combinations. *Journal of Agricultural Science*, Vol. 146, pp. 287–300. ISSN 0021-8596

Morgan, J. (1980). Osmotic adjustment in the spikelets and leaves of wheat. *Journal of Experimental Botany,* Vol. 31, pp. 655-665. ISSN 0022-0957

Moussa, H. & Abdel-Aziz, S. (2008). Comparative response of drought tolerante and drought sensitive maize genotypes to water stress. *Australian journal of Crop Science,* Vol.1, pp.31-36. ISSN 1835-2693

Nakajima, Y.; Yamamoto, T.; Muranaka, T. & Oeda, K. (1999). Genetic variation of petalord male-sterile cytoplasm of carrots revealed by sequence-tagged sites (STSs). *Theoretical and Applied Genetics,* Vol. 99, pp. 837-843, ISSN 0040-5752

Ortiz, R.; Sevilla, R.; Alvarado, G. & Crossa, J. (2008). Numerical classification of related Peruvian highland maize races using internal ear traits. *Genetic Resources and Crop Evolution*, Vol. 55, pp.1055–1064, ISSN 0925-9864

Pavlicic, J. & Trifunovic, V. (1966). A study of some important ecologic corn types grown in Yugoslavia and their classification. *Journal of Scientific Agricultural Research* 19, pp.44-62, ISSN 0354-5695

Pejić,̧ I.; Ajmone-Marsan, P.; Morgante, M.; Kozumplick,V.; Castiglioni, P.; Taramino, G. & Motto, M. (1998). Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. *Theoretical and Applied Genetics,* Vol. 97, pp. 1248-1255, ISSN 0040-5752

Pollak, L.M.& Scott, M.P. (2005). Breeding for grain quality traits. *Maydica*, 50, pp. 247-257, ISSN 0025-6153

Qualset, C.O.; Damania, A.B.; Zanatta, A.C.A. & Brush, S.B. (1997). Locally based crop plant conservation. In: *Plant Genetic Conservation: The in situ Approach,* Maxted N., Fordof Lloyd B. and Hawkes J.G. (eds), ISBN 0412637308,Chapman & Hall, London, pp. 160–175.

Raboy, V. (2009). Approaches and challenges to engineering seed phytate and total phosphorus. *Plant Sci.*, Vol. 177, pp. 281–296, ISSN 0168-9452

Radovic, G.; Muminovic, J.&Jelovac D. (2000). Local maize germplasm-potentially valuable breeding material. *Genetika*, Vol.32, no.3, pp.221-234, ISSN 0534-0012.

Rebourg, C.; Gouesnard, B. & Charcosset, A. (2001). Large scale molecular analysis of traditional European maize populations. Relationships with morphological variation. *Heredity,* Vol.86, pp. 574-587, ISSN 0018-067X

Rebourg, C.; Gouesnard, B.; Welcker, C.; Dubreuil, P.; Chastanet, M. &, Charcosset, A. (2003) Maize introduction into Europe: the history reviewed in the light of molecular data. *Theor Appl Genet,* Vol. 106, pp.895–903, ISSN 0040-5752

Reif, J.C.; Hamrit, S.; Heckenberger, M.; Shipprack, W.; Maurer, H.P.; Bohn, M. & Melchinger, A.E. (2005). Genetic structure and diversity of European flint maize populations determined with SSR analysis of individuals and bulks. *Theoretical and Applied Genetics,* Vol. 111, pp. 906-913, ISSN 0040-5752

Reynolds, T. L.; Nemeth, M. A.; Glenn, K. C.; Ridley,W.P.& Astwood, J. D. (2005). Natural variability inmetabolites in maize grain: Difference due to genetic background. *J. Agric. Food Chem.*, Vol. 53, pp.10061-10067, ISSN 0021-8561

Ribaut, J. (2006). *Drought adaptation in cereals.* Food Products Press, ISBN-13: 978-1-56022-27-4, New York, USA.

Ribaut, J.; Hoisington, D.; Deutsch, J.; Jiang, C.&Gonzales de Leon, D. (1996). Identification of quantitative trait loci under drought conditions in tropical maize. 1. Flowering parameters and the anthesis-silking interval. *Theoretical and Applied Genetics,* Vol. 92, pp. 905-914. ISSN 0040-5752

Richards,R.(1996). Defining selection criteria to improve yield under drought. *Plant Growth Regulatio*n, Vol. 20, pp. 57–166. ISSN 0167-6903

Rogers, J.S. (1972). Measures of similarities and genetic distances.pp. 145-153. In: M.R. Wheeler (Ed.), *Studies in Genetics*. VII. vol. 7213, Univ. of Texas Publications.

Ruta, N.; Liedgens, M.; Fracheboud, Y.; Stamp, P. & Hund A. (2010). QTLs for the elongation of axile and lateral roots of maize in response to low water potential. *Theoretical and Applied Genetics,* Vol. 120, pp. 621-631. ISSN 0040-5752

Saghai_Maroof, M.; Soliman, K.; Jorgensen, R. & Allard, R. (1984). Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proceedings of the National Academy of Sciences of the USA*, Vol. 91, pp. 8014–8018, ISSN 0027-8424

Sato, Y. (1998). PCR amplification of CMS-specific mitochondrial nucleotide sequences to identify cytoplasmic genotypes of onion (*Allium cepa* L.). *Theoretical and Applied Genetics,* Vol.96, pp. 367-370, ISSN 0040-5752

Schnable, P.S. & Wise, R.P. (1998). The Molecular Basis of cytoplasmic male sterility and fertility restoration. *Trends in Plant Science,* Vol. 3, pp. 175–180, ISSN 1360-1385

Sharp, R. & Davies, W. (1979). Solute regulation and growth by roots and shoots in water tressed maize plants. *Planta*, Vol. 147, pp. 43–49. ISSN 0032-0935

Sharp, R. & Davies, W. (1989). Regulation of growth and development of plants with a restricted supply of water. In: *Plants under stress*. H.G. Jone et al (Ed.), 71-93, Cambridge Univ. Press, ISBN 0521344239, Cambridge, England, UK.

Singh, S.P.; Gutierrez, J.J.A.; Molina, A.; Urrea, C. & Gepts, P. (1991). Genetic diversity in cultivated common bean. II. Marker-based analysis of morphological and agronomic traits. *Crop Science,* Vol. 31, pp. 23–29, ISSN 1835-2693

Smith, J.S.C.; Chin, E.C.L;. Shu, H.; Smith, O.S.; Wall, S.J.; Senior, M.L.; Mitchell, S.E.; Kresovich, S. & Ziegle, J. (1997). An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): comparisons with data from RFLPs and pedigree. *Theoretical and Applied Genetics,* Vol. 95, pp. 163-173, ISSN 0040-5752

Sredojevic, S. & Dragicevic V. (2009). The quantitative determination of antioxidative-protective and some storage substances of seeds and leaves-modification of methodological approach. *Journal of Scientific Agricultural Research*, 70, 249, pp. 13-20, ISSN 0354-5695.

Suharjo, U. (2007). Use of Polyethylene Glycol (PEG) 8000 for Rapid Screening of Potato (*Solanum tuberosum* L.) Genotypes for Water stress Tolerance: III: Root and shoot Growth. *Journal Akta Agrosia Edisi Khusas*, Vol. 1, pp. 11-18, ISSN 1410-3354.

Sweeny, P.M. & Danneberger, T.K. (1997). RAPD markers from perennial ryegrass DNA extracted from seeds. *HortScience,* Vol. 32, pp. 1212-1215, ISSN 0018-5345

van der Weele, C.; Spollen, W.; Sharp, R. & Baskin, T. (2000). Growth of *Arabidopsis thaliana* seedlings under water deficit studied by control of water potential in nutrient-agar media. *Journal of Experimental Botany*, Vol. 51, pp.1555-1562, ISSN 0022-0957

van Etten, J.; Lopez M.R.F.; Monterroso, L.G.M. & Samayoa, K.M.P.(2008).Genetic diversity of maize (Zea mays L. ssp. mays) in communities of the western highlands of Guatemala:geographical patterns and processes. *Genetic Resources and Crop Evolution,* Vol. 55, pp. 303–317, ISSN 0925-9864

Vancetovic, J.; Andjelkovic, V. & Shaboon, S. (2008). The identification of drought tolerant sources from the MRI gene bank. Modern variety breeding for present and future needs.In: *Proceedings of the 18th EUCARPIA General Congress*, 9-12.September 2008.Valencia-Spain, pp. 261.

Vancetovic, J.; Vidakovic, M.; Ignjatovic-Micic, D.; Nikolic, A.; Markovic, K. & Andjelkovic V. (2010). The structure of sterile cytoplasm types within a maize genebank collection. *Russian Journal of Genetics*, Vol.46, No.7, pp. 836–840, ISSN 1022-7954

Verslues, P.; Agarwal, M.; Katiyar-Agarwal, S.; Zhu, J. & Zhu, J. (2006). Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *Plant Journal,* Vol. 45, pp. 523–539, ISSN 0960-7412

Weerathaworn, P.; Soldati, A. & Stamp P. (1992b). Anatomy of seedling roots of tropical maize (*Zea mays* L.) cultivars at low water supply. *Journal of Experimental Botany,* Vol. 43, pp. 1015–1021, ISSN 0022-0957

Weerathaworn, P.; Soldati, A. & Stamp, P. (1992a). Shoot and root growth of tropical maize seedlings at different moisture regimes. *Maydica*, Vol. 37, pp.209-215, ISSN 0025-6153

Westgate, M. & Boyer, J.(1985). Osmotic adjustment and the inhibition of leaf, root, stem and silk growth at low water potentials in maize. *Planta*, Vol. 164, pp. 540-549, ISSN 0032-0935

Yancey, P.; Clark, M.; Hand,S.; Bowlus,R. & Somero, G. (1982). Living with water stress: evolution of osmolyte systems. *Science*, Vol. 217, pp. 1214-1222, ISSN 0036-8075

Yu, K. & Pauls, K.P. (1993). Rapid estimation of genetic relatedness among heterogeneous populations of alfalfa by random amplification of genomic DNA samples. *Theoretical and Applied Genetics,* Vol. 86, pp. 788-794, ISSN 0040-5752

Zabala, G.; Gabay-Laughnan, S. & Laughnan, J.R. (1997). The nuclear gene *Rf3* affects the expression of the mitochondrial chimeric sequence R in S-type male sterility in maize. *Genetics*, Vol. 147, pp. 847–860, ISSN 0016-6731

Zhu, K. (2002). Salt and drought stress signal transduction in plants. *Annual Review in Plant Biology,* Vol.53, pp.247-273, ISSN 1543-5008

Genetic Biodiversity and Phylogenetic Studies in Poplar by Means of the Metallothionein Multigene Family

Angela Cicatelli[1], Renato Lupi[2], Francesco Sala[2],
Han Yfan[3] and Stefano Castiglione[1]
[1]Dipartimento di Chimica e Biologia, Università di Salerno, Fisciano, Salerno
[2]Dipartimento di Biologia, Università di Milano, Milano
[3]Research Institute of Forestry
Chinese Academy of Forestry, Wan Shou Shan, Beijing
[1,2]Italia
[3]People's Republic of China

1. Introduction

Biodiversity can be defined as the totality of living forms present on earth, with the wealth of genetic information that they have (Grassi et al., 2006). It can be viewed at various levels, from ecosystems, down to species, populations, genomes and genes. Populations are representative of the species and may contain a large part of their genetic biodiversity. They are the basis of allopatric and sympatric speciation phenomena, due to a reduction of gene flow, caused by the presence of new geographical barriers, the former, and to auto- or allopolyploidy events, the latter.

During the last two decades many molecular tools based on PCR (Polymerase Chain Reaction) have been developed and used to calculate and evaluate, through bioinformatics, genetic biodiversity of all living organisms and of plants in particular (Lowe et al., 2004). These molecular tools comprise DNA sequencing of single genes, or of entire genomes (*e.g.* rice, poplar, wheat, grape, etc.), locus specific (SSRs and SNPs) and random genome analysis (AFLPs, RAPDs, etc.).

Poplar, after man and rice, was one of the first organisms with a fully sequenced plant genome (Tuskan et al., 2006). Moreover it has also been a subject study of many plant biologists and physiologists due to its commercial and economic importance, and to its relevance for riparian ecosystems. Many populations and germplasm collections of different poplar species have been investigated (Brundu et al., 2008; Castiglione et al., 1993; Castiglione et al., 2010; Fossati et al., 2003) using random and locus specific molecular analysis to evaluate their genetic biodiversity. Poplar is a perennial flowering plant, native to most of the northern hemisphere, known to play valuable ecological roles: in fact, it is considered a trustworthy health indicator of riparian ecosystems and a promising phytoremediation tree of polluted soils (Castiglione et al., 2009; Sebastiani et al., 2004; Yadav

et al., 2010). Poplars, aspens and cottonwoods are part of the genus *Populus* that, along with the two other related genera *Salix* and *Chosenia*, belongs to the Salicacae family. It is now generally accepted that the genus *Populus* comprises 29 species, identified on the basis of diagnostic morphological characters and assigned to six different sections, namely Abaso, Aigeros, Leucoides, Populus, Tacamahaca and Turanga (Dickmann, 2001). However, correct identification and classification of poplar species are made difficult by: (a) the very high variability of their morphological traits; (b) their strong tendency towards intra-section hybridization, especially in sympatric ranges, resulting in hybrids with mixed and broad-spectrum parental features; (c) the unusual possibility to inter-section hybridization (*e.g.,* between specimens belonging to the Aigeros and Tacamahaca sections - Dickmann, 2001). Moreover, it should be noticed that fertile F1 hybrids could backcross towards one/both parental species, or produce further generations, thus making species and hybrid identification even more problematic (Fossati et al., 2004; Lexer et al., 2005). Recently, several studies using molecular markers have been carried out in order to revise the assignment of poplar species to the above-listed sections and even to resolve their controversial taxonomy (Cervera et al., 2005; Hamzeh & Dayanandan, 2004). Indeed, due to discrepancies among phylogenetic trees reconstructed with nuclear ribosomal DNA (rDNA) or chloroplast genes, it was established that nuclear genome (nDNA) of *P. nigra* (Aigeros) is divergent from species belonging to Populus section, although its chloroplast DNA (cpDNA) is strongly associated to that of trembling poplars (Populus) (Hamzeh & Dayanandan, 2004). Thus, *P. nigra* should not be considered derived from the inter-section crossing between *P. deltoides* x *P. alba* (Smith, 1988), but might be classified in a brand-new and independent taxon (*i.e.*, Nigrae). Further studies, using Amplified Fragment Length Polimorfisms (AFLP), revealed that *P. mexicana*, the single species present in the Abaso section, is so divergent from all the other poplar sections that might be considered part of the *Salix* genus (Cervera et al., 2005). Moreover, AFLP data also showed that species belonging to the Populus section are characterized by the greatest levels of interspecific genetic variability, whereas the opposite trend has been observed for the sections Aigeros and Tacamahaca. Given that ancestral species have greater variation with respect to the derived ones, Populus is believed to be the oldest section, whilst Aigeros and Tacamahaca the most recent ones (Cervera et al., 2005; Eckenwalder, 1996). However, controversial conclusions have been drawn on the taxonomy of the Populus section using different molecular markers. On one hand, the analysis of rDNA sequences indicates that Populus section could be further sub-divided into two major clades: one comprising *P. alba*, *P. tremula* and *P. davidiana* (Korean aspen); the second one grouping together the American aspens *P. tremuloides* and P. *grandidentata* (Hamzeh & Dayanandan, 2004). On the other hand, AFLP data suggested that *P. tremula* and *P. tremuloides* are ecotypes still capable of hybridization and thus part of the same group, whilst *P. alba* represents a totally different taxon (Cervera et al., 2005). In addition, Tacamahaca section was surprisingly found to be a polyphyletic group, *i.e.*, a kind of meta-cluster comprising many small sub-groups, each containing closely related species (Hamzeh et al., 2004). Finally, American cottonwoods *P. trichocarpa* and *P. balsamifera*, which have been ever considered subspecies of a single taxon, were discriminated as different species when assessed by AFLP analysis (Cervera et al., 2005). Such inconclusive and contrasting results indicate that both rDNA and AFLP markers are not suitable to correctly classify poplar species and fully resolve their phylogenetic relationships. Indeed, it should be stressed that AFLP analysis is intended to get insight into the overall structure of genomes, thus being more properly used to determine interspecific

genetic variability. Moreover, the rDNA gene family is known to be affected by concerted evolution and gene conversion events, which may alter the phylogenetic signals carried by the analyzed sequences. Despite these limitations, molecular evolution of plants is mainly inspected using rDNA and/or cpDNA alone, which possibly lead to evident mistakes in phylogeny reconstruction due to biases present at the molecular level (Alvarez & Wendel, 2003; Doyle & Gaut, 2000; Small et al., 2004). Undoubtedly, the drawing of accurate and resolved phylogenetic trees may be supported using additional molecular markers, such as nuclear –and thus mendelian-inherited– single-copy genes and/or multigene families. In particular, multigene families are very attractive since they offer the possibility to sample independent loci (*i.e.*, not associated in linkage units) that share a deeply related evolutionary framework. This is a crucial feature as far as hybridization and introgression events are concerned: thus, the use of such nuclear genes should be considered for reconstructing the phylogeny of poplar and also for assessing the genetic variability within the studied genus. Actually, several analyses of plant gene families, with particular emphasis on *P. trichocarpa*, the poplar model system, have been carried out in the recent past in order to understand their evolutionary history, the functional diversification of their members and their detailed expression in poplar (Lan et al., 2009; Petre et al., 2011). At present, however, little information is available about phylogenetic relationships among poplar species by means of nuclear genes analysis (Fladung & Buschbom, 2009). Hence, in order to shed light on the unresolved problematic issues, we have undertaken a deep investigation of the nuclear-encoded metallothionein (MT), a gene family involved in response to plant stress. In particular, MTs are known to have protective roles against plant wounds of pathogenic origins and senescence (Kohler et al., 2004), and to be involved in metal homeostasis (Cobbett & Goldsbrough, 2002), being upregulated in foliar tissue, together with polyamines, to enhance poplar tolerance towards heavy metal-contaminated soils (Castiglione et al., 2007; Cicatelli et al., 2010; 2011), or in water deficit (Berta et al., 2009). Nonetheless, no comprehensive study on MT gene family has been performed so far, even if it shows interesting and potentially attractive features. In fact MT multigene family is characterized by: (a) a quite high but manageable number of members (six), at least in *P. trichocarpa* (personal observation); (b) a peculiar organization of specific cysteine residues, whose function is cooperatively binding metal ions (Kille et al., 1991), thus allowing easy arrangement of orthologous MTs into three types, namely MT1, MT2 and MT3 (Cherian & Chan, 1993; Cobbett & Goldsbrough, 2002); (c) presence of a full-exon coding for a specific region, called "spacer", which show remarkable levels of genetic variability (Buchanan-Wollaston, 1994; Zhou & Goldsbrough, 1994). Indeed these spacers could be rich of single-nucleotide polymorphism (SNP). SNP including single-base changes or indels (insertion or deletion), at specific nucleotide positions, has been shown to be the most abundant class of DNA polymorphism in many organisms (Brookes, 1999; Cho et al., 1999). SNP variation analysis and SNP marker development from candidate genes could provide valuable information regarding their evolution and effects on complex traits. Very recently, Fladung and Buschbom (2009) have used partial DNA sequence of six nuclear genes to draw phyologenetic relationships among a limited number of poplar species using SNPs. This study substantiate the importance of this class of molecular markers as a very promising molecular tool to study the evolution or measure the genetic variability of plants, in general, and within *Populus*, in particular. The present study is part of a wider and ample collaborative project between Italy and the People's Republic of China focused on poplar.

For this reason, particular interest was devoted to the analysis of Chinese poplar species belonging to the Populus section (Castiglione et al., 2010; Lexer et al., 2010). In the present study MT genes have been used to estimate phylogenetic relationships among 11 *Populus* ssp belonging to three poplar sections (Aigeros, Populus and Tacamahaca). Furthermore a natural population of 63 *P. alba* trees, collected along the banks of the river Sele (South Italy), was analysed by means of nuclear (10) and (3) chloroplast SSRs. The population genetic variability was evaluated using indices commonly employed to estimate genetic biodiversity (*e.g. h*, *Na*, *Ne*, *Ho*, *He* etc.) or genetic dissimilarity (Jaccard index, PCA). On the basis of the estimated genetic biodiversity, a subset of the population (5 specimens), characterized by very high genetic dissimilarity based on Jaccard index, underwent to DNA sequence analysis in order to identify new SNPs in MT genes and an exemplificative analysis is given for the isoforms of MT1a and MT3a. Based on our results, we can state that MTs are promising markers to shed light on both intra- and inter-section relationships among poplars, as well as to assess genetic variability within selected natural populations.

2. Material and methods, results and discussion

2.1 Material and methods

Poplar samples. Eleven different species, representing three distinct sections (Tacamahaca, Aigeros and Populus) of the genus *Populus*, were selected (Tab. 1) for phylogenetic study. Nine species were sampled for DNA sequencing, among those three were Italian (*P. nigra*, *P. alba* and *P. tremula*) and six Chinese (*P. pyramidalis*, *P. adenopoda*, *P. davidiana*, *P. serrata*, *P. bonatii* and *P. tomentosa*). For each species, fresh young leaves of 2-3 individuals were sampled and dried in *Silica Blue Gel* (Sigma-Aldrich Italia – Milano, Italy) prior to DNA extraction. The sequences of the remaining 2 balsam poplars, *i.e. P. trichocarpa* and *P. balsamifera*, were retrieved from publicly available databases (GenEMBL). It should be stressed that the selected Chinese poplars have never been characterized from a molecular point of view, and that their classification within the Populus section is troublesome due to peculiar morphological features (see the description given in the Chinese Floras Atlas at: http://www.efloras.org/flora_page.aspx?flora_id=2). To study the genetic biodiversity of a *P. alba* natural population, 63 individuals were sampled along the banks of the river Sele (Salerno - South Italy) in the springs of 2008 and 2009. Young leaves were collected from single individuals and then stored in absolute ethanol until use. Genetic biodiversity of white poplar natural population of the Sele river was compared with eight reference samples: two clones collected in Sardinia (b33SS, b4SS); further two clones part of a collection of Northern Italy clones and analysed during a previous research project focused on phytoremediation [AL22, AL35 - (Castiglione et al., 2009)], and four hybrids (*P.* x *canescens*) collected along the banks of the Ticino river in Northern Italy (# 1, # 13 , # 15, # 2 - Castiglione et al., 2010).

Abbreviations used for poplar species. *P. nigra*: *Populus nigra* L.; *P. adenopoda*: *Populus adenopoda* Maximowicz; *P. alba*: *Populus alba* L.; *P. pyramidalis*: *Populus alba* var. *pyramidalis* Bunge; *P. davidiana*: *Populus davidiana* Dode; *P. serrata*: *Populus ningshanica* C. Wang & S. L. Tung:; *P. bonatii*: *Populus rotundifolia* var. *bonatii* (H. Léveillé) C. Wang & SL Tung; *P. tomentosa*: *Populus tomentosa* Carr.; *P. tremula*: *Populus tremula* L.; *P. balsamifera*: *Populus balsamifera* L.; *P. trichocarpa*: *Populus balsamifera* L. *ssp. trichocarpa* (Torr. & Gray ex Hook.) Brayshaw.

DNA extractions. Total genomic DNA was extracted from dried poplar leaf (species reported in Tab. 1) using DNeasy Plant Mini kit (Qiagen; Milano, Italy), or "REDExtract-N-Amp Plant PCR Kit", (Sigma-Aldrich Italia, Milano, Italy) following the supplier instructions.

Species	Section	Source (geographic origin)
P. nigra L.	*Aigeros*	This study (Italy)
P. adenopoda Maximowicz	*Populus* (T)	This study (China)
P. alba L.	*Populus* (A)	This study (Italy)
P. alba var *pyramidalis* Bunge	*Populus* (A)	This study (China)
P. davidiana Dode	*Populus* (T)	This study (China)
Populus ningshanica C. Wang & S. L. Tung	*Populus*	This study (China)
P. rotundifolia var. *bonatii* (H. Léveillé) C. Wang & SL Tung	*Populus*	This study (China)
P. tomentosa Carr.	*Populus* (A)	This study (China)
P. tremula L.	*Populus* (T)	This study (Italy)
P. balsamifera L.	*Tacamahaca*	Trace Archive (http://www.ncbi.nlm.nih.gov/Traces/home/)
P. balsamifera L. ssp. *trichocarpa* (Torr. & Gray ex Hook.) Brayshaw	*Tacamahaca*	Poplar Genome Browser (http://www.plantgdb.org/PtGDB)

Table 1. The 11 analyzed poplar species. (T) or (A) refers to the sub-sections *Trepidae* and *Albidae*, respectively.

Identification of MT genes in poplar. The first step for our analysis consisted in the identification of genomic position of the metallothionein (MT) genes of *P. trichocarpa.* This was done through BLAST searches against the *P. trichocarpa* genome v2.0 (available at http://www.plantgdb.org/PtGDB and http://plants.ensembl.org/Populus_trichocarpa) using as queries the mRNAs of *P.* x *generosa* (*P. trichocarpa* x *P. deltoides*) retrieved from GenEMBL (MT1a: AY594295, MT1b: AY594296, MT2a: AY594297, MT2b: AY594298, MT3a: AY594299, MT3b: AY594300). Based on the identified sequences, we designed specific primers (available on request) for MT genes PCR amplification in the remaining poplar species. For each MT-gene, two primer pairs were designed, so that partially overlapping amplicons, suitable for assembling and analysing DNA sequences, were obtained. As only exception, MT2a was sequenced using only one primer pair due to its unusual exon-intron

structure (see the *Results* section). The primer pairs, designed on the basis of *P. trichocarpa* genome, were successfully used even for the amplification and sequencing of MT1b, MT3a and MT3b genes from a specimen of *Salix matsudana* L. growing at the Orto Botanico Cascina Rosa (Milano-IT). The remaining MT genes of the same *Salix* species were retrieved from GenEMBL (MT1a: EF157299, MT2a: EF157297, MT2b: EF157298).

PCR amplification and sequencing. MT genes were amplified in PCR reactions contained 200 µM dNTP, 1.5 mM MgCl$_2$, 2.5 µM of each primer, 1.5 U of *PolyTaq - Recombinant* 5 U/µl (PolyMed, Sambuca Val di Pesa, Italy), and 5 µl of Poly-Taq 10X Buffer in a total volume of 50 µL. The PCR thermal profile was as following : 94°C for 60 s, TA°C for 60 s TA= Annealing Temperature for each single primer pairs, available on request) , and 72°C for 90 s for 35 cycles. After purification and estimation of DNA quantity by agarose-electrophoresis, single-band amplified DNAs were sequenced using ABI Big Dye Terminator version 3.1 Cycle Sequencing Ready Reaction Kit (Applied Biosystem, Monza, Italy). PCR sequencing products were electrophoresed on an ABI 310 automated genetic analyzer (Applied Biosystem). Some of MT amplified products, showing double bands of different molecular weight, were cloned, according to the procedure provided by Clone JET PCR Cloning kit (Fermentas, Burlington, Ontario), and sequenced as above described. Each amplicon was sequenced at least two times, using the same PCR primers, and chromatograms were processed by careful visual inspection. In order to assemble MT-genes, the amplicons of each gene belonging to the same species were aligned using the ClustalW multiple sequence alignment software (Thompson et al., 1994). Aligned DNA sequences were verified by manual editing of the sequence alignments.

Gene annotation and analysis of sequence variability. The exon-intron structure and the protein-coding sequence (CDS) of the amplified genes were identified using the GenomeScan tool (http://genes.mit.edu/genomescan.html), which also allowed us to verify the presence of the canonical splicing sites. The CDS of MT-genes were aligned using the RevTrans program (http://www.cbs.dtu.dk/services/RevTrans), which takes into account the codon structure of the analyzed sequences. The six resulting alignments, which support the exon-intron structure identified with GenomeScan, were manually curated in order to remove the primer sequences. Variability analyses were carried out on a final dataset including 12 taxa (*i.e.*, 11 poplar species plus the outgroup *S. matsudana*) for each of the six MT-genes. Moreover, the amino acid sequences inferred from the translation of the MT-genes were aligned using MUSCLE program (http://www.ebi.ac.uk/Tools/msa/muscle/#). For each data set, the proper variability analysis was performed using the MEGA4 software (Tamura et al., 2007), *i.e.* we calculated the number of synonymous and non-synonymous substitutions for the nucleotide alignment, and number of amino acid invariant and variant sites for the protein alignment.

Phylogenetic data analysis. Phylogenetic analyses were performed on each of the six nucleotide dataset and on the concatenation of all MT-genes in a so-called "supergene" (1,224 sites with no gapped position). Only the last dataset proved to be enough informative to resolve the relationships among poplar species. As for inferred translation of MT-genes, due to both the very low number of sites and variability levels detected among the different poplar species (see *Results*), we decided to exclude protein sequences from the phylogenetic analysis. With respect to the nucleotide "supergene" dataset, model selection was performed using the on-line version of Modeltest software,

Genetic Biodiversity and Phylogenetic Studies in
Poplar by Means of the Metallothionein Multigene Family
119

(http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html), which showed that the most suitable model was GTR+G (Lanave et al., 1984; Saccone et al., 1990). The parameters of the model were: base frequencies: A 0.28, C 0.23, G 0.31, T 0.31; substitution rate matrix: (A–C) 1.672 (A–G) 0.312 (A-T) 0.275 (C–G) 0.710; (C–T) 1.251; (G–T) 1.0; gamma distribution parameter (alpha): 0.041. Phylogenetic analyses and bootstrapping were carried out using programs of the PHYLIP package (Felsenstein, 1993) and PHYML software (Guindon & Gascuel, 2003). In particular, phylogenetic analysis was performed using two different Maximum Likelihood (ML) procedures: a "Classical ML" and a "Hybrid ML" method. The Classical ML analysis was directly carried out by means of the PHYML software. On the other hand, the Hybrid ML method consisted of two steps: (a), calculation of ML distances by TreePuzzle software (Smith, 1988); (b) calculation of phylogenetic trees via the Neighbor Joining (NJ) method. NJ was performed using the NEIGHBOR software, one of the several phylogenetic tools belonging to the PHYLIP package. In the case of the Classical ML method, bootstrapping was implemented as follows: (a), generation of 1,000 replicates of the original data set, using SEQBOOT software; (b), bootstrap test by means of PHYML, activating "multiple data sets analysis" option. For the Hybrid ML method, bootstrapping was implemented as follows: (a) generation of 1,000 replicates of the original data set using SEQBOOT software (PHYLIP package); (b) bootstrap test, using the UNIX script "puzzleboot" (http://rogerlab.biochemistryandmolecularbiology.dal.ca/puzzleboot.php). Phylogenetic trees for each of the obtained distance matrix were calculated using NIGHBOR and CONSENSE software (PHYLIP package). Finally, the calculated phylogenetic trees using either the Classical or the Hybrid ML methods were visualized and manipulated using TreeView (Page, 1996) and TreeMe (http://en.bio-soft.net/tree/TreeMe.html).

SSR genotyping of a *P. alba* natural population. Molecular analysis on the natural white poplar population growing on the banks of the Sele river was performed by means of Simple Sequence Repeats (SSR). SSR assays was performed as described by Yin et al. (2004) and Van der Schoot et al. (2000) using ten nuclear (ORPM_30a, ORPM_30b, ORPM_312a, ORPM312b, ORPM_60, PMGC_2852, WPMS5, WPMS14, WPMS18, WPMS20) and three chloroplast DNA primers [CCMP2, CCMP6, CCMP10 – (Weising & Gardner, 1999)] These were chosen among those characterized by high polymorphism, absence of multiple bands and high discrimination power. PCR reactions were performed using a traditional two-primer approach, being one of the two primer fluorescent-labelled. Reactions were performed in 10 µl total volume containing: 2 µl of template DNA, 2 µl of reverse primer (1 µM, unlabeled), 2 µl of forward (1 µM, a third of which was labelled), 4 µl of RedExtract–N–Amp PCR Ready Mix (Sigma-Aldrich). For WPMS5 and WPMS14 primers, cycling conditions were as follows: initial denaturation at 94°C for 3 min; followed by 35 cycles of 1 min and 15 s at 94°C, 1 min and 15 s at annealing T (annealing T was 50°C for WPMS5 and 60°C for WPMS14), 1 min and 45 s at 72°C; and a final extension time of 10 min at 72°C. ORPM_30 region was amplified using the following thermal profile: 94°C for 3 min; followed by 30 cycles of 1 min at 94°C, 1 min at 55°C, 1 min at 72°C, and a final extension time of 10 min at 72°C. Remaining PCR amplifications were performed using: initial denaturation at 94°C for 3 min; followed by 15 cycles of 30 s at 94°C, 1 min at 62°C (annealing temperature was reduced by 0.5 degree per cycle); followed by 20 cycles (or 30 cycles for ORPM_312) of 30 s at 94°C, 1 min at 52°C, 1 min and 30 s at 72°C; and a final extension time of 7 min at 72°C. The SSR genotypes were analyzed on ABI-PRISM 310 Genetic Analyzer (Applied Biosystems, Monza, Italy), while fragment sizing was carried out

with Gene Mapper version 4.0 (Applied Biosystems) utilizing the internal 500 ROX Size Standard (Applied Biosystems). Population biodiversity was evaluated by number of alleles, frequency, observed and expected heterozygosity; within-population inbreeding coefficient F_{IS} for microsatellite loci was also estimated using the available free software package GenAlex6 (Peakall & Smouse, 2006) freely available at the following web site: http://www.anu.edu.au/BoZo/GenAlEx/new_version.php. Furthermore, some of the nuclear SSR markers were used to determine the genetic relationship of the Sele population compared to some selected specimens belonging to already surveyed *P. alba* populations (Brundu et al., 2008; Castiglione et al., 2009; Castiglione et. al., 2010), as described above. The molecular similarity of the defined four groups of poplars was assessed performing a PCA (Principal Component Analysis) by means of the NTSYS-pc program version 2.1 (Rohlf & Marcus, 1993), using as distance the Jaccard dissimilarity index (Jaccard, 1908). Finally, in the light of previous analyses, in order to maximize the observed genetic variability we picked out five clones from the Sele population, and carried out a molecular analysis of the SNPs (Single Nucleotide Polymorphisms) present in their MT-genes, searching for point mutations and indels in both exons and introns.

2.2 Results

MT Multigene family within the genus *Populus*. MT genes, sequenced or available in public databases for the chosen poplar species (Tab. 1), were analyzed in order to evaluate their sequence variability and, thus, their resolving power as molecular markers for assessing the phylogenetic relationships among cryptic species of the genus *Populus*. The amplification of purified PCR products of the MT CDS, followed by RevTrans alignment of all nucleotide sequences and their arrangement as a supergene, resulted in a data matrix of 1,224 characters (with no gapped positions) for all 11 poplar species plus the outgroup (*S. matsudana*). It should be noticed that the corresponding DNA sequence of two or more specimens of the same taxon were identical, e.g. *P. alba* (2). BLAST searches performed against the Poplar Genome Browser, using as queries mRNA sequence from *P.* x *generosa*, allowed the identification of the number and genomic position of all MT genes in the genome of model system *P. trichocarpa*. As expected, poplar MT genes are encoded by a multigene family encompassing six members, that is two genes for each MT type, as described by Koheler and co-workers (2004). It should be notice that the two members belonging to the same MT type, *e.g.* MT1a and MT1b, are neither found on the same chromosome, nor are organized as cluster like the well-known rDNA genes.

Exon-intron structure of poplar MT genes and intron variability. Poplar MT genes proved to have a well-conserved exon-intron structure: in fact, the two isoforms of the same MT gene show a CDS of equal length as well as the same number of exons/introns. According to Cobbet & Goldsbrough (2002) and Koheler et al. (2004), MT genes are encoded by short genes, composed of small exons and introns. Our analyses confirmed data present in literature, having for MT1, MT2b and MT3 mean length of 585, 473 and 616 bp, respectively. The MT genes code for small polypeptides with mean lengths of 73, 78 and 66 amino acids for MT1, MT2b and MT3, respectively. The only important exception is the MT2a gene, which is 293 bp long and has 2 exons coding for a polypeptide of 79 aa. This exon-intron structure is peculiar to MT2a, and it is neither an annotation nor sequencing artifact. In fact, all the MT2a genes identified by means of BLAST searches, or by sequencing, have the same structure in all the

analyzed poplar species (*i.e.* two exons and one intron). The molecular analysis performed on MT2a showed that it has a different CDS length (240 bp) with respect to MT2b CDS (237 bp) due to the presence of a triplet "GCG" (coding for a valine residue) at positions 148-150 with respect to MT2b. The deletion is located in a very short but variable region of the MT2a spacer domain, which had 12 out of 20 different residues compared to the MT2b homologous region (data not shown). Generally speaking, we could consider the CDS of MT-genes as formed by three exons: (a) the first coding for almost the entire Cys-rich domain; (b) the second coding for the majority of the spacer region; (c) and the third coding for the terminal part of the spacer plus the whole C-terminal domain. A remarkable sequence variability was also detected within the introns of the considered MT gene family. In particular, we found an SSR within intron I of MT3b, which showed a different and peculiar number of tandem repeats in each of the analyzed species. Moreover, we detected several insertion/deletions (indels) specific to different taxonomical rank (*e.g.*, sub-section), or related to the geographical origin of poplars exclusively in the introns of MT genes (data not shown).

Sequence variability in MT-genes and their products. Our analyses, focused on the CDS of MT-genes, showed that poplar MT-genes have a high level of nucleotide sequence identity, *i.e.* 96% (1,170/1,224 sites), as well as a very high level of amino acidic sequence identity, *i.e.* 93% (380/410 sites), see Table 2.

Gene	Nucleotide alignment			Amino acid alignment		
	No. sites	No. invariable sites	Invariable sites (%)	No. sites	No. invariable sites	Invariable sites (%)
MT1a	198	194	98	66	64	97
MT1b	210	203	97	70	68	97
MT2a	225	203	90	75	63	84
MT2b	219	212	97	75	74	99
MT3a	189	183	97	61	56	92
MT3b	183	175	96	63	55	87
Total	**1224**	**1170**	**96**	**410**	**380**	**93**

Table 2. Number (No.) of total and of invariable sites detected in the coding sequences of the MT-genes, based on a dataset of the 11 poplar species listed in Table 1. *Statistics referring to nucleotide sequences are on the left half of the table, whilst those referring to translated sequences are on the right half.*

A detailed variability analysis of metallothionein CDS showed that 29/54 (54%) of the variable sites are non-synonymous (Tab. 3).

Thus, based on the number of non-synonymous (N) and synonymous substitutions (S), we calculate the N/S ratio with respect to the different MT types. Interestingly, most of MT-genes are subject to positive evolutionary pressures, being N/S>1 for all genes with the

exception of MT1a (N/S=1, neutral selection) and MT1b (N/S<1, purifying selection). The N/S ratio, calculated separately for each nucleotide sequences coding for a given MT domain, showed that 33/54 (61%) variable sites are present in the spacer domain, thus indicating that this domain is the most variable region of MT-genes. This observation is also supported by the fact that the spacer domain contains 20/29 (69%) non-synonymous substitutions globally found in all MT-genes. Finally, the N/S ratio showed that the spacer region is subject to positive selection (N/S=1.54), whilst both N- and C-domains are, as expected, under slight purifying pressures (N/S=0.75).

Gene	N° variable sites	Codon position			Substitution type		dN/dS
		I	II	III	N° syn	N° non-syn	
MT1a	4	2		2	2	2	1
MT1b	7	2	1	4	4	3	0.75
MT2a	22	6	7	9	9	13	1.44
MT2b	7		1	6	6	1	2.33
MT3a	6	2	2	2	2	4	2
MT3b	8	3	2	3	2	6	3
Total	**54**	**15**	**13**	**26**	**25**	**29**	**1.16**

Table 3. Substitution types found in the coding sequences of the MT genes belonging to the 11 poplar species listed in Table 1. *Syn: synonymous substitution; non-syn: non-synonymous substitution.*

Phylogenetic analysis of the MT-supergene. As described in *Materials and Methods*, we carried out the phylogenetic analysis following two different procedures: a Classical and a Hybrid ML methods. We used these methods to calculate phylogenetic trees for each MT gene under study, testing the statistical reliability of the obtained trees by means of non-parametric bootstrapping (1,000 repetitions). The phylogenetic trees separately calculated for each of the MT genes were almost completely unresolved (data not shown). Therefore, we decide to concatenate the CDS of all six MT-genes in order to shape a so-called "supergene" encompassing a total of 1,224 characters. Both the Classical and the Hybrid ML methods resulted in comparable phylogenetic trees with respect to either topology and bootstrapping value of their internal nodes. Therefore, we just show the *consensus* phylogenetic tree obtained through the Classic ML method (Fig.1).

In general, the calculated phylogenetic tree is well-resolved since 4/9 (44%) internal nodes are supported by bootstrapping values higher than 75, and 7/9 (78%) are associated to

bootstrapping value higher than 50. In particular, it can be noticed that the six MT-genes arranged as supergene are able to group together the two balsam poplars *P. trichocarpa* and *P. balsamifera* with a good supporting value (65), and correctly distinguish section Tacamahaca from *P. nigra* (Aigeros) with a very high supporting value (92). As for the Populus section, it is reliably identified with a bootstrap value equal to 88. Moreover, poplars belonging to this section are further clustered into two distinct groups: (a) the first one, comprising the only Italian *P. alba* with the full bootstrapping value (100); (b) the second one, encompassing all the remaining poplar species. Cluster (b) is further sub-divided into two smaller and morphologically-coherent groups: (b.1) a sub-group that contains the Chinese white poplars, namely *P. alba* var. *pyramidalis* and *P. tomentosa*; and (b.2), another cluster that corresponds to all trembling under study (*P. tremula*, *P. adenopoda* and *P. davidina*), plus the two Chinese species whose classification was uncertain (*P. serrata* and *P. bonatii*).
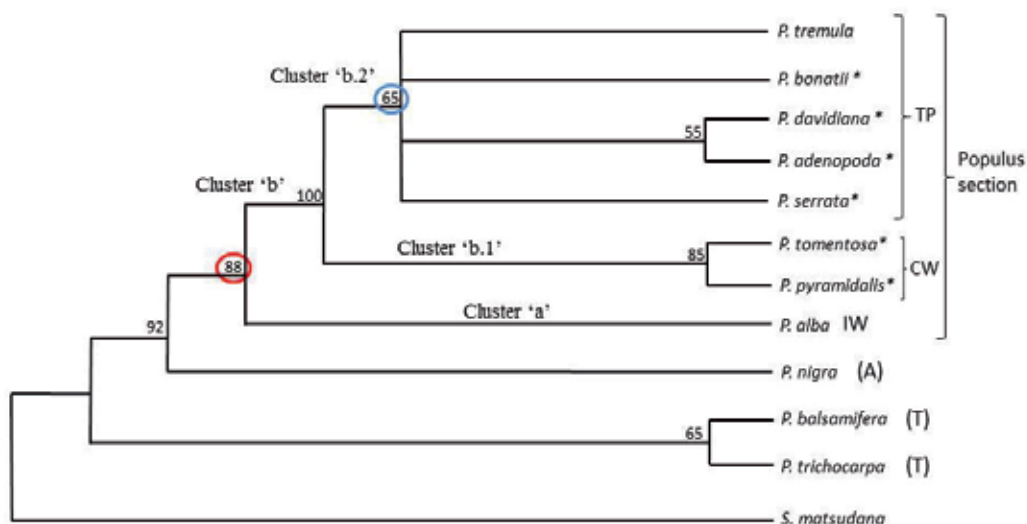
Fig. 1. Maximum likelihood reconstruction based on the "Classical ML method" of the nucleotide supergene encompassing the CDS of all six members of MT multigene family. *Bootstraping values are showed above each node, when greater than 50; unresolved nodes are represented as collapsed. (T): Tacamahaca section; (A) Aigeros section; IW: Italian white poplar; CW: Chinese white poplar; TP: trembling poplars. Nodes corresponding to the Populus section and to Trepidae sub-section are indicated with a red and blue circle, respectively. Chinese species are starred. The cluster names (a, b, b.1 and b.2) are those explained in the text. Sequences belonging to* S. mastudana *were used as outgroup.*

SSR analysis of the *P. alba* natural population. The number of alleles per chloroplast (Cp) SSR locus varied from three to five, with an average of four (Tab.4). Combining the data of the three Cp-SSR loci, eight different chloroplast haplotypes were detected in the *P. alba* population of the Sele river. Nei's (1973) gene diversity index per locus (*h*) varied from 0.447 to 0.566 (average of 0.512), which indicate moderately high level of variability (Tab. 4).

Genetic diversity within the Sele population was also estimated using 10 nuclear SSRs (nSSR see Tab. 5). The total data set included 84 alleles, whilst the number of alleles per locus ranged from 2 up to 17.

The total data set included 84 alleles, whilst the number of alleles per locus ranged from 2 up to 17. The most informative markers were ORPM312a, ORPM312b, ORPM30b and WPMS14 with more than 10 alleles for each single locus. Multilocus analysis of nSSR revealed a high level of gene diversity H_e = 0.58. The highest level of H_e was estimated for ORPM312b (0.852) locus, whilst the lowest for the locus ORPM30a (0.058). The mean observed heterozygosity (H_o) was 0.46. Since H_o was smaller than H_e for 7/10 analysed loci, an excess of homozygosity was observed. Finally, the inbreeding coefficient (Fis) for all loci was 0.18 (ranging from 0.040 to 0.752). Based on SSR markers, we also determined the genetic structure of the Sele population (63 individuals) using, as reference samples, eight white poplars representing a wide biodiversity molecular range for the Italian *P. alba*, as already stated in Material and Methods section. A PCA was carried out on the combined dataset comprising 71 samples (Fig 2), using the Jaccard dissimilarity index as genetic distance.

Allele\Locus	CCMP2	CCMP6	CCMP10
a	0.02	0.015	0.46
b	0.60	0.42	0.46
c	0.31	0.55	0.043
d		0.015	0.014
e			0.014
No. Alleles	3	4	5
h	0.447	0.522	0.566

Table 4. Genetic biodiversity, at the chloroplast level, of the *P. alba* natural population (63 individuals) of the Sele river. *Allelic frequency, number of alleles per locus, gene diversity index per locus (*h*) are reported.*

Allele\Locus	ORPM 30a	ORPM 30b	ORPM 312a	ORPM 312b	ORPM60	PGMC 2852	WPMS05	WPMS14	WPMS18	WPMS20
a	0.03	0.137	0.01	0.021	0.007	0.202	0.018	0.023	0.014	0.079
b	0.97	0.363	0.01	0.063	0.022	0.079	0.061	0.03	0.007	0.886
c		0.016	0.01	0.083	0.283	0.044	0.368	0.008	0.336	0.007
d		0.056	0.125	0.021	0.203	0.018	0.465	0.008	0.05	0.029
e		0.194	0.021	0.125	0.457	0.579	0.088	0.008	0.571	
f		0.129	0.073	0.042	0.029	0.026		0.068	0.021	
g		0.008	0.01	0.021		0.044		0.015		
h		0.016	0.01	0.021		0.009		0.591		
i		0.016	0.354	0.104				0.205		
l		0.04	0.042	0.208				0.015		
m		0.016	0.177	0.25				0.023		
n		0.008	0.094	0.042				0.008		
o			0.01							
p			0.021							
q			0.01							
r			0.01							
s			0.01							
No alleles	2	12	17	12	6	8	5	12	6	4
Ho	0.06	0.758	0.646	0.708	0.739	0.228	0.158	0.576	0.586	0.186
He	0.058	0.789	0.81	0.852	0.669	0.613	0.636	0.602	0.558	0.208
F$_{IS}$	0.031	0.04	0.203	0.168	0.105	0.628	0.752	0.043	0.051	0.109

Table 5. Genetic diversity, at the nuclear level, of the P. alba natural population (63 individuals) of the Sele river. Allelic frequency, number of alleles per locus, expected (He) and observed (Ho) heterozigosity, F$_{IS}$ values are reported.
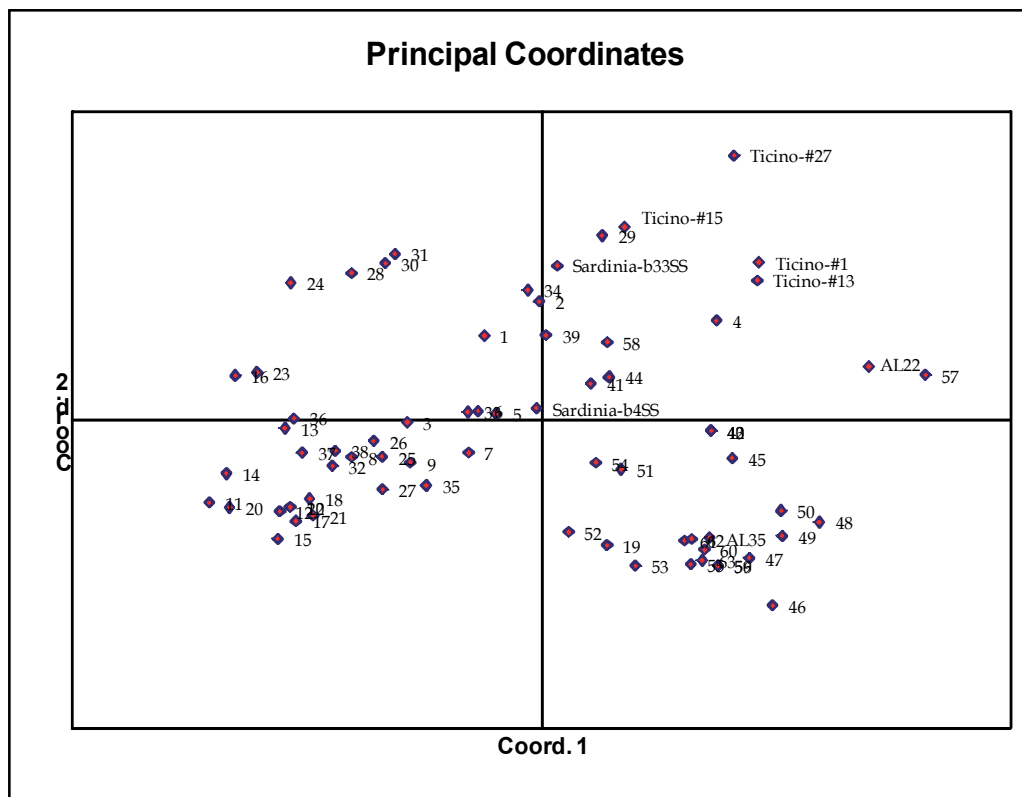
Fig. 2. Two-dimensional plot of the PCA performed on a combined data set comprising 71 poplar samples, *i.e.* 63 individuals from the Sele population (Southern Italy) plus eight reference poplars. *The poplars from river Sele are indicated with a number, whilst the reference sample were from: Sardinia (b33SS and b4SS) or Ticino (Ticino-#1, etc) natural populations, as well as from a collection of Northern Italy white poplar clones selected for pronounced heavy metal tolerance (AL22, AL35).*

As a result, we could assess the tendency of the Ticino samples to form an independent group, well-separated from all the remaining poplars. On the contrary, the Sele river population was more scattered and overlapping with the selected poplars sampled from Sardinia and Northern Italy.

Sequence analysis of MT-genes from the Sele population. Based on the previous SSR analyses, five genetically distinct poplar clones of the Sele river population were chosen for investigating the levels of SNP in all MT-genes. It should be stressed that the selected clones represented the most divergent samples of the data set, thus we expected to maximize the number of SNP found. Successful and reliable PCR amplifications were obtained for all 5 selected clones as single DNA bands with the notable exception of MT1b, which was not reproducibly amplified in all 5 clones (and for this reason not considered in the analysis) and of MT1a, which resulted in a double DNA bands for four out the five selected clones. Therefore, the double PCR product was cloned in *E. coli*, and several *E. coli* colonies were randomly chosen and sequenced to identify single base mutations. Overall, 104 SNPs, *i.e.* 56 point mutations and 48 indels, were identified on 1,456 sites and 18 different alleles, thus

being 7.1 SNPs found every 100 bp. To better illustrate what we found, we may describe the cases of the MT1a and MT3a genes. As for MT1a, the two fragments amplified via PCR were singularly cloned on the basis of their diverse molecular weight, and then sequenced to identify SNPs. In the first MT1a fragment, made up of 332 intronic and 166 exonic bp, a total of 30 SNPs (*i.e.*, 24 indels and 6 single-base changes) were identified; whilst in the second fragment, 27 SNPs (*i.e.*, 4 indels and 23 single-base changes) were found in 409 non-coding plus 79 coding bp. As previously stated above, the number of SNPs was higher in non-coding with respect to the coding region. In fact, 55/57 variable sites were located in the introns of the MT1a gene (9.5 SNPs per 100 bp), whilst only two SNPs were present in its exons (0.49 SNPs per 100 bp). Furthermore, two haplotypes (conventionally named B and C, data not shown) of the Sele population, related to MT1a gene, were characterized by a greater number of SNPs, mostly (>80%) consisting in insertion/deletion events. As for the MT3a gene of Sele poplar population, a single fragment of 476 bp, corresponding to 110 bp of coding and 366 bp of non-coding sequences, was cloned and sequenced: only 1 SNP was detected in its three exons (0.91 SNPs per 100 bp), and 2 were found in the non-coding regions (0.55 SNPs per 100 bp). As a conclusive remark, it should be mentioned that many different and short indels were observed comparing the five Sele clones to the *P. alba* specimens employed for phylogenetic analysis.

2.3 Discussion

Poplar phylogeny. This pilot study provides, for the first time, an extensive genetic survey of the metallothionein multigene family in different poplar species and in a natural population of *P. alba*, a member of the Populus section. The position along the *Populus* genome of the six genes of the MT family may support the indications provided by Brunner and co-worker (2000) regarding the genome duplication event occurred in *Populus* during the last 60 millions years. Here, it should be noticed that, in contrast to *Arabidopsis*, the genome of *Populus* has not been truncated (Kelleher et al., 2007; Tuskan et al., 2006), which means that for most of the genes still two copies are present in poplar genome. Coherently, we confirm that two isoforms (commonly indicated as "a" and "b") are present in the *P. trichocarpa* genome for each of the three MT types.This observation is very interesting but contributes in complicating the possible evolutionary pattern of this multigene family. In fact, the analysis of 89 MT genes retrieved from public data bases and belonging to 19 different plant species, ranging from basal bryophytes to higher dicotyledons [data not shown - (Lupi, 2007)] revealed that at least two different duplications are necessary to originate the full MT family. In particular, the more ancient event duplicated the ancestor gene into the MT3 and MT2 types, whilst a more recent duplication of MT2 likely gave rise to MT1 genes. However, some of the analyzed monocotyledon and dicotyledon species showed a different number of MT genes, suggesting that MT family underwent gain/loss of members via independent events in the diverse lineages, or even to degeneration of its various members [as occurred in *A. thaliana*; (Zhou & Goldsbrough, 1994)]. In the light of the above outlined evolutionary framework, it is not surprising that the evolution process is still acting on some of the poplar MT genes. In fact, as shown via the N/S ratio (Tab 3.), *i.e.* the ratio between non-synonymous and synonymous substitutions, MT1a and MT1b are under neutral and purifying selection, respectively, whilst the remaining four genes are subject to a mildly strong positive selection. This would suggest that MT1a and MT1b could have

already acquired a more specific and "frozen" role in the evolutionary process of the MT family with respect to the other four members. On the contrary, since evolution favors the occurrence of non-synonymous mutations along the MT polypeptide sequence, the other four genes might be still available to become more specialized into a somewhat different function always linked to metal homeostasis. Therefore, on the basis of our data, we could hypothesize that different paralogous genes belonging to the same multigene family may play for distinctive functions implied in plant metal metabolism and/or response to stress tolerance, as already documented by Kohler et al.(2004) and Castiglione et al. (2007, 2010 and 2011). The sequence analysis of the MT genes shows that exon-intron structure is well conserved within the genome of all analysed *Populus* spp (three exons and two introns) with the only exception of the MT2a gene. In this case, the gene presents only two exons and one intron. On the light of the typical MT gene structure found in both *Populus* and *Salix*, the most parsimonious explanation might be that an intron has been lost during the evolution process. Another interesting feature of poplar MT genes in regard to the structure of their CDS. In fact, each of three exons constituting the CDS corresponds to a well-defined functional region of the polypeptide, that is the Cys-rich N-terminal domain, the central spacer and the Cys-rich C-terminal domain. In particular, the number and position of the Cys-residues inside the two terminal domains represent molecular signatures useful to univocally classify the MT genes into the three types (MT1, MT2 and MT3). The presence of well-conserved Cys-residues strongly suggests that terminal domains can coordinate positively charged ions aimed at metal homeostasis and heavy metal tolerance, as observed in the case of different poplar species and natural populations (Castiglione et al., 2010). On the contrary, the spacer region shows a less conserved amino acidic sequence, thus suggesting that the about 40 amino acids that constitute this region may have no other function but linking the two terminal Cys-rich domains. As for molecular variability, only low levels for both nucleotide (96%) and amino acidic (93%) MT sequences were detected within the *Populus* genus. However, these results are not surprising since genes pooled from 11 closely related species were analyzed. As already stated, the spacer domain bears the vast majority (61%) of the variable sites, therefore it proved to be very useful to elucidate the evolutionary relationships among the analyzed poplar species. However, the regions showing the highest variability levels were, as expected, the introns (Lupi, 2007). For instance, the species-specific SSR identified within intron I of MT3b could have a practical application and therefore used to identify hybridization events either in cultivated hybrids of unknown origin, or to discriminate crosses among compatible poplar species in the early generations of natural populations. Interestingly, both section-specific and geographic-specific point substitutions were observed within MT CDS for the different groups of aspens, Italian white poplar and Chinese poplars. These mutations have been fundamental to calculate the phylogenetic tree based on the MT supergene (1,224 sites) and to discriminate the poplar sections and species analyzed. The phylogenetic tree obtained through the Classical ML (Fig. 1) method shows a monophyletic origin of the genus *Populus*. Such results are in contrast to what obtained by Hamzeh and Dayanandan (2004) using both chloroplast and nuclear rDNA sequences, but they are in agreement to what found by Cervera and co-workers (2005) using a different molecular approach (AFLP) in order to estimate phylogenetic relationships among poplars. Our analyses also provide a clear separation between the Tacamahaca and Aigeros (92 bootstrap value), as well as the Populus section (88), thus confirming the results obtained by Fladung and Bushbom (2009) in their pilot study, where five different poplar species belonging to the same three sections

were analyzed on the basis of SNP mutations detected in six nuclear genes. Moreover, the *P. trichocarpa* and *P. balsamifera* species were clearly distinguished and coherently grouped together in a cluster corresponding to the Tacamahaca section with a support value higher than 50 (65), thus corroborating the findings by Cervera et al. (2005). This is a remarkable result since the relationships between *P. trichocarpa* and *P. balsamifera* were not resolved in the phylogenetic analysis conducted by Hamzeh and Dayanandan (2004) using the ITS and rDNA, *i.e.* the two commonly most exploited markers for phylogenetic studies concerning plant organisms. The obtained ML tree also shows that Populus section is basal with respect to the remaining younger sections, Aigeros and Tacamahaca. Once again, our results are in accordance with those reported by Cevera et al. (2005) and also in agreement with the evolutionary pattern proposed by Eckenwalder (1996) on the basis of completely different markers, *i.e.* morphological and phenetic characters. As shown in Figure 1, poplar species belonging to the Populus section are clearly separated from the representative species of the other sections (red circle). Besides, the sub-clusters b1 and b2 are well separated from the Italian *P. alba* clones. It should also be mentioned that the Chinese white-poplar cluster of the Populus section is well-resolved (85 bootstrap value), so that *P. tomentosa* and *P. pyramidalis* actually appears as closely-related species that are likely paraphyletic with respect to the Italian *P. alba*. Albeit this result could depend, at least in part, on the geographical origins of the considered poplar species, a different and quite easy explanation for this paraphyletic relationships between the Italian *P. alba* and the Chinese white poplars could be assessed. In fact, although many Chinese taxonomists still consider *P. tomentosa* as a distinct species from *P. alba* (the so-called "Chinese *P. alba*"), Dickman (2001) stated that the correct nomenclature for this species is *P.* x *tomentosa* Carr: in other words, *P. tomentosa* might be an hybrid derived from a natural cross between *P. alba* and *P. adenopoda* that, in some cases, could be a tri-hybrid due to further introgression toward the *P. tremula* genome. On the other hand, *P. pyramidalis* (the "Bolleana poplar") may be a variety of *P. alba* (Dickmann, 2001), probably consisting in a single genetic unit with a columnar growth habit particularly appreciated for ornamental uses and line planting. Therefore, it is not unfeasible that *P. pyramidalis* could be a *P.* x *tomentosa* clone gathered, long time ago, from a Chinese natural population, and subsequently employed for wind breaking and landscape gardening. As for the four Chinese poplars of the Populus section *P. davidiana, P. adenopoda, P. bonatti* and *P. serrata*, usually considered trembling aspens, they are part of a poorly-resolved group (cluster b.2 in Fig. 1) in our phylogenetic tree. The only exception is represented by *P. davidiana* and *P. adenopoda* that are clustered together with a low supporting bootstrap value (55). Therefore, all four trembling poplar species provided by the Chinese author counterpart of this article for phylogenetic purposes seem to be part of a single species with very limited variations at the genomic level as compared with the Italian *P. tremula*. Indeed, on the basis of the low dissimilarity of both morphological characters and phenetic traits, Eckenwalder (1996) already argued to merge these four Chinese species in a single taxon. Therefore, we could state that both our molecular data on MT genes and the AFLP study, conducted by Cervera and co-workers (2005), strongly confirm the section revision proposed by Eckenwalder (1996), so that, from a genetic point of view, the four Chinese aspens here assessed and *P. tremula* could be considered a unique species.

Genetic biodiversity of a white poplar population along the river Sele. In this study, the genetic biodiversity present in a large natural population of *P. alba* along the Sele river was

estimated. Sele is a river located in south-western Italy originating from the Monti Picentini in Caposele. It flows through the region of Campania, in the provinces of Salerno and Avellino. Its delta is in the gulf of Salerno on Tyrrhenian sea. White poplar and black poplar are the dominant tree taxa along the banks in the area close to the outfall and the spring of the Sele, respectively. The natural and semi-natural banks of the river are surrounded by grazing lands for buffalo and fields of maize and vegetables, and are also the ideal sites for the spontaneous reproduction of white poplar natural populations. The Cp-SSR analysis performed on the 63 individuals of the Sele river population highlighted a number of alleles (per single locus and in total) and haplotypes comparable with that observed for the white poplar population of the Ticino river previously investigated by this research group (8 Vs 9 haplotypes, Brundu et al., 2008), and slightly higher than that of two populations of the Danube river (8 Vs 5, Lexer et al., 2005), but a little lower than that calculated for a limited number of white poplar specimens collected in different regions of the Mediterranean basin (8 Vs 10, Brundu et al., 2008). However, the higher number of haplotypes observed in the population of the Ticino river could be ascribed to the presence of several hybrids (*P.* x *canescens*) derived from the natural cross between *P. alba* and *P. tremula* (Fossati et al., 2004; Castiglione et al., 2010), being the last observation further confirmed by Lexer and co-workers in the case of two populations of the Danube and of one of the Ticino (2005, 2010). However, this shouldn't be the case of the Sele river population since PCA analysis clearly separates the white poplars of the Sele from the hybrids of the Ticino (see comments below). The mean *h* value of the Sele population was slightly higher in comparison with that calculated by Salvini et al. (2001) in the case of five populations of *P. tremula* distributed along the Italian peninsula (0.51 Vs 0.33). These moderately high level of both genetic variability and number of haplotypes with respect to the same *P. tremula* Italian populations [8 Vs 6 - (Salvini et al., 2001)] can possibly be attributed to the fact that the white poplar population here analysed is located in an area (the Cilento) considered as a glacial refugium and a hot spot of biodiversity for several angiosperm species (Cottrell et al., 2005; Fineschi & Vendramin, 2004; Grassi et al., 2009; Petit et al., 2002; Petit et al., 2003). The nSSR analysis carried out on the same population revealed that the expected heterozygosity was relatively high (*He* 0.460), but in accordance with those reported by Castiglione et al. (2010) in the case of three populations distributed along the banks of three rivers in northern and central Italy and with that calculated by Lexer et al. (2005) for two populations of *P. alba* (0.419 and 0.341), and two of *P. tremula* (0.466 and 0.483 - Lexer et al. 2005). Moreover the produced data were also comparable also with those of four different *P. tremuloides* populations collected in North American [0.460, 0.310, 0.560 and 0.410 – (Cole, 2005)]. The nSSRs were also employed to estimate the genetic diversity among the Sele population and the other 8 specimens, used as white poplar reference samples representative of an ample range of Italian *P. alba* biodiversity. In particular, the PCA plot (Fig. 2) clearly splits the Sele poplars from the Ticino hybrids, but not from the remaining reference samples (Sardinia and Northern Italian clones). Thus, this intermingled group does not reflect the geographic origins of the species growing in the Italian peninsula and in the Sardinia island, suggesting that the white poplar population of the Sele is highly variable and consistently different only with respect to the Ticino river representatives. Here, it should be mentioned that the Ticino white poplars are not genetically pure individuals, as in the case of other populations present in the European river basins (Lexer et al., 2005; van Loo et al., 2008), indeed they show different levels of introgression towards *P. tremula* as extensively

demonstrated by Fossati et al.(2004), Lexer et al. (2010) and Castiglione et al. (2010). Based on such considerations, the very high value of genetic diversity found for *P. alba* so far analyzed can be explained not only by the high degree of polymorphism typical of the surveyed SSR loci, but also by the reproductive features of poplar as dioecious species. In fact, poplar is strictly outcrossing and both pollen and seeds can be moved over large distance by wind, facilitating their dispersion even among geographically well separated populations (Castiglione et al., 2010).

To estimate the numbers of SNPs within the MT gene family, we used an innovative approach that proved to be extremely rapid and also cost effective. In fact, we made use of the already "in house" available information about the genetic biodiversity [previously estimated by means of nSSR (Innac, 2009)] on the river Sele white poplar population, so that the five most genetically divergent specimens were selected on the basis of the Jaccard dissimilarity index graphically plotted on a UPGMA dendrogram (data not shown). The purified DNA of the five chosen poplar genotypes were PCR amplified and sequenced to identify the SNPs present along the six MT genes. SNPs were detected in all MT gene sequences of the five chosen white poplar specimens. The only exception was MT2b, where no SNP was detected. This observation corroborated our phylogenetic sequence data, which revealed a very low number of variable sites in a MT2b multialignment comprising 11 different nucleotide sequences belonging to different poplar species. Interestingly, an uneven distribution of SNPs was observed among the diverse MT-gene isoforms, as in the specific case of MT1a and MT3a (see below), suggesting that the occurrence of SNPs varies among the different members of the MT multigene family. These observations suggest a different role of MT genes in metal homeostasis and in response to different stimuli and stresses (Berta et al., 2009; Castiglione et al., 2007; Cicatelli et al., 2010; 2011). The study on MT genes of *P. alba* Sele population showed an average of 7 SNPs and 0.9 SNPs per 100 bp in MT1a and MT3a genes, respectively. In comparison to other *Populus* spp., SNP frequencies were substantially greater than those observed in: *P. tremula*, [one SNP in every 60 bp and one of 208 bp - (Ingvarsson, 2005; Ingvarsson, 2008); *P. nigra* (one SNP in every 26 bp over the nine sequenced genes - (Chu et al., 2009)]; *P. balsamifera* [one SNP in 520 bp - (Breen et al., 2009)]. Nevertheless, a quite similar nucleotide variability at the intraspecific level (7.0%) was observed by Fladung and Buschbom (2009), over 3,221 bp corresponding to the sequences of six genes which are very important for both plant growth and metabolism, and which in addition proved to be suitable molecular tools for estimating genetic and phylogenetic relationships among poplar species.

3. Conclusions

In the light of the data here presented and of the above considerations, we can state that the results of this study on the MT multigene family are sufficient to resolve phylogenetic relationships among poplar species belonging to different sections: in fact, phylogenetic trees based on the MT supergene can be considered more reliable than those obtained by other authors using several type of different molecular markers as in the case of rDNA, ITS and AFLPs. Moreover the SNPs identified in different specimens of the white poplar population of the Sele river suggest a further investigation in other poplar species and populations to validate our findings, aimed particularly at the possible use of poplar for heavy metals phytoremediation purposes.

4. Acknowledgments

5. Note

This chapter of the book is dedicated to the memory of Prof. Francesco Sala (one of the co-authors), who died a few days ago. Francesco was a teacher of living for those who have known him professionally, and a brotherly friend for some of the authors. Francesco was also pleased with the fact that this chapter would have been released in 2012, so he could celebrate 50 years of scientific publications, unfortunately this has not been possible. However, his participation in the drafting of this chapter remains fundamental.

6. References

Alvarez, I. & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution*, Vol. 29, No. 3, pp. 417-434

Berta, M.; Giovannelli, A.; Potenza, E.; Traversi, M. L. & Racchi, M. L. (2009). Type 3 metallothioneins respond to water deficit in leaf and in the cambial zone of white poplar (*Populus alba*). *Journal of Plant Physiology*, Vol. 166, No. 5, pp. 521-530

Breen, A. L.; Glenn, E.; Yeager, A. & Olson, M. S. (2009). Nucleotide diversity among natural populations of a North American poplar (*Populus balsamifera*, Salicaceae). *New Phytologist*, Vol. 182, No. 3, pp. 763-773

Brookes, A. J. (1999). The essence of SNPs. *Gene*, Vol. 234, No. 2, pp. 177-186

Brundu, G.; Lupi, R.; Zapelli, I.; Fossati, T.; Patrignani, G.; Camarda, I. et al. (2008). The Origin of Clonal Diversity and Structure of *Populus alba* in Sardinia: Evidence from Nuclear and Plastid Microsatellite Markers. *Annals of Botany*, Vol. 102, No. 6, pp. 997-1006

Buchanan-Wollaston, V. (1994). Isolation of cDNA clones for genes that are expressed during leaf senescence in *Brassica napus* - Identification of a gene encoding a senescence-specific metallothionein-like protein. *Plant Physiology*, Vol. 105, No. 3, pp. 839-846

Castiglione, S.; Cicatelli, A.; Lupi, R.; Patrignani, G.; Fossati, T.; Brundu, G. et al. (2010). Genetic structure and introgression in riparian populations of *Populus alba* L. *Plant Biosystems*, Vol. 144, No. 3, pp. 656-668

Castiglione, S.; Franchin, C.; Fossati, T.; Lingua, G.; Torrigiani, P. & Biondi, S. (2007). High zinc concentrations reduce rooting capacity and alter metallothionein gene expression in white poplar (*Populus alba* L. cv. Villafranca). *Chemosphere*, Vol. 67, No. 6, pp. 1117-1126

Castiglione, S.; Todeschini, V.; Franchin, C.; Torrigiani, P.; Gastaldi, D.; Cicatelli, A. et al. (2009). Clonal differences in survival capacity, copper and zinc accumulation, and correlation with leaf polyamine levels in poplar: A large-scale field trial on heavily polluted soil. *Environmental Pollution*, Vol. 157, No. 7, pp. 2108-2117

Castiglione, S.; Wang, G.; Damiani, G.; Bandi, C.; Bisoffi, S. & Sala, F. (1993). Rapd Fingerprints for Identification and for Taxonomic Studies of Elite Poplar (Populus Spp) Clones. *Theoretical and Applied Genetics*, Vol. 87, No. 1-2, pp. 54-59

Cervera, M. T.; Storme, V.; Soto, A.; Ivens, B.; Van Montagu, M.; Rajora, O. P. et al. (2005). Intraspecific and interspecific genetic and phylogenetic relationships in the genus

Populus based on AFLP markers. *Theoretical and Applied Genetics*, Vol. 111, No. 7, pp. 1440-1456

Cherian, G. M. & Chan, H. M. (1993). Biological function of metallothioneins - a review, In: *Metallothionein III*, Suzuki, K. T., Imura, N., Kimura, M., (Ed.), pp. 87-109, Birkhauser, Basel

Cho, R. J.; Mindrinos, M.; Richards, D. R.; Sapolsky, R. J.; Anderson, M.; Drenkard, E. et al. (1999). Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nature Genetics*, Vol. 23, No. 2, pp. 203-207

Chu, Y.; Su, X.; Huang, Q. & Zhang, X. (2009). Patterns of DNA sequence variation at candidate gene loci in black poplar (*Populus nigra* L.) as revealed by single nucleotide polymorphisms. *Genetica*, Vol. 137, No. 2, pp. 141-150

Cicatelli, A.; Lingua, G.; Todeschini, V.; Biondi, S.; Torrigiani, P. & Castiglione, S. (2010). Arbuscular mycorrhizal fungi restore normal growth in a white poplar clone grown on heavy metal-contaminated soil, and this is associated with upregulation of foliar metallothionein and polyamine biosynthetic gene expression. *Annals of Botany*, Vol. 106, No. 5, pp. 791-802

Cicatelli, A.; Lingua, G.; Todeschini, V.; Biondi, S.; Torrigiani, P. & Castiglione, S. (2011). Arbuscular mycorrhizal fungi modulate the leaf transcriptome of a *Populus alba* L. clone grown on a zinc and copper-contaminated soil. *Environmental and Experimental Botany*, Vol. doi:10.1016/j.physletb.2003.10.071, No.

Cobbett, C. & Goldsbrough, P. (2002). Phytochelatins and metallothioneins: Roles in heavy metal detoxification and homeostasis. *Annual Review of Plant Biology*, Vol. 53, No. 159-182

Cole, C. T. (2005). Allelic and population variation of microsatellite loci in aspen (*Populus tremuloides*). *New Phytologist*, Vol. 167, No. 1, pp. 155-164

Cottrell, J.; Krystufek, V.; Tabbener, H.; Milner, A.; Connolly, T.; Sing, L. et al. (2005). Postglacial migration of *Populus nigra* L.: lessons learnt from chloroplast DNA. *Forest Ecology and Management*, Vol. 206, No. 1/3, pp. 71-90

Dickmann, D. I. (2001). An overview of the genus *Populus*, In: *Poplar Culture in North America*, Dickmann, D. I., Isebrands, J. G., Eckenwalder, J. E., Richardson, J., (Ed.), pp. 1-42, NRC Research Press, Ottawa, Ontario, Canada

Doyle, J. J. & Gaut, B. S. (2000). Evolution of genes and taxa: a primer. *Plant Molecular Biology*, Vol. 42, No. 1, pp. 1-23

Eckenwalder, J. E. (1996). Systematics and Evolution of *Populus*, In: *Biology of Populus and its implications for management and conservation*, Stettler, R. F., Bradshaw Jr, H. D., Heilman, P. E., Hinckley, T. M., (Ed.), pp. 7-32, NRC Research Press, Ottawa

Fineschi, S. & Vendramin, G. G. (2004). La diversità cloroplastica delle querce italiane: evidenze di una maggiore ricchezza genetica nelle popolazioni meridionali e insulari . *Forest Genetics*, Vol. 1, No. 2, pp. 82-87

Fladung, M. & Buschbom, J. (2009). Identification of single nucleotide polymorphisms in different *Populus* species. *Trees-Structure and Function*, Vol. 23, No. 6, pp. 1199-1212

Fossati, T.; Grassi, F.; Sala, F. & Castiglione, S. (2003). Molecular analysis of natural populations of *Populus nigra* L. intermingled with cultivated hybrids. *Molecular Ecology*, Vol. 12, No. 8, pp. 2033-2043

Fossati, T.; Patrignani, G.; Zapelli, I.; Sabatti, M.; Sala, F. & Castiglione, S. (2004). Development of molecular markers to assess the level of introgression of Populus tremula into P-alba natural populations. *Plant Breeding*, Vol. 123, No. 4, pp. 382-385

Grassi, F.; Minuto, L.; Casazza, G.; Labra, M. & Sala, F. (2009). Haplotype richness in refugial areas: phylogeographical structure of *Saxifraga callosa*. *Journal of Plant Research*, Vol. 122, No. 4, pp. 377-387

Genetic Biodiversity and Phylogenetic Studies in
Poplar by Means of the Metallothionein Multigene Family
133

Guindon, S. & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, Vol. 52, No. 5, pp. 696-704

Hamzeh, M. & Dayanandan, S. (2004). Phylogeny of Populus (Salicaceae) based on nucleotide sequences of chloroplast TRNT-TRNF region and nuclear rDNA. *American Journal of Botany*, Vol. 91, No. 9, pp. 1398-1408

Ingvarsson, P. K. (2005). Nucleotide polymorphism and linkage disequilbrium within and among natural populations of European Aspen (*Populus tremula* L., Salicaceae). *Genetics*, Vol. 169, No. 2, pp. 945-953

Ingvarsson, P. K. (2008). Multilocus patterns of nucleotide polymorphism and the demographic history of *Populus tremula*. *Genetics*, Vol. 180, No. 1, pp. 329-340

Innac, A. (2009). *Studio della biodiversità mediante marcatori molecolari e inquadramento cartografico di una popolazione naturale di pioppo bianco del fiume Sele (SA, Italia)*, Tesi sperimentale, Scienze Biologiche, Università degli Studi di Salerno, Fisciano, Italia

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Société Vaudoise des Sciences Naturelles*, Vol. 44, No. 12, pp. 223-270

Kelleher, C. T.; Chiu, R.; Shin, H.; Bosdet, I. E.; Krzywinski, M. I.; Fjell, C. D. et al. (2007). A physical map of the highly heterozygous *Populus* genome: integration with the genome sequence and genetic map and analysis of haplotype variation. *Plant Journal*, Vol. 50, No. 6, pp. 1063-1078

Kille, P.; Winge, D. R.; Harwood, J. L. & Kay, J. (1991). A Plant Metallothionein Produced in Escherichia-Coli. *Febs Letters*, Vol. 295, No. 1-3, pp. 171-175

Kohler, A.; Blaudez, D.; Chalot, M. & Martin, F. (2004). Cloning and expression of multiple metallothioneins from hybrid poplar. *New Phytologist*, Vol. 164, No. 1, pp. 83-93

Lan, T.; Yang, Z. L.; Yang, X.; Liu, Y. J.; Wang, X. R. & Zeng, Q. Y. (2009). Extensive Functional Diversification of the Populus Glutathione S-Transferase Supergene Family. *Plant Cell*, Vol. 21, No. 12, pp. 3749-3766

Lanave, C.; Preparata, G.; Saccone, C. & Serio, G. (1984). A new method for calculating evolutionary substitution rates. *Journal of molecular evolution*, Vol. 20, No. 1, pp. 86-93

Lexer, C.; Fay, M. F.; Joseph, J. A.; Nica, M. S. & Heinze, B. (2005). Barrier to gene flow between two ecologically divergent *Populus* species, *P alba* (white poplar) and *P tremula* (European aspen): the role of ecology and life history in gene introgression. *Molecular Ecology*, Vol. 14, No. 4, pp. 1045-1057

Lexer, C.; Joseph, J. A.; van Loo, M.; Barbara, T.; Heinze, B.; Bartha, D. et al. (2010). Genomic admixture analysis in European *Populus* spp. reveals unexpected patterns of reproductive isolation and mating. *Genetics*, Vol. 186, No. 2, pp. 699-U391

Lupi, R. (2007). *Marcatori molecolari per la filogenesi del genere* Populus *L.: analisi delle famiglie geniche MT, ACO e ACS*, Tesi sperimentale, Scienze Biologiche, Università degli Studi di Milano, Milano, Italia

Page, R. D. M. (1996). TreeView: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*, Vol. 12, No. 4, pp. 357-358

Peakall, R. & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, Vol. 6, No. 1, pp. 288-295

Petit, R. J.; Aguinagalde, I.; Beaulieu, J.; Bittkau, C.; Brewer, S.; Cheddadi, R. et al. (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science (Washington)*, Vol. 300, No. 5625, pp. 1563-1565

Petit, R. J.; Brewer, S.; Bordacs, S.; Burg, K.; Cheddadi, R.; Coart, E. et al. (2002). Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, Vol. 156, No. 1-3, pp. 49-74

Petre, B.; Major, I.; Rouhier, N. & Duplessis, S. (2011). Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar. *Bmc Plant Biology*, Vol. 11, No. doi 10.1186/1471-2229-11-33, pp.

Rohlf, F. J. & Marcus, L. F. (1993). A Revolution in Morphometrics. *Trends in Ecology & Evolution*, Vol. 8, No. 4, pp. 129-132

Saccone, C.; Lanave, C.; Pesole, G. & Preparata, G. (1990). Influence of Base Composition on Quantitative Estimates of Gene Evolution. *Methods in Enzymology*, Vol. 183, No. 570-583

Salvini, D.; Anzidei, M.; Fineschi, S.; Malvolti, M. E.; Taurchini, D. & Vendramin, G. G. (2001). Low genetic differentiation among Italian populations of *Populus tremula* L. (Salicaceae) estimated using chloroplast PCR/RFLP and microsatellite markers. *Forest Genetics*, Vol. 8, No. 1, pp. 81-87

Schoot, J.; Pospiskova, M.; Vosman, B. & Smulders, M. (2000). Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.)50. *Theoretical and Applied Genetics*, Vol. 101, No. 1/2, pp. 317-322

Sebastiani, L.; Scebba, F. & Tognetti, R. (2004). Heavy metal accumulation and growth responses in poplar clones Eridano (*Populus deltoides* x *P. maximowiczii*) and I-214 (*P.* x *euramericana*) exposed to industrial waste. *Environmental and Experimental Botany*, Vol. 52, No. 1, pp. 79-88

Small, R. L.; Cronn, R. C. & Wendel, J. F. (2004). Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany*, Vol. 17, No. 2, pp. 145-170

Smith, R. L. (1988). Phylogenetics of *Populus* L. (*Salicaceae*) based on restriction site fragment analysis of cpDNA, University of Wisconsin, Madison, Wisconsin, USA

Tamura, K.; Dudley, J.; Nei, M. & Kumar, S. (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, Vol. 24, No. 8, pp. 1596-1599

Thompson, J. D.; Higgins, D. G. & Gibson, T. J. (1994). Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Research*, Vol. 22, No. 22, pp. 4673-4680

Tuskan, G. A.; DiFazio, S.; Jansson, S.; Bohlmann, J.; Grigoriev, I.; Hellsten, U. et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*, Vol. 313, No. 5793, pp. 1596-1604

van Loo, M.; Joseph, J. A.; Heinze, B.; Fay, M. F. & Lexer, C. (2008). Clonality and spatial genetic structure in Populus x canescens and its sympatric backcross parent P-alba in a Central European hybrid zone. *New Phytologist*, Vol. 177, No. 2, pp. 506-516

Weising, K. & Gardner, R. (1999). A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome*, Vol. 42, No. 1, pp. 9-19

Yadav, R.; Arora, P.; Kumar, S. & Chaudhury, A. (2010). Perspectives for genetic engineering of poplars for enhanced phytoremediation abilities. *Ecotoxicology*, Vol. 19, No. 8, pp. 1574-1588

Yin, T. M.; DiFazio, S. P.; Gunter, L. E.; Riemenschneider, D. & Tuskan, G. A. (2004). Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. *Theoretical and Applied Genetics*, Vol. 109, No. 3, pp. 451-463

Zhou, J. M. & Goldsbrough, P. B. (1994). Functional homologs of fungal Metallothionein genes from *Arabidopsis*. *Plant Cell*, Vol. 6, No. 6, pp. 875-884

Analysis of Genetic Diversity and SSR Allelic Variation in Rubber Tree (*Hevea brasilensis*)

Suping Feng[1,2], Yaoting Wu[2], Weiguo Li[3], Fei Yu[1] and Jingyi Wang[1]
[1]Key Laboratory of Tropical Crop Biotechnology
Institute of Tropical Bioscience and Biotechnology
Chinese Academy of Tropic Agricultural Science
[2]Qiongzhou University
[3]Key Laboratory of Rubber Biology of Ministry of Agriculture
Rubber Research Institute,Chinese Academy of Tropical Agricultural Sciences
China

1. Introduction

Rubber tree, *Hevea brasiliensis*, belongs to the family of Euphorbiaceae, originated from Amazon Trends. The family has ten varieties (*H. brasiliensis, H. nitida, H. pauciflora, H. spruceana, H. benthamiana, H. camporum, H. microphylla, H. rigidifolia, H. guianensis, H. comargcana*) and four variation varieties (*H. guianensis var. luter, H. guianensis var. marginata, H. parciflora var coriacea, H. nitida Mart var. toxicadendroides*). *H. brasiliensis* is the most economically important member of the genus *Hevea*, because its economic importance and its sap-like extract (as latex) can be collected and is the primary source of natural rubber. There are many rubber tree germplasm resources in Brazil, Malaysia, Indonesia, India, French, and China.

Simple sequence repeat (SSR) marker has been used as an ideal molecular marker to investigate the genetic diversity because of its multi-allelic nature, reproducibility, co-dominant inheritance, high abundance and extensive genome coverage (Gupta & Varshney, 2000) in many crops, such as rubber tree (Lekawipat et al., 2003; Saha et al., 2005; Gouvêa et al., 2010), wheat (Liu et al., 2007; Hao et al., 2006), bean (Choi et al., 1999), barley (Brantestam et al., 2007), cole (Hasan et al., 2006), jowar (Marco et al., 2007), triticale (Tams et al., 2004), rice (Song et al., 2003) and coffee (Aggarwal et al., 2007).

Several researchers have investigated the genetic diversity of rubber tree by using molecular markers (Lekawipat et al., 2003; Saha et al., 2005; Lam et al., 2009; Gouvêa et al., 2010; Oktavia et al., 2011), but there was no report about the polymorphism of lower repeats SSR markers and the loci and the flanking area variation of SSR markers in rubber tree. SSR marker used to detect the alleles by PAGE gel after PCR has slight restrictions in distinguishing the fragments as length or size homoplasy (Estoup et al., 1995; Grimaldi & Crouau-Roy, 1997; Angers & Bernatchez, 1997). However, sequencing of repeats and flanking regions can help detect the difference of the alleles exactly (Xie et al., 2006; Feng et al., 2008).

In this study, 16 primer pairs from genome and EST-SSR amplified across the popular cultivars cultivated in China, wild accessions and interspecies. There were three main objectives: (1) detect the genetic diversity and relationships between cultivars and wild accessions in rubber tree, (2) to investigate the polymorphism of low repeat SSR marker, and (3) analyze the loci variations in rubber tree.

2. Materials and methods

2.1 Plant materials and SSR markers

Forty-five cultivars which are the main cultivars in China, 11 wild accessions from Brasil and 3 related species were used in this study (Table1). Fresh leaves were collected in bronze period from Rubber Research Institute, Chinese Academy of Tropic Agricultural Science (Danzhou) and stored at -20℃ after washing by pure water. Leaf genomic DNA was extracted following the CTAB protocol (Venkatachalam et al., 2002).

Sixteen SSR primer pairs (Table 2) were used, including 10 EST-SSRs (SSR from expressed sequence tag) (Feng et al., 2009; Unpublished), 6 genomic-SSRs (http://tropgenedb.cirad.fr/en/rubbertree.html). The four effective EST-SSR primer pairs were (1) HBE280 (gga cac ctg gag caa aat ag ＆ tat gct tcg atg tat att cac agt: $[(gaaa)_4]$)；(2) HBE301 (ggc ata caa gaa aaa aat tt ＆ taa gga ttg acg gct acg: $[(cagcaa)_5]$); (3) HBE316 (cga caa cca gga act tac c ＆ aaa caa ctg cgg agg att: $[(tctgt)_4]$); (4) HBE329 (cca aaa caa ggg aaa tca c ＆gac cga gac gct tag ttc: $[(aga)_9]$); All primers were synthesized by the Shanghai Sangon Biological Engineering Technology & Services Co., Ltd.

No.	materials	scientific name	Pedigree/source
1	RRIM600	*H. brasiliensis*	Tjir 1×PB86
2	Zhenxuan1		PB28/59×RRIM60
3	Dafeng117		RRIM513×PR107
4	Reyan88-13		RRIM600×PilB84
5	Baoting155		RRIM600×PR107
6	Haiken2		PB86×PR107
7	Wenchang217		Haiken1×PR107
8	Reyan7-18-55		RRIM600×PR107
9	IAN873		PB86×PR107
10	PB86		original clone
11	Tianren31-45		original clone
12	Reyan7-33-97		RRIM600×PR107
13	Gunagxi6-68		original clone
14	PB5/51		PB56×PB24
15	GT1		original clone
16	PR107		original clone

No.	materials	scientific name	Pedigree/source
17	Yunyan68-273		GT1×PR107
18	Reyan217		RRIM600×PR107
19	Dafeng99		PB86×PR107
20	Reyan7-20-59		RRIM600×PR107
21	Dafeng95		PB86×PR107
22	Hekou3-11		original clone
23	Tjir 1		original clone
24	Hongxing 1		original clone
25	Haiken 1		Unknown
26	RRIM712		RRIM605×RRIM71
27	RRIM513		Pil B16×Pil A 44
28	Haiken 6		PB86×PR107
29	Baoting 3410		RRIM600×PR107
30	Wenchang 7-35-11		PB5/51×PR107
31	Baoting 911	*H. brasiliensis*	RRIM600×PR107
32	Tianren 93-114		Tianren31-45×Hekou3-11
33	Wenchang 193		PB5/51×PR107
34	Wenchang 11		RRIM600×PR107
35	Daling 68-35		Qiaozhi42-67×PB86
36	PB 28/59		Unknown
37	PB 5/63		PB56×PB24
38	Yunyan 277-5		PB5/63×Tjir 1
39	Reyan 8-333		Reyan88-13×Reyan217
40	Baoting 235		RRIM600×PR107
41	Reyan8-79		Reyan88-13×Reyan217
42	Reken525		IAN873×RRIM803
43	Reken523		IAN873×PB260
44	Yunyan77-2		GT1×PR107
45	PB217		PB5/51×PB6/9
46	AC/T/15/114		Brazil
47	AC/S/1037/297		Brazil
48	RO/CM/1163/17		Brazil
49	RO/C/923/176		Brazil
50	MT/C/1017/88		Brazil
51	MT/IT/1634/209		Brazil
52	AC/AB/1554/16		Brazil
53	MT/C/210/28		Brazil
54	MT/IT/1634/192		Brazil
55	RO/A/725/121		Brazil
56	MT/IT/1430/118		Brazil
57	Sebao rubber	*H. spruceana*	Brazil
58	Guangye rubber	*H. nitida*	Brazil
59	Bianqin rubber	*H. benthamiana*	Brazil

[1] Rubber Cultivation Research Institute, Chinese Academy of Tropic Agricultural Science, Danzhou

Table 1. Plant materials used for analysis of genetic biodiversity and loci variation

2.2 PCR amplification and detection of fragments

All primers were amplified in the TaKaRa PCR Thermal Cycler Dice, each PCR reaction consisted of: 2 µl of 10×buffer, 0.25 µl of 10 M dNTPs, 1 µl each of forward and reverse primer (20 µmol), 2µl of template leaf genomic DNA (20 ng/µl), 0.15 µl of Taq polymerase (5 U/µl) (TAKARA Biotechnology (Dalian) Co. Ltd), ddH$_2$O added to a total reaction volume of 20 µl. The PCR reaction profile was pre-denatured at 94 ℃ for 2 min followed by 30 cycles of 94 ℃ for 30 sec, annealing temperature for 45 sec and 72 ℃ for 1 min and finally, 72 ℃ for an extension of 5 min. To ensure precision and reproducibility of fragments, DNA samples were amplified and analyzed at least twice from each individual sample.

2.3 Cloning and sequencing of SSR alleles

Five EST-SSR markers (HBE008, HBE063, HBE164, HBE187, and HBE199) were selected to investigate SSR loci variation. Of the five primer pairs, the repeat motif of HBE008 and HBE187 was (CT) n, HBE164 and HBE199 was (AG) n. The selected alleles were amplified, recovered, purified, cloned and sequenced.

Alleles from these SSR loci were cut from the dried PAGE gels and used as templates for a new round of PCR amplifications. Each of these alleles was directly cloned into the pGEM-T Easy Vector (Promega, USA) according to the manufacturer's instructions, and transformed into *Escherichia coli* DH5 α cell. The positive clones were sequenced using the ABI PRISM 3730 sequencer. To obtain reliable sequences, at least three clones per allele were sequenced. The nucleotide sequences were aligned using Clustal X (http://www.ftp-igbmc.u-strasbg.fr/pub/ClustalX/) to compare the amplified SSR alleles with the SSR-containing ESTs to investigate the loci variation.

2.4 Data analysis

Sixty-two genomic DNA of cultivars and wild accessions were amplified with 16 primer pairs, and visualized on PAGE gel with silver stain. These bands were recorded as "1" for presence, "0" for absence, "999" for missing data. Number of alleles, observed heterozygosity (*H*o) and power of discrimination (PD) were calculated for each locus. *H*o was calculated as the number of genotypes which were heterozygous at a given locus divided by the total number of genotypes surveyed at that locus. PD was calculated as 1-$\sum G^2ij$ (Kloosterman et al., 1993), where *G*ij is the frequency of the *j*th genotype for the *i*th locus summed across all alleles at that locus. Genetic similarity (GS) between any two pairs of the 56 cultivars and wild accessions was calculated from the alleles across the 16 SSR loci using the Jaccard similarity coefficients (Sneath & Sokal, 1973). A dendrogram was constructed with the un-weighted pair group method with arithmetic averages (UPGMA) on the basis of the similarity coefficients. All these analyses were performed with NTSYS-pc 2.10 software package.

3. Results

3.1 Polymorphic analysis of EST- and gSSR markers

Sixteen SSR primer pairs, of which ten from EST-SSRs and six from genome, could successfully amplify, expected products across 45 cultivars, 11 wild accessions and 3 related species (Table

2). A total of 43 alleles were obtained from 10 EST-SSR primer pairs across cultivars and wild accessions, with an average of 4.3 alleles and H_o = 0.488. Ten alleles were amplified by HBE280, followed by HBE199 with 5 alleles, and HBE164 and HBE316 with 2 alleles each, respectively. Six alleles were amplified by HBE008 and HBE280 in wild accessions, respectively, followed by HBE199 with 5 alleles and HBE316 with 2 alleles. A total of 30 alleles were obtained from 6 gSSR primer pairs, with an average of 5 alleles and H_o=0.743. Six alleles were amplified by M197 and MnSOD, respectively in cultivars, 6 alleles by MnSOD as well as in wild accessions. And the other primer pairs were amplified 4 or 5 alleles in cultivars and wild accessions, respectively. HBE280, M197 and MnSOD were the most informative.
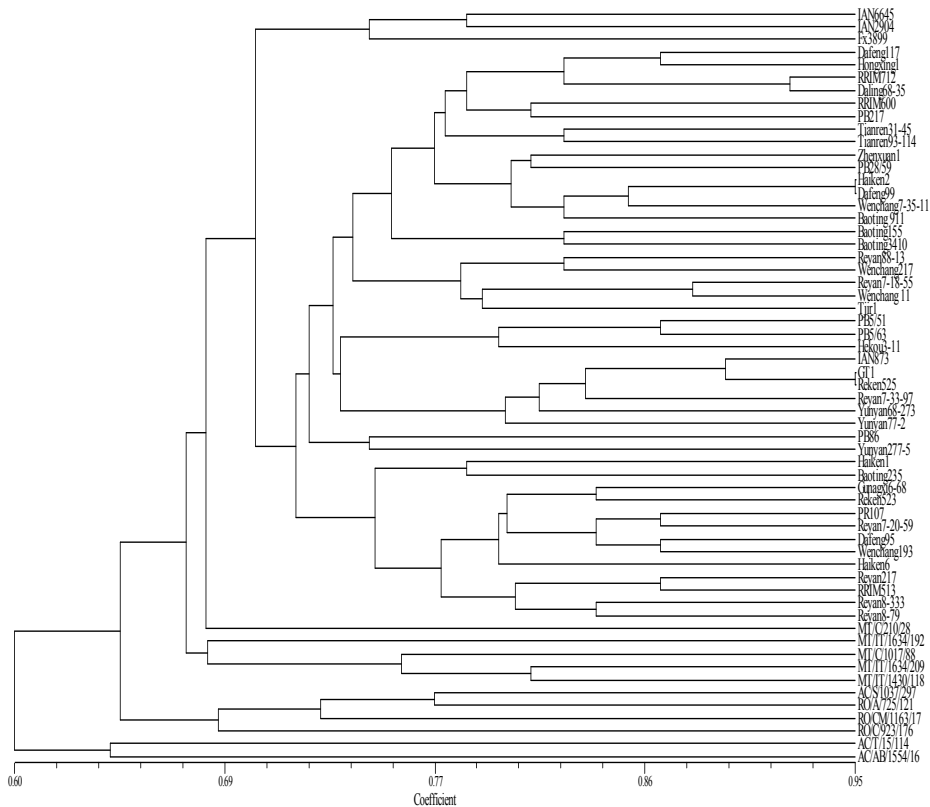
Fig. 1. UPGMA dendrogram of genetic relationship among cultivars and wild accessions of rubber tree based on analysis using ten EST- and six genomic SSR markers.

For cultivars, the ten EST-SSR markers produced a total of 40 alleles with an average of 4 alleles, H_o = 0.495 and PD = 0.591 per locus; and the six gSSRs produced a total of 30 alleles with an average of 5 alleles, H_o = 0.739 and PD = 0.681 per locus (Table 2). For the wild accessions, the ten EST-SSR markers produced a total of 38 alleles with an average of 3.8 alleles, H_o = 0.455 and PD = 0.616 per locus; and the six gSSRs produced a total of 28 alleles with an average of 4.7 alleles, H_o = 0.758 and PD = 0.685 per locus (Table 2). HBE280 was the most informative among the EST-SSRs, and MnSOD was the most informative among the genomic SSR markers.

Of the 16 primer pairs, only PD value of HBE280 was larger than 0.8 (0.834). PD values of the ten EST-SSRs ranged from 0.420 to 0.834 and the average was 0.597; and the PD values of the six genomic-SSRs ranged from 0.587 to 0.772 and the average was 0.689.

3.2 Analysis of genetic diversity

The Jaccard similarity coefficient for the 16 SSR markers was used to analysis the genetic similarities (GS) among the 45 cultivars and 11 wild accessions. In comparison with the 56 accessions, GS values of cultivars ranged from 0.51 (Reken523 and PB5/63) to 0.95 (Reken525 and GT1, Haiken2 and Dafeng99). GS values between cultivars and wild accessions ranged from 0.48 (AC/T/15/114 and Wenchang217) to 0.79 (MT/IT/1430/118 and Tianren31-45).

UPGMA cluster analysis based on GS values for comparisons among all samples was used to construct a dendrogram (Fig. 1) with the cophenetic value of 0.854, indicating a high level of reliability. The 45 cultivars and 11 wild accessions were clustered into two categories. Wild accessions were distinguished from cultivars on the level of similarity coefficient 0.68. On the level of similarity coefficient 0.73, cultivars were divided into four groups, IAN6645,IAN2904 and Fx3899 in group I, PB86 and Yunyan277-5 in group Ⅲ, Haiken1, Baoting235, Gunagxi6-68, Reken523, PR107, Reyan7-20-59, Dafeng95, Wenchang193, Haiken6, Reyan217, RRIM513, Reyan8-333 and Reyan8-79 in group Ⅳ. Group Ⅱ contained mostly cultivars except those in group I , Ⅲ and Ⅳ. On the level of similarity coefficient 0.77, Group Ⅱ can be divided into six sub-groups. Group Ⅳ can be divided into three sub-groups on the level of similarity coefficient 0.78.

All accessions could be distinguished from each other on the level of similarity coefficient 0.95 except Haiken2 and Dafeng99, GT1 and Reken525.

3.3 Detection of loci variation

Substantial sequence variation was found in the 61 sequences of 5 EST-SSR loci. The insertion, deletion, transition and transversion can be found in these sequences (Fig. 2).

The original EST sequence (the RefEST) ID of HBE008 is EC609578.1 in Genebank and the SSR repeat loci is $(CT)_{21}(AT)_{14}$. In comparison with EC609578.1, 1~6 bp deletions were happened at the AT repeat loci of *H. nitida,* IAN2904, Fx3899, AC/T/15/114 and MT/IT/1634/192, and transversions were also observed at the CT and AT repeat loci of Fx3899, *H. spruceana* and AC/T/15/114, the CT repeat lost 1~3 bp and C was replaced by G or A in sequences of *H. nitida,* IAN2904, MT/IT/1634/192, AC/T/15/114, *H. spruceana* and Fx3899.

The original EST sequences (the RefEST) ID of HBE063 is EC607362.1, and the SSR repeat loci is $(GA)_{16}$. In comparison with EC607362.1, insertions were observed from the repeat sequences of Zhenxuan1, Hekou3-11, Yunyan77-2, Reyan88-13, Baoting155 and RRIM712, and deletions were observed at all other sequences except RRIM600, and transversions were also observed from the flanking regions of Hekou3-11, Reyan88-13, Baoting155, Fx3899, RRIM712, RO/C/923/176, MT/IT/1634/209 and MT/IT/1430/118 too.

The original EST sequences (the RefEST) ID of HBE164 is EC603146.1, and the SSR repeat loci is $(AG)_6$. In comparison with EC603146.1, no variation was observed from all repeat

regions, and transversions were observed from the flanking regions of all sequences and insertions were found in the flanking sequences of Dafeng117 and AC/T/15/114. The original EST sequences (the RefEST) ID of HBE187 is EC601817.1, and the SSR repeat loci is (CT)$_6$. In comparison with EC601817.1, no variation was observed from all repeat regions, but deletions were observed from the flanking sequences of RRIM712, Dafeng99, Hongxing1, AC/T/15/114, MT/IT/1634/192, MT/IT/1430/118, Reken525, IAN873, IAN2904 and *H. nitida*. However, transversions were found in the flanking regions of RRIM712, AC/T/15/114, MT/IT/1634/192, MT/IT/1430/118, *H. nitida* and Reken525.
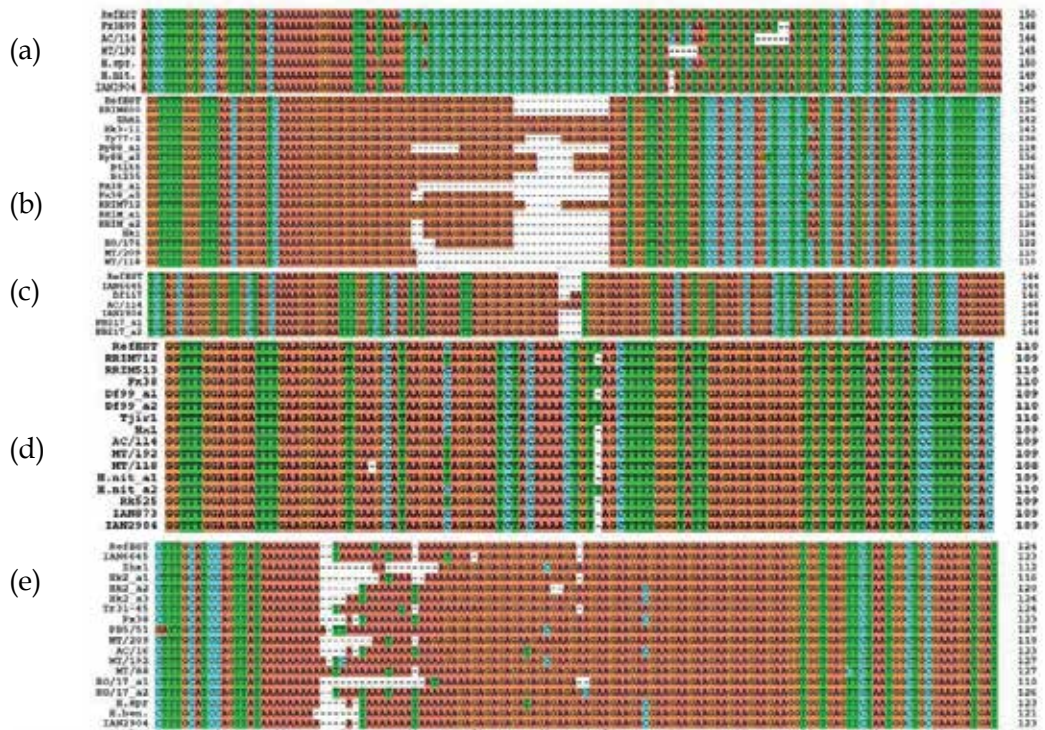
Fig. 2. Sequences obtained using 5 EST-SSRs markers amplifying across *H. brasiliensis*, wild *H. brasiliensis* and interspecies. (a), (b), (c), (d) and (e) (from the top down) represent the markers HBE008, HBE063, HBE164, HBE187 and HBE199, respectively; RefESTs in (a), (b), (c), (d), (e) represent the accession numbers: EC609578.1, EC607362.1, EC603146.1, EC601817.1 and EC600890.1 in NCBI database, respectively. The suffix 'a1', 'a2' and 'a3'represent the allele numbers. AC/114, MT/192, RO/176, MT/209, MT/118, RO/17, MT/88, AC/16, Zhx1, Ry88, Bt155, Hk3-11, Hk1, Bt235, Yy77-2, Fx38, Df117, Df99, Hx1, Rk525, Hk2, Tr31-45, RRIM, H.spr., H.nit., H.ben., represent AC/T/15/114, MT/IT/1634/192, RO/C/923/176, MT/IT/1634/209, MT/IT/1430/118, RO/CM/1163/17, MT/C/1017/88, AC/AB/1554/16, Zhenxuan1, Reyan88-13, Baoting155, Hekou3-11, Haiken1, Baoting235, Yunyan77-2, Fx3899, Dafeng117, Dafeng99, Hongxing1, Reken525, Haiken2, Tianren31-45, RRIM513, *H. Spruceana*, *H. Nitida*, *Hevea. Benthamiana*, repectively.

Sequence ID	Marker	SSR motif	No of alleles			H_o			PD		
			Cultivars	Wild accessions	Total	Culti-vars	Wild access-ions	Total	Culti-vars	Wild access-ions	Total
EST-SSR markers											
EC609578.1	HBE008[a,f]	(CT)21(AT)14	4	6	7	0.674	0.455	0.633	0.669	0.641	0.699
EC609548.1	HBE010[a]	(AG)14	3	3	3	0.456	0.636	0.490	0.634	0.642	0.637
EC607362.1	HBE063[a,f]	(GA)16	3	4	4	0.278	0.273	0.277	0.734	0.734	0.590
EC603146.1	HBE164[a,f]	(AG)6	2	3	3	0.304	0.636	0.366	0.371	0.649	0.493
EC601817.1	HBE187[a,f]	(CT)6	3	3	3	0.413	0.273	0.387	0.430	0.357	0.420
EC600890.1	HBE199[a,f]	(A)11(AG)9*(AGA)11*	5	5	5	0.870	0.727	0.843	0.737	0.787	0.748
EC605179.1	HBE280[b]	(GAAA)4	10	6	10	0.696	0.545	0.668	0.840	0.799	0.834
EC602073.1	HBE301[b]	(CAGCAA)5	3	3	3	0.391	0.182	0.352	0.439	0.556	0.488
EC600641.1	HBE316[b]	(TCTGT)4	2	2	2	0.348	0.364	0.351	0.427	0.391	0.429
CB376545.1	HBE329[b]	(AGA)9	3	3	3	0.522	0.455	0.510	0.628	0.602	0.636
Average			4.0	3.8	4.3	0.495	0.455	0.488	0.591	0.616	0.597
Genomic SSR markers											
AF221698	M127[c,d,e]	(GA)13	5	5	5	1.000	1.000	1.000	0.750	0.710	0.751
AF221711	M197[c,d,e]	(GA)10	6	4	6	0.609	0.636	0.614	0.716	0.560	0.681
AF221705	M508[c,d,e]	(TAA)21 + (TA)6	4	4	4	0.435	0.727	0.489	0.541	0.690	0.587
AF221707	M613[c,d,e]	(GA)23+(GA)8	4	4	4	0.565	0.455	0.544	0.608	0.602	0.626
AF383936	mHBCIR425[c,d,e]	(CA)6	5	5	5	0.826	0.727	0.808	0.711	0.744	0.719
G73377	MnSOD[c,d,e]	(GA)16	6	6	6	1.000	1.000	1.000	0.759	0.802	0.772
Average			5	4.67	5	0.739	0.758	0.743	0.681	0.685	0.689
Average(all markers)			4.38	4.13	4.56	0.587	0.568	0.549	0.625	0.642	0.632

[a]Feng et al., 2009

[b]Feng et al.,data not shown

[c]Seguin et al., 1996

[d]Lespinasse et al., 2000

[e]Seguin et al.,2002

[f]Used for amplified sequence analysis

Table 2. No. of alleles, observed heterozygosity (H_o) and power of discrimination (PD) of EST- and genomic SSR markers among *H. brasiliensis*

The original EST sequences (the RefEST) ID of HBE199 is EC600890.1, and the SSR repeat loci is $(A)_{11}(AG)_9*(AGA)_{11}$. In comparison with EC600890.1, deletions were observed from the AG repeat regions of IAN6645 and Haiken2, and transversions or transitions were also observed from the AG repeat regions of Zhenxuan1, PB5/51, MT/IT/1634/192, AC/AB/1554/16 and *H. spruceana.* Deletions, insertions, and transversions or transitions were found in the flanking regions of all sequences.

4. Discussion

4.1 Polymorphism of SSR marker

Feng et al. (2009) developed 87 EST-SSR primer pairs by analyzing the NCBI database, and the criteria of searching SSR-ESTs was using mononucleotide repeats ≥ 10, dinucleotide to hexanucleotide repeats ≥ 6. But in this study, the search criteria of EST-SSR was changed and 108 novel EST-SSR primer pairs were developed by the same method previous mentioned, of which 4 were selected to analysis the genetic diversity of cultivated varieties and wild accessions. HBE280 appeared to have the most informative despite its repeat number was 4 and the other three novel EST-SSR primer pairs (HBE301, HBE316) also showed different polymorphisms, which were similar to the results of Steinkellner et al. (1997) studied in oak. The four novel primer pairs could be useful as molecular markers in the future study for rubber tree.

Generally, higher-order repeat motifs, which refers to repeat motif more than three bases as well as the compound repeat, have lower polymorphism than lower-order repeat ones (Dreisigacker et al., 2004; Feng et al., 2009). In this study, higher polymorphism was observed in HBE280 [(gaaa)$_4$] which contained only four times repeats, and this may be due to more A/T content in the repeat unit GAAA/CTTT which deduced a replication origin during DNA replication, and the mismatch cannot be repaired easily because of the existence of rows of (A/T) n.

4.2 The polymorphism of EST- and gSSR

SSR molecular markers have been widely used to distinguish crop genotypes (Sun et al., 1999; Virk et al., 1999; Eujayl et al., 2001; Eujayl et al., 2002) and study genetic diversity (Song et al., 2003; Hao et al., 2006; Liu et al., 2007; Caniato et al., 2007). In recent years, a large number of EST- and gSSR molecular markers for many crops were developed (Davierwala et al., 2001; Varshney et al., 2005; Aggarwal et al., 2007). However in rubber tree study, the earliest SSR markers developed by Lespinasse et al. (2000) were from the genomic DNA. Recently, Feng et al. (2009) developed 87 EST-SSR markers from NCBI database. Because of the conservative sequences of ESTs, The level of polymorphism in EST-SSR was lower compared to that of genomic-SSR marker (Eujayl et al., 2002; Leigh et al., 2003; Gonzalo et al., 2005; Yang et al., 2005; Pinto et al., 2006). In this study, gSSRs produced more polymorphisms than EST-SSR as well as PD was slightly higher, despite HBE280 was the most informative marker which belonged to EST-SSRs.

All the SSR primer pairs could be amplified successfully through the 3 related species, and there were no significant differences for transferability between EST- and gSSR, and this was different from other reports (Liewlaksaneeyanawin et al., 2004; Feng et al. 2009). This may due to the relevant lower number of SSR primer pairs.

4.3 Diversity analysis and genetic relationship within/between cultivars and wild accessions

Crop genetic diversity is the basis of genetic improvement of crops, and the study was of great significance in the collection, preservation, evaluation and utilization of crop germ plasm resources. In the past few years, genetics investigations of *H. brasiliensis* have been

studies (Seguin et al., 1996; Chevallier et al., 1998; Seguin et al., 2002), Lekawipat et al. (2003) used twelve microsatellite markers to detect DNA polymorphism among 108 accessions of *H. brasiliensis* inclusive of 40 cultivated (Wickham) clones and 68 wild accessions (1981 Amazonian accessions) collected from Amazon forest, and they found wild accessions were more polymorphic than cultivated Wickham clones and could be divided into three clusters, depending on the geographical origin of collection areas such as Acre, Rondonia and Mato Grosso state. In this study, the similar results were also reached and 16 EST-and gSSR molecular markers were used to detect the genetic diversity within/between 45 cultivars planted in China and other countries, of which 33 were cultivars of China, and 11 wild accessions. The results showed that wild accessions were more polymorphic than cultivars, and the analysis of rubber tree germplasm resources by SSR markers consisted with the pedigree analysis approximately, a true reflection of genetic variation and relationships for rubber tree. For example, IAN6645, IAN2904 and FX3899 were clustered together, and they are good clones from Brazil, in which IAN6645 is a descendant from an FX43.655 [FX213 (F4542×AVROS183) ×AVROS183] ×PB86 cross, IAN2904 from FX516 (F4542×AVROS363) ×PB86, FX3899 from F4542×AVROS363, which all have the same parent F4542. The clustering result consisted with the pedigree analysis. PB5/51 and PB5/63 were clustered together, which are high-yielding clones in Malaysia, and they are the descendant from a common PB56×PB24 cross and the clustering result consisted with the pedigree analysis as well; Haiken2, Dafeng99 and Wenchang7-35-11 were clustered together, and the similarity coefficient between Haiken2 Dafeng99 was 0.95, and they are the descendant from a common PB86×PR107 cross. Wenchang7-35-11 from PB5/51×PR107, these three wind-resistant and high-yielding clone have the same paternal PR107, which is a wind-resistant and high-yielding clone, and the high-yielding maternal PB86 and PB5/51, the clustering result consisted with the pedigree analysis as well. IAN873, GT1, Reken525 and Reyan7-33-97 were clustered together; these results consisted with Feng et al. (2009). At the same time the descendants from the same cross would not been clustered in the same category or nearest position, such as Dafeng99 and Dafeng95 bred by Haiken Dafeng Farm, Hainan, from the same PB86×PR107 cross, and the similarity coefficient between them was 0.72, but they were clustered in group Ⅱ and Ⅳ respectively, more variations would be in F1 hybrid generation. Baoting155, Reyan7-18-55, Reyan7-33-97, Reyan7-20-59, Baoting3410, Baoting911, Wengchang11 and Baoting235 are descendants from the same RRIM600×PR107 cross, but the similarity coefficient between them varied from 0.57 to 0.88, and they were clustered in different group or sub-group. The reasons of inconsistent with the pedigree may be due to that rubber trees are cross-pollinated crops and there is a long-term natural hybridization among the population. The variations between descendants from the different species cross may be from their parents or separation.

With the development of the forest breeding, the consistency of bred varieties increased gradually, resulting in the narrow basis of genetic breeding. Therefore, the reasonable utilization of wild germplasm resources may be an effective way to breeding improvement and widen genetic basis (Benong, 2002; Aidi-Daslin, 2002; Clement-Demange, 2002; Varghese, 2002). In our study, the genetic diversity of wild accessions was higher than the cultivars, which consisted with Besse et al. (1993, 1994). Similarity coefficient between cultivars ranged from 0.51 to 0.95, with an average of 0.73, and the variations of the descendants from those crosses with high similarity coefficient would be limited. It was difficult to breed breakthrough varieties. The similarity coefficient between the cultivars and

wild accessions was 0.48~0.81, which may provide possibility for the selection of elite parents with improved genetic basis (Varghese, 2002).

It was difficult to distinguish Haiken2 from Dafeng99, GT1 from Reken525 on the level of similarity coefficient of 0.95. Feng et al. (2009) also failed to distinguish GT1 from Reken525 on the level of similarity coefficient of 0.96 by 87 EST-SSR markers, however, according to Feng et al. (2008), there were both the same repeat number and structure in GT1 and Reken525, and the point mutations were found in the flanking regions by sequencing in the HBE156 loci .

The methods of detecting genetic diversity, whether on the level of morphology, cytology (chromosome), physiology, biochemical or even now molecular method, each had its own advantages and limitations in theory or practice, and they cannot be replaced with each other completely. Therefore, using different method reflecting by agronomic characterization or molecular data, the decision should be based on one or several methods selected in detecting and evaluation of the genetic diversity. Analysis of genetic diversity may be help to select the elite and fit parents for improving breeding efficiency.

4.4 Detection of loci variation

As exactly as most crops (Fraser et al., 2004; Jung et al., 2005; Aggarwal et al., 2007), Feng et al. (2009) reported that DNRs were the main repeat motif, and AG / TC were the predominant DNRs; In this study, 61 sequences were recovered from 5 loci, of which 44 sequences belonged to AG / TC repeats.

The changes of flanking regions might lead to SSR loci variation, Feng et al. (2009) found that there were point mutations and deletions occurred in the flanking regions in rubber tree, and allelic variations were due to the most frequent InDels (insertions and deletions) in maize flanking regions (Matsuoka et al., 2002). On the contrary, Xie et al. (2006) pointed out that there were no insertion or deletion mutation occurred in AG/CT repeat loci flanking regions for almond, and the similarity results were found in *A. thaliana* (Symonds & Lloyd, 2003). In this study, of the four AG/CT repeat loci in flanking regions, no insertions or deletions mutation occurred in HBE008, however in HBE164, Dafeng117 and AC/T/15/114 were found AA and GAAA insertion respectively, and deletions occurred in HBE187 and HBE199.

Gutierrez et al. (2005) reported that variation was mainly due to the change of the repeat number and insertion, deletion mutation and base substitution in *Medicago truncatula*, insertion and deletion mutation also led to sequence variation (Feng et al., 2009); Symonds & Lloyd (2003) pointed out that interruptions in the repeat regions of most SSR loci were associated with shortening of the original repeat length in *A. thaliana*. In rubber tree, substitutions can be found in many alleles, and a complete and long-repeat sequence was divided into several smaller repeats or become relatively short. For example, CT repeats were shortened for C was replaced by G or A in HBE008, and AT repeats were divided into smaller ones for A was replaced by C as well, in HBE199, A was replaced by C or T led the repeat units into several smaller ones.

Dieringer & Schlötterer (2003) assumed that a length-independent mutation process operates on short SSR loci; and Xie et al. (2006) believed that the point mutation in flanking

sequence might lead to a new SSR repeat unit. But in rubber, interestingly, point mutations frequently occurred in flanking sequence of different loci, in HBE063, A was replaced by C, in HBE187, A by G, in HBE87, C by T, and in HBE199, A/T and T/A were replaced with each other. Long repeat sequences are more frequently targets for mutation (Johannsdottir et al., 2000), Symonds & Lloyd (2003) and Xie et al., (2006) have proved that a SSR motif with more repeats should provide an even more efficient substrate for rapid mutation rate in comparison with SSR motifs containing fewer repeats. In rubber tree, more variations occurred in HBE063 supported the above view.

5. Conclusion

A total of 43 alleles were obtained from 10 EST-SSR and 30 alleles from 6 genomic SSR (gSSR) primer pairs across cultivated and wild accessions; and HBE280, M197 and MnSOD were the most informative SSR markers. All the cultivated and wild accessions were clustered into two big groups. On the level of similarity coefficient 0.68, wild accessions were distinguished from cultivars. Sixty one sequences were sequenced from 5 EST-SSR loci. In comparison with the original EST sequences, insertion, deletion, conversion and transversion mutation occurred in SSR repeats and flanking regions, and long repeat sequences had more variations, and point mutation frequently occurred in flanking regions indicated the new SSR loci in rubber tree.

6. Acknowledgment

7. References

Aggarwal, R.K., Hendre, P.S., Varshney, R.K., Bhat, P.R., Krishnakumar, V. & Singh, L. (2007) Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theor Appl Genet*, Vol. 114, no. 2, (November 2006), pp. 359-372, ISSN: 0040-5752

Aidi-Daslin, (2002) Report on the evaluation and utilization of the 1981 IRRDB *Hevea* germplasm in Indonesia. Proceedings of *IRRDB Plant Breeding, Agronomy and Socio-Economic Joint Workshop*, Malaysia & Indonesia, 28 Aug-7 Sept, 2002

Angers, B. & Bernatchez, L. (1997) Complex evolution of a Salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Molec. Biol Evol*, Vol. 14, no. 3, (November 1996), pp. 230-238, ISSN: 0737-4038

Benong, M. (2002) Status on the evaluation and utilization of the 1981 IRRDB *Hevea* germplasm. Proceedings of *IRRDB Plant Breeding, Agronomy and Socio-Economics Joint Workshop*, Malaysia & Indonesia, 28 Aug-7 Sept, 2002

Besse, P., Seguim, M. & Lebrun, P. (1993) Ribosomal DNA variations in wild and cultivated rubber tree (*Hevea brasiliensis*). *Genome*, Vol. 36, no. 3, (December 1993), pp. 1049-1057, ISSN: 0831-2796

Besse, P., Seguin, M., Legrun, P.,Chevallier, M.H., Nicolas, D. & Lanaud, C. (1994) Genetic diversity among wild and cultivated populations of *Hevea brasiliensis* assessed by

nuclear RFLP analysis. *Theor Appl Genet*, Vol. 88, no. 2, (November 2004), pp. 199-207, ISSN: 0040-5752

Brantestam, A.K., Bothmer, R.V., Dayteg, C., Rashal, I., Tuvesson, S. & Weibull, J. (2007) Genetic diversity changes and relationships in spring barley (Hordeum vulgare L.) germplasm of Nordic and Baltic areas as shown by SSR markers. *Genet Resour Crop Evol*, Vol. 54, no. 4, (January 2007), pp. 749-758, ISSN: 0925-9864

Caniato, F.F., Guimaraes, C.T., Schaffert, R.E., Alves, V.M.C., Kochian, L.V., Borem, A., Klein, P.E. & Magalhaes, J.V. (2007) Genetiv diversity for aluminum tolerance in sorghum. *Theor Appl Genet*, Vol. 114, no. 5, pp. 863-876, ISSN: 0040-5752

Chevallier, M.H. (1998) Genetic variability of *Hevea brasiliensis* germplasm using isozyme markers. *Journal of Rubber Research*, Vol. 3, no. 1, pp. 42-53, ISSN: 1511-1768

Choi, I.Y., Kang, J.H., Song, H.S. & Kim, N.S. (1999) Genetic diversity measured by simple sequence repeat variations among the wild soybean, *Glycine soja*, collected along the riverside of five major rivers in Korea. *Genes Genet Syst*, Vol. 74, no. 4, pp. 169-177, ISSN: 1880-5779

Clement-Demange, A (2002) Report on the IRRDB 1981 Hevea germplasm by Cirad (France). Proceedings of *IRRDB Plant Breeding, Agronomy and Socio-Economics Joint Workshop*, Malaysia & Indonesia, 28 Aug-7 Sept, 2002

Davierwala, A.P., Ramakrishna, W., Chowdari, V., Ranjekar, P.K. & Gupta, V.S. (2001) Potential of (GATA)n microsatellites from rice for inter-and intraspecific variability studies. *BMC Evolutionary Biology*, Vol. 1, pp. 7

Dieringer, D. & Schlötterer, C. (2003) Two distinct modes of microsatellite mutation processes: evidence from the complete genomic sequences of nine species. *Genome Res.*, Vol. 13, pp. 2242-2251, ISSN: 1088-9051

Dreisigacker, S., Zhang, P., Warburton, M.L. & Ginkel, M.V. (2004) SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different megaenvironments. *Crop Sci.*, Vol. 44, no. 2, pp. 381-388, ISSN: 0011-183X

Estoup, A., Garnery, L., Solignac, M. & Gournuet, J.M. (1995) Microsatellite variation in honey bee (*Apis mellifera L.*) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics*, Vol. 140, no. 1, pp. 679-695, ISSN: 0016-6731

Eujayl, I., Sorrells, M., Baum, M., Wolters, P. & Powell, W. (2001) Assessment of genotypic variation among cultivated durum wheat based on EST-SSRs and genomic SSRs. *Euphytica*, Vol. 119, no. 1-2, (December 2004), pp. 39-43, ISSN: 0014-2336

Eujayl, I., Sorrells, M.E., Baum, M., Wolters, P. & Powell, W. (2002) Solation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theor Appl Genet*, Vol. 104, no. 2-3, (February 2002), pp. 399-407, ISSN: 0040-5752

Feng, S.P., Li, W.G., Huang, H.S., Wang, J.Y. & Wu, Y.T. (2009) Development, application and cross-species/genera transferability of EST-SSR makers for rubber tree (*Hevea brasiliensis*). *Mol Breeding*, Vol. 23, no. 1, (August 2008), pp. 85-97, ISSN: 1380-3743

Feng, S.P., Li, W.G., Wang, J.Y. & Wu, Y.T. (2008) The elementary application of SSR loci variation for *Hevea brasiliensis* taxonomic relationships. *Journal of Biotechnology*, Vol. 136S, pp. S625, ISSN: 0168-1656

Fraser, L.G., Harvey, C.F., Crowhurst, R.N. & Silva, H.N. (2004) EST-derived microsatellites from Actinidia species and their potential for mapping. *Theor Appl Genet*, Vol. 108, no. 6, (February 2004), pp. 1010–1016, ISSN: 0040-5752

Gonzalo, M.J., Oliver, M., Garcia-Mas, J., Monfort, A., Dolcet-Sanjuan, R., Katzir, N., Arus, P. & Monforte, A.J. (2005) Simple-sequence repeat markers used in merging linkage maps of melon (*Cucumis melo* L.). *Theor Appl Genet*, Vol. 110, no. 5, (February 2005), pp. 802-811, ISSN: 0040-5752

Gouvêa L.R.L., Rubiano L.B., Chioratto A.F., Zucchi M.I. & Gonçalves P.S. (2010) Genetic divergence of rubber tree estimated by multivariate techniques and microsatellite markers. *Genetics and Molecular Biology*, Vol. 33, no. 2, pp. 308-318, ISSN: 1415-4757

Grimaldi, M.-C. & Crouau-Roy, B. (1997) Microsatellite allelic homoplasy due to variable flanking sequences. *J. Mol. Evol.*, Vol. 44, no. 3, (March 1997), pp. 336-340, ISSN: 0022-2844

Gupta, P.K. & Varshney, R.K. (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. *Euphytica*, Vol. 113, no. 3, (November 2004), pp. 163-185, ISSN: 0014-2336

Gutierrez, M.V., Vaz Patto, M.C., Huguet, T., Cubero, J.I., Moreno, M.T. & Torres, A.M. (2005) Cross-species amplification of Medicago truncatula microsatellites across three major pulse crops. *Theor Appl Genet*, Vol. 110, no. 7, (April 2005), pp. 1210-1217, ISSN: 0040-5752

Hao, C.Y., Zhang, X.Y., Wang, L.F., Dong, Y.S., Shang, X.W. & Jia, J.Z. (2006) Genetic diversity and core collection evaluations in common wheat germplasm from the Northwestern Spring Wheat Region in China. *Molecular Breeding*, Vol. 17, no. 1, (February 2006), pp. 69-77, ISSN: 1380-3743

Hasan, M., Seyis, F., Badani, A.G., Pons-Kuhnemann, J., Friedt, W., Luhs, W. & Snowdon, R.J. (2006) Analysis of genetic diversity in the *Brassica napus L.* gene pool using SSR markers. *Genetic Resources and Crop Evolution*, Vol. 53, no. 4, (October 2005), pp. 793-802, ISSN: 0925-9864

Johannsdottir, J.T., Jonasson, J.G., Bergthorsson, J.T., Amundadottir, L.T., Magnusson, J., Egilsson, V. & Ingvarsson, S. (2000) The effect of mismatch repair deficiency on tumourigenesis; microsatellite instability affecting genes containing short repeated sequences. *Int J Oncol.*, Vol. 16, no. 1, pp. 133-139, ISSN: 1019-6439

Jung, S., Albert, A., Jesudurai, C., Tomkins, J. & Main, D. (2005) Frequency, type, distribution and annotation of simple sequence repeats in *Rosaceae* ESTs. *Functional & Integrative Genomics*, Vol. 5, no. 3, (March 2005), pp. 136-143, ISSN: 1438-793X

Kloosterman, A.D., Budowle, B. & Daselaar, P. (1993) PCR-amplification and detection of the human D1S80 VNTR locus. Amplification conditions, population genetics and application in forensic analysis. *Int J Leg Med*, Vol. 105, no. 5, (April 2005), pp. 257-264, ISSN: 0937-9827

Lam L.V., Thanh T., Chi V.T.Q. & Tuy L.M. (2009) Genetic Diversity of *Hevea* IRRDB'81 Collection Assessed by RAPD Markers. *Mol Biotechnol*, Vol.42, no. 3, (March 2009), pp. 292–298, ISSN: 1073-6085

Leigh, F., Lea, V., Law, J., Wolters, P., Powell, W. & Donini, P. (2003) Assessment of EST- and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. Euphytica, Vol. 133, no. 3, (October 2004), pp. 359-366, ISSN: 0014-2336

Lekawipat, N., Teerawatanasuk, K., Rodier-Goud, M., Seguin, M., Vanavichit, A., Toojinda, T. & Tragoonrung, S. (2003) Genetic diversity analysis of wild germplasm and

cultivated clones of Hevea brasiliensis Müell. Arg. using microsatellite markers. *Journal of Rubber Research*, Vol. 6, pp. 36-47, ISSN:1511-1768

Lespinasse, D., Rodier-Goud, M., Grivet, L., Leconte, A., Legnate, H. & Seguin, M. (2000) A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. Theor Appl Genet: Vol. 100, no. 1, (January 2000), pp. 127-138, ISSN: 0040-5752

Liewlaksaneeyanawin, C., Ritland, C.E., El-Kassaby, Y.A. & Ritland, K. (2004) Single-copy, species-transferable microsatellite markers developed from loblolly pine ESTs. *Theor Appl Genet*, Vol. 1109, no. 2, (March 2004), pp. 361-369, ISSN: 0040-5752

Liu, J.C., Liu, L., Hou, N., Zhang, A. & Liu, C.G. (2007) Genetic diversity of wheat gene pool of recurrent selection assessed by microsatellite markers and morphological traits. *Euphytica*, Vol. 155, no. 1-2, (December 2006), pp. 249-258, ISSN: 0014-2336

Marco, M., Luca, P.i., Nadia, L.S.B. & Vincenzo, V. (2007) Genetic diversity assessment in Somali sorghum (*Sorghum bicolor (L.*) Moench) accessions using microsatellite markers. *Biodivers Conserv*, Vol. 16, no. 6, pp. 1715–1730, ISSN:0960-3115

Matsuoka, Y., Mitchell, S.E., Kresovich, S., Goodman, M. & Doebley, J. (2002) Microsatellites in Zea-variability, patterns of mutations, and use for evolutionary studies. *Theor Appl Genet*, Vol. 104, no. 2-3, (February 2002), pp. 436-450, ISSN: 0040-5752

Oktavia F. & Kuswanhadi L.M. (2011) Genetic relationship of Wickham and IRRDB 1981 rubber population based on RAPD markers analysis. *HAYATI Journal of Biosciences*, Vol. 18, no. 1, pp. 27-32, (March 2011), ISSN: 2086-4094

Pinto, L.R., Oliveira, K.M., Marconi, T., Garcia, A.A.F., Ulian, E.C. & deSouza, A.P. (2006) Characterization of novel sugarcane expressed sequence tag microsatellites and their comparison with genomic SSRs. *Plant Breeding*, Vol. 125, no. 4, pp. 378-384, ISSN:1439-0523

Saha T., Roy C.B. & Nazeer M.A. (2005) Microsatellite variability and its use in the characterization of cultivated clones of *Hevea brasiliensis*. *Plant breeding*, Vol. 124, no. 1, (June 2008), pp. 86-92 ISSN: 1439-0523

Seguin, M., Besse, P., Rodier-Goud, M., Nicolas, D., Lebrun, P. & Lespinasse, D. (1996) *Hevea* molecular genetics. *Plantations Rech. Dev.*, Vol. 3, no. 2, pp. 77-88

Seguin, M., Gay, C., Chen, X.T. & Rodier-Goud, M. (2002) Microsatellite markers for genome analysis of rubber tree (Hevea spp.). *The IRRDB Symposium 2001 Biotechnology & Rubber Tree*, 25-28 September 2001. Montpellier, France (http://tropgenedb.cirad.fr/cgi-bin/generic/tree/rubbertree?name=4&class=Reference)

Sneath, P.H.A. & Sokal, R.R. (1973) Numerical taxonomy. Freeman W.H. and Company (eds) ISBN: 0-7167-0697-0, San Francisco, USA

Song, Z.P., Xu, X., Wang, B.,Chen, J.K. & Lu, B.-R. (2003) Genetic diversity in the northernmost *Oryza rufipogon* population estimated by SSR markers. *Theor Appl Genet*, Vol. 107, no. 8, (February 2004), pp. 1492-1499, ISSN: 0040-5752

Steinkellner, H., Fluch, S., Lexer, C., Turetschek, E., Streiff, R. & Glossl, J. (1997) Identification and characterization of microsatellite marker in *Quercus petraea*. *Plant Mol. Biol.*, Vol. 33, no. 6, (October 2004), pp. 1093-1096, ISSN: 0167-4412

Sun, G., Diaz, O., Salomon, B. & Bothmer, R.V. (1999) Genetic diversity in *Elymus caninus* as revealed by isozyme, RAPD and microsatellite markers. *Genome*, Vol. 42, no. 3, pp. 420-431, ISSN: 0831-2796

Symonds, V.V. & Lloyd, A.M. (2003) An analysis of microsatellite loci in *Arabidopsis thaliana*: mutational dynamics and application. *Genetics*, Vol. 165, (November 2003), pp. 1475-1488, ISSN: 0016-6731

Tams, S.H., Bauer, E., Oettler, G. & Melchinger, A.E. (2004) Genetic diversity in European winter triticale determined with SSR markers and coancestry coefficient. *Theor Appl Genet*, Vol. 108, no. 7, (February 2004), pp. 1385–1391, ISSN: 0040-5752

Varghese, Y.A. (2002) Management of the 1981 IRRDB germplasm collection in India. Proceedings of *IRRDB Plant Breeding, Agronomy and Socio-Economic Joint Workshop*, Malaysia & Indonesia, 28 Aug-7 Sept, 2002

Varshney, R.K., Graner, A. & Sorrells, M.E. (2005) Genic microsatellite markers in plants: features and applications. *Trends Biotech*, Vol. 23, no. 1, (January 2005), pp. 48-55, ISSN: 0167-7799

Venkatachalam, P., Thomas, S., Priya, P., Thanseem, I., Gireesh, T.,Saraswathyamma, C.K. & Thulaseedharan, A. (2002) Identification of DNA polymorphism within the clones of *Hevea brasiliensis* (Muell.) Arg. using RAPD analysis. *Indian J Nat Rubb Res*, Vol. 15, pp. 172-181

Virk, P.S., Poomi, H.S., Syed, N.H. & Kearsey, M.J. (1999) Fast and reliable genotype validation using microsatellite markers in Arabidposis thaliana. *Theor Appl Genet*, Vol. 98, no. 3-4, (March 1999), pp. 462-464, ISSN: 0040-5752

Xie, H., Sui, Y., Chang, F.Q., Xu, Y. & Ma, R.C. (2006) SSR allelic variation in almond (*Prunus dulcis* Mill.). *Theor Appl Genet*, Vol. 112, no. 2, (December 2005), pp. 366-372, ISSN: 0040-5752

Yang, X.Q., Liu, P., Han, Z.F., Ni, Z.F., Liu, W.Q. & Sun, Q.X. (2005) Comparative analysis of genetic diversity revealed by genomic-SSR, EST-SSR and pedigree in wheat (*Triticum asetivum* L.). *Acta Genetica Sinica*, Vol. 32, no. 4, (April 2005), pp. 406-416, ISSN: 0379-4172

Yu F., Wang B.H., Feng S.P., Wang J.Y., Li W.G. & Wu Y.T. (2011) Development, characterization, and cross-species/genera transferability of SSR markers for rubber tree (*Hevea brasiliensis*). *Plant Cell Rep*, Vol. 30, no. 3, (October 2010), pp. 335–344, ISSN: 0721-7714

Genetic Diversity and Transferability of Rubus Microsatellite Markers to South American *Rubus* Species

Marta L. Marulanda*, Ana M. López and Marcela Uribe
Plant Biotechnology Laboratory, Universidad Tecnológica de Pereira
Apartado Aéreo, Pereira
Colombia

1. Introduction

Rubus L. (Rosaceae) is grown extensively worldwide, in altitudes ranging from 0 to 4500 m above sea level. Found in six continents, this genus is reported to contain approximately 800 species due to biological processes such as hybridization, apomixis, and polyploidy that weaken species boundaries (Thompson, 1995). *Rubus* has been divided into 12 subgenera of which only a few species have been domesticated (Table 1).

Subgenus	Number of species
Anoplobatus (Focke) Focke	6
Chamaebatus (Focke) Focke	5
Chamaemoras (Hill) Focke	1
Comaropsis (Rich.) Focke	2
Cylactis (Raf.)Focke	14 (4 series)
Dalibarda (L.) Focke	5
Dalibardastrum Focke	4
Idaeobatus (Focke) Focke	117 (9 sections)
Lampobatus Focke	10
Malachobatus (Focke) Focke	115 (7 sections)
Orobatus (Focke)	19
Rubus L. (= *Eubatus* Focke)	132 (6 sections)

Table 1. Subgenera and number of species of *Rubus* (Focke, 1910, 1911, 1914).

The subgenus *Idaeobatus* contains the "raspberries" that are distributed in the Northern Hemisphere, mainly Asia, Africa, Europe, and North America. The subgenus *Rubus* includes species found in Europe, Asia, and North America, whereas the subgenus *Orobatus* is

---

* Corresponding Author

exclusive to South America. Species representative of the subgenera *Rubus, Orobatus*, and *Idaeobatus* are found in the Colombian and Ecuadorian Andes (Ballington et al., 1993). Vargas (2002) reported six species in Colombia's central Andes: *R. bogotensis*, *R. glaucus*, *R. macrocarpus*, *R. nubigenus*, *R. porphyromallus*, and and *R. urticifolius*. Romoleroux (1992) also reported the existence of nine edible species in Colombia of a total of 44 species reported, and that due to natural crossing, up to 500 varieties can be identified. Besides its ecological relevance, for example as an invasive weed, this genus is an economically important fruit crop for small farmers in the northern Andes, mainly because of the production of *Rubus glaucus* L., commonly known as the Andean blackberry (Fig. 1).

Fig. 1. Plants of *Rubus glaucus* L. growing on the Andes.

Of flowering plant species, the genus *Rubus* is one of the most complexes because of its hybridization, polyploidy, agamospermy, and loss of the concept of species (Alice & Campbell, 1999). Polyploidy and hybridization prevails in the subgenus *Rubus*. Only subgenera *Idaeobatus, Dalibarda*, and *Anoplobatus* are predominantly diploid, whereas *Dalibardastrum, Malachobatus*, and *Orobatus* are exclusively polyploid (Thompson, 1995, 1997). Hybridization occurs in *Rubus* between closely related species (Kraft et al., 1996; Naruhashi, 1979, 1990; Steele & Hodgdon, 1963, 1970) and, in some cases, between subgenera (Alice et al., 1997; Gustafsson, 1942; Jennings, 1978; Weber, 2003).

This species presents traits of two subgenera—*Idaeobatus* and *Rubus*—possibly being a fertile amphidiploid or allotetraploid (n=7, 4x=28) (Delgado et al., 2010; Thompson, 1997). Sympatrically with *R. glaucus*, some other wild *Rubus* species are found in the Andean cordillera along with the introduced and cultivated Eurasian *R. idaeus* L. Because hybridization is a common process that affects species of any genus (Randell et al., 2004), it is reasonable to believe that gene flow is currently taking place between species of this genus.

Studies on the genetic diversity of *Rubus* have been carried out in temperate species, such as *Rubus idaeus* (Graham & Mcnicol, 1995; Graham et al., 1997; Parent & Fortin, 1993) and *Rubus occidentalis* (Parent & Page, 1998), and Asian species (Amsellem et al., 2000). These

works used RAPD, RFLP, and SCAR markers as well as SSR (Antonius-Klemola, 1999). ITS are also been used to study hybrids of *R. idaeus* and *R. caesius* (Alice et al., 1997) .These markers made it possible to confirm the genetic origin of the hybrids and further phylogenetic and evolutionary studies in *Rubus* (Alice, 2002). Recently, major advances have been achieved worldwide in the use of molecular markers in temperate species of *Rubus*, such as DNA fingerprinting to study and characterize genotypes, development of linkage maps, marker-assisted selection, and mapping of QTLs (Antonius-Klemola, 1999; Graham et al., 2002).

To date, molecular markers such as randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP) (Marulanda et al., 2007), and microsatellites (simple sequence repeats, SSRs) (Marulanda et al., 2011) have been used to study the genetic diversity of the Andean blackberry. Previous work carried out by Marulanda et al. (2007) and Marulanda & López (2009) on the genetic diversity of Colombian blackberries identified high phenotypic and molecular plasticity in the *R. glaucus* species known as the "Castilla" blackberry in Colombia's central Andean area. Other wild *Rubus* species present in the Andean region are found near farms where the "Castilla" blackberry is commercially grown. These plants were also submitted to morphological, agronomic, and molecular characterizations using AFLP and SSR molecular markers (Marulanda & López, 2009). A genomic library enriched for microsatellite sequences was recently developed for *R. glaucus.*

This chapter presents the results of the molecular characterization of wild and cultivated *Rubus* species collected in the central Andes of Colombia using SSR markers from other *Rubus* species available in Genbank, together with 11 microsatellite markers isolated from *R. glaucus* and characterized in 39 samples of *Rubus* (Table 2). Intra- and inter-specific differences between *R. glaucus* and its wild relatives were established, generating not only information on the current status of populations, their uses, and distribution, but also information considered crucial to launch a breeding program for *R. glaucus*.

2. Materials and methods

2.1 Plant material and DNA extraction

A total of 39 *Rubus* samples were collected at altitudes ranging from 1800 to 2455 m above sea level in the central Andes of Colombia (between 1° 42´10.7 ´´ and 6° 99¨44´´N and 72° 92´80´´ and 76° 25´ 35.9´´ W), and placed on silica gel (1:10, plant tissue: silica gel) (Table 2). DNA was isolated using the Plant DNAeasy Miniprep kit (QIAGEN®), following the manufacturer's instructions. Several samples did not show any DNA after the isolation procedure so it was necessary to reprocess these samples following the Doyle & Doyle (1990) protocol. In all cases, samples were purified using the protocol described by Castillo (2006).

2.2 Analysis with SSR markers

A total of 36 microsatellite sequences from other *Rubus* species, *R. idaeus* (23 primer pairs) (Series RhM, RiM and Rubus) and *R. occidentalis* (2 primer pairs) (Series mRaCIRRI), and 11 microsatellites from *R. glaucus* were used (Table 3). The microsatellite named as "Rg" was

developed using a genomic library enriched for microsatellite sequences from a cultivated genotype of *R. glaucus*, following the protocol described by Billotte et al. (1999).

Code	Latitude (N)	Longitude (W)	Cultivated/Wild	Altitude (m.a.s.l.)	Characteristics
CVM 1	4°48′06′′	74°24′49′′	*R. glaucus* Cultivated	2455	Thorny
CVM 2	4°47′37′′	74°25′27′′	*R. glaucus* Cultivated	2337	Thorny
CVM 3	4°79′33′′	74°42′68′′	*R. glaucus* Cultivated	2288	Thorny
CVM 4	4°09′09′′	74°23′27′′	*R. glaucus* Cultivated	2007	Thorny
CVM 6	6°99′44′′	72°98′80′′	*R. glaucus* Cultivated	2157	Thorny
CVM 7	6°99′44′′	72°98′80′′	*R. glaucus* Cultivated	2157	Thornless
CVM 8	6°59′39.1′′	72°59′13′′	*R. glaucus* Cultivated	2176	Thorny
CVM 9	7°00′38′′	72°58′39′′	*R. glaucus* Cultivated	2357	Thorny
CVM 10	4°40′00.8′′	76°16′06.6′′	*R. glaucus* Cultivated	2049	Thorny
CVM 11	4°13′23.8′′	76°25′35.9′′	*R. idaeus* Raspberry	2100	Thornless
CVM 12	4°13′23.8′′	76°25′35.9′′	*R. glaucus* Cultivated	2380	Thorny
CVM 13	4°13′23.8′′	76°25′35.9′′	*R. glaucus* Cultivated	2380	Thorny
CVM 15	6°14′52.3′′	75°24′08.5′′	*R. glaucus* Cultivated	2150	Thorny
CVM 18	6°09′15.4′′	75°23′00.1′′	*R. glaucus* Cultivated	2000	Thornless
CVM 19	6°09′15.4′′	75°23′00.1′′	*R. glaucus* Cultivated	2000	Thorny
CVM 20	6°09′15.4′′	75°23′00.1′′	*R. glaucus* Cultivated	2000	Thorny
CVM 22	1°57′45.2′′	76°14′17.4′′	*R. glaucus* Cultivated	1840	Thorny
CVM 25	1°51′41.8′′	76°21′21.9′′	*R. glaucus* Cultivated	2140	Thorny
CVM 26	1°53′42.8′′	76°18′45.5′′	*R. glaucus* Cultivated	1850	Thorny
CVM 27	1°42′10.7′′	76°12′08.3′′	*R. glaucus* Cultivated	2000	Thorny
CVM WILD	_____	_____	*R. glaucus* Wild	_____	Wild
CVM A	4° 44′45.1′′	75° 36′39.6′′	*R. glaucus* Cultivated	2000	Thornless
CVM B	4° 39′7′′	75° 35′26.3′′	*R. glaucus* Cultivated	2014	Thorny
CVM C	4°38′36′′	75°28′41,5′′	*R. glaucus* Cultivated	2300	Thornless
CVM D	4°48′99.2′′	75°41′86′′	*R. glaucus* Cultivated	1950	Thorny
CVM E	5° 2′2.7′′	75° 27′10.5′′	*R. glaucus* Cultivated	1800	Thorny
95	4°21′22.5′′	75°42′10.1′′	*R. glaucus* Wild	1805	Wild
107	4° 33′51.9′′	75° 39′14.7′′	*R. urticifolius*	1800	Wild
106	4°39′7′′	75° 35′26.3′′	*R. urticifolius*	1574	Wild
97	4° 41′23.3′′	75°37′33.8′′	*R. glaucus* Wild	1997	Wild
86	4° 38′36′′	75° 29′9.5′′	*R. glaucus* Wild	2430	Wild
22	5° 4′ 35.8′′	75° 32′ 31.1′′	*R. glaucus* Cultivated	1947	Thorny
37	5° 0′50.3′′	75° 32′11.7′′	*R. urticifolius*	1777	Wild
44	4° 45′ 00.2′′	75° 36′ 39′′	*R. urticifolius*	1879	Wild
MSA 1	4°52′15.0′′	75°37′32.4′′	*R. glaucus* Cultivated		Thornless
MSA 2	4° 44′45.1′′	75° 36′39.6′′	*R. glaucus* Cultivated	1850	Thornless
MSA 3	4°38′36′′	75°28′41.5′′	*R. glaucus* Cultivated		Thornless
MSA 4	4°11′36.1′′	75°48′14.6′′	*R. glaucus* Cultivated	2000	Thornless
MSA 5	5°0′50.3′′	75°32′11.7′′	*R. glaucus* Cultivated	1825	Thornless

Table 2. Samples of *Rubus* species, accessions of *R. glaucus*, and collection sites.

Amplification reactions were performed in a final volume of 12.5 µl, with 5 ng genomic DNA, 0.3 µM of each primer, 1X reaction buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 200 µM of each dNTP, 2 mM $MgCl_2$ and 1.0 unit of Taq DNA polymerase. The PCR was performed according to the following parameters: 94 °C for 4 min, 10 cycles of 94 °C for 30 sec, 65 °C (−1 °C/cycle) for 30 sec and 72 °C for 1 min; 35 cycles of 94 °C for 15 sec, annealing temperature (°C) for 30 sec and 72 °C for 1 min; and 72 °C for 5 min.

2.3 Statistical analysis

The allelic diversity of the SSR was evaluated by determining polymorphism information content (PIC) value, as described by Bonstein et al. (1980) and cited and modified by Anderson et al. (1993), as described in Equation 1:

$$_n\text{PIC}_i = 1 - / \sum_{j=1} P_{ij^2} \tag{1}$$

Where P_{ij} is the frequency of the jth pattern, i is the sum, and n are the patterns. To measure the utility of the marker systems, average heterozygosity and expected and observed heterozygosity were calculated. The partitioning of genetic variation within and among the groups by the SSR marker system was achieved by analysis of molecular variance (AMOVA) using the same software. Analyses were performed in GenAlex (Peakall & Smouse, 2006) and Arlequin v.3.5 (Excoffier & Lischer, 2010).

3. Results and discussion

3.1 Genetic diversity and variability

A total of 41 loci and 133 alleles were detected. The number of alleles observed for each locus ranged from 2 to 6, with an average of 4.6 alleles per locus. The PIC value varied between 0.07 and 0.61 (average 0.48), and the discriminating power (D) ranged from 0.05 to 0.52 (average 0.27). Observed heterozygosities (Ho) were 0.078–1.0 (average 0.84) and expected heterozygosities (He) were 0.07-0.582 (average 0.473).The highest PIC value (0.61) was found in the Rg-D7 locus, which presented a high number of alleles (5). To compare the efficiency of markers to identify varieties, the D value of each primer was estimated. The highest D value (0.52) was also found in the Rg-D7 locus (Table 3).

Similar results were reported by Castillo (2006), who used 12 SSRs to analyze an extensive collection of North American *Rubus* (raspberry) germplasm. Results indicated from 3 to 16 alleles per locus, with an average of eight alleles per locus and a total number of alleles of 96. More recently, Flores et al. (2010) isolated 12 microsatellites from SSR-enriched genomic libraries of *R. idaeus.*

Another measure of genetic variability is the presence of exclusive alleles per loci and genotype (Tables 4 and 5). The genotypes presenting the highest number of exclusive alleles are listed in Table 4, with *R. idaeus* genotype CVM11 ranking highest, which could be attributed to the fact that most of the SSRs used are derived from *R. idaeus*. Of the evaluated microsatellites, those of the series "Rubus" were one of the most polymorphic groups and detected the highest number of alleles in the study population. It should be mentioned that exclusive alleles also appear in wild genotypes and in genotypes 106 and 107, which belong to the species *R. urticifolius*. The loci in which the private alleles were detected are very important for genotype differentiation, particularly in the case of the thornless genotypes.

Fig. 2 presents the results of the principal coordinates analysis. There is no clear differentiation of genotypes based on collection sites; however, the genotypes belonging to the species *R. urticifolius* (37, 44, 106, and 107) are separate from the species *R. glaucus*, both cultivated and wild. Likewise, the species *R. idaeus* is separate from the species *R. urticifolius*

	Locus	Motif	GenBank accession no.	Primer sequences (5'-3')	Size (bp)	T (°C)	Number of loci	PIC*	A*	D*	He*	Ho*
1	RhM018[3]	(CTT)6	FJ194447	F-CACCAATTGTACACCCAACAAC / R-GATTGTGAGCTGGTGTTACCAA	379	52	1	0.3510	2	0.5094	0.279	0.435
2	RhM023[3]	(CAT)5	FJ194449	F-CGACAACGACAATTCTCACATT / R-GTTATCAAGCGATCCTGCAGTT	196	52	1	0.4864	2	0.3388	0.463	0.835
3	RhM011[3]	(TC)18	FJ194446	F-AAAGACAAGGCGTCCACAAC / R-GGTTATGCTTTGATTAGGCTGG	280	58	2	0.5419	3	0.1448	0.530	0.969
4	RhM001[3]	(CA)7	FJ194444	F-GGTTCGGATAGTTAATCCTCCC / R-CCAACTGTTGTAAATGCAGGAA	232	50	1	0.1328	3	0.2148	0.117	0.078
5	RhM021[3]	(TC)6	FJ194448	F-CAGTCCCTTATACGATCCAACG / R-GAACTCCACCATCTCCTCGTAG	282	50	1	0.5461	3	0.219	0.536	1.000
6	RhM003[3]	(TG)10	FJ194445	F-CCATCTCCAATTCAGTTCTTCC / R-AGCAGAATCGGTTCTTACAAGC	200	48	1	0.5151	3	0.047	0.512	1.000
7	RiM017[3]	(TG)6	FJ194453	F-GAAACAGGTGGAAAGAAACCTG / R-CATTGTGCTTATCGATGGTTTCG	194	58	1	0.5869	3	0.2148	0.556	0.897
8	RiM015[3]	(ATC)5	FJ194452	F-CGACACCGATCAGAGCTAATTC / R-ATAGTTGCATTGGCAGGCTTAT	350	58	1	0.4462	2	0.4451	0.367	0.610
9	Rubus 76b[2]	(CT)5(CT)4	—	F-CTCACCCGAAATGTTCAACC / R-GGCTAGGCCGAATGACTACA	190-210	55	1	0.6004	4	0.2463	0.567	0.897
10	Rubus 16a[2]	(AT)8(GT)11	—	F-TGTTGTACGTGTTGGGCTTT / R-GGGTGTTTGCCAGTTTCAGT	169	55	3	0.3461	3	0.0588	0.533	1.000
11	Rubus 116a[2]	(CT)12(T)10	—	F-CCAACCCAAAAACCTTCAAC / R-GTTGTGGCATGGCCTTTTAT	299	55	2	0.5297	3	0.053	0.520	0.969
12	Rubus 105b[2]	(AG)8	—	F-GAAAATGCAAGGCGAATTGT / R-TCCATCACCAACACCACCTA	165-173	55	2	0.5924	6	0.2171	0.571	0.948
13	Rubus 137a[2]	(TG)8(TA)4	—	F-TGTGAGCAGAGTGAAGGAGCTA / R-AGCATTATTCGGCGCAGTTTT	198	55	1	0.4995	2	0.0465	0.496	0.969
14	Rubus 259f[2]	(CT)4 (AG)8	—	F-TGGCACAAGAAGCCTGTAAC / R-TCCCATATCCCTCAGCATTC	265	55	1	0.5717	4	0.1727	0.552	1.000
15	Rubus 98d[2]	(GAA)5(GA)10	—	F-GGCTTCTCAATTTGCTGTGTC / R-TGATTTGAAATCGTGCGGTTA	198	55	2	0.5880	4	0.2558	0.544	0.861
16	mRaCIRRI1G3[1]	(GA)28	AF205116	F-CTCTACAAAAGGATCTGCATGA / R-CAGCAAAAGTGAAATGGTTCA	195-265	55	2	0.5473	3	0.2566	0.524	0.881
17	mRaCIRRIV2A8[1]	(CA)12(CT)11	AF261693	F-TAAAAGGCGCAACAGTCG / R-AGACACAGAAAACAGCATCG	191-237	55	2	0.5604	3	0.23	0.548	0.959
18	Rubus 285a[2]	(TC)9	—	F-TCGAGAAGCTTGCTATGCTG / R-GGATACCTCAATGGCTTTCTTG	138-167	55	1	0.5497	4	0.2462	0.543	1.000

#	Marker	Repeat motif	GenBank	Primer sequences (5'–3')	Size (bp)	Ta	N	He	A	PIC	Ho	D
19	Rubus 145a[2]	$(GT)_7$	——	F-TGTCCCAGCTTTCTGCTTTC R-GGCATCTGTCGCGTAAAAAT	131	55	1	0.5365	3	0.2348	0.527	0.969
20	Rubus 262b[2]	$(AG)_{15}$	——	F-TGCATGAAGGCGATATAAAGG R-TCCGCAAGGGTTGTATCCTA	217-225	55	1	0.5497	4	0.2348	0.543	1.000
21	Rubus 107a[2]	$(AG)_8$	——	F-GCCAGCACCAAAAACCTACA R-TTTCACCGTCAAGAAGAAAGC	179	55	1	0.5990	5	0.2348	0.581	0.969
22	Rubus 251a[2]	$(GA)_{10}$	——	F-GCATCAGCCATTGAATTTCC R-CCCACCTCCATTACCAACTC	157-283	55	1	0.5949	3	0.3826	0.582	0.969
23	Rubus 270a[2]	$(GA)_{10}$	——	F-GCATCAGCCATTGAATTTCC R-CCCACCTCCATTACCAACTC	167-185	55	1	0.5485	3	0.2347	0.543	1.000
24	Rubus 2a[2]	$(GT)_{12}\text{-}G\text{-}(GT)_8$	——	F-TGAGGGAAGAAGAGGCAAGA R-CACGTGTGACCCCAATGATA	175-180	55	1	0.5485	3	0.2294	0.543	1.000
25	Rubus 194h[2]	$(GA)_{12}$	——	F-TGTGTTGTTCTCTGCAACCA R-AGCCCTTACTTTTCCTGCAA	100-110	55	1	0.4940	2	0.2347	0.488	0.891
26	Rg-A6[4]	$(GAA)_4$	——	F-AGCGCAAGGACTTCTACC R-TTCAGCTCCGTAGTAGC	282-293	55	1	0.4969	2	0.1889	0.492	0.922
27	Rg-A12-1[4]	$(GA)_{23}$	HQ637525	F-TCATGTTCATGTTCGTGTG R-CTAACACGCGATAGAATAGG	387-432	48	1	0.3573	4	0.4438	0.296	0.137
28	Rg-A12-2[4]	$(AC)_8$	HQ637525	F-GCCGGGCATTCTTCTTGCTTAC R-GCCGGTTCGTGACTCAGACAG	179-194	58	1	0.4905	2	0.2294	0.472	0.863
29	Rg-B7[4]	$(CA)_{11}$	HQ637519	F-TACGGCGTGGCACTAACAC R-TCAGAGCACTCCAAGAAGG	150-171	58	2	0.3711	4	0.4651	0.304	0.183
30	Rg-B8[4]	$(AG)_{23}$	HQ637493	F-CTGATGAGTAGCTGCTTTCTC R-GTCTGAAATAGCACATCGAC	185-230	52	1	0.0762	2	0.1889	0.070	0.078
31	Rg-B9[4]	$(CA)_9$	HQ637500	F-TTCAGAAACCTCATGCAGAGCC R-TTCACTCAGGAGGAGGAGGTG	348-367	57	1	0.4309	2	0.5391	0.392	0.566
32	Rg-D2[4]	$(GA)_{16}$	HQ637534	F-AAACCCTAATTTGTTGAC R-CCCAAGCTAGATTTCTCTC	418-449	48	1	0.0747	2	0.1889	0.070	0.078
33	Rg-D7[4]	$(GA)_{25}$	HQ637494	F-AACCATCGGTGTCGACCTC R-GGCTCTTGACCGGTAACTTAG	61-110	55	1	0.6115	5	0.5155	0.563	0.891
34	Rg-F3[4]	$(GA)_{17}(GT)_9$	HQ637516	F-GCCAATGAAACGCGAAAGAC R-GCCTTCACTCATATCATTCTCC	210-265	60	1	0.5644	4	0.2348	0.550	0.969
35	Rg-F7[4]	$(CT)_{21}$	HQ637527	F-GCAGGTGTGTCTATGTGATATG R-ATTCCCAACGACTACATAATTC	290-331	55	1	0.5267	3	0.2348	0.511	0.891
36	Rg-F8[4]	$(GA)_{14}$	HQ637499	F-CATGCTGGCTTTGCTGCTC R-TGCTTACGGCGTGGACTAACAG	267-294	60	1	0.2810	4	0.3981	0.245	0.123

[1] *R. alceifolius* (Amsellem et al., 2001); [2] *R. idaeus* (Graham et al., 2002, 2004) ; [3] *R. idaeus* and blackberry "Marion" (Castillo, 2006); [4] *R. glaucus* (Marulanda et al., 2010 (data not published).

* PIC = polymorphic information content; A = allele number; D = discrimination power; He = expected heterozygosity; Ho = observed heterozygosity.

Table 3. Microsatellite sequences and characteristics of each SSR used to evaluate *Rubus* materials.

and from cultivated and wild genotypes of *R. glaucus*. Genotype CVM Wild of the species *R. glaucus* is notably separate from the genotypes of the same species.

Locus	Private alleles	Frequency
RhM003	2	0.125
RhM001	2	0.250
Rubus 76b	1	0.125
Rubus 98d	1	0.333
Rubus 98d	1	0.333
mRaCIRRI1G3	2	0.125
mRaCIRRIV2A8	5	0.125
Rubus 128a	3	0.125
Rubus 262b	3	0.125
Rubus 107a	5	0.250
RgA12-1	4	0.125
RgB7	2	0.125
RgB7	1	0.125
RgB7	2	0.125
RgF3	1	0.250
RgF7	1	0.250
Rubus 116a	1	0.250
Rubus 105b	2	0.100
Rubus 105b	5	0.100
Rubus 105b	4	0.200
RgA12-1	2	0.100
Rubus 98d	3	0.222

Table 4. Summary of private alleles.

Sample	No. of loci with private alleles	Loci with private alleles
R. glaucus cultivated (CVM 10)	1	RgB7
R. idaeus (CVM 11)	14	RhM003, RhM001, Rubus 76b, Rubus 98d, Rubus 98d, mRaCIRRI1G3, mRaCIRRIV2A8, Rubus 128a, Rubus 262b, Rubus 107a, RgA12-1, Rg B7, RgF3, RgF7
R. glaucus cultivated (CVM 13)	1	RgB7
R. glaucus wild (CVM WILD)	3	Rubus 116a, Rubus 105b, Rubus 105b
R. glaucus cultivated (22)	1	RgA12-1
R. urticifolius (107)	1	Rubus 98d
R. urticifolius (106)	1	Rubus 98d

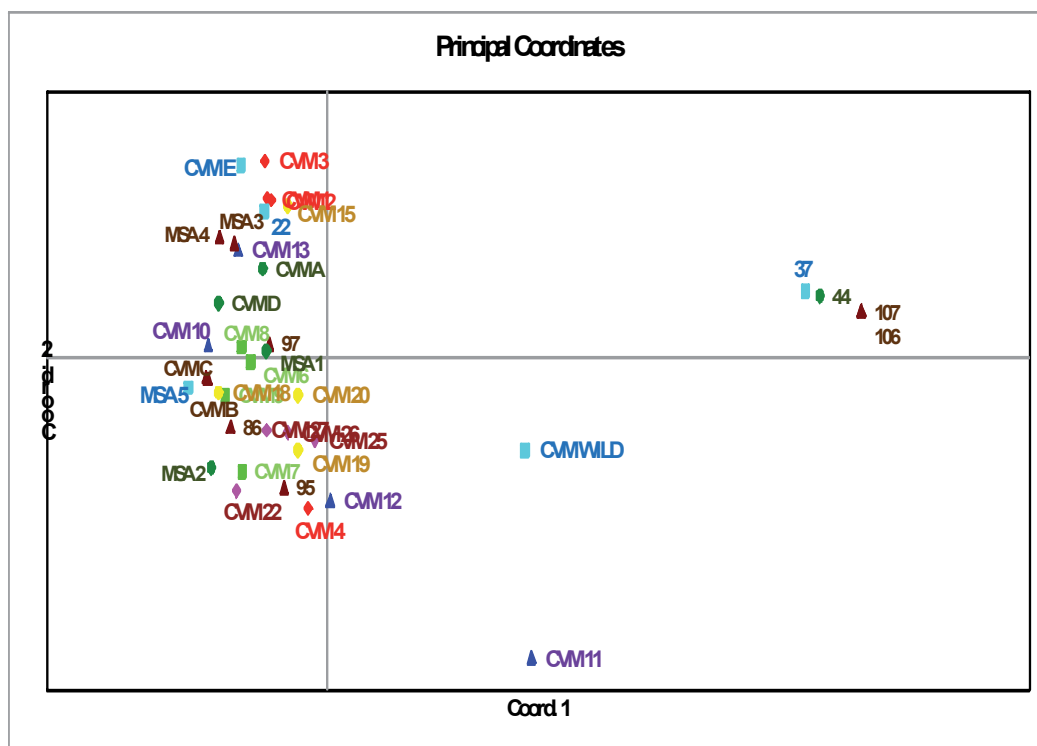Table 5. List of samples with one or more private alleles.

Fig. 2. Principal coordinates analysis among *Rubus* genotypes based on genetic distance .

In addition to the genetic diversity measurements already mentioned, the analysis of molecular variation (AMOVA) revealed 98%variability among all genotypes and 2% variation between populations (Michalakis & Excoffier, 1996) (Table 6). These data agree with those observed in the principal coordinates analysis, where variation is mostly attributed to individuals variation. Kollmann et al. (2000) concluded that genetic variability in *Rubus* is determined by the plant propagation system and demonstrated that there is an effect of cross-pollination between polyploid *Rubus* species. This type of crossbreeding influences seed and fruit quality positively, while increasing ploidy levels and taxonomic proximity.

Source of variation	df*	Sume of squeres	Variance components	Estandar Variation	% of variation
Among populations	7	18.067	2.581	0.054	2%
Within populations	31	72.005	2.323	2.323	98%
Total	38	90.071		2.377	100%

*df= degrees of freedom

Table 6. Summary of molecular variation (AMOVA) among *Rubus* genotypes.

3.2 Transferability of microsatellite sequences

Microsatellites RhM and RiM developed from *R. idaeus* showed cross-species amplification in *R. glaucus* genotypes and in other two wild species of *Rubus,* with 2 and 3 alleles per locus. High D values were found for loci RhM018 (0.5094) and RiM015 (0.4451). The "Rubus" microsatellite markers, also from *R. idaeus*, were the most polymorphic and presented the highest number of alleles per locus: Rubus 105b with 6 alleles and Rubus 107a with 5. The microsatellite markers from *R. alceifolius* showed 3 alleles and D values ranging from 0.23 to 0.2566. The *R. glaucus* microsatellites (Rg series) amplified for other wild species, with the number of alleles ranging from 2 to 5 per locus.

The cross-species amplification data suggested that microsatellites developed for *R. idaeus*, *R. alceifolius*, and *R. glaucus* can be potentially useful for genetic diversity studies of different *Rubus* species. In the case of conservation programs, they should prove useful for characterizing natural populations and germplasm collections, as well as for determining the degree of relatedness between individuals or groups of accessions.

The microsatellites developed by Amsellem et al. (2001) to study *R. alceifolius*, subgenus *Malachobatus,* that grows in Southeast Asia were used for the characterization of Andean *Rubus*, the transferability and applicability of microsatellites of *R. alceifolius* to study and evaluate the diversity of *Rubus* species in the American Andes were demonstrated; results were similar to those obtained in Asian species.

Amsellem et al. (2001) also observed amplification from 3 to 4 alleles per individual in the species *R. alceifolius*, confirming the suspicion that this is a tetraploid species. Based on the analysis carried out by Amsellem et al. (2001), the present study produced between 3 and 5 alleles for *R. glaucus* and between 2 and 4 alleles for *R. urticifolius*., suggesting that both cultivated and wild materials of *R. glaucus* have ploidy levels greater than those of *R. urticifolius*. This polyploidy was also described by Hall (1990), who explained that *Rubus* species used in plant breeding programs have produced euploid and aneuploid hybrids and that diploid, triploid, tetraploid, hexaploid, septaploid, octaploid, and nonaploid cultivars have been selected, most of them tetraploids.

4. Conclusion

Wild forms are also usually found at sites where *Rubus* species are cultivated, particularly in forest clearings, along roadsides, and on hillsides. Both cultivated and wild forms have the potential for interacting in different ways with cultivated materials. Cultivars can influence the genetic diversity of natural populations through gene transfer by pollen and wild populations are a potential source of genetic material for improvement programs.

This evaluation of the status of genetic resources of the species *R. glaucus* and related wild species serves to provide guidelines for conservation and breeding efforts aiming to promote the development of cultivated species important for the rural economies of South America's Andean region.

Using microsatellites from other *Rubus* species has proven to be a very useful strategy to differentiate between wild and cultivated *R. glaucus* genotypes, as well as between thorny and thornless cultivars.

The development of a genomic library enriched with microsatellites and the design of microsatellite sequences for the Andean specie *Rubus glucus,* is allowing a deeper comprehension of the genetic variability existing among cultivated and wild genotypes as well as the relationships between the cultivated specie and the wild relatives.

The Analysis of molecular variation (AMOVA) showed a higher variability distributed between genotypes than between populations, which agrees with the results obtained in the principal coordinates analysis.

5. Acknowledgment

6. References

Alice, L.A. (2002). Evolutionary relationships in *Rubus* (Rosaseae) based on molecular data. *Acta Horticulturae*, Vol.585, pp. 79–83, ISSN 0567-7572

Alice, L.A & Campbell, C.S (1999) Phylogeny of *Rubus* (Rosaceae) based on nuclear ribosomal DNA internal transcribed spacer region sequences. *American Journal of Botany*. 86:1:81-97

Alice, L.A.; Eriksson, T.; Eriksen, B. & Campbell, C.S. (1997). Intersubgeneric hybridization between a diploid raspberry, *Rubus idaeus*, and a tetraploid blackberry, *R. caesius* (Rosaceae). *American Journal of Botany*, Vol.84, pp. 171, ISSN 0002-9122

Amsellem L., Dutech C., Billotte N. (2001). Isolation and characterization of polymorphic microsatellite loci in *Rubus alceifolius* Poir. (Rosaceae), an invasive weed in La Reunion Island. *Molecular Ecology Notes*, Vol.1, pp. 33–35, ISSN 1471-8278

Amsellem, L.; Noyer, J.L.; Le Bourgeois, T.; Hossaert-Mckey, M. (2000). Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology*, Vol.9, pp. 443–455, ISSN 0962-1083

Anderson, J.A.; Churchill, G.A.; Autrique, J.E.; Tanksley, S.D. & Sorrells, M.E. (1993). Optimizing parental selection for genetic linkage maps. *Genome*, Vol.36, pp. 181-186, ISSN 0831-2796

Antonius-Klemola. K. (1999). Molecular markers in *Rubus* (Rosaceae) research and breeding. *The Journal of Horticultural Science & Biotechnology*, Vol.74, No.2, pp. 149–160, ISSN 1462-0316

Ballington, J. R.; Luteyn, M.M.; Thompson, K.; Romoleroux, K. & Castillo, R. (1993). *Rubus* and *Vacciniaceous* germplasm resources in the Andes of Ecuador. *Plant Genetic Resources Newsletter*, Vol.93, pp. 9–15, ISSN 1020-3362

Billotte, N.; Lagoda, P.J.L.; Risterucci, A.M. & Baurens, F.C. (1999). Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits*, Vol.54, No.4, pp. 277–288, ISSN 0248-1294

Bonstein, D.; White, R. L.; Skolnick, M.E. & Davis, R.W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. *American Journal of Human Genetics*, Vol.32, pp. 314–331, ISSN 0002-9297

Castillo, N.R. (2006). Fingerprinting and genetic stability of *Rubus* using molecular markers. M.Sc. Thesis. Oregon State University, USA

Delgado, L.M.; Uribe, M. & Marulanda, M.L. (2010). Estandarización de una técnica citogenética para el conteo de cromosomas mitóticos en *Rubus glaucus* Benth. *Scientia et Technica*, Vol.46, pp. 74–80, ISSN 0122-1701

Doyle, J. J. & Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, Vol.12, pp. 13–15

Excoffier, L. & Lischer, H.E.L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, Vol.10, pp. 564–567, ISSN 1755-098X

Flores, N. R.; Reed, B.M.; Graham, J.; Fernandez-Fernandez, F. & Bassil, N.V. (2010). Microsatellite markers for raspberries and blackberries. Journal of the American Society for Horticultural Science, Vol.135, No.3, pp. 271–278, ISSN 0003-1062

Focke, W.O. (1910-1914) Species *Ruborum*. *Bibliotheca Botanica*, Vol.72, pp. 1-274, ISSN 0067-7892

Graham, J. & Mcnicol, R.J. (1995). An examination of the ability of RAPD markers to determine the relationships within and between *Rubus* species. Theor. Appl Genet 90: 1128-1132.

Graham, J., Squire, G.R. Marshall, B. & Harrison, R.E. (1997). Spatially dependent genetic diversity within and between colonies of wild raspberry *Rubus idaeus* detected using RAPD markers. Molecular Ecology 6(11): 1001 – 1008.

Graham, J.; Smith, K.; Mackenzie, K.; Jorgenson, L.; Hackett, C.A. & Powell, W. (2004). The construction of a genetic linkage map of red raspberry (*Rubus idaeus* subsp. *idaeus*) based on AFLPs, genomic SSR and EST-SSR markers. Theoretical and Applied Genetics, Vol.109, pp. 740–749, ISSN 0040-5752

Graham, J.; Smith, K.; Woodhead, M. & Russell, J. (2002). Development and use of SSR markers in *Rubus* species. *Molecular Ecology*, Vol.2, pp. 250–252, ISSN 0962-1083

Graham, J.; Squire, G.R.; Marshall, B.Y. & Harrison, R.E. (1997). Spatially dependent genetic diversity within and between colonies of wild raspberry *Rubus idaeus* detected using RAPD markers. *Molecular Ecology*, Vol.6, No.11, pp. 1001–1008, ISSN 0962-1083

Gustafsson, A. (1942). The origin and properties of the European blackberry flora. *Hereditas*, Vol.28, pp. 249–277, ISSN 0018-0661

Hall, H.K. (1990). Blackberry breeding, In: *Plant Breeding Reviews* Vol. 8, J. Janick (Ed.), 249–312, Timber Press, ISBN 9780881921731, Portland, OR

Jennings, D.L. (1978). The blackberries of South America—an unexplored reservoir of germplasm. Fruit Varieties Journal, Vol.32, pp. 61–63, ISSN 0091-3642

Jennings, D.L. (1988). Raspberries and Blackberries: Their Breeding, Diseases and Growth, Academic Press, ISBN 0123842409, New York

Kollmann, J.; Steinger, T. & Roy, B.A. (2000). Evidence of sexuality in European *Rubus* (Rosaceae) species based on AFLP and allozyme analysis. *American Journal of Botany*, Vol.87, No.11, pp. 1592–1598, ISSN 0002-9122

Kraft, T. & Nybom, H. (1995) DNA fingerprint and biometry can solve some taxonomic problems in apomictic blackberries (*Rubus* sub *Rubus*) Watsonia 20:329-43

Kraft, T.; Nybom, H. & Werlemark, G. (1996). DNA fingerprint variation in some blackberry species *Rubus* sug. *Rubus*, Rosaceae. *Plant Systematics and Evolution*, Vol.199, pp. 93–108, ISSN 0378-2697

Marulanda, M.L. & López, A.M. (2009). Characterization of thornless *Rubus glaucus* in Colombia. *Canadian Journal of Pure & Applied Sciences*, Vol.3, No.2, pp. 927–937, ISSN 1715-9997

Marulanda, M.L.; López, A.M. & Aguilar, S.B. (2007). Genetic diversity of wild and cultivated *Rubus* species in Colombia using AFLP and SSR markers. *Crop Breeding and Applied Biotechnology*, Vol.7, pp. 243–253, ISSN 1518-7853

Marulanda M.L.; López, A.M. & Uribe, M. (2011). Molecular characterization of *Rubus* spp using SSR markers. *Genetic Molecular Research* (GMR) Brasil. (In press)

Michalakis , Y. & Excoffier, L. (1996). A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics, 142, 1061-1064.

Naruhashi, N. (1979). Notes on Japanese *Rubus* (3). *Journal of Phytogeography and Taxonomy*, Vol.27, pp. 38–40, ISSN 0388-6212

Naruhashi, N. (1990). Rubus 3 semi-nepalensis, a new natural hybrid from Nepal Himalaya. *Journal of Japanese Botany*, Vol.65, pp. 186–191, ISSN 0022-2062

Parent J. G. & M. G. Fortin. (1993). Identification of raspberry cultivars by random amplified polymorphic DNA (RAPD) analysis. *Canadian Journal of Plant Science* 73(4): 1115 – 1122.

Parent, J.G. & Page, D. (1998). Identification of raspberry cultivars by sequence characterized amplified region DNA analysis. *HortScience*, Vol.33, No.1, pp. 140–142, ISSN 0018-5345

Peakall, R. & Smouse, P. (2006). GENALEX 6: Genetic analysis in Excel. Population genetics software for teaching and research. *Molecular Ecology Notes,* Vol.6, pp. 288–295, ISSN 1471-8278

Randell, R.A.; Howarth, D.G. & Morden, C.W. (2004). Genetic analysis of natural hybrids between endemic and alien *Rubus* (Rosaceae) species in Hawaii. *Conservation Genetics*, Vol.5, No.2, pp. 217–230, ISSN 1566-0621

Romoleroux, K. (1992). Rosaceae in the paramos of Ecuador, In: *Paramo: An Andean Ecosystem under Human Influence*, H. Balslev & J.L. Luteyn (Eds.), 85-94, Academic Press, ISBN 0124604420, London

Steele, F.L. & Hodgdon, A.R. (1963). Hybridization of *Rubus hispidus* and *R. setosus*. *Rhodora*, Vol.65, pp. 262–270, ISSN 0035-4902

Steele, F.L. & Hodgdon, A.R. (1970). Hybrids in *Rubus* subgenus *Eubatus* in New England. *Rhodora,* Vol.72, pp. 240–249, ISSN 0035-4902

Thompson, M.M. (1995). Chromosome numbers of *Rubus* cultivars at the National Clonal Germplasm Repository. *HortScience*, Vol. 30, pp. 1453–1456, ISSN 0018-5345

Thompson, M.M. (1997). Survey of chromosome numbers in *Rubus* (Rosaceae: Rosoideae). *Annals of the Missouri Botanical Garden*, Vol. 84, pp. 128–164, ISSN 0026-6493

Vargas, W. (2002). *Guía Ilustrada de las Plantas de las Montañas del Quindío y los Andes Centrales*. Universidad de Caldas and Corporación Autónoma Regional del Quindío, ISBN 958-8041-38-4, Manizales, Colombia

Weber, C.A. (2003). Genetic diversity in black raspberry detected by RAPD markers. *HortScience*, Vol.38, No.2, pp. 269–272, ISSN 0018-5345

Genetic Diversity of Seed Orchard Crops

Murat Ertekin
Bartın University, Faculty of Forestry
Turkey

1. Introduction

Different climate types enable local or exotic plantations to be established. Thus, the opportunity to establish stands of high quality and quantity is increasing. In establishing a plantation, it is necessary to select appropriate species, obey improvement principles, consider nursery and plantation techniques, and take into account economic and social issues. Selecting genetically high-quality seed sources and performing plantation improvement studies are of particular importance (Ürgenç, 1982; Alptekin, 1986; Ertekin, 2006). Genetic diversity among seeds is an important aspect of plantation forestry, especially when using improved seeds (e.g., seed orchard crops). Many currently established plantations originated from seed orchard crops. The genetic structure of these plantation forests is the same as the original seed orchard structure; therefore, the rich genetic diversity of natural forests is lacking in new plantation forests. This narrow genetic diversity increases the risk of mass deaths, insect and fungal diseases, and lowers resistance to climate change in the future forest.

Seed orchards are important seed sources and are essential for global tree improvement programs and studies. The primary objective of seed orchards is to produce genetically improved seeds, but they also function as a breeding population. Seed orchards have been established in many countries to produce improved seeds. The first studies using seed orchards were performed in 1934 using a vegetative production technique (Larsen, 1956). Since then, seed orchards have become an important global seed source (Zobel et al., 1958; Faulkner, 1975; Wright, 1976; Zobel and Talbert, 2003).

The genetic quality of seed orchard crops and the vitality and performance of the resulting trees depend on many factors, including clone fertility, genetic diversity, mating design, and combining ability (Hosius et al., 2006; Ertekin, 2010). In seed orchards, some clones produce more flowers or pollen than others. Also, certain genotypes rarely mate because they flower out of synchronization, or because the male and female gametes of different clones contribute differently to the specific seed crops (sexual asymmetry [A_s]) (Müller-Starck et al., 1982; Zobel and Talbert, 2003). Thus, seed orchard crops are generally derived from a limited number of trees. It is reasonable to assume that the seed lot from one crop year does not represent the total gene pool of the respective orchard. It is common to have a small portion of the orchard parents contribute a disproportionately large amount to the orchard crop (El-Kassaby et al., 1989; El-Kassaby and Cook, 1994). This unequal contribution leads to an increase in genetic relatedness and a loss of genetic diversity in seed crops (Kang, 2001).

Therefore, determining the genetic composition of seed orchard crops is important. When establishing new plantations and performing genetic research studies, genetic parameters such as genetic relatedness, inbreeding, and genetic diversity should be measured and monitored in seed orchard crops.

2. Seed orchard

A seed orchard is defined as an area where seeds are mass-produced to increase the genetic quality as quickly and inexpensively as possible (Zobel et al., 1958). It is also defined as a plantation of selected clones or progeny that are isolated or managed to avoid or reduce pollination from outside sources and produce frequent, abundant, and easily harvested seed crops (OECD, 1974; Feilberg and Soegaard, 1975).

The concept of using seed orchards to produce genetically superior seeds was first employed in Europe at the beginning of the last century. In 1906, Gunnar Andersson of Sweden pioneered the vegetative breeding of forest trees. In 1909, Oppermann, Andersson, and Hesselman used elite seeds collected from small natural stands for vegetative breeding and early plantation establishment. These early plantations have been a crucial part of present-day reforestation efforts. The first clonal seed orchards, on the island of Java in the Netherlands, were established to increase the abundance of *Cinchona ledgeriana*, a major source of quinine, in 1880 (Feilberg and Søegaard, 1975). In 1918, Sylvén drew attention to the choice of origin. In 1922, Fabricius coordinated use of the seed orchard with the Forest Trees breeding program. In 1923, Oppermann used seed orchards to propagate *Larix eurolepis*. According to Feilberg and Søegaard (1975), the first forest tree seed orchard was established in 1931 in England with an *L. eurolepis* hybrid. In 1934, Syrach-Larsen established seed orchards using the vegetative technique (Larsen, 1956). At this point, establishing clonal seed orchards as seed sources became globally important (Faulkner, 1975; Wright, 1976; Zobel and Talbert, 2003).

In Europe, establishing clonal seed orchards began immediately after World War II. In Sweden, *Pinus sylvestris* and *Picea abies* clonal seed orchards were established in the 1950s. In Denmark, an *L. eurolepis* clonal seed orchard was established in 1946. In Hungary, *P. sylvestris*, *Larix* spp., *Pinus nigra*, and *Picea* spp. clonal seed orchards were established in 1951. In the United States, *Pinus taeda*, *Pinus elliotti*, and *Pinus echinata* clonal seed orchards were established in 1957. In Finland, clonal seed orchards of *P. sylvestris* and species of *Picea* and *Betula* were established in 1960. In Canada, a *Pseudotsuga* clonal seed orchard was established in 1966. In 1970, *Cryptomeria japonica* and *Pinus densiflora* clonal seed orchards were established in Japan. In 1953, a *Pinus radiata* clonal seed orchard was established in New Zealand.

Seed orchards are commonly categorized by the first, second, or more advanced generation depending on how many cycles of improvement they have undergone (Zobel and Talbert, 2003). Seed orchards are also grouped by origin type (e.g., tree seed orchard, elite seed orchard, or hybrid seed orchard) according to the purpose or establishment form (Boydak, 1979). Generally, global seed orchards are first generation orchards initiated by parents whose genetic worth is unknown and in which the trees are generally closely spaced to allow for rouging of poor genotypes while maintaining a fully function seed orchard (Zobel and Talbert, 2003).

The optimal number of individuals in an orchard allows for rouging of the poorer genotypes, maintains the desired spacing, and maximizes seed production by having enough high-quality trees with adequate pollination and ensuring a minimum of relatedness (Zobel and Talbert, 2003). Older seed orchards are composed of a smaller number (20-30) of clones, while recently established orchards contain a much larger number of clones and aim to have a combination of maximal breeding progress and strong genetic diversity (Gagov et al., 2004). The correct number of clones to deploy in an orchard is an important consideration (Lindgren and Prescher, 2005). Generally, in orchard design the goal is to minimize selfing, maximize out-crossing and mating of all genotypes, allow for simple and easy establishment and management, and allow for any number of clones/families to fit into the completely randomized block and seeding designs (Schmidt, 1991). Seed orchards should be located within the natural range of the species; however, warmer climates can be advantageous for seed maturation and earlier flowering. Physical isolation from pollen contamination is also beneficial (Sarvas, 1970; Kang, 2001).

3. Genetic diversity in seed orchards

Genetic diversity refers to the richness of genetic information in the gene pool of a specific species. The genetic diversity of a species is shaped by the frequency of genetic change (migration, mutation, or isolation). In different ecosystems, the same species may exist with different genetic constitutions. The size of the gene pool in these populations (according to their degree of inheritance) may be wide or narrow. Forest trees typically have high levels of genetic diversity compared with other species. Also, genetic diversity within a species is often higher than that between populations. Recently, DNA or isozyme markers have been used to analyze genetic diversity. For example, a study performed in Turkey based on fourteen different isoenzyme analyses of black pine identified 92.6% of the total genetic diversity within the population (Yüksek, 1997).

Seed orchards must reflect the genetic diversity of the original population and be sufficiently large to maintain genetic diversity for future generations. If the genetic diversity of a seed orchard crop is maximized, all clones must contribute equally. Unequal gametic contributions can result from an absence of flowering synchronization among the clones (Matziris, 1993). Strong genetic diversity in seed orchard crops can only be attained when all parents contribute equally to the gamete pool.

4. Genetic gain in seed orchards

The primary issues affecting the genetic characteristics of seeds are genetic diversity and genetic gain (Stoehr et al., 2004). Genetic gain is directly related to the genetic diversity and degree of genetic structure transferred from parent to progeny. Kang (2001) stated that there are long- and short-term considerations in practical tree breeding programs. In the short term, forestry practices should result in productive stands that can tolerate changing environmental pressures for the duration of the rotation. Long-term concerns include maintaining reservoirs for genetic variability, which is required for current breeding populations. Short-term genetic gain is typically maximized in clonal forestry and seed orchards. For long-term breeding strategies, breeding and base populations should be managed for sustainable genetic diversity.

The genetic gain obtained from seed orchards is broader than that obtained from natural forests. By using plus trees to establish seed orchards, one can achieve genetic gain in orchard crops. In orchards, the increased combining ability of plus trees increases the genetic gain, which is derived from the additive variance in the referencing population (Kang, 2001). Wright (1976) stated that genetic gain could reach 30% in seed orchards using plus trees. Weir and Zobel (1975) reported a genetic gain from first- to second-generation seed orchards of 35%, exceeding the previously reported values of 10-20%. Matziris (1999) reported that *Pinus halepensis* clonal seed orchards had an increased volume ratio of 21.25%. Öztürk (2003) reported that by using plus trees to select for optimal mass, an 8.1% increase in size could be attained. Moreover, by reducing the number of clones in an orchard and further selecting for optimal characteristics, an additional 5% gain in size could be achieved.

Maximum flowering synchronization and equality, combined with minimal inbreeding and self-pollination, are important for seed orchards to reach their theoretically expected genetic gain and to achieve genetic diversity (Kang, 2001). In particular, flowering synchronization (Fig. 1) among clones in a seed orchard is important for the genetic composition of orchard crops since it affects the genetic exchange between clones. If seeds are collected from clones that lack synchronization, genetic diversity will be below the ideal level, resulting in a panmictic equilibrium (El-Kassaby and Askew, 1991; Kang and Lindgren, 1998).

Fig. 1. Optimal flowering synchronization period: maximum pollen accepting period (left) and maximum pollen shedding period (right).

5. Measuring of genetic diversity in seed orchards

5.1 Parental balance and maleness index

The number of female and male flowers, the number of cones, and the conversion rates of flowers to cones in a seed orchard is important information for forest genetic and tree

improvement studies (Matziris, 1997; Kang and Lindgren, 1998; Kang, 2000; Gömöry et al., 2000; Choi et al., 2004; Ertekin, 2006). In seed orchards, it has been shown that some clones produce large numbers of flowers, while others produce very few. Many studies have shown that a small number of clones are often responsible for a large part of clone production in seed orchards. Also, some clones produce more male than female flowers. For example, Ertekin (2006) observed that 33% of the total clone population accounted for 62% of all male flowers and 49% of all female flowers. Also, nine clones accounted for 50% of the total clone production on average for two years in a black pine seed orchard in Turkey (Figs. 2-4). Also, Johsson et al. (1976) showed that in a *P. sylvestris* seed orchard, 25% of the total clones accounted for 62.1% of all male flowers produced and 50.8% of all female flowers produced. Nikkanen and Velling (1987) observed that in a *P. sylvestris* seed orchard, 19% of the total clones accounted for 50% of all male flowers produced, while 35% of the total clones accounted for 50% of all female flowers produced. Kang (2000) found that in a *P. densiflora* seed orchard, 25% of the clones accounted for 37.3% of all male flowers produced and 48.1% of all female flowers produced. Yazdani and Fries (1992) found that in a *Pinus contorta* seed orchard, 23% of the clones accounted for 50% of the total female flower population.

Fig. 2. Clonal contributions to male flower production in a back pine seed orchard in Turkey (Ertekin, 2006).

For seed orchard crops, parental balance curves are used to characterize high- or low-flowering clones. Ertekin (2010) generated parental balance curves in a three-year study of black pine seed orchards in Turkey. As seen in Fig. 5, the parental balance curves varied between flowerings and years, and did not improve with age. The observed curves for female and male flowering deviated significantly from the ideal situation, and the orchard's clones contributed equally to the gamete pool. Thus, specific clones may consistently produce high- or low-flowering clones based on genetic tendencies. Maternal and paternal genetic contributions can be explained by parental balance curves. A

cumulative contribution curve is often used to quantify fertility variation in forest populations (Griffin, 1982; El-Kassaby and Reynolds, 1990; Adams and Kunze, 1996). Parental balance in seed orchard crops is commonly summarized using cone-yield curves (Griffin, 1982). Using this method, seed orchard clones are ranked from high to low yield, and cumulative percentage calculations are plotted against the total number of clones (Chaisurisri and El-Kassaby, 1993).

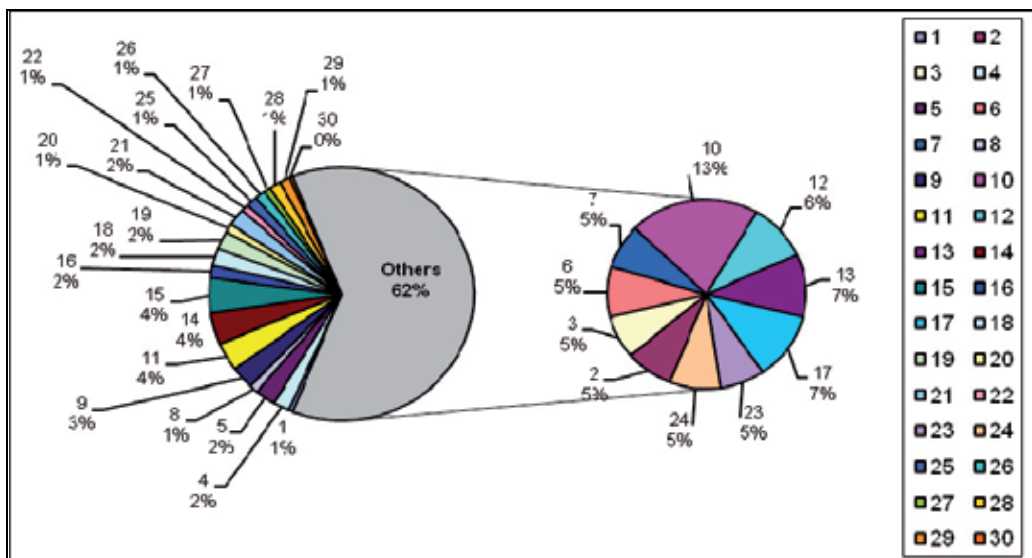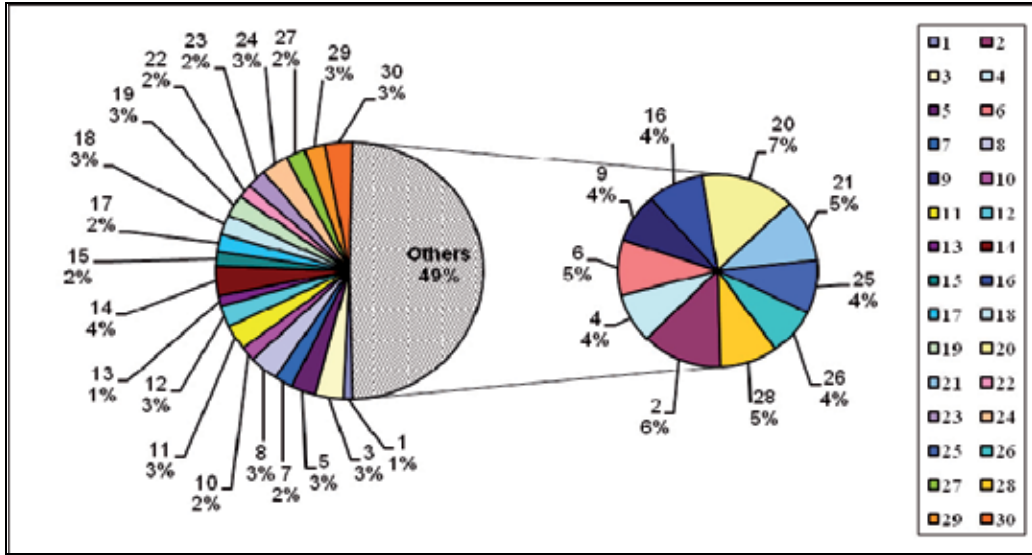Fig. 3. Clonal contributions to female flower production in a back pine seed orchard in Turkey (Ertekin, 2006).
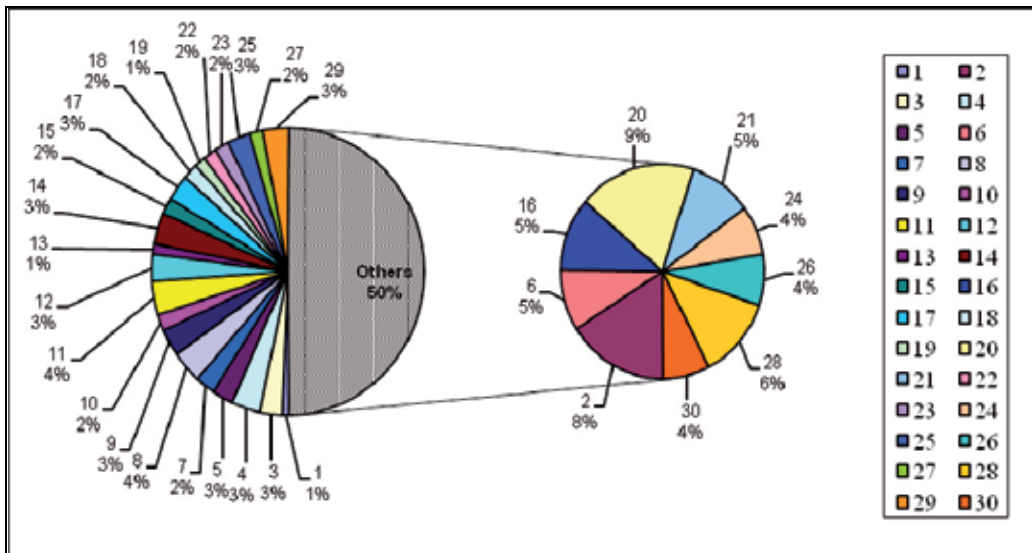

Fig. 4. Clonal contributions to cone production in a back pine seed orchard in Turkey (Ertekin, 2006).
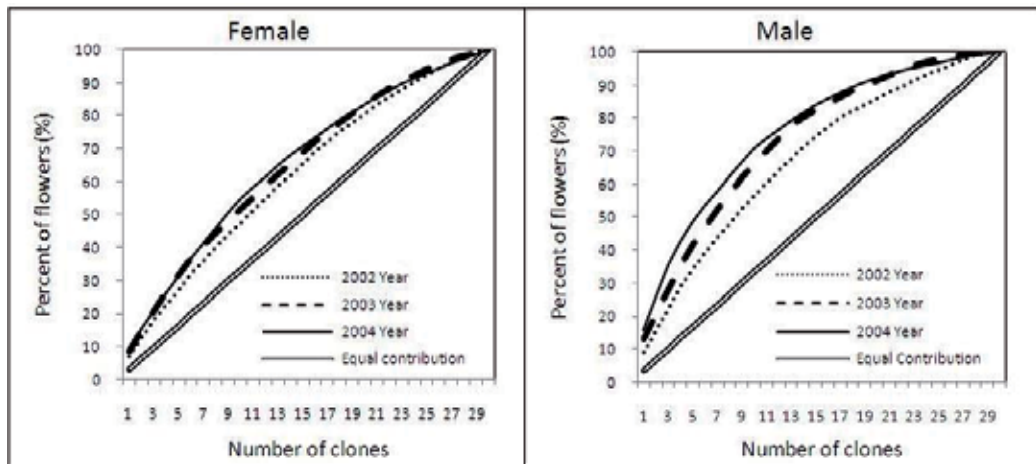
Fig. 5. Parental balance curves for female and male flowers with equal clone contributions (Ertekin, 2010).

To characterize A_s among the clones, the M_i was used. The M_i is defined as the proportion of a clone's reproductive success that is transmitted through its pollen (Kang, 2000). The M_i (based on flower production) was estimated as follows (Choi, 2004):

$$M_i = \frac{m_i}{(m_i + f_i)} \tag{1}$$

where m_i and f_i are the number of male and female flowers in i^{th} clones, respectively.

A high M_i indicates that a clone is contributing more paternally than maternally (Choi, 2004). Most studies have reported an M_i ranging from 0.2 to 0.8 (Burczyk and Chalupka, 1997; Kang, 2000). An M_i near 0.5 indicates nearly equal female and male fertility. If a small number of clones accounts for a large percentage of the male flowers produced in an orchard, self-pollination may be increased. In this case, sexual asymmetry may reduce the number of homozygotes in an orchard from that expected under Hardy-Weinberg equilibrium and mask the effects of inbreeding (Kang, 2000). To avoid this situation, management activities such as supplemental mass pollination, hormonal treatments, or pruning ramets from high-producing clones could be implemented in these orchards (Ertekin, 2010).

5.2 Fertility variation and status number

Fertility variation between clones provides information about genetic diversity in a seed orchard. Fertility variation is measured using the coefficient of variation (CV) and sibling coefficient (Ψ). According to Kang (2001), ψ provides more accurate seed orchard crop genetic diversity information than the CV. The CV (standard deviation divided by the mean) was calculated as follows:

$$CV = \frac{s}{M} = \sqrt{\frac{N\left(N\sum_{i=1}^{N} P_i^2 - 1\right)}{N-1}} \tag{2}$$

where s is the standard deviation, M is the mean, N is the number of the clones, and P_i is the fertility (female or male) of clone i.

The female and male Ψ (Ψ_f and Ψ_m, respectively) were calculated as follows:

$$\psi_f = N \sum_{i=1}^{N} \left(\frac{f_i}{\sum f_i} \right)^2 = CV_f^2 + 1 \qquad (3)$$

$$\psi_m = N \sum_{i=1}^{N} \left(\frac{m_i}{\sum m_i} \right)^2 = CV_m^2 + 1 \qquad (4)$$

where N is the census number and CV_f and CV_m are the CVs for female and male flower production among the clones, respectively.

Generally, fertility variation is small initially then increases in subsequent years (Matziris, 1993; Nikkanen and Ruotsalinen, 2000). Kang et al. (2003) found that fertility differences were slightly larger for males than for females, and suggested that a CV of 100% would be typical in good- or moderate-flowering years for a mature seed orchard.

Status numbers were calculated based on the fertility variation of female and male parents (Lindgren et al., 1996), respectively, as follows:

$$N_{s(f)} = \frac{1}{\sum_i^N f_i^2} \qquad (5)$$

$$N_{s(m)} = \frac{1}{\sum_i^N m_i^2} \qquad (6)$$

where $N_{s(f)}$ and $N_{s(m)}$ are the N_s for female and males, respectively.

The N_s was calculated as follows, according to the equation of Kang (2001):

$$N_s = \frac{4N}{[\Psi_f + \Psi_m + 2 + 2r\sqrt{(\Psi_f - 1)(\Psi_m - 1)}]} \qquad (7)$$

where r is the correlation coefficient between female and male flower production.

If the N_s values for flowering and the number of years were greater than ten, the depletion of genetic diversity in the following generation due to genetic drift and fertility variation would be small, as reported by Kang (2001). Consequently, small numbers of clones produce most of the flowers or seeds in the orchard, resulting in a loss of genetic diversity. Kang et al. (2005) stated that a loss in genetic diversity was expected due to the accumulation of relatedness or fertility variation. A reduction in genetic diversity in orchard crops affects the level of genetic diversity in seedlings and, subsequently, in the plantation forest.

5.3 Relative status number and group coancestry

N_r was calculated as the ratio of N_s to N as follows:

$$N_r = \frac{N_s}{N} \tag{8}$$

N_s was defined as half the inverse of the coancestry group (Lindgren et al., 1996) as follows:

$$N_s = \frac{0.5}{\Theta} \quad \Theta = \frac{1}{2N_s} \tag{9}$$

The expected genetic diversity of seed crops from first-generation seed orchards can be measured relative to the group coancestry of a reference population (Kang, 2001). The relative genetic diversity values, calculated for three years, were high in this seed orchard. Kang and Lindgren (1998) reported that the relative genetic diversity (compared to reference populations) was quite high in first-generation seed orchards.

5.4 Gene diversity and sexual asymmetry

Genetic diversity among seed orchard crops is significantly influenced by the relatedness of orchard clones, parental fertility variation, and pollen contamination. Based on a study of a black pine seed orchard, the top ten cone-producing clones accounted for 33% of all cones produced in year 1, 48% in year 2, and 58% in year 3 (Ertekin, 2006). Based on seed orchard studies, the overall genetic diversity of all clones in the orchard is not represented; thus, the genetic diversity is narrow. Matziris (1993) reported that in an abundant cone production year there was a decrease in the differences between cone-producing clones, and that there was an expansion of the genetic base of seeds from seed orchards.

The expected genetic diversity of seed orchard crops (GD) from first-generation seed orchards can be measured relative to the group coancestry of a reference population. GD is a function of group coancestry (Lacy, 1995) and is inversely proportional to the status number (Kang, 2001), as follows:

$$GD = 1 - \frac{1}{2N_s} \tag{10}$$

A_s was also estimated (Kang and Mullin, 2007) as follows:

$$N_a = \frac{N_{s(f)} + N_{s(m)}}{2} \qquad A_s = \frac{\left(N_{s(c)} - N_a\right)}{N_a} \tag{11}$$

where $N_{s(f)}$, $N_{s(m)}$, and $N_{s(c)}$ are the N_s for females, males, and clones, respectively, and N_a is the arithmetic mean of the measures ($N_{s[f]}$ and $N_{s[m]}$).

Kang and Mullin (2007) stated that A_s theoretically ranges from zero to one, and that an A_s value of zero indicates perfect sexual symmetry. Variation in fertility can be compensated for by intentionally adjusting the number of ramets to manage the orchard. Equal seed harvests or mixing of seeds from consecutive years can be used to reduce the impact of

fertility variation on the diversity of seed orchard crops, since mixing seed crops reduces fertility variation (Kang et al., 2003; Varghese et al., 2000; Ertekin, 2010).

6. Pollen contamination

The flowering phenology of clones in seed orchards allows us to determine the genetic diversity of orchard crops. Xie and Knowles (1994) stated that some clones (those known to be early- or late-flowering) should be removed from seed orchards. Using this method, the risk of self-pollination and pollen contamination can be reduced. Pollen contamination, which directly affects the genetic diversity of seeds, is a major source of gene migration in seed orchards (Adams and Birkes, 1989).

By examining the flowering phenology of specific clones, some variations were identified in the bud burst, pollen shedding, and pollen acceptance of male/female flowers. Similar observations have been made by groups working in various seed orchards, including *P. nigra* (Matziris, 1994; Lario et al., 2001), *P. abies* (Eriksson et al., 1973; Skrøppa and Tutturen, 1985; Nikkanen, 2001), *P. sylvestris* (Jonsson et al., 1976; Gömöry et al., 2000), *Pinus brutia* (Keskin, 1999), and *Cunninghamia lanceolata* (Zhuowen, 2002) seed orchards. In addition, varying flower development times were observed on the northern or southern branches of the same trees (Ertekin, 2006).

Reynolds and El-Kassaby (1990) reported that the most important indicator of genetic diversity in orchard crops is the flowering quality and harmony of the flowering periods (parental balance). Gömöry et al. (2003) stated that 15% of the total clones in a *P. sylvestris* seed orchard flowered early or late; thus, these clones were not pollinated with the others. These early or late flowering clones pollinated other trees outside the orchard. Thus, pollen contamination is increased in these seed orchards. Pakkanen et al. (2000) found that that the pollen contamination rate was 69-71% in a *P. abies* seed orchard.

In a *P. brutia* seed orchard in Turkey, genetic contamination of the orchard by seeds was detected and pollen contamination rates were very high (estimated at 85.7%) (Kaya, 2001). Also, the growth of female flowers before male flowers has been reported in numerous studies. For example, Zhuowen (2002) observed that female flowers were present five to six days before male flowers in a *C. lanceolata* seed orchard. Parantainen and Pulkkinen (2003) stated that female flowers opened three days before male flowers, but ended at the same time in *P. sylvestris* seed orchards. Yazdani and Fries (1992) noted a three- to four-day difference in female and male flower activity in a *P. contorta* seed orchard. Also, Ertekin (2006) observed female flower growth three to four days before male flower growth in a *P. nigra* seed orchard.

The duration of the pollen accepting or shedding periods is as important as the synchronization of flowering for pollen contamination and the genetic diversity of clones.

Matziris (1994) observed that the duration of pollen acceptance was between two and eight days in a *P. nigra* seed orchard. Nikkanen (2001) noted that the duration of the pollen shedding period was five to eight days, while the pollen acceptance period was five to ten days, in a *P. abies* seed orchard. Ertekin (2006) observed that the duration of the pollen shedding period was six to nine days while the pollen acceptance period was six to eight days in *P. nigra*. Therefore, the fertilization period in seed orchards is generally short.

However, within this time period pollen gene exchange in an indoor or outdoor orchard will occur among receptive clones. Pollen contamination also occurs during this time.

Although pollen contamination in seed orchards can be reduced by isolating the orchard, it cannot be completely eliminated. Kang (2001) observed that many of the orchard seeds had unknown fathers from outside sources.

7. Selfing and inbreeding

Selfing is a mating process that occurs by self-pollination among parents of the same genotype. In many forest trees, self-compatibility is necessary for selfing to occur, and it can result in reduced seed germination and growth (Zobel and Talbert, 2003). Selfing in natural populations occurs at a higher frequency than in seed orchards. In seed orchards, ramets of the same clone are scattered and not as close to neighboring trees. Thus, inbreeding is very low because of the mating design. Squillace and Goddard (1982) observed that the yield of selfed seed orchards averaged only 2.5%, lower than the estimates for trees in natural stands. They also noted that selfing had a significant effect on orchard crops since approximately 9.5% of all seeds were self-fertilized. Also, if the self-pollinated offspring survive, the genetic gains are reduced since they are less vigorous than outcrosses. Moreover, Sarvas (1962) stated that inbreeding could result in homozygous lethal genes, causing embryo collapse and empty seeds.

Selfing is especially important in pine, spruce, and fir species since they are monoic. These species possess certain mechanisms to prevent selfing. For example, male and female flowers occur at different locations on the crown in monoic species. Male flowers occur on the lower crown while female flowers are on the upper crown. Franklin (1971) stated that selfing tended to be greater in the lower crown than the mid or upper crown. However, female and male flowers of seedlings produced by vegetative techniques occur on the end of the same shoot (Fig. 6). In this situation, selfing is unavoidable.

Seed orchards are established to produce genetically superior seeds for plantations. In these orchards, selfing can occur by mating among ramets of the same clone, as well as by self-pollination within individual ramets. Selfing reduces the genetic value of orchard crops. Generally, the selfing rate can be estimated based on isozyme markers (Burczyk, 1991; Harju, 1995; Squillace and Goddard, 1982). If the selfing rate is >10%, the selfed offspring need to be identified and either removed from the orchard or treated by artificial pollenization (El-Kassaby and Ritland, 1986; Kang, 2001).

8. Hermaphroditism in seed orchard

Hermaphroditism (bisexual flowers) means that an organism that has morphologically female and male reproductive organs (strobili) on the same flower bud. In nearly all gymnosperms, male and female strobili occur on the same tree (monoic). Male and female flowers occur at different parts of the crown (female flowers occur on the upper crown and male flowers on the lower crown); moreover, male flowers are borne in clusters at the base of the twig bud while the female flowers are borne in one conelet at the bud apex. However, hermaphroditism occurs in various trees (Zobel and Goddard, 1954; Chamberlain, 1966; Burley, 1976; Matziris, 2002).

In hermaphroditic flowers, the female flowers are at the top while the male flowers are below (Fig. 6). According to Matziris (2002), the cause of this anomaly is not well understood. It has been observed more frequently in plantations of exotic species, possibly due to the lack of adaptability of a species or genotype to a new exotic environment. However, the benefits of hermaphroditism are unknown; thus, further research is required.

Fig. 6. Hermaphroditic flowers of black pine clones at the pollen shedding stage; a female flower (top) has formed on the male flower (base). Also, the selfing risk is quite high.

9. Conclusion

The level of genetic diversity in seed orchards is important to plantation forests because high levels will increase resistance to pests, diseases, or climate change. Some of the clones in seed orchards that produced high numbers of female flowers did not mature to cones because of insect damage, physiological stress, or a lack of fertility synchronization. Generally, clonal variation increased with increases in male flowering. However, if fewer clones produced a large proportion of the male flowers, genetic diversity decreased. If a species is often wind-pollinated, a large amount of pollen will come from outside the seed orchard, widening the genetic diversity but decreasing the genetic gain. Therefore, artificial pollination, thinning, or pruning orchard management techniques can be used. For thinning, low-flowering clones are removed. For pruning, crowns are modified based on the flowering type. These orchard management techniques can increase genetic diversity in seed orchard crops. Strong genetic diversity in seed orchard crops can only be attained when all parents contribute similarly to the gamete pool, which virtually never occurs. Generally, a small portion of the orchard parents contribute a disproportionately large amount to the orchard crop (El-Kassaby et al., 1989; El-Kassaby and Cook, 1994), leading to an accumulation of genetic relatedness and a loss of genetic diversity in seed crops (Kang, 2001). Variations in fertility can be compensated for by intentionally adjusting the number of

ramets. Equal seed harvests, or mixing seeds from consecutive years, can be used to reduce fertility variation.

10. References

Adams. W.T. & Birkes, D.S. (1989). Mating Patterns in Seed Orchards, *Proceedings of 20th Southern Forest tree Improvement Conference*, pp.75-86, Charleston, South Carolina, USA, June 26-30, 1989

Adams. G.W. & Kunze, H.A. (1996). Clonal Variation in Cone and Seed Production in Black and White Spruce Seed Orchards and Management Implications. *Forestry Chronicle*, 72, pp. 475-480

Alptekin, Ü. (1986). Geographic Variations in Anatolian black pine (*Pinus nigra* Arnold. ssp. *pallasiana* Lamb. Holmboe), Ph.D Thesis, Istanbul University, pp.170, İstanbul, Turkey

Boydak, M. (1979). Seed Gardens as Advanced Seed Sources. *Journal of the Faculty of Forestry, Istanbul University*, vol. 29-2, pp. 89-111, Istanbul, Turkey

Burczyk, J. (1991). The Mating System in a Scots Pine Clonal Seed Orchard in Poland. *Annales des Sciences Forestieres*, 48, pp. 443-451

Burczyk, J. & Chalupka, W. (1997). Flowering and Cone Production Variability and its Effect on Parental Balance in a Scots Pine Clonal Seed Orchard, *Annales des Sciences Forestieres,* vol. 54, 2, pp. 129-144

Burley, J. (1976). Genetic System and Genetic Conservation in Tropical Trees, in: *Tropical Trees*, S. Burley and B. T. Styles (Ed.), pp. 85–99, London, England

Chaisurisri, K. & El-Kassaby, Y.A. (1993). Estimation of Contribution to Cone and Seed Crops in a Sitka Spruce Seed Orchard, *Annales des Sciences Forestieres*, 50, pp. 461-467

Chamberlain, C. (1966). *Gymnosperms, Structure and Evolution*, Dover Publications, pp. 484, New York, United State

Choi, W.Y.; Kang, K.S.; Han, S.U. & Hur, S.D. (2004). Estimation of Heritabilities and Clonal Contribution Based on The Flowering Assessment in Two Clone Banks of *Pinus koraiensis* Sieb. et Zucc., *Proceedings of USDA Forest Service*, RMRS-P-32, pp. 172-180

El-Kassaby, Y.A. & Ritland, K. (1986). The Relation of Outcrossing and Contamination to Reproductive Phenology and Supplemental Mass Pollination in a Douglas-Fir Seed Orchard. *Silvae Genetica,* 35, pp. 240-244

El-Kassaby, Y.A.; Fashler, A.M.K. & Crown, M. (1989). Variation in Fruitfulness in a Douglas-Fir Seed Orchard and its Effect on Crop-Management Decisions. *Silvae Genetica,* 38: pp. 113-121

El-Kassaby, Y.A. & Askew, G.R. (1991). The Relation Between Reproductive Phenology and Reproductive Output in Determining the Gametic Pool Profile in a Douglas-Fir Seed Orchard, *Forest Science,* 37, pp. 827-835

El-Kassaby, Y.A. & Reynolds, S. (1990). Reproductive Phenology, Parental Balance, and Supplemental Mass Pollination in a Sitka-Spruce Seed Orchard. *Forest Ecology and Management*, 31, pp. 45-54

El-Kassaby, Y. A. & Cook, C. (1994). Female Reproductive Energy and Reproductive Success in a Douglas-Fir Seed Orchard and its Impact on Genetic Diversity. *Silvae Genetica*, 43, pp. 243-246

Eriksson, G.; Jonsson, A. & Lindgren, L. (1973). Flowering in a Clonal Trial of *Picea abies* Karst, *Studia Foresttalia Suecica*, Nr 110, Stockholm, pp. 45

Ertekin, M. (2006). Clonal Variations in Flowering, Cone Production and Seed Characteristics of Black Pine (*Pinus nigra* Arnold.Subsp. *pallasiana* (Lamb.) Holmboe) Seed Orchard Originated From Yenice-Bakraz, Ph.D Thesis, Zonguldak Karaelmas University, pp.191, Zonguldak, Turkey

Ertekin, M. (2010). Clone Fertility and Genetic Diversity in Black Pine Seed Orchard. *Silvae Genetica*, 59, pp. 145-150

Faulkner, R. (1975). *Seed Orchards*, Forestry Commission Bulletin, No.54, pp. 149, London, England

Feilberg, L. & Søegaard, B. (1975). Historical Review of Seed Orchards, in. *Seed Orchards*, Forestry Commission Bulletin, No.54, pp. 1-8, London, England

Franklin, E.L. (1971). Pollen Management in Southern Seed Orchard, *Proceeding of 11th SFTIC*, p.218-223, Atlanta, Georgia, USA

Gagov, V.; Bergmann, F. & Hosius, B. (2004). *Seed orchards as a tool in tree improvement and conservation of forest gene resources*, Mitteilungen der Forschungsanstalt für Waldökologie und Forstwirtschaft Rheinland-Pfalz Nr, 52, pp. 59–70

Gömöry, D.; Bruchánik, R. & Paule, L. (2000). Effective Population Number Estimation of Three Scots Pine (*Pinus Sylvestris* L.) Seed Orchards Based on An İntegrated Assessment of Flowering, Floral Phenology, and Seed Orchard Design. *Forest Genetics*, 7 (1), pp. 65-75.

Gömöry, D.; Bruchánik, R. & Longauer, R. (2003). Fertility Variation and Flowering Asynchrony in *Pinus Sylvestris*: Consequences for the Genetic Structure of Progeny in Seed Orchards, *Forest Ecology and Management*, 174, pp. 117-126.

Griffin, A.R. (1982). Clonal Variation in Radiata Pine Seed Orchards. I. Some Flowering, Cone and Seed Production Traits. *Australian Forest Research*, 12, pp. 295-302

Harju, A. (1995). Genetic Functioning of Scots Pine Seed Orchards, Ph.D Thesis, University of Oulu, Acta Universitatis Ouluensis 271, Finland.

Hosius, B.; Leinemann, L.; Konnert, M. & Bergmann, F. (2006). Genetic Aspects of Forestry in the Central Europe, *European Journal of Forest Research*, 125, pp. 407–417

Jonsson, A.; Ekberg, I. & Eriksson, G. (1976). Flowering in a Seed Orchard of *Pinus sylvestris* L. *Studia Foresttalia Suecica*, No.135, pp. 38.

Kang, K.S. (2000). Clonal and Annual Variation of Flower Production and Composition of Gamate Gene Pool in a Clonal Seed Orchard of *Pinus densiflora*, *Canadian Journal of Forest Research*, 30(8), pp. 1275-1280

Kang, K.S. (2001). Genetic Gain and Gene Diversity of Seed Orchard Crops, Ph. D. Thesis, Swedish University of Agricultural Science, Acta Universitatis Agriculturae Sueciae, Silvestria 187, pp. 75, Umeå, Sweden

Kang, K.S. & Lindgren, D. (1998). Fertility Variation and its Effect on the Relatedness of Seeds in *Pinus densiflora*, *Pinus Thunbergii* and *Pinus koraiensis* Clonal Seed Orchards, *Silvae Genetica*, 47, 4, pp. 196–201

Kang, K.S.; Bila, A.D.; Harju, A.M. & Lindgren, D. (2003). Fertility Variation in Forest Tree Populations. *Forestry*, 76, pp. 329–344

Kang, K.S.; El-Kassaby, Y.A.; Chung, M.S.; Kim, C.S.; Kang, Y. J. & Kang, B.S. (2005). Fertility Variation and Genetic Diversity in a Clonal Seed Orchard of *Cryptomeria japonica*. *Silvae Genetica*, 54, pp. 104-107

Kang, K.S. & Mullin, T.J. (2007). Variation in Clone Fertility and its Effect on the Gene Diversity of Seeds From a Seed Orchard of *Chamaecyparis obtusa* in Korea, *Silvae Genetica*, 56, pp. 134-137

Kaya, N. (2001). Estimation of the Mating System and Genetic Contamination a Clonal Seed Orchard of Pinus Brutia Ten. in Asar-Antalya, Ph. D. Thesis, Akdeniz University, pp. 80, Antalya, Turkey

Keskin, S. (1999). Clonal Variation in Flowering and Cone Characteristics in a *Pinus brutia* Seed Orchard, Technical Bulletin of Southwest Anatolia Forest Research Institute, No. 9, pp. 96, Antalya, Turkey

Lacy, R.C. (1995). Clarification of Genetic Terms and Their Use in the Management of Captive Populations. *Zoo Biology,* 14, pp. 565-578

Lario, F.; Merlo, E.; Peñuelas, J. & Gil, L. (2001). Variabilidad Clonal de La Fenología Reproductiva Y Producción Floral. Participación Clonal En Un Huerto Semillero De *Pinus nigra A*rnold *Salzmannii* (Dunal) Franco, Proceeding of *Actas del III Congreso Forestal Español.* 2001. Granada. Mesa, 3, pp. 539–545, Spain

Larsen, C.S. (1956). *Genetics in Silviculture*, Oliver & Boyd., pp. 224, Edinburgh, Scotland

Lindgren, D.; Gea, L.D. & Jefferson, P.A. (1996). Loss of Genetic Diversity Monitored by Status Number, *Silvae Genetica*, 45, pp. 52–59

Lindgren, D. & Prescher, F. (2005). Optimal Clone Number for Seed Orchards with Tested Clones. *Silvae Genetica*, 54, 2, pp. 80-92

Matziris, D. (1993). Variation in Cone Production in a Clonal Seed Orchard of Black Pine. *Silvae Genetica,* 42,(2/3), pp. 136–328

Matziris, D. (1994). Genetic Variation in the Phenology of Flowering in Black Pine. *Silvae Genetica*. 43, pp. 321-328

Matziris, D. (1997). Variation in Growth, Flowering and Cone Production in a Clonal Seed Orchard of Aleppo Pine Grown in Greece, *Silvae Genetica*, 46 (49), pp. 224–228

Matziris, D. (1999). Genetic Variation and Realized Genetic Gain from Aleppo Pine Tree Improvement, *Silvae Genetica*, 49, 1, pp. 5–10

Matziris, D. (2002). Hermaphrodism in Black Pine, *Silvae Genetica*, 51, 2–3, pp. 130-131

Müller-Stark, G. (1982). Tracing External Pollen Contribution to the Offspring of Scots Pine Seed Orchard, Proceeding of IUFRO Meeting on Breeding Strategies Including Multiclonal Varieties. Lower Saxony FRI, Dep. of For.Tree Breeding, pp.176, Escherode, Germany

Nikkanen, T. (2001). Reproductive Phenology in a Norway Spruce Seed Orchard, *Silva Fennica*, 35 (1), pp. 39-52

Nikkanen, T. & Velling, P. (1987). Correlations Between Flowering and Some Vegetative Characteristics of Grafts of *Pinus Sylvestris*, *Forest Ecology and Management*, 19, pp. 35-40

Nikkanen, T. & Ruotsalainen, S. (2000). Variation in Flowering Abundance and Its Impact on The Genetic Diversity of the Seed Crop in a Norway Spruce Seed Orchard, *Silva Fennica*, 34 (3), pp. 205-222

OECD (1974). Establishing an OECD Scheme Fort the Control of Forest Reproductive Material Moving in International Trade, C (74)29, pp. 23

Öztürk, H. (2003). Estimating breeding values of the Turkish red pine (*Pinus brutia* Ten.) clones by using open pollinated progeny trials, Ph. D. Thesis, Akdeniz University, pp. 219, Antalya, Turkey

Pakkanen, A., Nikkanen, T. & Pulkkinen, P. (2000). Annual Variation Pollen Contamination and Outcrossing in a *Picea abies* Seed Orchard, *Scandinavian Journal of Forest Research*, 15, pp. 399-404

Parantainen, A. & Pulkkinen, P. (2003). Flowering and Airborne Pollen Occurrence in a *Pinus sylvestris* Seed Orchard Consisting of Northern Clones, *Scandinavian Journal of Forest Research,* 18, pp. 111-117

Reynolds, S. & El-Kassaby, Y.A. (1990). Parental Balance in Douglas-fir Seed Orchards- Cone Crop vs. Seed Crop, *Silvae Genetica*, 39, 1, pp. 40-42

Sarvas, R. (1962). Investigations on the Flowering and Seed Crop of *Pinus sylvestris*. Communications Instituti Forestalis Fenniae, 53, pp. 198

Sarvas, R. (1970). Establishment and Registration of Seed Orchards, *Folia Forestalia*, 89, pp. 24

Schmidt, L. (1991). *Seed Orchards, Guidelines on Establishment and Management Practices,* UNDP/FAO Regional Forest Tree Improvement Project, Field Manual No.4, RAS/91/004, pp. 23

Skrøppa, T. & Tutturen, R. (1985). Flowering in Norway Spruce Seed Orchards, *Silvae Genetica,* 34, 2-3, pp. 90-95

Squillace, A.E. & Goddard, R.E. (1982). Selfing in Clonal Seed Orchards of Slash Pine. *Forest Science*, 28, pp. 71-78

Stoehr, M.; Webber, J. & Woods, J. (2004). Protocol for Rating Seed Orchard Seedlots in British Columbia: Quantifying Genetic Gain and Diversity. *Forestry*, 77, pp. 297-303

Ürgenç, S. (1982) *Forest Tree Breeding, Istanbul University, Faculty of Forestry,* No. 2836/293, pp. 414, Istanbul, Turkey

Varghese, M.; Nicodemus, A.; Nagarajan, B.; Siddappa, K.R.S.; Bennet, S.S.R. & Subramanian, K. (2000). *Seedling Seed Orchards for Breeding Tropical Trees,* Scroll Press, Institute of Forest Genetics and Tree Breeding, Coimbatore, pp. 126, India

Weir, R.J. & Zobel, B.J. (1975). *Advanced-Generation Seed Orchards*, Forestry Commission Bulletin, No.54, pp. 118-127, London, England

Wright, J.W. (1976). *Introduction to Forest Genetics*, Academic Press, pp. 463, New York, United State

Xie, C.Y. & Knowles, P. (1994). Mating System and Effective Pollen Immigration in a Norway Spruce (*Picea abies* (L.) Karst) Plantation. *Silvae Genetica*, 43, pp. 48-52

Yazdani, R., & Fries, A. (1992). Flover Abundance, Phenology and Pollination Pattern in a *Pinus contorta* Seed Orchard. *Pinus Contorta* from Untamed Forest to Domasticated Crop, *Proceeding of Meeting of IOFRO WP 2.02.06 and Frans Kempe Symposium*, August   24-28, pp. 375–387, Sweden

Yüksek, F. (1997). The Genetic Diversity of Black Pine Populations in Turkey, Ph. D. Thesis, Karadeniz Technical University, pp. 77, Trabzon, Turkey

Zhuowen, Z. (2002). Differences in Flowering Characteristic among Clones of *Cunninghamia lanceolata* (LAMB.) Hook, *Silvae Genetica*, 51, 5–6, pp 206–210

Zobel, B. & Goddard, R. (1954). *Pine Flowering and Seed Ripening in Texas*. Tex. For. Serv. Res. Note, No. 8, pp. 10

Zobel, B.J. & Talbert, J. (2003). *Applied Forest Tree Improvement*, John Wiley&Scons, pp. 505 New York, England

Zobel, B.J.; Barber, J.; Brown, C.L. & Perry, T.O. (1958). Seed Orchard; their concept and management, *Journal of Forestry*, 56, pp. 815- 825

Genetic Diversity of Bulgarian
Phaseolus vulgaris L. Germplasm Collection

N. B. Tomlekova

Maritsa Vegetable Crops Research Institute
Bulgaria

1. Introduction

The common bean, *Phaseolus vulgaris* (L.) *Savi*, is a member of the genus *Phaseolus*, belonging to subtribe *Phaseolinae*. It in turn is a member of the tribe *Phaseoleae* of the subfamily *Papilionoideae* and the family *Leguminosae (Fabaceae)* (van Eseltine, 1931; in Lackey, 1977; Maréchal et al., 1978). The neotropical genus *Phaseolus* includes about 50 to 60 species of wild bean (Toro et al., 1990), most of them distributed in Mesoamerica, five of which are cultivated: *P. vulgaris* (L.) *Savi* (common), *P. dumosus* Macfady (tepari), *P. coccineus* L. (multicolored, scarlet-coloured), *P. lunatus* L. (Lima) and *P. polyanthus Green* (Debouck, 1989; Caicedo et al., 1999; Delgado-Salinas et al., 1999; Gutiérrez Salgado et al., 1995). The last four species have much narrower commercial importance than *P. vulgaris*. In the USA less than five economically important varieties of common bean occupy half the area of this culture and at the same time it is from there that more than 10% of world production originates (Pachico, 1989). The vast narrowing of genetic variability is an indication of the risk of potential loss of a species and emphasizes the importance of its evaluation.

1.1 Increasing the genetic diversity in *P. vulgaris* L.

Various ways to increase the genetic diversity of *P. vulgaris* are known:

1.1.1 Interspecies hybridization

Studies of interspecies hybridisation have been carried out by Baudoin et al. (1985); Hucl & Scoles (1985); Federici & Waines (1988); Cabral & Crocomo (1989); Katanga (1989); Baudoin & Katanga (1990) and Smartt (1990), as cited in Maquet et al. (1997).

Harlan et al. (1971) with modifications (Smartt, 1990, as cited in Maquet et al., 1997) distinguish three main centres of origin:

- Primary gene pool, the equivalent to the concept of a species, whose members are free to cross-fertilize, and consists of cultivated and wild forms of the species.
- Secondary gene pool, the equivalent of all the species that can give viable or partially fertile hybrids with the primary one.
- Tertiary gene pool includes types among which artificial hybridization is possible, but the resulting hybrids are sterile, abnormal or lethal.

The genus Phaseolus is characterized by two main poles (Marshal et al., 1978). The first one includes the complex *P. vulgaris* - *P. coccineus*, and at the opposite end is *P. lunatus*. Interspecies hybridization has allowed to specify the proximity of a large number of species in this genus. The results in this aspect are summarized in several publications (Lorz, 1952; Dhaliwal et al., 1962; Fozdar, 1962, as cited in Baudoin, 1981 & Maquet, 1995).

Sources of genetic diversity of *P. vulgaris* are wild species of the genus *Phaseolus,* but also cultivated ones:

- *P. polyanthus* easily hybridizes with the other four cultivated species and is considered the ancestor of the common bean. Its value lies in its excellent taste.
- *P. coccineus* is difficult to hybridize with other *Phaseolus* species, but hybridizes successfully with *P. vulgaris*. It is of interest with its taste qualities, great potential for yield, resistance to low temperatures, viral, bacterial and fungal diseases. Because of these valuable traits, it is used in the breeding process in Bulgaria for gene transfer into *P. vulgaris* varieties, which enriches the Bulgarian working collection (Zagorcheva et al., 1983a, 1983b; Poryazov et al., 1988).
- *P. lunatus* is difficult to hybridize with the common bean. Lima beans are mainly grown for the green-coloured seeds. It is of interest for its excellent taste, great potential for yield and disease resistance.
- *P. acutifolius*, cultivated mainly for mature grain, is valuable for its resistance to bacterial blight and drought tolerance. It hybridizes with great difficulty with common bean (Poryazov et al., 1988, as cited in Kumanov et al., 1988). Successful hybridization has been achieved in Bulgaria using the *in vitro* methods for embriocultures (Nikolova et al., 1986; Rodeva, 1988).

1.1.2 The introduction of genotypes

The introduction from the primary centres of origin in the collections of crop plants is preceded by identification and characterization of sources and donors of valuable characters for selection, which is very important in organizing pre-breeding work (Konarev, 1998b). In targeted selection of accessions in breeding, it is especially important to establish and implement markers that are genetically linked to important economic traits.

1.1.3 Selection of local accessions

All data described support the American origin of the genus *Phaseolus*. The cultivated forms of beans were distributed around the world shortly after the discovery of America by Columbus in 1492. In Europe, a secondary gene pool, the path of introduction of bean species passes through Spain, Turkey, and only then to Bulgaria, about 300 years ago. The pathways of distribution of beans into and across Europe were very complex, with several introductions from the New World that were combined with direct exchanges between European and other Mediterranean countries. Currently, *P. vulgaris* has pantropical distribution, with a distinct genetic diversity for Europe, Africa and Asia.

The lack of homogeneity of the habitat determines the unevenness in the distribution of genetic diversity within species (Vavilov, 1987). There are many examples of differences in the variability distribution of specific traits, especially the resistance to diseases and stress. What matters is knowing the extent of diversity and the distribution of specific alleles

(Marshall et al., 1975; Maquet et al., 1999). Accessions introduced centuries ago have adapted to our latitudes and are thus regarded as Bulgarian. In the process of adaptation of the introduced forms to specific latitudes and altitudes, it is possible that spontaneous mutations, allowing better adaptability to the new environment of development, are retained. The presence of a vast genetic diversity adequately reflects the environmental conditions and guidelines for breeding.

1.1.4 Induced mutagenesis

Induced mutagenesis is a powerful source of genetic diversity. Mutational techniques lead to the improvement of important economic characters in *P. vulgaris* varieties useful in breeding. The induction of mutations for increasing the genetic basis or variability is not a new approach in agriculture, but it could become revolutionary if modern biological techniques are used, such as development of DNA-based molecular markers to map and/or clone these mutations (Fofana, 1995; Angioi et al., 2010). This could not only help molecular characterization of mutations, but also anable the marker-assisted selection (MAS). For many years and many efforts to increase natural genetic resources and to improve cultivated forms, breeding lines have been established in the working collection by applying these methods (Poryazov, 1990; Sofkova & Yankova, 2008; Tomlekova, 2010).

1.1.5 Genetically modified organisms

Genetic engineering employs techniques of targeted transfer of genes encoding valuable economic characters, among which the most frequent are resistance to pathogens and pests. Transformed genotypes are means of solving economically important breeding issues and a tool to increase the genetic diversity of a species. Molecular characterization and cloning of valuable genes is a source for the implementation of their transfer into other closely related or more taxonomically distant species. Based on studies, Shade (1994) and Ishimoto (1996) (as cited in Tomlekova, 1998) have conducted the transfer of a gene from *P. vulgaris* in *Pisum sativum* and *Vigna vulgaris*, which has a key role in determining resistance to the bruchid beetle *Callosobruchus maculatus*.

1.2 Levels of characterization of genetic diversity in *P. vulgaris* L.

Various data confirming variability can be collected at phenotypic level of the plant, as well as at various protein and nucleotide levels (electrophoresis of storage proteins and/or enzymes, total DNA restriction profile, chloroplast DNA and/or mitochondrial DNA, DNA sequencing) (Singh, 1991; Fofana et al., 2001; Tomlekova et al., 2009). Regarding the choice of marker used (morphological and/or molecular), there are ongoing disputes (Hills, 1987; Moritz et al., 1990; Donoghue et al., 1992; CIAT Report Information, 1990b).

1.2.1 Morpho-physiological characters

Material from the world collection of *Phaseolus*, stored in CIAT, has been characterized in detail and its genetic diversity has been calculated (Frankel, 1987). There have been observed special characters that allow to identify a given accession introduced. The data encompassed within the collection (the so-called "passport data") include basic information on the origin and place of collecting the accessions. The characterisation employs gene

markers which are mostly monogenic or linked genes, such as for the type of germination and vegetative appearance. The characters must be easy to identify and should be stable to the impact of a great number of external factors. In this way they promote the classification and description of the forms (Leakey, 1988).

The assessment includes identification of characters that are important for the potential of species, i.e. those that can be introduced into breeding programs. Most often they are polygenic and are influenced by environmental variations. The most sensitive of all characters are the ones related to the yield and the yield itself, tolerance to stress factors, e.g. tolerance to drought and high temperature stress (Nikolova et al., 2003; Krasteva et al., 2004; Sofkova & Petkova, 2007; Petkova et al., 2010), and resistance to pests and diseases (Poryazov, 1990; Yankova et al., 2011). To standardize the different characteristics and assessments, the International Plant Genetic Resources Institute (IPGRI) has published classifiers for many species (van Slotten, 1987). They can be improved according to the needs of a particular study (van Schoonhoven et al., 1987, as cited in Zoro, 1999). The observations must be coded and collected in a database in a way so as to facilitate their use in statistical analyses (Vanderborght, 1988). To study the structure of genetic diversity of *Phaseolus*, different statistical methods have been used (Singh et al., 1991a). The analysis of the major components allows to differentiate the accessions of *P. vulgaris* into two groups, Mesoamerican and Andean (Singh et al., 1991a), because they have different morpho-physiological characters. Singh et al. (1991b) grouped them into six races: three Mesoamerican races (Durango, Jalisco, Mesoamerica), and three Andean races (Chile, Nueva Granada, Peru). A prevalence of the Andean type (76%) was first detected by Gepts & Bliss (1988), and was then confirmed by Lioi (1989) in an analysis of a large collection from Italy, Greece and Cyprus (66%), by Logozzo et al. (2007) from abroad European collections (76%), and by others, from Portuguese and Spanish samples. Some researchers such as Adams et al. (1988) in Malavi, Castineiras et al. (1991) in Cuba and Gil de Rona (1992) in Spain have used statistical methods to estimate the variability and to explore the genetic structure of their national collections in order to give direction to their future breeding programs. The use of morpho-physiological characters is not without inconvenience. They respond to the environment and are often correlative, which reduces the number of discriminant characters. Moreover, morphologically similar varieties may differ at molecular level. Dominance in morphological characters of wild forms is a bottleneck in the search for polymorphism in a population. To know the exact genotype of an individual, proceeding from a certain morphological character, is difficult because it often depends on many genes that may have pleiotropic effects, of epistasis and/or can be influenced by the surrounding environment. These shortcomings have forced scientists to look for other markers to better characterize the genotype studied. It is assumed that the similarity or difference at the phenotypic level reflects similarity and difference at the genetic level. Protein and nucleic acid markers do not depend or rely less on the environment and the former often are codominant. According to Mayr (1992, as cited in Maquet, 1995), using only the morphological characters is helpless in cases such as "twin-species" (two different species with similar morphological traits) or in cases of multiple distinct intraspecies variants. The identification of accessions requires knowledge of the genetic basis of morphological diversity.

1.2.2 Variability of seed storage proteins

The most frequently used organ for conducting biochemical research is the seed as it is a well-defined stage in the plant life cycle and has a high concentration of proteins (Konarev,

1993). In bean seeds, globulins account for the largest proportion followed by albumins, and the lowest are glutelins. Electrophoresis on polyacrylamide gel (SDS/PAGE) in the presence of SDS clearly distinguishes between Lima bean and common bean (Manen et al., 1981; Baudoin et al., 1991). *P. lunatus* is indisputably different from *P. vulgaris, P. coccineus* and *P. acutifolius* on account of the absence of phaseolin established by immunoelectrophoresis (Kloz, 1962, 1971, as cited in Maquet, 1995) and by SDS-PAGE (Derbyshire et al., 1976, as cited in Maquet, 1995). In the seeds of *P. vulgaris* the globulin fraction predominates, with phaseolin being the most abundant. The albumin fraction is the second largest, including lectins and other albumins (Gottschalk et al., 1983). Two types of electrophoretic analysis are employed in beans. One is based on total proteins as a criterion for grouping, which is used with *P. polyanthus*, and the other on individual proteins like phaseolin and arcelin in *P. vulgaris.*

Our working collection includes later generations, breeding lines, developed, highly resistant to *Acanthoscelides obtectus* – a bruchid beetle that is widespread in the latitudes of Bulgaria (Yankova et al., 2004; Yankova, 2010; Sofkova & Yankova, 2011). The study of storage proteins is of interest in the fight against this pest, as the substance that has a major role in the mechanism of resistance – a lectin-like substance, located in the area of phytohemaglutinins, corresponds to arselin, which determines the resistance to *Callosobruchus maculatus* and *Zabrotes subfasciatus* and can be isolated from proteinograms (Dobie et al., 1990; Tomlekova, 1998).

The vicilin-like globulin, called phaseolin, isolated from bean seeds, is well studied. It includes a group of related proteins whose genetic control has already been characterized (Chrispeels, 1978; Brown et al., 1981; Hall et al., 1977; Ma et al., 1978). In the sites of origin of *P. vulgaris*, the 'S' type has been found to prevail for the Mesoamerican (M) gene pool, initially described in the variety 'Sanilac', presented in cultivated and wild genotypes and 'M' phaseolin type in wild forms, where 'I' and 'J' types are also found (Gepts et al., 1986a; Koenig et al., 1989b). In the Andean (A) gene pool, 'T' phaseolin dominates, first described in the variety 'Tendergreen' (Gepts et al., 1986a; Gepts et al., 1986b), but 'H', 'C' and 'CH' types in wild and cultivated forms are also found (Koenig et al., 1989b). In the Mediterranean region (Gepts et al., 1988; Lioi, 1989) there are reports on 'T', 'S', 'C' and 'B' types, the latter being characteristic of wild and cultivated accessions. Along with the phaseolin variants such as 'Sb' and 'Sd', identified by two-dimensional electrophoresis (2D-SDS/PAGE/IEF), genetically similar to type 'S' (Gepts et al., 1986a); 'M$_1$', 'M$_2$', 'M$_3$', 'M$_4$', 'M$_5$', 'M$_6$' (Romero-Andreas et al., 1985); 'Ca', 'Nu', 'Ko' close to phaseolin 'T' (Tohme et al., 1995) and other variants representing its heterogeneity, so far, in total there are about 30 different known types of phaseolin in the cultivated representatives of *P. vulgaris*, while in the wild ones they are about twice as many (Debouck, personal communication). Research in molecular biology on phaseolin complements its characterization (Hall et al., 1977; Slightom et al., 1983; Talbot et al., 1984; Lawrence et al., 1990; Kami et al., 1995). The subsequencing analysis has shown that this division into two groups is at the level of nucleic acids, at which 'S' and 'T' types can be distinguished by a single polymorphism using Southern hybridization (Talbot et al., 1984; Nodari et al., 1992). Phaseolin DNA sequencing reveals that the gene phaseolin subfamilies are two, α and β, and that they differ by the presence or absence of a short direct repeat of the phaseolin stretch of DNA, consisting of 27 bp (base pairs); α and β are similar to each other in the 'S' and 'T' phaseolin types. They only differ in the α gene type by an additional repeat, made up of 15 bp, which

is found in 'T' phaseolin and is absent in 'S' (Slightom et al., 1985). They are responsible for the synthesis of protein complexes observed in the SDS-PAGE phoregrams. The presence of these repeats in combination with the post-translational glycosylation is responsible for the different 'T' phaseolin profile of gels in one-dimensional SDS-PAGE (Lioi et al., 1984; Sturm et al., 1987). In addition, the sequential analysis of the phaseolin type, characterizing the large groups according to the centres of origin of *P. vulgaris*, may enable the tracking of the potential sites for amino acid substitutions to enrich the methionine content of phaseolin (Gepts et al., 1984; Ma et al., 1978). By modifying phaseolin it is possible to achieve variability and because its proportion is significant among the total protein content of seeds, in the long run this approach would lead to overcoming the nutritional deficit present in the family *Leguminosae*. The amino acid content of protein determines its fraction composition. The nutritional value of beans depends on it. In general, dicarbonic amino acids bean varieties are present in the largest quantities, while the lowest are the sulfur-containing. The proportions of aspartic and glutamic acids, arginine, lysine, leucine and threonine are significant. A slight deficit of valine and isoleucine has been established in some bean species. The deficiency of methionine and cysteine is significant, their content being relatively higher in *P. lunatus* than in other representatives of the genus *Phaseolus* (Stoyanova, 1989). The phaseolin marker is most often used to characterize wild or cultivated populations by detecting new phaseolin types. Thus the identification of the phaseolin type is used as a geographic and evolutionary marker for studying variability, grouping of the germplasm of the type and studying the geographical distribution of groups. The phaseolin marker is also used to search for phaseolin recombinants whose existence could demonstrate combining the high productivity of the 'S' type with, for example, disease resistance or other characteristic of the 'T' type (Gepts et al., 1986b; Tomlekova et al., 2001; Genchev, 2011). The biochemical characterization of the accessions and above all the identification of their phaseolin type complete the overall characteristics of the breeding materials and the collections in the genetic banks. The domestication of the common bean was conducted in separate areas where accessions with small or large seeds prevail. The differentiation of the centres of origin of the species in seed size does not provide distinct differences, i.e. it does not always serve as a discriminant character. In contrast, the genetically determined biochemical character - phaseolin type of seeds is a reliable distinguishing factor.

1.2.3 Polymorphism of enzymes

Enzymes used as another type of protein markers, are characterized by catalytic power and specificity. Enzyme electrophoresis uses both the protein nature and specificity of enzymes. Contrary to storage proteins, enzymes should not be denatured. The ubiquity of enzymes and their role in metabolism is the reason for analysis of many tissues (cotyledons, roots, leaves, pollen, etc.). Isoenzymes are convenient markers in the study of plant taxonomy (Crawford, 1990; Murphy et al., 1990), evolution of cultivated plants (Doebley, 1990), genetics of populations (May, 1992) and conservation of plant genetic resources (Schaal et al., 1991; Gepts, 1995). The presence of genetic diversity is the basis of plant breeding. The description is also required for proper upbuilding the genebanks, for performing efficient selection of genotypes maintained in them, building a 'core collection'. Detailed studies have been conducted in predominantly autogamous populations of different species (Loveless et al., 1984). The first studies of *P. lunatus* were conducted with seven enzyme

systems of 20 populations originating from the central valley of Costa Rica by Maquet et al. (1996). At the level of populations, the studies with enzyme markers in this species were conducted by Zoro et al. (2003). The genetic diversity of the following species has been studied: *Arenaria uniflora* by Wyatt et al. (1992); *Bromus tectorum* by Novak et al. (1991); *Ceratophyllum demersum* and *C. echinatum* by Les (1991); *Datura stramonium* (Motten et al., 1992); *Eichhomia paniculata* by Barrett & Shore (1990); *Phaseolus acutifolius* by Schinkel & Gepts (1989); Koenig & Gepts (1989); *P. lunatus* by Wall et al. (1975, as cited in Brown, 1978); Baudoin (1991); Zoro (1994, 1997, 1998, 1999, 2003); *P. vulgaris* by Koenig & Gepts (1989); *Setaria glauca* by Wang et al. (1995b); *Sorghum bicolor* by Ollitrault, (1987); *Sphodromantis viridis* by Wang et al. (1995a).

The absence of natural areas in the distribution the species *Phaseolus vulgaris* and the transition from traditional to intensive agriculture in recent decades have contributed to the significant loss of genetic diversity. The threat of genetic erosion in common bean has become ever more real. Any analysis conducted to assess genetic diversity can help to solve the problems of species selection (Chase et al., 1991; Nienhuis et al., 1995; Forfana et al., 1997; Maquet et al., 1997).

Mendel's genetic basis of isoenzymes is a source for their numerous applications. The easy extraction of these proteins facilitates the research of a huge number of species, and the possibility of defining the homo- and heterozygote ones among them makes them irreplaceable markers in crop genetics. By using isozymes, intra- and inter- crop variability can be characterized and the degree of similarity among accessions can be evaluated (Hamrick et al., 1991; Schaal et al., 1991a).

In systematics, an enzymatic analysis defines the taxonomic relations in one group (generally the different species in a genus) by taking into account the percentage of common alleles (Crawford, 1990). West & Garber (1967) and Bassiri & Adams (1978) proved the existence of a close similarity between *P. vulgaris* and *P. coccineus* as well as the possibility to develop hybrids of these two species. Despite their cross-ability, there is a problem with rejecting the genome that originates from the male gametophyte (Wall, 1968; Wall & Wall, 1975; Guo et al., 1989). By using isozymes, a breeding program can be performed for hybrids that are close to common beans but in addition possess a part of the *P. coccineus* genome.

The greatest advantage of enzyme electrophoresis is the possibility to associate the observed phenotype (the profile of the bands on the gel) to a genetic basis. Proper analysis of the gels depends not only on the quality of enzyme expression but also on the genetic determination of the enzyme system. The aim is to identify how these different isozymes translate the genome variability. The differences derive from: the expression products of many alleles of one and the same gene with a single locus; the expression products of many genes located on equally many loci; the different molecules resulting from the changes in the conformation of one and the same protein molecule; the molecules synthesized by one gene or a group of genes, which have undergone various posttranslational modifications (Ryder et al., 1980, as cited in Maquet, 1995).

Morphological markers allow to classify wild and cultivated plant species, but when it is necessary to classify individuals of the same species, for example from different populations (as is the case of the below presented study), descriptive capacity of morphological markers is quite limited (Roux, 1987). The Mendelian genetic basis of the enzymes determines their

multiple applications. Their easy extraction allows a large number of individuals to be studied. The possibility to detect homo- and heterozigotes make isozymes essential markers in the genetics of populations, as the dominance of morphological characters impedes their characterization. Isozymes are useful in the characterization of intra- and interpopulation variability and the degree of similarity (Hamrick, 1990; Schaal et al., 1991; Sozinov, 1995). Through them the populations can be grouped to follow the dispersal and colonization of one species, to calculate their dynamics (Barrett et al., 1990). Isozymes markers enable the study of population variability depending on the surrounding environment (conservation *in situ*) or according to the conditions of multiplication and regeneration of the collections (conservation *ex situ*) (Konarev, 1998a, 1998b). By following the dynamics, it is possible to clarify the factors responsible for the maintenance and loss of polymorphism within populations. In taxonomy, especially for species in a genus, enzyme analysis determines the taxonomic relationships in a taxonomic group, based on the percentage of common alleles (Crawford, 1990). Lefort-Buson et al. (1985), Roux (1987) and Murphy et al. (1990) summarize the different uses of isozymes in breeding. The enzymatic analysis has allowed the identification of certain species of the *Phaseolus-Vigna* complex to be improved (West et al., 1967a; Bassiri et al., 1978a; Jaaska & Jaaska, 1988), and the study of the species *P. coccineus* (West & Gabet, 1967b; Bassiri et al., 1978; Pinero et al., 1988). Despite the possibility of a crossover between these two species, there is rejection of the genome of the male gametophyte (Wall, 1968; Wall et al., 1975; Guo et al., 1989). Isozymes allow that the program for selection of generations, closely related to common beans and possessing a part of the *P. coccineus* genome, not to be eliminated too quickly.

The number of enzymes that can be tested is about 100 of the 2000 that are known, which is a small number compared with the overall genome of the plant. Very often only 10-30 enzymes have good electrophoretic separation (Zoro, 1994; Zoro, 1999).

All this is indicative of the difficulties encountered in the introduction of new enzyme systems in a particular species. Enzyme electrophoresis does not fully establish variability. Of all substitutions at the gene level, only about 30% result in an amino-acid substitution with a change of the net charge of protein (Pasteur et al., 1987). Many triplets encode one and the same amino acid. The substitution of one amino acid by another with the same charge would not normally have any effect on the net charge of protein. Therefore, identical electrophoretic mobility does not necessarily mean the same DNA sequence. In all cases, however, the quantity of unmarked variability would be proportional to the amount of the marked one (Crawford, 1990). For a given species, the detection of rare alleles would not change significantly the frequency of common alleles and genes that were monomorphic and will probably remain such.

Biochemical and morphological data have supported the existence of two gene pools in *Phaseolus vulgaris* species. In recent years, combinations of phaseolins and various morphological characters have been intensively used to analyze the structure of the European common bean and the presence of representatives of both gene pools. A prevalence of the Andean type (76%) in the European collections was first detected by Gepts & Bliss (1988), and was then confirmed by Lioi (1989) in an analysis of a large collection from Italy, Greece and Cyprus (66%). It was also confirmed in an analysis of a large collection from Portuguese and Spanish samples by Rodiño et al. (2001), Ocampo (2002), and Rodiño et al. (2003); and in representative European collections by Logozzo et al. (2007).

In some self-pollinating plants, genetic variability can be too low to characterize a species. The number of polymorphic loci and the number of alleles per locus are limiting in the case of restricted gene flow. The low level of polymorphism in these autogamous species makes these markers inadequate to identify the genetic structure of the species. The search for allozymes to be used as molecular markers makes it possible to trace the variability between individuals within the populations and between the populations, irrespective of environmental factors. The fact that in other genotypes or related species of *P. vulgaris* there are allozyme variants of the enzymes described is a reason for conducting this study. This evaluation would allow to better understand the factors responsible for the availability or loss of polymorphism. In connection with the expected results, we cannot but mention the disadvantages of isozymes as markers of genetic diversity. In cases of restricted gene flow in some autogamous plants, genetic variability may be too low to characterize a particular species. The absence of polymorphism in self-pollinating species can render these markers inadequate to identify the genetic structure of the populations studied. To avoid difficulties associated with the use of biochemical and morphological markers, researchers have turned to DNA markers in order to obtain a direct evaluation of the genetic variability.

Koenig & Gepts (1989a) suggest the use of DNA-markers to characterize the genetic diversity of common beans as a consequence of the low variability assessed using isozymes. Enzyme electrophoresis does not fully establish variability. Only about 30% of the substitutions at the gene level result in an amino acid substitution with a change in the net charge of protein (Pasteur et al., 1987). Many triplets encode one and the same amino acid. The substitution of one amino acid by another with the same charge would normally have no effect on the net charge of protein. Therefore, the identical electrophoretic mobility does not necessarily mean the presence of the same DNA sequence. In all cases, however, the quantity of unmarked variability would be proportional to the amount of the marked one (Crawford, 1990). Considering above all the high cost of consumables and equipment needed to perform DNA analyses for a large-size population and the fact that so far in Bulgaria enzyme systems have not been applied in studying the variability of *P. vulgaris* they were used for the purpose of the present study.

1.2.4 DNA molecular markers

The direct use of DNA ensures accuracy and speed of the techniques due to their specificity, while there is no requirement to wait for the manifestation of a property at a certain age, and also is not affected by the surrounding conditions. With the development of new techniques for DNA testing, the study of the genome is achieved in its entirety. Techniques are not intended to detect characters and to draw conclusions on the activity of genes. They aim to isolate DNA from samples that are representative of the studied populations. Therefore, they consist in determining the differences in the compositions of the DNA samples analyzed that are to serve as markers for variability. They reveal the genetic diversity at the level of genetic information. Once DNA has beens isolated, it can be analyzed either by fragmentation by enzymes, involving restriction fragment length polymorphism (RFLP - Restriction Fragment Length Polymorphism), or by amplification (PCR-based techniques) and DNA sequencing. For example, in studies using *Phaseolus*, based on the use of Randomly Amplified Polymorphic DNA - RAPD (Nienhuis et al., 1995; Fofana et al., 1997), Inter-Simple Sequence Repeats - ISSR, Amplified Fragment Length Polymorphism - AFLP (Svetleva et al., 2006), diversity is calculated not only according to

geographical origin and botanical status, but also depending on the cultivated group. Based on the organization of genetic diversity observed in *P. lunatus*, Fofana et al. (1997) offer a hypothesis on the process of domestication of this species.

A summary of molecular marker capabilities indicates that the molecular markers have various applications and utilities, being indicated for different objectives and/or populations. It is advisable that markers that have low cost are applied properly, i.e. in fingerprinting and/or variability studies, suggestions point favourably to microsatellites and AFLP, or RFLP when synteny information is available (Park et al., 2000; Svetleva et al., 2003; Torres et al., 2004).

Depending on species and related breeding problems for solving, additional techniques may be used, such as DNA microarrays, single nucleotide polymorphisms (SNPs) and ELISA. General analytical approaches such as MAS, tagging, genotyping for germplasm surveys, heterosis groups, cytogenetic stocks, pedigree analysis, gene bank inventory/collections are useful for gaining further information about the characters of interest. The development of such data will provide support for utilization of more powerful techniques for physical mapping such as bacterial artificial chromosomes (BACs), tranformation-competent artificial chromosomes (TACs) and gene isolation (Working material, 2001).

International organizations such as the International Plant Genetic Resources Institute (IPGRI), el Centro Internacional de Agricultura Tropical (CIAT) and the International Institute of Tropical Agriculture (IITA) have undertaken numerous studies to establish and maintain gene banks, compiling catalogues and formulae of database, based on up-date and dissemination of information for the genetic resources, input and information analysis, curation database through data collection. To increase the efficiency of utilization of plant genetic resources in improving the existing and developing new varieties, the study of genetic diversity has an important role, and so does the proper organization of the crop collection (Konarev, 1998a). An important stage in the process of formation of the collection is the identification and registration of the accessions and the molecular markers are very helpful. As noted, for the identification and registration of genetic diversity DNA and protein markers are used. For a number of reasons (co-dominant nature of inheritance, and also inexpensive and simple techniques, well reproducible results), the latter are employed with priority in studies of *P. vulgaris*.

Part A

A study of genetic diversity of *P. vulgaris* L. by total storage proteins

The aim of the present study was to investigate the *P. vulgaris* genetic diversity in the Bulgarian collection through identification of phaseolin types, with a view to assist gene bank management and plant breeding.

2. Materials and methods

In this study, 409 common bean genotypes were investigated. Among them 235 local accessions from the gene bank of the Institute of Plant Genetic Resources in Sadovo, collected by 21 expeditions since 1977, were included. The accessions were sampled from 69 locations (Fig. 1) in order to cover the whole territory of the country and to be representative of the

geographical and ecological areas of Bulgaria. Another set of genotypes consisted of 127 breeding lines and varieties resulting from breeding programmes and available in the working collections of the Maritza Vegetable Crop Research Institute, Plovdiv, and of the Dobroudza Agricultural Institute, General Toshevo. The remaining 47 recently introduced accessions during the last ten years came from different gene banks outside the country; such accessions are integrated as parental genotypes in hybridization programmes with local accessions and breeding lines and varieties from the Maritsa and Dobroudza Institutes working collections. In order to compare the electrophoretic profiles of the Bulgarian accessions, controls ('witnesses') for the phaseolin types 'S', 'T', 'Ca' and 'C', were obtained from CIAT, Cali, Colombia, recognized as the worldwide repository bank of *Phaseolus* genus.

2.1 Basic method - 1D-SDS/PAGE

Identification of phaseolin type was performed in 10 single ripe seeds from randomly selected plants of each accession by one-dimensional SDS/PAGE (Koenig et al., 1970). Two protein molecular weight markers, purchased in 1998 from Merck, MMW–1 – Standard mixture IV (Art. No.15791) and MMW-2 - Standard mixture VIII (Art. No.11536), were used for phaseolin zone determination on the proteinograms. Protein types were identified thanks to referential accessions: ICA Pijao - for 'S' type phaseolin, CDRK - for 'T' type, G2362 - for 'Ca' type, and G21194 - for 'C' type.

2.2 Assessment the intensity of protein fragments

For the phaseolin identification of variants with close electrophoretic phenotypes, a separation to 1D-SDS/PAGE proteinograms was performed. Proteinograms were photographed with a gel documenting system using the Ultra Violet Products Gel Documentation System, Version 7.00 Science Park, Milton Road, Cambridge CB4 4FH, England. Consecutively Gel Work-1D-Advanced, Version 3.01 (http://www.statsoft.com) software product was applied. The intensity of each fraction of the phaseolin zone on the proteinograms obtained from different accessions was measured and compared. Intensity measuring is a method developed by the authors and has not been published yet. This was also useful to distinguish the phaseolin types 'T' and 'Ca'.

The calculations for each fraction of the proteinograms were made including:

- Rf relative mobility of the molecules (distance from start to band relative to length of phoregram)
- RV (Raw Volume – intensity expressing in relative units protein levels)
- The molecular weight was measured by the method of Weber & Osborne (1967, in Acquaah, 1997)
- Factor (Kn) to equalize protein levels for each proteinogram (calculated by the RV of a band with the same intensity at individual starts (reference) and their overall RV)
- Construction of graphs for comparing the intensity of phaseolin fractions.

2.3 Mathematical procedures

To be grouped in clusters data from the proteinogram analysis, the following procedure was applied: the Euclidean distances between genotypes were calculated using the data from the analysis of storage proteins with the formula:

$$E(X,Y) = ((X_1 - Y_1)^2 + (X_2 - Y_2)^2 + \ldots (X_K - Y_K)^2)^{0.5}$$

The matrix was formed from the distances and the distances thus determined were classified according to proximity. Statistica, Version 6.0. software was applied. Each genotype was involved in the assessment of the intraspecies diversity through a representative profile (> 90%) in an analysis of 10 single seeds.

To calculate the diagrams representing the proportion of accessions from the Andean and Mesoamerican gene source, in each climate zone the standard procedures by using Microsoft Excel Help were implemented.

3. Results and discussion

The accessions available in our collections were found in four different climatic zones (Fig. 1). The most appropriate climatic zone, i.e. the Highlands zone, was represented by the lowest number of accessions in Bulgaria. This can be explained by a different adaptation of common bean accessions (only domesticated) in a secondary diversity centre, compared with the adaptation of accessions (both wild and domesticated) occurring in the primary diversity centre. The breeding "pressure" (both natural and anthropic) is different between Bulgaria and the neotropical regions. More details on the geographical distribution of these localities are given in Tomlekova et al. (1999).

3.1 Phaseolin types

The screening of 409 common bean local accessions, breeding lines and varieties from the Bulgarian collection allowed us to identify four phaseolin types: 'S', 'C', 'Ca' and 'T'. Phaseolin 'S' corresponds to the Mesoamerican gene pool while 'C', 'Ca' and 'T' types correspond to the Andean gene pool (Fig. 1).

The proteinograms of the total seed storage proteins include the fractions corresponding to the phaseolin types, which identification was described in Tomlekova (1999) and Tomlekova et al. (2009). No difference in phaseolin type was observed from single ripe seeds belonging to each Bulgarian accession except few cases within the breeding lines and varieties. The different phaseolin types identified in plants from the group of the breeding lines and varieties corresponded to the same types of phaseolin observed in the parents. The phaseolin fractions cover all types assessed in the national collection. Phaseolin fractions on the proteinograms are characterized by molecular weights between 45 000 D and 51 000 D (Brown et al., 1981; Osborn, 1988).

In Fig. 2 and Fig. 3, the three groups of Bulgarian common beans (local accessions; breeding lines and varieties; and introductions) are compared according to phaseolin types. The percentage of the phaseolin types is differently distributed in each one of the groups. 'T' type is prevalent in local accessions while 'Ca' type prevails over the 'C' and 'T' in breeding lines and varieties. The 'C' type proportion is low, especially for the breeding lines and varieties from the Andean gene pool (7%), while in the group of local forms, it represents 25%. 'S' phaseolin type covers 40% and 43%, respectively of the collection made of local accessions on one part, and of breeding lines and varieties, on the other part. The remaining material belongs to the Andean gene pool. The accessions from the A predominate in both local accessions (60%) and breeding lines and varieties (58%). 'T' type predominates in local

forms from the Andean gene pool, while 'Ca' type is the most frequent among breeding lines and varieties from the same gene pool. The difference in the predominating phaseolin types is obvious and the difference of distribution frequencies of all the phaseolin types in the two groups (local accessions, and breeding lines and varieties), reflects the preferences of farmers and breeders for genotypes bringing specific characters. The 'T' type is associated with the white seed coat colour that is preferred by the Bulgarian consumers. The introduced accessions of common bean are characterized with their high proportion of 'S' type. The common bean introductions had a stronger restricting effect on the genetic variability: the meso-American group was presented from an only 'S' phaseolin type. A greater diversity of the accessions from the Andean gene pool, in the introduced collection was established 'T', 'C' and 'Ca', however it was presented mostly from phaseolin 'T' type.
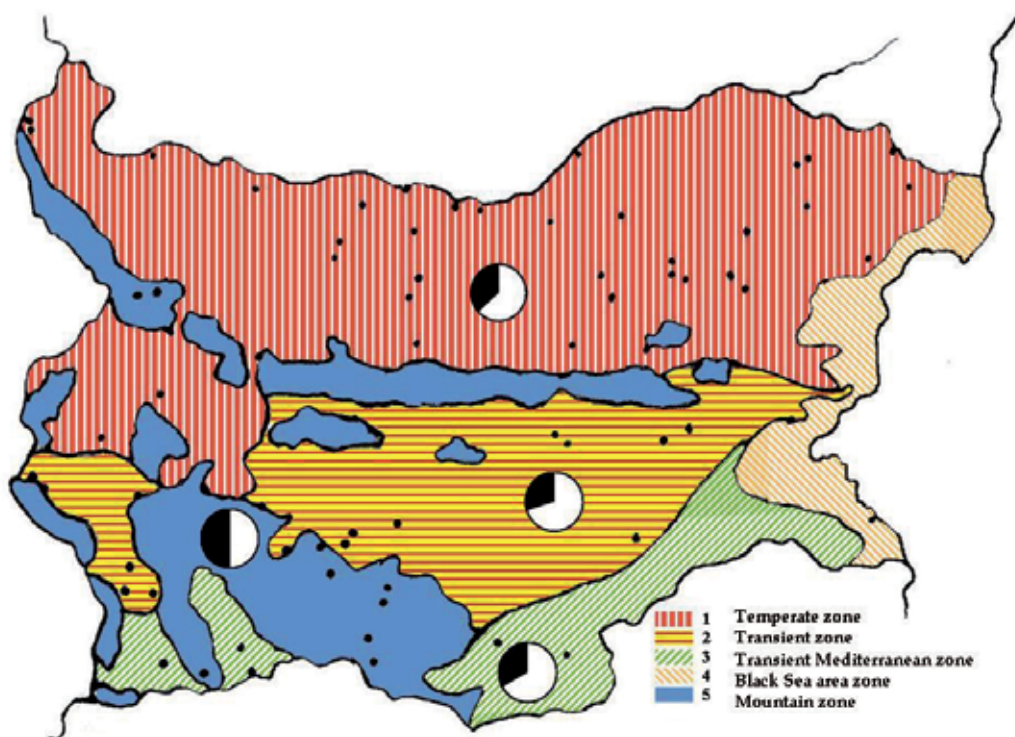
Fig. 1. Map of Bulgaria with marked climatic zones and localities of collection of *P. vulgaris* accessions and diagrams of the share allocation of the phaseolin types (Andean and Mesoamerican groups, white and black colour, respectively).
* The points mark the locations of collecting sample in the territory of Bulgaria.

Classification of the accessions in two basic groups, corresponding to the established both gene pools - Andean (from the region of the Andy) and meso-American (from middle-America) was made on the basis of phaseolin type identification from the electrophoregrams of storage proteins of the dry seeds from common bean accessions (Fig. 2). Data from our study showed that 60% of the Bulgarian accessions came from the Andean gene pool (A) while 40% came from the Mesoamerican one (M). According to a preliminary study (Tomlekova, 1999), the domination of the Andean representatives over

the Mesoamerican ones was expected. However, the results of this study conducted approximately 15 years later later showed a relatively fair proportion of accessions from the Mesoamerican group, which can be explained by the recent introduction of accessions characterized predominantly by 'S' phaseolin type (66%). Among the accessions introduced mostly during the last 10-15 years into our working collections, the Andean phaseolin types represented a lower proportion. Presence of 'S' phaseolin is frequently associated with resistances to important diseases which explain their introduction in the country for breeding purpose and hybridization with local materials (Koenig et al., 1989b). The breeding efforts towards desease and pest resistance to overcome this major problems in crop species (i.e., in common bean), have altered the ration between A and M groups.
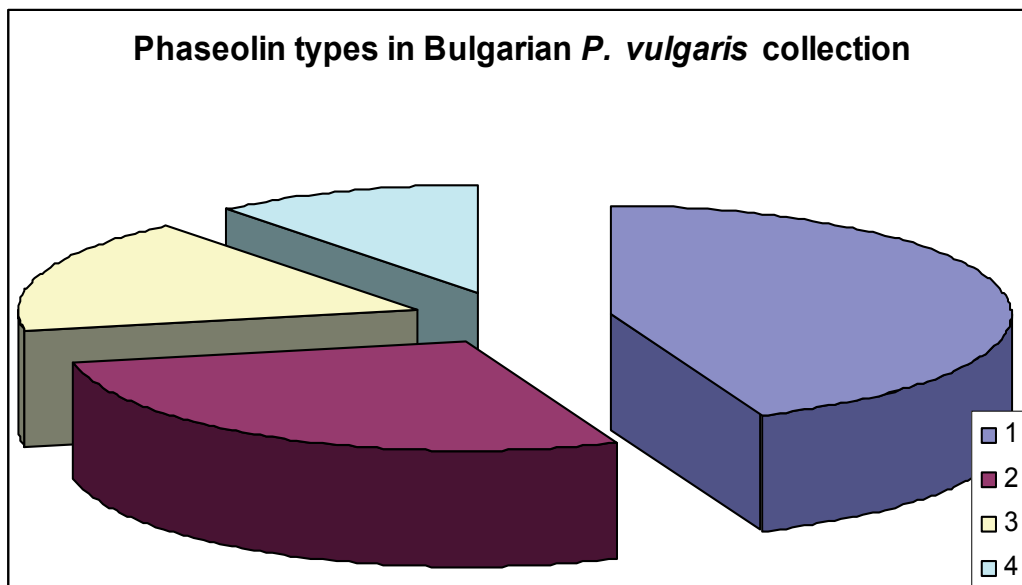

Fig. 2. Proportion of phaseolin types in Bulgarian *P. vulgaris* collection;
* 1- 'S', 2- 'T', 3- 'Ca', 4- 'C'

On the basis of phaseolin determination, results showed a low variability level in the analyzed Bulgarian accessions. The latter represent a very limited genetic diversity, by comparison with the total variability displayed among the material distributed in the primary centre of common bean, i.e. the New World. The low level of genetic diversity is a result of the significant "founder effect", a term introduced by Ladizinsky (1985), in the secondary domestication centres of *P. vulgaris*, such as Bulgaria. Common bean was introduced more than 300 years ago in the country and this food legume was converted as a traditional crop. At the beginning, the level of genetic variability was low due to the limited number of introduced accessions. This reduction of genetic diversity was reinforced because of the lack of adaptation in some materials to the new ecological conditions. Additional factors might have strongly restricted the genetic variability of Bulgarian accessions, such as a specific consumer preference (for large and white seeds), occurrence of diseases (*Bean Common Mosaic Virus*, *Xanthomonas campestris* pv. phaseoli (Common

Bacterial Blight), *Pseudomonas syringae* pv. phaseolicola (Halo Blight) and insects, such as *Acanthoscelides obtectus* (Bean weevil) and the breeding systems of the species (autogamous floral biology).
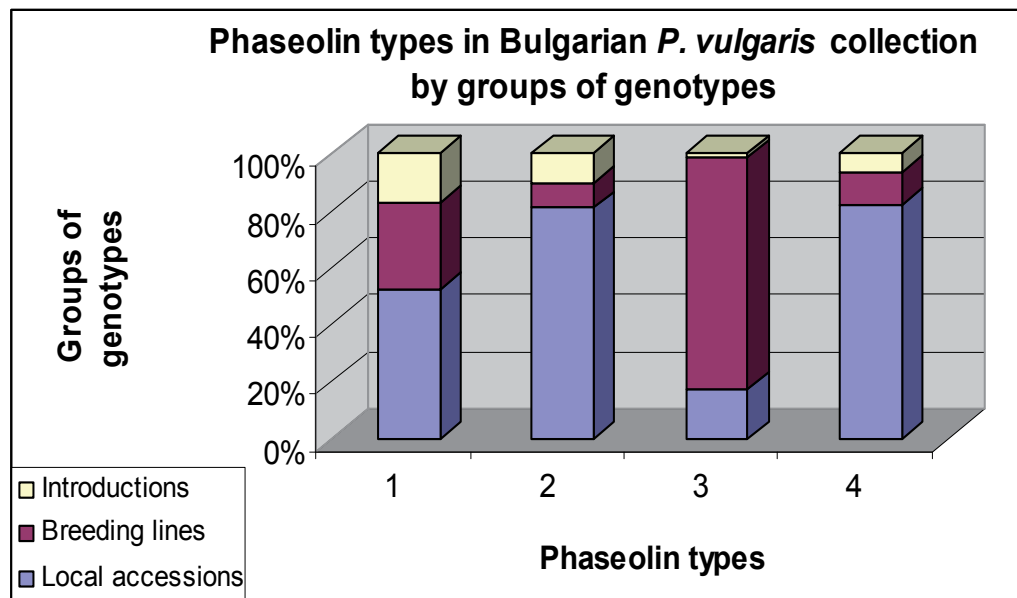

Fig. 3. Proportion of phaseolin types established in Bulgarian *P. vulgaris* collection by groups of genotypes; * 1- 'S', 2- 'T', 3- 'Ca', 4- 'C'

In this context, it might also be relevant to consider the domestication phases in *P. vulgaris*. First, there was a decrease in genetic diversity, but accompanied by an increase in morphological variability (Gradinarov (1939), Vishnevsky (1940), Rachinsky (1968), Ganeva (1983), Stoilova (1998), as cited by Genchev, 2007). This restricting effect was, on the basis of our own results, stronger among the Mesoamerican group represented by one phaseolin type, compared with the Andean group represented by three different phaseolin types. This confirms the data published by Gepts (1990) for the primary domestication centre of the species that can also be used as a plausible explanation for the larger genetic diversity of the Andean accessions in Bulgaria. In a second phase, there was a reduction in genetic diversity explained by the linkage occurring between the phaseolin locus and other loci, influencing the phenotype and submitting the genotype to breeding "pressure" (Maquet, 1995).

3.2 Patterns of storage proteins

To facilitate the analyses, the fractions of the total protein patterns were grouped into zones (Fig. 4):

Zone A; ~ 200 kD of MMW. This zone includes the slowest moving fractions of the phoregrams which have different Rf and RV (Fig. 4). Three variants were observed: Ax - in accessions having 'S' phaseolin (Fig. 4c, i); Ay - in accessions with 'T' phaseolin (Fig. 4a, b, d,

e, f, g, h, j, m; Az – rarely found in some breeding lines (Fig. 4l, m) (not clearly visible in the figure because of low intensity).

Zone B; MMW from ~ 114 to ~ 130 kD. Two variants were observed: A 2 options: Bx - in accessions having 'S' type phaseolin (Fig. 4c, i, k, l, m; By - in accessions with 'T', 'Ca' and 'C' type (Fig. 4a, b, d, e, f, g, h, j, n, o). There is a well-defined correlation, without exception, between Bx profile and 'S', and between and By 'T', 'Ca' and 'C' phaseolin type. Its visualization depends to the greatest extent on the proper application of the method. (The more intense stain in the sample in Fig. 4e, in zone B, marked with an arrow, is an artefact.) Bx and By fractions have the same Rf and differing RV, i.e. the differences in both variants originate from their different expression expressed by the intensity of staining of the bands. The intense staining of the area shows its relatively high share of the accessions having 'S' phaseolin.
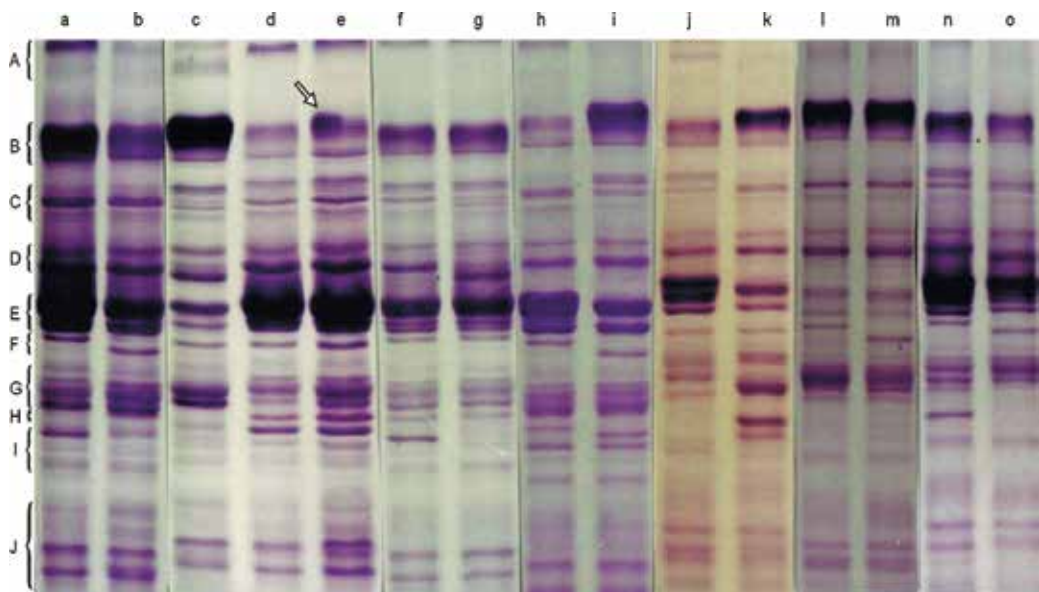
Fig. 4. Proteinograms of *P. vulgaris* accessions in the Bulgarian collection: a) predominant profile of the variety 'Oreol' and b) a rare profile of the variety 'Oreol'; c) 'Vulkan'; d) predominant profile of the variety 'Thrakiiski' and e) a rare profile of variety 'Thrakiiski'; f) predominant profile of line 'Mastilen 11b' and g) a rare profile of line 'Mastilen 11b'; h) accession E6091; i) line 213-17-1; j) line 536; k) line 538; l) line 563-6; m) line 563-5; n) line 564-3; o) line 5642. * As the phoregrams does not represent a stereotype, differences of the order of millimeters are quite permissible.

The difference in the significantly higher intensity of the Bx area in the accessions with 'S' type is obvious and can be used to distinguish them from those with 'T', 'Ca' and 'C' type (By). This area can help identify the type of phaseolin.

Zone C; MMW ~ 75 - ~ 97 kD. In Rf, RV and lower band number 5 variants were distinguished. When comparing genotypes with the reference accessions, no relation with the phaseolin type was found. For example, Fig. 4c, k, l, m (variety 'Vulcan', and lines 538-S, 563-6-S, 563-5-S) shows the configuration of zone C, typical of the 'S' phaseolin type. The profile of C zone is the same in the variety 'Mastilen 11b' (the rare profile) that is with 'Ca' phaseolin (Fig. 4g), and in E6091 (Fig. 4h) with 'C' phaseolin. The characteristic profile of zone C is not associated to the phaseolin type.

Zone D; MMW from ~ 60 to ~ 70 kD. It was observed in two variants: Dx and Dy vary in intensity of the last two bands in it. In types 'S' and 'C' (the fastest line is intense) and Dx variant (Fig. 4c, g, o); in 'T' and 'Ca' variant Dy is presented, where the most intense is the penultimate band (Fig. 4a, b, d, e, f, h, i, j, k, l, m, n). Furthermore, the Dx and Dy configurations cannot be associated with a phaseolin type or centres of origin.

Zone E; Phaseolin zone; MMW - 45-51 kD. The variants (phaseolin types) identified in the studied material are four: 'S' (Fig. 4c, i, k, l, m), 'T' (Fig. 4j), 'Ca' (Fig. 4a, b, g, h, n, o), 'C' (Fig. 4d, e). (The designation used was introduced in studies of species in centres of origin.)

Area F; MMW - 41-42 kD. Immediately after the phaseolin zone two bands were observed. Their different intensity divides the material into two groups. They migrate close one after another and the accessions are present in the following two variants: One of the fractions is much more intense than the other and in some accessions it is the one that is more high-molecular. In others the fraction with lower molecular weight has a stronger expression. Variation in this zone is independent of the phaseolin type as well as of other bands.

Zone G; Zone of lectins (phytohemagglutinins); MMW - ~ 33,5 - ~ 41 kD. Variation is expressed by differences in Rf and RV of the fractions. In most cases it is associated with characteristic electrophoretic phenotype of the zone, which depends on phaseolin. This correlation is strictly observed in the accessions from the local germplasm. But in the breeding lines the phaseolin type and lectins combine more independently. There are cases of such a profile of zone G, which are not observed in the used witnesses. They cannot be associated with phaseolin. As can be seen in Fig. 8k, l, m, o, these configurations are present in the accessions with 'S' and 'Ca' phaseolin. It is characteristic of some of the groups of lines numbers: 556, 564, 566 and all studied groups of lines 561, 563, bred for resistance to bacterial diseases. The gene families that encode phaseolin and phytohemagglutinin are not linked (Brown et al., 1981). But in the cultivated forms there are characteristic profiles of the fractions of the lectin group which are associated with a particular profile of phaseolin (Lioi, 1991).

Zone H; MMW ~ 23 - 24 kD. It usually consists of one line, which varies in Rf in different accessions. The cases where the band is the least mobile were denoted with variant H1 (Fig. 4a, b, c), with H3 where it is the fastest moving (Fig. 4d, e, f, g, i, j), H2 with medium mobility is shown in Fig. 4h, l, m, n, o. In one and the same phoregram two versions are rarely seen: H12 and H23, in which there are two bands simultaneously.

Zone I; MMW - ~ 25 - ~ 28 kD. The first band from this zone in some accessions has strong intensity (Fig. 4a, d, e, f, h, i, k, n), while in others the intensity is weak (Fig. 4b, c, g, j, l, m, o). The intensity of this band is not related to the phaseolin type or to other band of the phoregram. Beside the described variants that are found in the reference accessions,

phoregrams are observed that are different from them, such as breeding line 538 (Fig. 4k) in zones G and I.

Zone J; MMW - 21,5 - 14 kD. It is composed of a great number of bands. Three different configurations of the bands Jx, Jy and Jz were observed. They differ in the absence (in Jx) or presence of one band with a different Rf value in accessions with Jy and Jz variants. Most often, the Jx profile is associated with 'S' phaseolin (Fig. 4b). In the breeding materials, Jx was observed in accessions containing 'C' as well as 'T' and 'Ca' phaseolin (Fig. 4a, d, f, g, k, n). For the accessions of Andean origin the Jy profile is typical (Fig. 4b, h, i, j, k, l, m). Although less frequent, it is found also in 'S'. The third Jz profile was observed in a small number of breeding forms possessing different phaseolin types (Fig. 4e).

In the variation of zones C, F, H, J, intrapopulation polymorphism was observed (variety 'Oreol': a predominant profile (Fig. 4a) and a rare profile (Fig. 4b); variety 'Trakiiski': predominant profile (Fig. 4d) and rare profile (Fig. 4e); breeding line 'Mastilen 11b': predominant profile (Fig. 4f) and rare profile (Fig. 4g).

The examples shown in Fig. 4 cover all different configurations of zones established. In the phoregrams of the accessions a much large number of combinations of the zones described are found.

Grouping of accessions

The screening of the electrophoretically fractionated total SDS-PAGE profile of storage proteins enabled to established polymorphism in the accessions and to group them. In 409 *P. vulgaris* accessions analyzed 116 groups were defined. This means that with regard to the researched spectrum of storage proteins, a third of the proteinograms differed from each other. Data from the identification of the phaseolin type of the accessions was the first grouping criterion and from the analysis of the overall profile of their storage proteins was the second grouping criterion (Table. 1). In column 2 are accessions indicated in the graphs of the clustering, presented in the same section (Fig. 5). (The choice of accessions in participating in the graphs is random within all presenting the same profile. The names and numbers of the accessions, varieties and lines in the dendrogram are abbreviated, while in Table 1, column 2 they are written in full). These accessions represent groups comprising different numbers of accessions specified in column 3. In each of these groups the accessions have completely identical phoregrams, i.e. they do not differ among themselves by any of the fractions of their storage proteins. Between the groups with the same phaseolin type there is diversity between other fractions of the proteinograms. The variability in these fractions does not correlate with any other character (or more precisely, there is no evidence of research of such dependency in the available literature and from the authors' observations). This means that in the grouping in this study, an independent (neutral) character is used - a marker of diversity, which is important when developing a core collection in a genetic bank. It can be used in so-called hierarching of the material is therefore important for the pre-breeding work. The polymorphism described in the analysis of proteinograms according to zones can serve as a basis for future research, which will look for correlation between the electrophoretic phenotype of the bands and characters that are target for the breeding of common beans.

No	Accession presented in dendrogram	Accessions with identical proteinograms to the representative one
Groups of accessions with 'T' phaseolin type divided by identical profiles:		
1	Ruets	E1426, 90E126, 90E298, 90E596-1, 90E596-7, 90E596-8, 90E596-9, 90E596-10, 3E96008, 3E96009, 3E96018, 90E297-4, 90E297-5, CDRK (witness)
2	3E96017	
3	190A-3	4E96068
4	3E96039	90E060, 91E048, 91E050, 91E053
5	E6779	E6823, 90E192, 190a-5, 83E1-200-1, 83E1-200-2, 83E1-220-1, 83E1-220-2, 83E1-220-3, 83E1-220-4, 83E1-220-5, 83E1-220-10
6	90E169	90E283
7	91E137	91E153, 3E96010, 3E96011
8	Hitovo 1	4E96078, 4E96086
9	Ilinden	Dunav-1, 90E139, 3E96027
10	89E085	E6192, 90E128, 90E129, 91E297, 3E96013, 3E96021, 3E96022, 3E96025, 3E96029, 4E9692
11	90E688	
12	566-1-1-1-2-1-3-2(7)	
13	566-1-1-1-1-1-1-1(1)	566-1-1-2-1-2-3-3(2), 566-1-1-1-4-2-1-1(4), 566-1-1-1-2-1-3-2(3)
14	566-1-1-1-4-2-2-1(5)	
15	566- 1-1-1-4-2-2-1(6)	
16	BP-1	
17	90E191	
18	83E229	90E303, 91E055, 91E209, 91E212, 91E282, 84-34-11
19	536-1-1-1-3	83E120, 90E222, 90E248, 90E250, 91E060, 4E96072
20	559-2-1-4-1-3-2-1(6)	
Groups of accessions with 'Ca' phaseolin type divided by identical profiles:		
21	Fiesta	
22	551-1-1-1-5-1-2(126)	90E116, 85-20-14, 551-1-1-1-5-1-3(127)
23	522-1-2-2-1	564-2-5-2-1-4-2-1(2)
24	Zornitsa	
25	Oreol	Perun
26	521-1-1-1-2	
27	529-1-1-1-1	
28	551-1-1-1-4-1-1(123)	551-1-1-1-4-1-1(124)
29	612	E7585, 90E125, 90E312, 92E050, 92E056, 91E299, 3E96014, 3E96015, Trudovets
30	534-1-1-1-3	
31	550-1-4-2-3-3-1-1(141)	
32	Padesh 1	E6087, 89E502, 4E96070, 559-2-1-1-2-3-2-1(2), 559-2-1-1-2-3-2-1(4), 559-2-1-4-1-3-2-1(5)
33	90E673	
34	E6122	90E679-1, 90E679-2, 90E679-3, 90E679-4, 90E679-5(X-5)
35	554-1-5-1-1-2-3-3(131)	
36	90E602	90E713, 91E171, 91E220, 559-2-1-4-1-3-2-1, 559-2-1-1-1-1-2-3-2, 559-2-1-1-2-3-2-1, 559-2-1-4-1-3-2-1-1, 559-2-1-4-1-3-2-1-2, 559-2-1-4-1-3-2-1-3, Cher Starozagorski

37	Zarya	
38	559-2-1-1-1-1-2-3-2(7)	
39	559-2-1-1-1-1-2-3-1(3)	
40	525-1-3-1-1	
41	558-2-2-1-3-2-5-1	559-1-1-1-1-1-1-3(1)
42	556-1-3-3-5-1-3-2(11)	Pokrovnik 7, 556-1-1-1-2-1-1-1(12)
43	556-1-1-1-2-1-1-1-2(5)	556-1-1-1-2-1-1-1-2(24), 556-1-1-1-5-2-2(25)
44	564-2-5-2-5-3-3-2(3)	564-2-5-2-5-3-3-1(4)
45	564-2-5-2-1-4-2-1(9)	
46	3E96016	564-2-5-1-3-3-1-1(1), 564-2-5-1-3-3-1-3(7)
47	Mastilen 11б	90E092-1-1-4, 90E092-4-4-1, 90E092-4-4-2
48	Veritsa	
49	564-2-5-2-5-3-3-1(5)	564-2-5-2-5-3-3-3(10), 564-2-5-1-3-3-1-2(11)
50	564-2-3-2-2-5-2-1(6)	564-2-3-2-2-5-2-2(8)
51	216-73-1	216-77-1
52	554-1-5-1-4-2-3-2(129)	85-20-60
53	554-1-5-1-4-2-1-1(128)	
54	554-2-2-1-3-1-3-1(132)	
55	Hotovo-2	
56	614	608
57	90E255-2	90E255-4, 90E255-8, 90E255-10
58	91E289	4E96074
59	93E007	90E609, 91E139, 91E205, 4E96091, 91E307
60	93E009	
61	89E185	
Groups of accessions with 'C' phaseolin type divided by identical profiles:		
62	Radomir 2	E5780, 91E054, 91E278, 91E280, 91E285, 91E316, 93E003, 93E011, 3E96032, 3E96034, 3E96035, 4E96082, 4E96083, 4E96084, 4E96085, 4E96089, Samuranovo
63	213-42	213-43
64	95E002	
65	Nikos	
66	4E96081	
67	Trakiiski	
68	E6195	
69	E6091	
70	Contender (witness)	
71	4E96079	
72	555-1-2-6-2-2-4-2	
73	620	
Groups of accessions with 'S' phaseolin type divided by identical profiles:		
74	194-1	194-2
75	556-1-3-3-4-3-2-1(19)	556-1-3-3-4-5-2-3(20), 556-1-3-1-4-5-2-2(21)
76	563-1-1-3-5-3-3-2(6)	
77	563-1-2-2-1-2-1-3(3)	
78	556-1-3-1-3-2-1-3(16)	

79	556-1-3-1-5-4-1-1(28)	
80	561-1-3-3-4-4-3-3(1)	561-1-3-3-1-2-3-2(2)
81	563-1-1-1-2-1-1-3(2)	563-1-2-2-1-2-1-3(4), 563-1-2-2-1-3-3-3(5), 563-1-2-1-4-4-2-1(7), 563-1-2-1-4-4-3-1(8), 563-1-2-3-2-1-2-1(9)
82	190A-4	
83	556-1-1-1-5-5-1-1 (26)	556-1-3-2-3-2-2-1(27)
84	556-1-3-1-5-4-1-3(14)	556-1-3-1-3-2-1-3(15)
85	556-1-3-3-4-5-2-1(13)	
86	556-1-3-1-5-4-1-1(1)	556-1-3-1-3-2-1-3(2), 556-1-3-3-4-5-2-1(3), 556-1-3-3-1-2-2-1(4), 556-1-3-3-4-5-2-1(6), 556-1-3-1-5-4-1-3(7)
87	556-1-3-4-2-2-1-1(22)	
88	550-1-4-3-1-3-2-1(143)	
89	550-1-4-2-3-3-1-3(142)	
90	550-1-4-1-4-3-1-1(32)	
91	563-1-2-1-4-1-3-1(1)	
92	556-1-3-2-3-2-2-1(17)	
93	Tarnovo 13	
94	556-1-1-1-2-1-1-1(18)	
95	556-1-1-1-5-5-1-2(8)	556-1-3-1-4-5-2-3(9), 556-1-3-1-5-4-1-1(10)
96	556-1-3-3-1-2-2-1(23)	
97	92E053	92E055
98	90E198	190A-1, 190A-2, 190A-6,
99	Prisad	90E243, 90E247, 90E597, 90E604, 90E617, 90E635, 90E687, 90E274, 92E054, 3E96019, Dobrudzhanski 5, Rusenski ran, Dobrudzhanski 7, Plovdiv 164, Planinets
100	Ludogorie	
101	Prelom	90E231, 90E240, 90E270, 90E277, 90E282, 90E288, 90E296-1, 90E296-2, 90E296-3, 90E296-4, 90E305, Garmen, Ruse-17
102	Balgari	90E187, 90E644-1, 90E644-2, 90E644-3, 91E314, 3E96038, 82-11-16
103	Abritus	
104	538-3-3-1-4	
105	Vulkan	E6083, 90E100, 90E138, 90E245, 90E715, 3E96020, Bistrenski, ICA Pijao (witness)
106	Dunavets	E6085, E7380, E7381, 90E228, 90E286, 90E301, 90E309, 90E319, 90E390, 91E281, 91E288, 95E003, 93E008, 4E96075, 4E96076, Samokov-3, E6787, Kristal-137, Ruse-2, 82-7-11-12
107	4E96071	
108	91E292	89E120, 89E405, 89E504, 90E251, 90E257, 93E005, 4E96067, 4E96080, 4E96087
109	Desislava	
110	91E304	
111	89E129	
112	91E044	90E107, 90E659, 3E96040, 4E96088
113	213-17-1	213-27-1
114	95E001	83E118, 89E116, 89E560, 91E072, 91E129, 3E96036, 4E96069, 4E96073

*№ - serial number of the group

Table 1. Groups of proteinograms of total storage proteins in Bulgarian *P. vulgaris* accessions

For proteinograms of the various accessions a hierarchical structure was observed that is illustrated with a dendrogram (Fig. 5). The representatives of the 114 groups listed (Table 5, column 2) are given in the overall dendrogram, composed of all accessions tested. The grouping carried out on the basis of this analysis provides mathematical explanation of the degree of similarity and difference. In the dendrogram the accessions are at different distances from one another depending on the number of differing fractions.
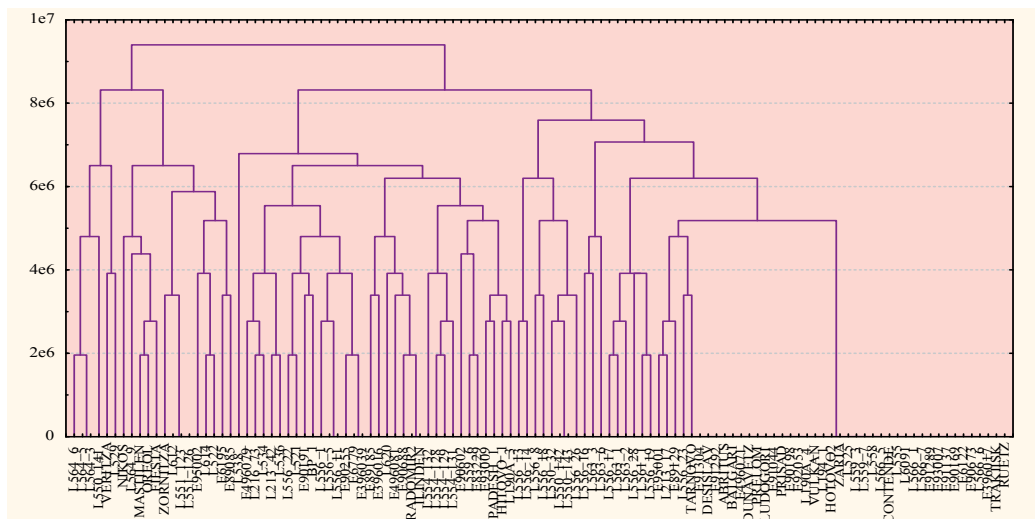

Fig. 5. Graph of cluster analysis on data from phoregrams of *P. vulgaris* L. accessions from the Bulgarian collection

According to the genetic potential they have, proteinograms can be grouped into two main clusters (Fig. 6).
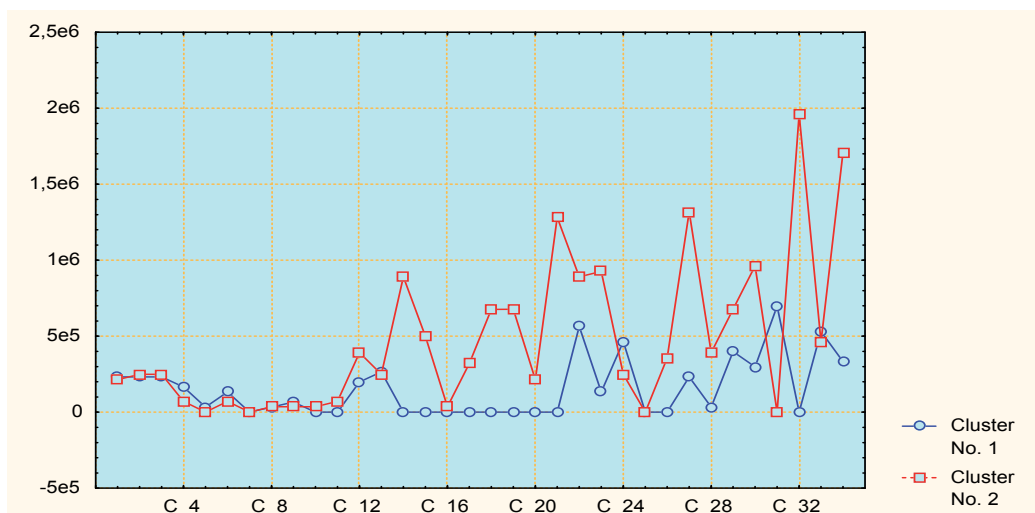

Fig. 6. Clusters in *P. vulgaris* L. accessions in the Bulgarian collection based on data from analysis of phoregrams of storage proteins

The graphs in Fig. 5 and 6 give an idea of the distance between clusters. In each basic cluster, there is distribution in subgroups. In each group, they have similar properties, but in the subgroups the accessions are united in strict differentiation. The presence of clusters means that there is a reason for grouping the accessions (objects), which is assumed to be genetically determined. Otherwise, the distribution of the objects would be totally chaotic, which means that the hierarchical structure, observed in Fig. 5 and Fig. 6, would be entirely absent. If grouping is done on the basis of data only on the phaseolin type and the zones correlating with it, it should be expected that the only cluster is composed of accessions containing 'S' type, and the other one of 'T', 'Ca' and 'C ' type, as their distribution according origin centre is. This dependence is followed, though not strictly, in the native accessions in which typical configurations of the zones corresponding to phaseolin type are observed. In most of the breeding material the dependence is not obeyed. In the cluster composed mostly of accessions with 'S' type, which attributes them to the Mesoamerican origin centre, there are included some breeding forms containing phaseolin typical of the Andean origin centre. The opposite case is also observed. The lack of genetic link between zones A, C, E, G, J and the phaseolin type explains the peculiarities in the grouping in two clusters.

The presence of clustering can be interpreted as a direct indication of the general behavior of the studied storage proteins, i.e. each of the analyzed fraction is in some relation with the other, which is the reason for the more detailed study of the experimental material. That is why within the four groups of accessions, differentiated by phaseolin type, grouping was carried out according to the fractions of the total pattern of storage proteins.

Grouping of material simultaneously according to both attributes facilitates the selection of accessions to participate in hybridizations, as well as the selection of accessions to be maintained in genebanks, avoiding duplication and reducing the amount of breeding work, costs and time. Inside the cluster, the accessions have a similar genetic potential and origin. Using accessions from one cluster in a possible breeding, differences will occur in terms of the investigated character, but they will not be statistically significant. Therefore, out of each group the most typical representative is selected, e.g. the most widely used and well-established variety.

Farmer's preferences and consumer requirements were the two reasons explaining the occurrence of a specific seed character. During the 20th century, breeding efforts concentrated on market varieties which were important to the region. Hybridizations were made between closely related genotypes resulting in homogeneous breeding lines. Gene pools and groups within them can be recognized in cultivated common bean on the basis of the protein marker. Our investigation confirmed that phaseolin, a useful indicator of genetic origin, has been a breeding-neutral character during the process of introduction and adaptation of bean accessions in Bulgaria.

4. Conclusions

Four phaseolin types are distinguished in the investigated material from the Bulgarian common bean collection: 'S', 'T', 'Ca' and 'C'. 'T' phaseolin type is the most predominant among the group of the local Bulgarian accessions, which represents 66% from the Andean gene pool. 'Ca' type is the most frequent in the breeding lines and varieties, covering 79%

from the Andean gene pool. The introductions have 'S' phaseolin from the Mesoamerican group; and the Andean one is presented mostly from phaseolin 'T' within 'T', 'C' and 'Ca' types.

The analyses carried out allow the formation of groups and clusters, because of variability in the germplasm. Divided according to phaseolin types and according to the other fractions from the phoregrams of the storage proteins, the accessions are grouped in different number of clusters: in type 'T' the clusters are 3; in type 'Ca' the clusters are 4; in type 'C' - 4, in type 'S' - 3. The accessions in different clusters have the same potential of the studied character, therefore the breeding work should include one representative of each cluster. It is advisable that the genebanks maintain one representative from each group that participates in the clustering.

Part B. A study of genetic diversity of *P. vulgaris* L. by isoenzyme markers

The aim of this study was to establish a methodology for characterizing the genetic diversity of large common bean collection by analysis of enzyme systems. Another aim was to analyse the genetic aspects of allozyme variation in genotypes cultivated in the Plovdiv agricultural region of Bulgaria, which has appropriate conditions for growing green bean (French bean) including the linkage relationships for certain loci to other species.

5. Material and methods

5.1 Plant material and sampling

The studied common bean *P. vulgaris* genotypes - 7 varieties ('Fiesta', 'Oreol', 'Perun', 'Veritsa', 'Starozagorski cher', 'Trakiiski', 'Nikos'), 3 accessions (Acc.6091, Acc.6787, Acc.82201206) and 4 breeding lines (BL) (BL620, BL614, BL612, BL608), were grown for green bean in the experimental field of Maritsa Vegetable Crops Research Institute, Plovdiv. Of the above genotypes, an average of 50 cotyledons of 4 to 6-day-old germinated mature seeds were analyzed in order to identify polymorphic loci at species level and between individual plants of the selected genotypes. The breeding lines were obtained by crosses: 608 ='Oreol' x 82201206; 612 ='Fiesta' x 6091; 614 ='Fiesta' x 6787; 620 ='Perun' x 6091.

5.2 Enzyme assays

For electrophoretic variation, 31 readable and reproducible enzyme loci resolved from 15 enzyme systems assayed by electrophoresis, were analyzed: aconitate hydratase (ACOH, E.C. 4.2.1.3. lyase, alcohol dehydrogenase (ADH, E.C. 1.1.1.1 oxidoreductase), cytosol aminopeptidase (CAP, E.C. 3.4.11.1), dihydrolipoamide dehydrogenase (DDH, E.C. 1.8.1.4), a non specific enzyme endopeptidase (END, E.C. 3.4.-.-, hydrolase), colorimetric and fluorimetric esterases (cEST and fEST, E.C. 3.1.1.-, hydrolases), non specific, glucose-6-phosphate isomerase (GPI, E.C. 5.3.1.9, isomerase, glucose-6-phosphate dehydrogenase (G6PDH, E.C. 1.1.1.49, oxidoreductase), glutamate dehydrogenase (GDH, E.C. 1.4.1.2, oxidoreductase), isocitrate dehydrogenase (IDH, E.C. 1.1.1.42, oxydoreductase), malate dehydrogenase (MDH, E.C. 1.1.1.37, oxidoreductase), malate dehydrogenase (NADP+) (MDHP, E.C. 1.4.1.40, oxidoreductase), phosphoglucomutase (PGM, E.C. 5.4.2.2, isomerase), shikimate dehydrogenase (SKDH, E.C. 1.1.1.25, oxidoreductase), and superoxide dismutase (SOD, E.C. 1.15.1.1, oxidoreductase. The choice of enzyme systems for analysis was dictated

by published data (Zoro et al., 1999, 2003) about high variability in different plant species that can serve as genetic markers for polymorphism, as well as for identification of genotypes. Loci were labeled sequentially, with those migrating closest to the anodal end designated as number 1. In the procedure of manifestation of the enzyme activity a 'witness' was placed in the initial and final positions of each gel. The 'witness' was one of the analyzed accessions used as a referent one, so that differences were standardized and laboratory errors were eliminated. The allozyme from this genotype was designated 100, and all other allozymes were assessed according to their relative migration distance.

The extraction of the material was conducted according to the procedure introduced by Maquet (1995) to study the *P. lunatus* species. The extraction buffer, consisting of potassium phosphate (Sigma-Aldrich # P-2222) with pH 7.0, containing 3% of sucrose (Sigma-Aldrich # S-8501), 5% PVP-40, 0.1 M KH_2PO_4, 0.05% Triton X-100 (Sigma-Aldrich # T-8532) added *ex tempore* and 14 mM 2-mercaptoethanol (Sigma-Aldrich # M-6250) also added *ex tempore*, allows the retaining of whole catalytic activity of the enzyme and inhibition of the interfering substances such as phenol, toxins and phenoloxidases. The ratio between the cotyledonic plant tissue and the extraction buffer was 1:2. Stored at -20 °C, the extract remained active for one year.

The solid carrier for the horizontal electrophoresis was 10% starch (Sigma-Aldrich # S-4501) gel, prepared by a procedure described by Wendel & Weeden (1990) and Maquet et al. (1993), and optimized by Zoro (1994).

The electrophoretic separation, procedure of loading the gel, migration of the fragments was performed in three different electrode systems, presented in the Results part, Table 1 according to the procedures previously described:

i.   0.065 M histidine-0.019 M citric acid (H-C), pH 6.1 (Kazan et al., 1993) prepared in a ratio of 1:10 mixture of solution buffer system H-C:H_2O.
ii. 0.030 M lithium-0.19 M borate (Li-Bo/T-C), pH 8.1/0.05 M Tris-0.008 M citrate, pH 8.4 (Murphy et al., 1990) prepared in a ratio of 1:9 mixture of solution buffer system Li-Bo/T-C:H_2O.
iii. 0.040 M citric acid-0.068 M morpholine (M-C), pH 6.1 (Wendel & Weeden, 1989) prepared in a ratio of 1:20 mixture of solution buffer system M-C:H_2O.

The chemicals were purchased by Sigma-Aldrich: L-histidine (# H-8000) and citric acid (# C-7129), boric acid (# B-0252) and lithium hydroxide (# L-4256), Tris (# T-1503) and citric acid (# T-8657).

Preferably, the organ is always separated at the same physiological stage for maximum avoidance of any possible modifications of enzymes. The organ used is young, since the enzymatic activity is at its peak. The preparation of gels, electrophoresis buffers, the loading of samples and the conditions for separation were developed by Bushuk et al. (1978), Bassiri et al. (1978), Aebersold et al. (1987), Pasteur et al. (1987), Stuber et al. (1988), Charlionet (1990), Hames (1990), Murphy et al. (1990), Wendel & Weeden (1990), Acquat (1993). To make an enzyme visible, the gel is immersed in a liquid reaction medium with a specific substrate for the enzyme system studied and various chemical compounds whose activity is needed (a coenzyme, ion, etc.). The staining recipes are provided in the work of Shaw et al. (1970), Harris et al. (1977), Trouslot et al. (1980), Gottlieb (1981, 1982), Soltis et al. (1983), Hussain et al. (1988), Guo et al. (1989), Weeden (1990), Gabriel et al. (1992), Hussain (1988), Waugh et al. (1992).

Different technical parameters for carrying out electrophoresis were tested: currency voltage and duration of fraction movement in the electric field. Of all voltage and amperage values tested for fraction movement, a constant voltage of 4-10 V.cm^{-1} and duration of electrophoresis 5-17 h were accepted (Table 2).

6. Results and discussion

6.1 Introduction of the methods for enzyme assays

As a result of the systems tested, in the migration of the protein extract the best separation by the optimal buffers and conditions for electrophoresis of the enzyme systems studied in cotyledon tissues of germinated seeds in genotypes from the Bulgarian collection of *P. vulgaris* are shown in Tables 1 and 2.

Enzyme	Electrophoresis buffer system		
	Histidine-Acid Citric	Litium-Borate/Tris-Citrate	Morpholin-Acid Citric
ACOH		**	-
ADH	*	***	-
CAP	***		
DDH	***	***	*
END		***	
F., C. EST		**	
GDH		***	
G6PDH		-	**
GPI	***		
IDH	***		
MDH	***		**
MDHP		*	***
PGM	***	***	***
SKDH		-	***
SOD		**	

Table 1. Electrophoresis buffer systems of studied enzyme systems
Notes: *** - perfect resolution (optimal conditions), ** - very good resolution, * - good resolution, 0 – no resolution (conditions for resolution are not created).

The buffer systems employed in the assays were: continuous histidine-citrate (pH 6.1) for ADH, CAP, DDH, GPI, IDH, MDH, PGM, and PGDH, and discontinuous lithium-borate (pH 8.1)/Tris-citrate (pH 8.4) for ACOH, ADH, DDH, END, cEST, fEST, GDH, PGM, SKDH, and SOD. Morpholin-citrate (pH 6.1) and lithium-borate (pH 8.1) buffer systems were also tested and presented the best resolution for G6PDH, MDH, MDHP, PGM, SKDH. The techniques for gel electrophoresis and histochemical staining procedures are those reported elsewhere (Zoro Bi et al., 1999).

The combination of factors coupled for the selected buffers that allowed good resolution of the isoenzymes in cotyledon tissue of common bean are summarized in Table 2.

Parameters	Buffer system		
	Histidine-Acid Citric	Litium-Borate/Tris-Citrate	Morpholin-Acid Citric
Voltage (constant)	10 V.cm^{-1}	4 V.cm^{-1}	10 V.cm^{-1}
Electrophoresis duration	5 h	17 h	5 h

Table 2. Conditions for electrophoresis buffer system for separation of isozymes in cotyledons of *P. vulgaris*
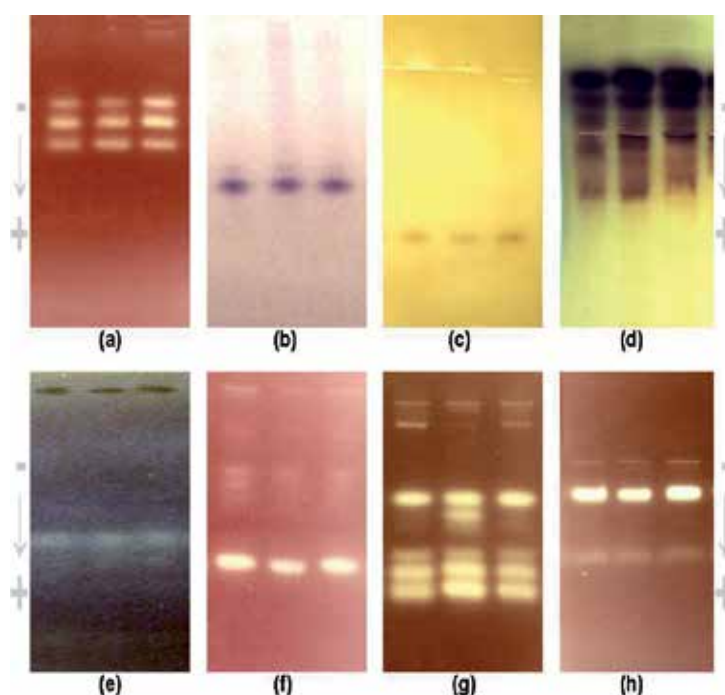
6.2 Monomorphic enzymes

Fig. 1. Zymogrammes of the monomorphic enzyme systems in French bean genotypes grown in the region of Plovdiv:

a. Alcohol dehydrogenase (ADH). Profiles 1, 2, 3: ADH-1$^{100/100}$; ADH-2$^{100/100}$.
b. Cytosol aminopeptidase (CAP). Profiles 1, 2, 3: CAP$^{100/100}$.
c. Endopeptidase (END). Profiles 1, 2, 3: END$^{100/100}$.
d. Esterases - colorimetric (cEST). Profiles 1, 2, 3: cEST$^{100/100}$.
e. Esterases - fluorimetric (fEST). Profiles 1, 2, 3: fEST$^{100/100}$.
f. Isocitrate dehydrogenase (IDH). Profiles 1, 2, 3: IDH$^{100/100}$.
g. Malate dehydrogenase (MDH). Profiles 1 & 3: MDH-1$^{100/100}$; MDH-2$^{100/100}$; MDH-3$^{100/100}$. Profile 2: MDH-1$^{100/100}$; MDH-2$^{100/84}$; MDH-3$^{100/100}$.
h. Malate dehydrogenase (NADP+) (MDHP). Profiles 1, 2, 3: MDHP$^{100/100}$.

6.2.1 Alcohol dehydrogenase (ADH)

The ADH zymograms in Fig. 1a reveal three zones of activity. Each of the anode (ADH-$1^{100/100}$) and cathode (ADH-$2^{100/100}$) bands shows one allelic variant. The band in an intermediate position comigrates between ADH-1 and ADH-2. In the analyses this band was accepted as a product of inter-gene interaction. This is confirmed by the more intense colouration manifested by this band on the gels. A change in the electrophoresis conditions (sequential electrophoresis) can be expected to cause the intermediate band to be divided into two parts. After the analysis of the enzymograms, there were established two active loci encoding ADH, manifested in *P. vulgaris*. In the studied genotypes, a single allozyme phenotype was observed for both loci. Alcohol dehydrogenase in the studied accessions grown in the region of Plovdiv is a monomorphic isoenzyme. In neither case was polymorphism detected. Profiles in *P. vulgaris* were similar to those observed in *P. coccineus* (Wall & Wall, 1975), *P. lunatus* (Zoro, 1999), *Pennisetum glaucum* (Banuett-Bourrillon & Hague, 1979), *Camellia* (Wendel & Parks, 1984), *Ciser arietinum* (Tuwafe et al., 1988). According to literature data on the genus *Phaseolus*, the natural hybrids observed allow to assume the existence of a dimeric quaternary structure of the isozyme (Zoro, 1999). The results for these species showed that two loci, *Adh*-1 and *Adh*-2, encoded ADH. Each locus encoded codominant alleles which were dimeric and viable. The studies of *P. vulgaris* revealed monomorphism in the zone controlled by *Adh*-1 (polymorphism in *P. coccineus* and *P. lunatus*) and polymorphism in the zone of the *Adh*-2 products. Polymorphism of alcohol dehydrogenase was observed in wild and cultivated populations.

In many plant species where different allelic variants were identified, the distribution of genotypes was traced in their disintegration with generations. It has been established that in some species there is a genetic linkage between the *Adh* loci (Banuett-Bourrillon & Hague, 1979; Brown, 1980), while in others they were not related (Freeling & Schwartz, 1973; Torres, 1974; Tanksley & Jones, 1981).

The identified phaseolin types of the accessions included in this study proved that they belong to different centres of origin of the species *P. vulgaris*. Andean and Mesoamerican representatives have different ADH alleles, which can be used to determine their affiliation to the respective centre of origin (Koenig & Gepts, 1989; Maquet et al., 1993). The absence of heterozygotes observed in the studied material could be due to various reasons: strong autogamy, crossbreedings aimed at homogamy, breeding favouring homozygous individuals. Since similar studies of Bulgarian material have not been performed until now, there is no possibility that breeding by enzyme systems was previously made. However, ADH may correlate with an economic character used in the selection of the studied forms, which has kept an allelic variant of the enzyme analyzed.

The genetics of alcohol dehydrogenase has been studied in a much greater number of higher plants than the genetics of other enzymes (Tanksley, 1983; Weeden, 1983). In the species *Zea mays, Triticum aestivum, Hordeum vulgare, Pennisetum typhoides, Solanum lycopersicum, Helianthus annuus, Stephanomeria exigua, Lupinus angustifolius* two *Adh* genes have been described. Under aerobic conditions, ADH isoenzymes can be found in seeds (embryo and/or endosperm) and pollen. In mature seeds, most frequently either of the two alcohol dehydrogenases prevails. The gene responsible for most of its activity in mature seeds is usually the same one that encodes ADH in pollen. Different *Adh* genes can be induced or not by anaerobiosis. In monocotyledonae species, all *Adh* genes are induced by

environmental stimuli. It has been established that in dicotyledonae species only one of the two genes can be activated by anaerobic stress. In maize, both genes are independent, but in other species they are closely related due to their origin by tandem duplications.

6.2.2 Cytosol aminopeptidase (CAP)

With electrophoretic separation the activity of cytosol aminopeptidase was demonstrated (Fig. 1b). Following interpretation of the zymograms obtained from the studied material, monomorphism was established for the enzyme CAP. There is a single band with the same mobility (Rf) of the gels in all investigated seeds. The zymograms are characterized by a single locus called *Cap*. Similar results for *P. lunatus* were found by Zoro (1994), who also established one monomorphic locus in the study of wild populations of this species. In later investigations of the polymorphism of *P. lunatus*, Maquet (1995, 1997) found four different alleles in the aminopeptidase cytosol locus. Some of the alleles are endemic for the accessions of the Andean gene pool and may serve to distinguish them from accessions belonging to the Mesoamerican gene pool. The natural hybrids observed in the genus *Phaseolus* allow to assume the existence of a dimeric quaternary structure of CAP isozyme (Maquet, 1995). The dimeric nature of this enzyme system is shown in *Pisum sativum* (Scandalios & Espiritu, 1969) and *Abies balsama* (Neale & Adams, 1981). In studying the species *P. acutifolius*, Schinkel & Gepts (1989) identified three isozymes of leucine aminopeptidase. Koenig & Gepts (1989) established two isozymes of the same enzyme system in *P. vulgaris*.

6.2.3 Endopeptidase (END)

The gels treated for manifestation of endopeptidase activity display a single band (Fig. 1c). The identified locus has migrated towards the anode. The identical electrophoretic phenotype shows that the species are homozygotic at the locus analyzed, which controls the expression of endopeptidase End^{100}/End^{100}. In the examined *P. vulgaris* crop one allele END^{100} with monomorphic pattern is detected. The researched natural hybrids of other types of the genus *Phaseolus* enable the evidence of a monomeric quarterly structure of the isozyme (Maquet, 1995). Maquet (1995) has detected four codominant alleles of the locus of END, which he relates to the different origin sources of the *Lima* bean accessions. Zoro (1999) examined a locus that possesses three codominant alleles that control the endopeptidases which represent monomeric proteins in wild types of *P. lunatus*. Similar results have also been achieved with soya (Griffin & Palmer, 1987).

The selected accessions are distinguished by morphologic and economic criteria, which vary according to the environmental conditions. The search for alozymes, aiming to use them as molecular markers, allows tracing the variability among species within a crop frame and among different crops, regardless of the environment factors. The fact that alozyme variants of endopeptidase are detected in other species close to *P. vulgaris* gives grounds for carrying out this research. This judgment allows better understanding the factors responsible for the loss of polymorphism. The breeding in these cultivated genotypes by certain agricultural characters has been done for many generations and this might have indirectly influenced the limitation of the alozyme variants of endopeptidase. An analysis of such a relation with the present work and available literature sources cannot be made, because the END locus does not feature in published gene maps of the species. On the other hand, the number of the

polymorphic loci as well as the number of the alleles for each locus is restricted in case of a narrowed gene flow. In some autogamous plants the genetic variability can be too low to characterize a certain pattern. The absence of polymorphism in self-pollinating species makes these markers inadequate to be able to define the genetic structure of the species. The detected lack of polymorphism in endopeptidase, as well as in the above discussed alcohol dehydrogenase and cytosol aminopeptidase enzyme systems, confirms their narrowed genetic base. These results in the chosen accessions do not exclude the chance of finding variability in the analysis of other isoenzyme systems. Literature references on endopeptidase are scarce. From the complex of the *Phaseolus* genus, which has five crop species, the endopeptidase enzyme system was analysed in *P. lunatus*, in which it can serve to establish the inter- and intra- polymorphism of crops (Maquet, 1995; Maquet et al., 1996; Maquet & Baudoin, 1997; Maquet et al., 1997; Zoro, 1999).

6.2.4 Esterases (EST)

Nonspecific esterases were manifested - colorimetric (cEST) (Fig. 1d) and fluorimetric (fEST) (Fig. 1e). Their use as markers, as is this case, does not require prior knowledge what type of esterases the analysis refers to. They are divided into two types according to the method for their determination. In the analyses conducted polymorphism was not established. In cEST, a double band migrating to the cathodic side of the gel is observed. This picture shows the second band, which was accepted to be a degradation product, i.e. one locus was determined whose activity in the studied genotypes is manifested through one allele ('100'). In the samples obtained at the witness, allele '84' was obtained, which was interpreted as "effect of positioning". A second zone of activity was expected, migrating to the anode and representing the activity of another locus. However, it was not well visualized on all gels and because the absence of good reproducibility it was not interpreted. In such analyses in *P. vulgaris* (Wall & Wall, 1975) as well as other species of the family *Leguminosae*, such as *P. coccineus* (Wall & Wall, 1975), *P. lunatus* (Maquet, 1995; Zoro, 1999), *Ciser arientum* (Kazan et al., 1993) two loci in two allelic variants of the second locus were identified, which produced two allozyme phenotypes. In the locus near the anode, heterozygous individuals were identified that showed the monomeric nature of the quaternary structure in *P. lunatus* (Lima bean) (Zoro, 1999).

Fluorimetric (fEST). Two zones of activity were observed, but only one was reproducible. The only one band on Fig. 1e) was considered as a product of one locus (*fEst*). One allele *fEst*[100] was identified at this locus. In some cases, there appear slightly coloured bands located at the anode side. These bands were accepted as an artifact. The results thus interpreted showed the monomer nature of the quaternary structure of proteins in *P. vulgaris*. Fluorimetric esterases were not successfully established in all samples and were therefore excluded from the analyses for assessment of polymorphic loci. Through numerous studies conducted by Weeden & Marx (1984); Aebersold et al. (1987); Weeden & Wendel (1989); Murphy et al. (1990); Zoro (1999), where the isozyme phenotype of the heterozygous individuals is characterized by three bands, it is assumed that genes in *P. lunatus* encode dimeric proteins.

6.2.5 Glutamate dehydrogenase (GDH)

In the studied *P. vulgaris* genotypes a single *Gdh* monomorphic locus was observed, with allele GDH[100], which is subject to interpretation for the purposes of our study (results not

shown in figure). In addition to the interpretable band, in the enzymograms there was one more accompanying band (trace). In studies conducted on *P. lunatus* by Zoro (1994), two zones of activity (monomorphic for this species were observed, each of which was encoded by a single gene controlling the isozymes of GDH. By extending the time of germination, activity of a second gene locus can be expected to be expressed.

6.2.6 Isocitrate dehydrogenase (IDH)

In the common bean genotypes analyzed, three zones of activity of IDH were identified, two of which are poorly stained (Fig. 1f). The cathodic zone has the lowest enzyme activity. There were no variations in any of the zones. Based on the results obtained, the quaternary structure of the IDH isozyme cannot be expressed. In other studies, these isozymes have been identified as dimeric proteins (Gurie & Ledig, 1978; Kiang & Gorman, 1985; Weeden & Lamb, 1987; Simonsen & Heneen, 1995). Due to poor visualization of two of the zones in the statistical analyses, only the locus encoding the most highly mobile band of IDH-1 was used.

6.2.7 Malate dehydrogenase (MDH)

Zymograms of MDH are shown in Fig. 1g. Many bands were observed in this enzyme system, but only five of them were typical of homozygous individuals. Some of these five bands were poorly stained. Three loci control the expression of these bands. The three bands comigrated in the intermediate zone. One of these bands, the closest to the anodic one, was considered to be the product of *Mdh*-2, and the other two were products of other loci or artifacts. *Mdh*-1 was represented by two alleles, but in our case by one homozygous allozyme phenotype. Experiments using extraction buffer containing ascorbic acid have revealed that the isozymes encoded by *Mdh*-1 *Mdh*-2 were located in the cytoplasm, while those of *Mdh*-3 were mitochondrial. In fact, the samples extracted using a standard buffer represented all isozymes, while in extraction using a solution with a very low pH only the activity of *Mdh*-3 was manifested. The poorly stained uninterpretable bands might be products of the activity of isozymes localized in other cellular compartments such as peroxisome and chloroplast, in which MDH activity was observed (Gietl, 1992). This hypothesis was verified by the fact that isozymes localized in different cellular compartments were usually encoded by different genes and differed in their electrophoretic mobility, the kinetics of the reaction they catalyze and the nature of their inhibitor (Newton, 1983). Due to the complexity of the MDH zymograms, the dimeric nature of this enzyme system was not clearly established in some species of the genus *Phaseolus*, such as *P. lunatus* (Zoro, 1999). In other species their dimeric nature was clarified (Goodman et al., 1980; Weeden & Wendel, 1989; Murphy et al., 1990). The fact that no band of interaction between the products of the *Mdh*-1 *Mdh*-2 was observed pointed out that some heterodimers were inactive (Zoro, 1999). To obtain information about the genetic determinism of the MDH isozymes, it is necessary to carry out crosses between the alternative homozygotes.

6.2.8 Malate dehydrogenase (NADP+) (MDHP)

This enzyme also called malic enzyme (ME), malate NADP+ dependent enzyme represents a single interpretable band (Fig. 1h), a product of the activity of one gene locus. No polymorphism was observed. Weeden (1984) found correlation between Me and RbcS-

Rubisco (ribulose biphosphate carboxylase) loci, which manifested themselves to be interconnected (r>30 cM). Therefore, it was one of them that we picked for our investigation.

ME catalyzes the metabolism of malic acid to pyruvate. ME activity was found in all plant organs, and it increases during fruit development in many plants, where it is an important determinant of flavour. In tomato, four malate dehydrogenase genes (*Mdh*-1, 2, 3, 4) are known, several of which have been placed on the genetic map. In mapping populations generated between different species, the number and location of *Me* genes has not been reported.

Linkage analysis using Mapmarker indicated the *Me*-1 locus was located on chromosome 5, between the RFLP markers TG379 and TG23 (Chetelat, 1999).

The exceptionally wide cross (F$_1$ hybrid from *Solanum lycopersicum cv. VF36* x *S. lycopersicoides LA2951*) has proven a rich source of isozyme variation, allowing the determination of map locations of several previously unmapped genes, including *Mdh*-1, *Dia*-1, -2, -3, and –4 (Chetelat et al., 1997).

6.2.9 Superoxide dismutase (SOD)

The enzyme was visualized simultaneously with dihydrolipoamide dehydrogenase (DDH) in Fig. 2a. On the same gels, the activity of DDH was visualized as dark bands. SOD-1 was represented by two white bands '123' and '100'. It was assumed that each zone was determined by one gene: *Sod*-1 and *Sod*-2, which determined the SOD isozymes. The second locus was difficult to interpret. All these bands were monomorphic.

To assess the genetic diversity it is necessary to express the polymorphic loci and alleles for each locus. In self-pollinating species such as *P. vulgaris*, their number is small. This fact restricts the use of allozymes as molecular markers. In this case, the studied material, which is otherwise multifarious in terms of morphological and economic characters – yield, resistance to diseases and pests, is monomorphic in terms of the electrophoretic profile of the above studied enzymes. These enzymes cannot be used to assess the heterogeneity of the studied population of common bean species grown in the region of Plovdiv. Despite the monomorphism of the enzyme systems in the research material, it could be used in the breeding process. Of particular interest is the search for an economically important character, difficult to identify or determine, whose gene(s) and the enzyme encoding gene(s) are inherited together. The preference of a particular character in the breeding process can lead to "fixing" a population (variety, accession) toward the enzyme. When in such a correlation is established, the isoenzyme may serve as a marker.

The low variability detected throuth isozymes suggests the use of DNA-markers to characterize the genetic variability of the green common bean. Enzyme electrophoresis does not show variability to its full extent. Only about 30% of the substitutions at a genetic level result in substitution of amino acid with altered clear load of the protein (Pasteur et al., 1987). Many triplets encode one and the same amino acid. Normally, replacing one amino acid with another bearing the same charge would not impact the clear charge of the protein. Therefore, the identical electrophoretic mobility does not necessarily mean one and the same sequence of DNA. In all cases, however, the quantity of the unmarked variability would be proportionate to the quantity of the marked one (Crawford, 1990).
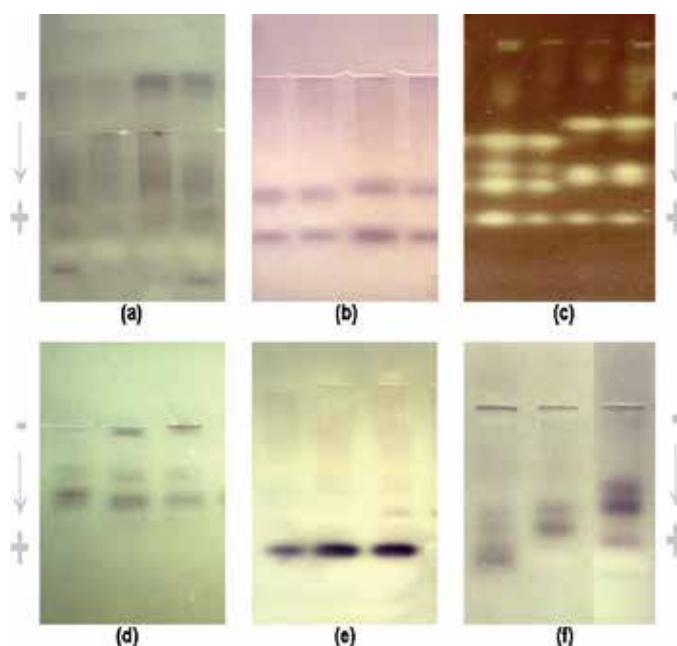
6.3 Polymorphic enzymes

Fig. 2. Zymogrammes of the polymorphic enzyme systems in French bean genotypes grown in the region of Plovdiv:

a.   Aconitate hydratase (ACOH). Profiles 1, 2: ACO-1$^{100/100}$; ACO-2$^{100/100}$; Profiles 3, 4: ACO-1$^{100/100}$; ACO-2$^{107/107}$.
b.   Superoxide dismutase (SOD) of common bean genotypes. Profiles 1, 2, 3: SOD-1$^{100/100}$; SOD-2$^{123/123}$. Dihydrolipoamide dehydrogenase (DDH). Profiles 1, 2, 3: DDH-1$^{100/100}$; DDH-2$^{100/100}$. Profile 4: DDH-1$^{90/90}$; DDH-2$^{100/100}$.
c.   Glucose-6-phosphate isomerase (GPI). Profiles 1, 2: GPI-1$^{100/100}$; GPI-2$^{100/100}$; GPI-3$^{100/100}$; Profiles 3, 4: GPI-1$^{100/100}$; GPI-2$^{96/100}$; GPI-3$^{96/96}$.
d.   Glucose-6-phosphate dehygrogenase (G6PDH). Profile 1: G6PDH-1$^{94/94}$; G6PDH-2$^{94/94}$; G6PDH-3$^{94/94}$; Profile 2: G6PDH-1$^{100/100}$; G6PDH-2$^{100/100}$; G6PDH-3$^{100/100}$; Profile 3: G6PDH-1$^{100/100}$; G6PDH-2$^{100/100}$; G6PDH-3$^{94/94}$.
e.   Phosphoglucomutase (PGM). Profile 1: PGM-1$^{100/100}$; PGM-2$^{100/100}$; Profile 2: PGM-1$^{100/100}$; PGM-2$^{100/100}$; Profile 3: PGM-1$^{100/100}$; PGM-2$^{67/67}$.
f.   Shikimate dehydrogenase (SKDH). Profile 1: SKDH-1$^{100/100}$; SKDH-2$^{133/133}$; Profile 2: SKDH-1$^{133/133}$; SKDH-2$^{196/196}$; Profile 3: SKDH-1$^{100/133}$, SKDH-2$^{133/196}$.

6.3.1 Aconitate hydratase (ACOH)

The zymograms of ACOH are characterized by two zones of activity, regarded as products of two loci (*Aco*-1 and *Aco*-2) (Fig. 2b). The product of locus *Aco*-1 is monomorphic, while in locus *Aco*-2 polymorphism was observed between the populations - two alleles producing the allozymes ACO-2^{100} and ACO-2^{107}. No heterozygous individuals were found in the

populations studied. In the study of the species *P. lunatus* (Zoro, 1999) three different allozymes have been identified, products of three alleles at locus *Aco*-2, and in the heterozygous individuals two bands. This observation indicates the monomeric nature of this enzyme system. Similar results were seen in species of the genus *Sorghum* (Morden et al., 1988) and in *Cicer arietinum* (Kazan et al., 1993). These authors not only observed the two loci, in addition they have proved that the activity of *Aco*-1 is limited in the cytoplasmic fraction, while the expression of *Aco*-2 is associated with the mitochondrial extracts. If the research continues at the level of cellular fractionation, that could clarify the results, because most often it is cytoplasmic enzymes that are polymorphic (Weeden & Gottlieb, 1980).

Studies of the genetic diversity of species of the genus *Phaseolus* often involve examination of diaphorase (DIAP) as an enzyme representing polymorphism. According to Nodari et al. (1992), who have published the related mapping, the alleles encoding DIAP isozymes (diaphorase) and ACO-2 are in the fifth group of linked genes. Therefore, our study focused only on the enzyme ACO.

According to literature data, in common bean the DIAP enzyme system has a quaternary structure (Sprecher, 1988). In 1990, Murphy et al. concluded that the structure of the enzyme can be monomeric or dimeric depending on the species observed.

6.3.2 Dihydrolipoamide dehydrogenase (DDH)

In the conditions of our study, two zones of enzyme activity are reproducible (DDH-1 and DDH-2), encoded by two loci (*Ddh*-1 and *Ddh*-2). Zone DDH-2 was observed as monomorphic. There is a concomitant band. Polymorphism was found in locus *Ddh*-1. Fig. 2a shows the two alleles of locus *Ddh*-1: DDH-1$^{100/100}$ and DDH-1$^{90/90}$.

From the heterozygous phenotype DDH-1$^{90/100}$, established in single individuals in two populations, the monomeric nature of the enzyme was accepted as likely, barring the assumption that there may be a band of interaction that is not manifested in the gels. At the level of this study, without any crossings or survey of the segregation of the character, it is not possible to make a definitive conclusion about the monomeric enzyme nature.

6.3.3 Glucose-6-phosphate isomerase (GPI)

For GPI four zones of activity were observed: a, b, c and d (Fig. 2c). The zymogram is characterized by two cytosolic loci, named *Gpi*-1 and *Gpi*-2, respectively, and one plastid locus *Gpi*-3, which migrates more slowly. These three groups identified in our study were detected in studies of various plant organisms carried by Weeden & Wendel (1989), in apple conducted by Weeden & Lamb (1987), in many *Leguminosae*, such as in *P. lunatus* (Maquet, 1995; Zoro, 1999). In our case, no interaction was detected between cytosolic loci, as in the studies of *P. lunatus*, where the two zones comigrated. The most anodic band was monomorphic in the populations studied here. In contrast, in *P. lunatus* (Maquet, 1995; Zoro, 1999) the same band was polymorphic and comigrates with the next electrophoretic mobility band. In the populations of this study, polymorphism was found in the second band, for which the following genotypes (allelic states) have been established: *Gpi*-2$^{100/100}$ and a small number of cases *Gpi*-2$^{99/100}$ (the latter are not shown in Fig. 2c). The null allele described by Zoro (1999) (absence of allele for locus *Gpi*-2) was not established in this study. There was a correlation between the variation of the second zone and the variation in the

fourth one, where two alleles, '100' and '96', were observed. The subunits of *Gpi*-3, the authors believed, they formed electrophoretically active heterodimers with the cytosolic loci. This contradicted the observations in Lima bean, where there was interaction established between *Gpi*-1 and *Gpi*-2 and the band in the intermediate position was taken to represent inter-gene heterodimer between these two genes. Observations on natural hybrids in the studied material of *P. lunatus* gave reason to Maquet (1995) and Zoro (1999) to assume that the quaternary structure of this enzyme was dimeric. Heterozygous individual was not found in this study, having allozyme phenotype characterized by three bands. One such observation would confirm the hypothesis about the dimeric structure of this enzyme system. However, in other taxa it has been demonstrated that the GPI isozymes have a dimeric structure (Guries & Ledig, 1978; Harry, 1986; Weeden & Lamb, 1987).

6.3.4 Glucose-6-phosphate dehydrogenase (G6PDH)

There are two polymorphic zones of activity on gels stained for G6PDH, one of which is well manifested and can be interpreted (Fig. 2d). The electrophoretic profile of this enzyme system is influenced by the power of germination, therefore it is not easily interpreted in the conditions of this study (sprouted seeds are tested). The electrophoretic variants of the level of each zone of activity segregate independently of those in the other zone. The result suggests that two loci control the expression of the G6PDH isozymes: *G6pdh*-1 and *G6pdh*-2. Both alleles producing the allozyme phenotypes G6PDH[94] G6PDH[100] and are identified in locus *G6pdh*-1. At the level of this locus heterozygous phenotypes were observed, characterized by two bands, '100' and '94' (the figure shows only one allozyme). These results indicate monomeric quaternary structure of the isozymes in representatives of *P. vulgaris*. According to Murphy et al. (1990), many quaternary structures have been suggested for the G6PDH isoenzymes.

6.3.5 Phosphoglucomutase (PGM)

Figure 2e shows the two zones of activity identified on the gels. They are regarded as products of two loci: *Pgm*-1 and *Pgm*-2. The zone closer to the anode is monomorphic in the studied populations. In the slower migrating zone two alleles, '100' and '67', are identified. Heterozygous individuals are not established. In *P. lunatus* the heterozygotes are characterized by two bands, thus confirming the expected phenotype of monomeric enzymes. The monomeric structure of the PGM isozymes has been repeatedly confirmed (Harry, 1986; Shore & Barrett, 1987; Weeden & Wendel, 1989).

6.3.6 Shikimate dehydrogenase (SKDH)

In the analysis of samples of all the gels, two comigrating zones of activity were observed (Fig. 2f). Both areas of activity can be controlled by a single locus, one of them being an artifact. However, they can also be products of expression of two genetically related genes, *Skdh*-1 and *Skdh*-2. Three allelic variants, producing allozymic phenotypes are identified: SKDH[100], SKDH[133] and SKDH[196]. The following genotypes were observed: in homozygous individuals - SKDH[100/100], SKDH[196/196] and SKDH[133/133], in heterozygous individuals – SKDH[100/133], SKDH[133/196] and SKDH[100/196]. Such profiles are identical to those of dimeric enzymes. Therefore, it is assumed that this enzyme system is controlled by one locus (*Skdh*) or two genetically linked loci (*Skdh*-1 and *Skdh*-2) having codominant alleles. Similar results were seen

in *P. lunatus* (Zoro, 1999), but in this species only two allozymic forms of the enzyme were observed. For the representatives of *P. lunatus* a monomeric quaternary structure of the enzyme was established. Because of the fact that the anodic band in the homozygotes SKDH[100/100] or SKDH[133/133] opposes the cathodic band of the homozygotes SKDH[196/196] or another respective one, in *P. lunatus* the heterozygous phenotype is characterized by three bands (Zoro, 1999). Weeden & Gottlieb (1980a) observed two forms of SKDH in *Clarkia williamsonii*: one cytoplasmic and one chloroplastic form. But the nature of genetic relationship, existing between the loci expressed by both forms of SKDH, has not been studied until now.

The results obtained for SKDH can be interpreted as follows: *P. vulgaris* has two genes - *Skdh*-1 and *Skdh*-2. These two genes produce polypeptide subunits that interact to form active tetrameric isozymes. The subunits combine randomly to form inter-gene homodimers and three inter-gene heterotetramers that in homozygous plants manifest five bands. The fact that the fifth line is of very weak intensity does not support such interpretation of the results and makes it highly unlikely. The crossover between alternative homozygotes of loci 1 and 2, and the analysis of the segregation of heterozygotes could contribute more information to the genetic determinism of the SKDH isozymes.

The enzyme systems studied were chosen to be products of unconnected genes, often representing different areas of the genome in the genus *Phaseolus*. Moreover, through the adopted research methods of the enzyme systems above described, the polymorphism of *Phaseolus vulgaris* is covered more completely as the loci encoding them are located on chromosomes other than those on which are the phaseolin-encoding loci.

The monomorphic and polymorphic isozymes established in this study, and respective alleles, are shown in Table 3.

Designation of the loci and alleles studied, direction of migration of the isoenzymes and polymorphism of the enzyme systems found in the study of our genotypes and established for representatives of the species according to data in literature are shown in Table 3. For comparison, literature data for *P. lunatus* was used about studies by Maquet (1995) and Zoro (1999). From the 24 analyzed enzyme loci and the screening of the 14 genotypes and accessions, eleven expressed polymorphisms and a total of 43 alleles were observed, most loci having a common allele. The proportion of polymorphic loci varied from 0% (i.e., Acc.6787) to 13.04% (i.e., BL614), with a mean of 6.15%. Accession 6787 is completely fixed in all loci analysed – zero genetic diversity. In the other genotypes studied there is an average of 1.43 polymorphic loci. The distribution of genotypes at loci showed a significant deficiency of heterozigots. "Pressure" exerted by the breeding is connected probably with polymorphic loci. Only 26% of 15 enzyme systems were polymorphic, which explains the origin of their very narrow size region, as well as the nature of the enzymes studied (all with the exception of CAP, participate in vital functions of plant organisms where the mutations are rarely observed). The sample of 45 to 51, average 49.50 individuals covers available alleles in the population, although in homozygous form.

For the cultivated *P. vulgaris* genotypes studied, data from isozyme electrophoresis indicated genetic variability mainly within the individuals of the accessions and the genotypes, i.e. at the interpopulation level. The values of allelic richness, heterozygosity and interpopulation gene flow were low contrary to the expected one. We recommend preserving the small number of the observed rare alleles at interpopulation level as a resource of maintenance of genetic

diversity that could be an appropriate management method. The observed substantial level of genetic uniformity threatens the studied genotypes. As a result of the obtained data, we also recommend a reintroduction of genotypes having rare alleles at regular time intervals to avoid the endangerment of a genetic drift. The genetic knowledge of the germplasm resources in the target species might directly result in better conservation and management control of the implications for the plant breeding.

Enzyme	Locus number	Locus designation	Migration direction	Alleles observed	Polymorphism / Number of alleles (%) observed in:	
					P. vulgaris	*P. lunatus*
ACOH	2	*Aco*-1 *Aco*-2	A	100 100, 107	Yes / 3	Yes / 4
ADH	2	*Adh*-1 *Adh*-2	A	100 100	No / 2	Yes / 4
CAP	1	*Cap*	A	100	No / 1	Yes / 4
DDH	2	*Ddh*-1 *Ddh* -2	A	100, 90 100	Yes / 3	-
END	1	*End*	A	100	No / 1	Yes / 4
EST	3	c*Est*-1 c*Est*-2 f*Est*-1	A C A	100 100 100	No / 3	Yes / 6
GDH	1	*Gdh*	A	100	No / 1	No / 2
G6PDH	2	*G6pdh*-1 *G6pdh*-2	A	100, 94 100, 94	Yes / 4	Yes / 5
GPI	3	*Gpi*-1 *Gpi*-2 *Gpi*-3	A	100 100 100, 96	Yes / 4	Yes / 5
IDH	3	*Idh*-1 *Idh*-2 *Idh*-3	A	100 100 100	No / 3	Yes / 5
MDH	3	*Mdh*-1 *Mdh* -2 *Mdh* -3	A	100 100, 84 100	Yes / 4	Yes / 6
MDHP	1	*Mdhp*	A	100	No / 1	-
PGM	2	*Pgm*-1 *Pgm*-2	A	100 100, 67	Yes / 3	Yes / 5
SKDH	3	*Skdh*-1 *Skdh*-2 *Skdh*-3	A	100, 196 100, 133 100, 133, 196	Yes / 7	Yes / 3
SOD	2	*Sod*-1 *Sod*-2	A	100, 123 100	No / 3	No / 2
N = 15	N = 31	N = 31		N = 43	P = 26%	P = 10.32 %

Table 3. Summarized data of enzyme locus of cotyledon tissues from the studied *P. vulgaris* genotypes, compared to literature data on *P. lunatus*.
Notes: N, number of established loci; A, anodic migration; C, cathodic migration; P, polymorphism; %, percentage of polymorphic loci; Yes – detected; No – non detected polymorphism. PL, number of polymorphic loci; PLP (%), percentage of polymorphic loci

Factors that may cause such a result are most often the impact of a narrow gene flow, the predominantly autogamous reproductive system possessed by the species, the breeding "pressure" by surrounding microenvironment. The analysis of the genotype composition allows to assume that the studied populations were influenced by factors that have fixed genes encoding enzyme systems.

The manifestation of the homogeneity of the selected group of representatives of *P. vulgaris* contributes to the assessed low level of genetic diversity in the working collection. From the analyses carried out in the gene pools bank of wild forms, compared with the cultivated ones, a vast preponderance is evident in the parameters used for assessment of variability, including those calculated on the basis of enzyme systems. A comparison within the same species, common bean, promotes the development of a strategy to increase its genetic diversity. This in turn will allow an increase in the genetic basis for improving the economic performance of the lines of the working collection and solving the breeding issues posed by them.

Despite the importance of the methods *ex situ* for conservation and breeding of the "big crops" (Falk, 1990), there are many problems related with their application. The problems mainly concern the inadequate procedures for selection of material for sampling and the lack of representativeness of the material maintained in genebanks (Altieri et al., 1987; Brown et al., 1997). The methods for collecting field samples have been thoroughly explored in research using models and theory of population genetics (Ewens, 1972; Marshall et al., 1975; Oka, 1975; Crossa, 1989; Yonezawa et al., 1989; Falk, 1991; Crossa et al., 1993).

Enzyme electrophoresis is one of the techniques frequently applied for analyses in studies requiring screening analyses of the genetic diversity of several loci in large numbers of individuals. Further efforts should be devoted to analysing additional accessions and natural populations of this food legume from the region by the introduced non-expensive, easy of implementation and resolution methods. Genetic tests from segregation data and linkage relationships between polymorphic loci should be investigated to check current results.

7. Conclusions

In this paper the previously described procedures were optimised and applied for analysing common bean accessions and genotypes.

Twenty four loci of 15 enzyme systems studied in cotyledons of French bean from the Plovdiv region were determined. A total of 43 alleles were scored at these loci and in 26% of all a polymorphism was detected.

The data obtained for monomorphic and polymorphic isozyme loci contributed to characterize the genetic basis and the genetic variation within *P. vulgaris* accessions and genotypes that could be used in gene bank to maintain as much genetic diversity as possible, in breeding for crosses and new varieties developing.

8. Acknowledgment

9. References

Adams, M. & Martin, G. (1988). Genetic structure of bean landraces in Malawi, *in* P. Gepts (Ed.), *Genetic ressources of Phaseolus beans; their maintenance, domestication, evolution, and utilization*, Kluwer Academic Publishers, Dordrecht, Holland, pp. 355-374

Aebersold, P., Winans, G., Teel, D., Milner, G. & Utter, F. (1987). *Manual for starch gel electrophoresis: a method for the detection of genetic variation*, U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration N.O.A.A., Report № N.M.F.S. 61, Seattle, WA, U.S.A., pp. 19

Altieri, M. & Merrick, L. (1987). *In situ* conservation of crop genetics of a pioneer tropical tree species: patch dynamics, seed dispersal, and seed banks, *Evolution* 50 (30): 1155-1166

Angioi, S., Rau, D., Attene, L., Nanni, E., Bellucci, G., Logozzo, V., Negri, P., Spagnoletti, L. & Zeuli Papa, R. (2010). Beans in Europe: origin and structure of the European landraces of *Phaseolus vulgaris* L., *Theoretical and Applied Genetics* 121 (5): 829-843

Apostolova, E., Svetleva, D. & Genchev, D. (2005). Diversity in size and shape of the Bulgarian common bean genotypes (*Phaseolus vulgaris* L.), *Annual Reports of BIC* 48: 6-8

Acquaah, G. (1997). *Practical protein electrophoresis for genetic research*, Dioscoredes Press, Portland, OR.

Banuett-Bourrillon, F. & Hague, D.R. (1979). Genetic analysis of alcohol dehydrogenase isozymes in pearl millet (*Pennisetum typhoides*), *Biochemistry & Genetics* 17 (5/6): 537-552

Barrett, S. & Shore, J.S. (1990). Isozyme variation in colonizing plants, *in* D.E. Soltis & P.S. Soltis (eds.), *Isozymes in Plant Biology*, Chapman and Hall, London, England, 106-126

Barrett, S. & Kohn, J. (1991). Genetic and evolutionary consequences of small population size in plants: implications for conservation, *in* D. A. Falk and K. E. Holsinger [eds.], *Genetics and conservation of rare plants*, Oxford University Press, Oxford, UK, pp. 3-30

Bassiri, A. & Adams, M. (1978). An electrophoretic survey of seedling isozymes in several *Phaseolus species*, *Euphytica* 27: 707-720

Baudoin, J.-P., Degreef, J., Hardy, O., Janart, F. & Zoro, I. (1998). Development of an *in situ* conservation strategy for wild Lima bean *(Phaseolus lunatus* L.) populations in the central valley of Costa Rica, *in* S.J. Owens & P.J. Rudall [eds.], *Reproduction Biology,* Royal Botanic Garden Press, Kew, England, 417-426

Baudoin, J.-P., Schmit, V. & Wathelet, B. (1991). Observations on seed protein electrophoretic patterns of some species within the genus Phaseolus, *Annual Report of Bean Improvement Cooperative* 34: 85

Baudoin, J.-P. (1981). Observation sur quelques hybrides interspécifiques avec *P. lunatus* L. *Buletin des Recherches Agronomiques de Gembloux, Belgique*, pp. 359

Beebe, S., Rengifo, J., Gaitan, E., Duque, M.C. & Tohme, J. (2001). Diversity and origin of Andean landraces of common bean, *Crop Science* 41: 854–862

Beebe, S., Skroch, P.W., Tohme, J., Duque M.C., Pedraza, F. & Nienhuis, J. (2000). Structure of genetic diversity between common bean landraces of Middle American origin based on correspondence analysis of RAPD, *Crop Science* 40: 264–273

Brown, A., Brubaker, C. & Grace, J. (1997). Regeneration of germplasm samples: wild versus cultivated plant species, *Crop Scienc*e 37 (1): 7-13

Brown, A.D.H. (1980). Genetic bases of alcohol dehydrogenase in *Hordeum spontaneum*, *Journal Heredity* 71: 127-128

Brown, A.H.D. (1978). Isozymes, plant population genetic structure and genetic conservation, *Theoretical and Applied Genetics* 52: 145-157

Brown, J., Ma, Y., Bliss, F. & Hall, T. (1981). Genetic variation in the subunits of globulin-1 storage protein of French bean. *Theoretical and Applied Genetics* 59: 83-88

Brownly, K.A. (1977). *Theory and methods of statistic analysis in science and technology*, 'Nauka', Moscow, pp. 404 /*in Russian*/

Caicedo, A., Gaitan, E., Duque, M., Toro, O., Debouck, D. & Tohme, J. (1999). AFLP fingerprinting of *Phaseolus lunatus* L. and related wild species from South America, *Crop Science* 39: 1497–1507

Castineiras, L., Esquivel, M., Lioi, L., & Hammer, K. (1991). Origin diversity and utilization of the Cuban germplasm of common bean (*P. vulgaris* L.), *Euphytica* 57: 1-8

Chacyn, S., Pickersgill, M. & Debouck, D. (2005). Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races, *Theoretical and Applied Genetics* 110: 432–444

Chase, C., Ortega, V. & Vallejos, C. (1991). DNA restriction fragment length polymorphisams correlate with isozyme diversity in *Phaseolus vulgaris* L., *Theoretical and Applied Genetics* 81: 806-811

Chetelat, R., Adams, F. & Adams, D. (1999). *Mae-1*, a *malic enzyme* coding gene on chromosome 5, *Tomato Genetics Cooperative* 49: 12-14

Chrispeels, M., Bollini, R., & Harris, N. (1978). Biosynthesis, accumulation and catabolism of storage proteins in legume seeds, *Berichte der Deutschen Botanischen Gesellschaft* 535-551

CIAT Report Informe (1990b). La Riqueza Oculta del Frijol Silvestre: Collección de Germoplasma del CIAT, 26-31

Crawford, D. (1990). *Plant Molecular Systematics-Macromolecular Approaches*, John Wiley & Sons, Inc., New York, U.S.A., pp. 388

Crossa, J. (1989). Methodologies for estimating the sample size required for genetic conservation of outbreeding crops, *Theoretical and Applied Genetics* 77: 153-161

Debouck, D., Castillo, T. & Tohme, J. (1989). Observations on little-known *Phaseolus* germplasm of Ecuador, *Plant Genetic Resources Newsletter (FAO/IBPGR)* 80: 15-21

Delgado-Salinas, A., Turley, T., Richman, A. & Lavin, M. (1999). Phylogenetic analysis of the cultivated and wild species of *Phaseolus* (Fabaceae), *Systemes of Botany* 24: 438–460

Dobie, P., Dendy, J., Sherman, C., Padgam, J., Wood, A. & Gatehouse, A.M.R. (1990). New sources of resistance to *A. obtectus* and *Z. subfasciatus* bokeman (*Coleoptera: Bruchidae*) in mature seed of five species of *Phaseolus*, *Journal of Products Stored Research* 4: 177-186

Doebley, J. (1990). Isozyme evidence and the evolution of crop plants, *in* Soltis D.E. & Soltis P.S. (eds), *Isozymes in plant biology*. Chapman and Hall, London, England, pp. 165-191

Donoghue, M. & Sanderson, M. (1992). The suitability of molecular and morphological evidence in reconstructing plant phylogeny, *in* Soltis P.S. & D.E. Soltis & J.J. Doyle (eds.), *Molecular Systematics of Plants*, Chapman and Hall, London, England, pp. 340-368

Evans, A.M. (1980). Structure, variation, evolution, and classification in *Phaseolus*, *in* R.J. Summerfield & A.H. Bunting (eds.), *Advances in legume science*, Royal Botanic Garden, Kew, England, pp. 337-347

Ewens, W.J. (1972). The sampling theory of selectively neutral alleles, *Theoretical Population Biology* 3: 87-112

Falk, D. (1991). Joining biological and economic models for conserving plant genetic diversity, *in* Falk, D.A. & Holsinger, K.E. (eds), *Genetic and conservation of rare plants*, Oxford University Press, Oxford, UK, pp. 209-223

Fofana, B., du Jardin, P. & Baudoin, J.-P. (2001). Genetic diversity in the Lima bean (*Phaseolus lunatus* L.) as revealed by chloroplast DNA (cpDNA) variations, *Genetic Resources & Crop Evolution* 48: 437–445

Fofana, B., Vekemans X., du Jardin, P. & Baudoin, J.-P. (1997). Genetic diversity in Lima bean (*Phaseolus lunatus* L.) as revealed by RAPD markers, *Euphitica* 95: 157-165

Frankel, O. (1987). Characterization and evolution what, why, how and whom? *Tenth anniversary of IBPGR/SEAN*, June 1977-June 1987. Special issue on characterization and preliminary evaluation of crop genetic resources, N. Chomchalow (ed.), I.B.P.G.R. Rome, Italy, 9-14

Freeling, M. & Schwartz, D. (1973). Genetic relationships between the multiple alcohol dehydrogenase of maiz, *Biochemistry & Genetics* 8 (1): 27-36

Genchev, D. (2011). Plasticity and stability of the productivity of new common bean cultivars (*Phaseolus vulgaris* L.), *Annual Report of the Bean Improvement Cooperative* 54: 70-71

Gepts, P. & Debouck, D. (1991). Origin, domestication, and evolution of the common bean (*P. vulgaris* L.), *in* A. van Schoonhoven & O. Voysest (eds.), *Common beans: research for crop improvement*, C.A.B. International, U.K., 7-53

Gepts, P. & Bliss, F.A. (1988). Dissemination pathways of common bean (*Phaseolus vulgaris*, *Fabaceae*) deduced from phaseolin electrophoretic variability. II. Europe and Africa, *Economical Botany* 42: 86–104

Gepts, P. & Bliss, F.A. (1986a). Phaseolin variability among wild and cultivated bean germplasm, *Journal of Heredity* 76: 447-450

Gepts, P., Osborn, T., Rashka, K. & Bliss, F. (1986b). Phaseolin protein variability in wild forms and landraces of the common bean (*P. vulgaris* L.): evidence for multiple centers of domestication, *Economical Botany* 40: 451-468

Gepts, P. & Bliss, F. (1984). Enhanced available methionine content associated with higher phaseolin levels in common bean (*P. vulgaris* L.) seeds, *Annual Report of Bean Improvement Cooperative* 27: 19-20

Gepts, P. (1995). Genetic markers and core collections, *in* Hodgkin, T., Brown, A.H.D., van Hintum, T.J.L., Morales, E.A.V. (eds), John, Wiley & Sons, *Core collections of plant genetic resourses*, Chichester, UK, pp. 127-146

Gepts, P. (1993). The use of molecular and biochemical markers in crop evolution studies, *in* M.K. Hecht (ed.), *Evolutionary biology*, Plenum Press, New York, 27: 51-94

Gepts, P. (1990). Biochemical evidence bearing on the domestication of *Phaseolus (Fabaceae)* beans, *Economical Botany* 44 (3): 28-38

Gepts, P. (1988a). A Middle American and an Andean common bean gene pool, *in* P. Gepts (ed.), *Genetic resources of Phaseolus beans; their maintenance, domestication, evolution, and utilization*. Kluwer Academic Publishers, Dordrecht, Holland, 375-390

Gepts, P. (1988b). Phaseolin as an evolutionary marker, *in* Gepts P. (ed.), *Genetic resources of Phaseolus beans; their maintenance, domestication, evolution, and utilization*, Kluwer Academic Publishers, Dordrecht, Holland, pp. 215-241

Gil, de Rona, F. (1992). Variation of *P. vulgaris* in the Northwest of the Iberian Peninsula, *Plant Breeding* 109(4): 313-319

Gottschalk, W. & Müller, H. (1983). Seed Proteins, *Biochemistry Genetics Nutritive Value*, pp. 531

Griffin, J. & Palmer, S. (1987). Inheritance and linkage studies with five isozyme loci in soybeans, *Crop Science* 27: 885-892

Guo, M., Mok, D. & Mok, M. (1989). Isozyme banding accessions and embryo development in interspecific crosses of *Phaseolus, The Journal of Heredity* 80: 29-32

Gutiérrez-Salgado, A., Gepts, P. & Debouck, D.G. (1995). Evidence for two gene pools of the lima bean, *Phaseolus lunatus* L., in the Americas, *Genetic Resources in Crop Evolution* 42: 15–28

IBPGR-FAO (1988). Progress report on genetic diversity of lima beans (*Phaseolus lunatus*). *Plant Genet Resources Newsletter* 73 (74): 57

Hall, T., McLeester, R. & Bliss, F. (1977). Equal expression of the material and paternal alleles for polypeptide subunits of major storage protein of the bean *P. vulgaris, Plant Physiology* 59: 11-22

Hamerick, J. (1990). Isozymes and analyses of genetic structure in plant populations, *in* D.E. Solyis & P.S. Soltis (eds.), *Isozymes and Plant Biology*, Sinauer Assoc. Ins. Sunderland, Massachussetts, U.S.A., pp. 43-63

Hamrick, J., Murawski, D. & Loveless, M. (1991). Correlation between species characters and allozyme diversity: implications for conservation biology, *in* Falk D. and K. Holsinger (eds.), *Genetics and conservation of rare plants*, Oxford University Press, Oxford, UK, 75-86

Harlan, J.R. & De Wet, J.M.J. (1971). Toward a rational classification of cultivated plants, *Taxon* 20: 509-517

Haufler, C.H. (1990). Proteins I: Isozyme electrophoresis, *in* D. M. Hillis & C. Moritz (eds.), *Molecular systematics*, Sinauer Associates, Inc., Sunderland, Massachusetts, U.S.A., pp. 45-126

Hillis, D. (1987). Molecular versus morphological approaches to systematics, *Annual Review of Ecology, Evolution & Systematics* 18: 23-42

Hussain, A., Bushuk, W., Ramirez, H. & Roca, W. (1988). A Practical Guide for electrophoretic analysis of Isozymes and Proteins in Cassava, *Field Beans and Forage Legumes Working Document* 140, CIAT

Isola, M. & Franzoni, L. (1993). Studies on the proteolytic enzymes of potato tuber, *Plant Physiology & Biochemistry* 31 (2): 169-174

Jaaska, V. & Jaaska, V. (1988). Isoenzyme variation in the genera *Phaseolus* and *Vigna* (*Fabaceae*) in relation to their systematics: Aspartate aminotransferase and superoxide dismutase, *Plant Systematics and Evolution* 159: 145-159

Kami, J., Becerra, Velasquez, V., Debouck, D. & Gepts, P. (1995). Identification of presumed ancestral DNA sequences of Phaseolin in *P. vulgaris. Proceedings of National Academic of Sciences of U.S.A.* 92: 1101-1104

Kaplan, L. (1981). What is the origin of the common bean? *Economic Botany* 35: 240-254

Kazan, K., Muehlbauer, F., Weeden, N. & Ladizinski, G. (1993). Inheritance and linkage relationships of morphological and isozyme loci in *Chickpea* (*Cicer arientinum* L.), *Theoretical and Applied Geneics* 86: 417-426

Kephart, S. (1990). Starch Gel Electrophoresis of Plant Isozymes: A Comparative Analysis of Techniques, *American Journal of Botany* 77 (5): 693-712

Krasteva, L., Nikolova, V., Petkova, V. & Poryazov, I. (2004). Study of local gene pool of garden bean *(Phaseolus vulgaris* L.) in relation to breeding, Dobroudja Agricultural Institute, General Toshevo, *Scientific Works, Field Crops Studies* 1 (2): 220-225 /in Bulgarian/

Koenig, R., Singh, S. & Gepts, P. (1990). Novel phaseolin types in wild and cultivated common bean (*P. vulgaris, Fabaceae*), *Economical Botany* 44: 50-60

Koenig, R. & Gepts, P. (1989a). Allozyme diversity in wild *P. vulgaris*: further evidence for two major centers of genetic diversity, *Theoretical and Applied Geneics* 78: 809-817

Koenig, R. & Gepts, P. (1989b). Segregation and linkage of genes for seed proteins, isozymes, and morphological traits in common bean (*P. vulgaris* L.), *The Journal of Heredity* 80: 455-459

Koenig, R., Stegemann, H., Francksen, H. & Paul, H. (1970). Protein subunits in the potato virus N group, Determination of the molecular weights by polyacrylamide electrophoresis, *Biochemical & Biophysical Acta* 207: 184-189

Konarev, V. (1998a). Molecular biological studies of gene pool of cultivated plants in VIR (1967-1997), *Theoretical bases in selection* 1983, M (I), pp. 187

Konarev, A. (1998b). Use of molecular markers in work with genetic resources, *Selkolhoziaiskoi Biology* 5: 3-25

Konarev, A. (1993). Molecular biological aspects of applied botanics, genetics and selection, in: *Theoretical Bases in Selection* M:I *I* Ed. by VG. Konarev, vol. 1, Moscow, Kolos, pp. 447 /in Russian/

Kumanov, B., Poryazov, I., Uzunova, E., Popov, D., Kostov, D. & Velev, B. (1988). *Legumes*, Zemizdat, Sofia, pp. 157

Kwak, M. & Gepts, P. (2009a). Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae), *Theoretical and Applied Geneics* 118: 979–992

Kwak, M., Kami, J.A. & Gepts, P. (2009b). The putative Mesoamerican domestication center of *Phaseolus vulgaris* is located in the Lerma-Santiago Basin of Mexico, *Crop Science* 49: 554–563

Lackey, J. (1977). A revised classification of the tribe *Phaseoleae* and its relation to canavanine distribution, *Botanical Journal of the Linnean Society* 74: 163-178

Ladizinsky, G. (1985). Founder effect in crop plant evolution, *Economical Botany* 39:191-199

Lawrence, M.J., Marshal, D.F. & Davies, P. (1995). Genetics of genetic conservation. II. Sample size when collecting germplasm, *Euphytica* 84: 89-99

Leakey, C.L.A. (1988). Genotypic and phenotypic markers in common bean, *in* P. Gepts, (ed.) *Genetic Resources of Phaseolus Beans; their Maintenance, Domestication, Evolution, and Utilization*, Kluwer, Dordrecht, the Netherlands, pp. 245-327

Lefort-Buson, M. & Vienne, D. (1985). *Les distances génétiques. Estimations et Applications.* INRA, Paris, France, pp. 181

Les, D.H. (1991). Genetic diversity in the monoecious hydrophile *Cerato-phyllum* (Ceratophyllaceae), *American Journal of Botany* 78: 1070-1082

Lioi, L. & Bollini, R. (1984). Contribution of processing events to the molecular heterogeneity of four banding types of phaseolin, the major storage protein of *P. vulgaris* L., *Plant Molecular Biology* 3: 345-535

Lioi, L. (1989). Geographical variation of phaseolin patterns in an old world collection of *Phaseolus vulgaris*, *Seed Science Technology* 17: 317–324

Logozzo, G., Donnoli, R., Macaluso, L., Papa, R., Knupffer, H. & Spagnoletti, Zeuli, P.L. (2007). Analysis of the contribution of Mesoamerican and Andean gene pools to European common bean (*Phaseolus vulgaris* L.) germplasm and strategies to establish a core collection, *Genetics Resources & Crop Evolution* 54: 1763–1779

Loveless, MD & Hamrick, J.L. (1984) Ecological determinants of genetic structure in plant populations, *Annual Revue of Ecological Systems* 15: 65-95

Ma, Y. & Bliss, F. (1978). Seed proteins of common bean, *Crop Science* 18: 431-437

Manen, J. & Otoul, E. (1981). Etudes électrophorétiques et détermination des fractions protéiques principales chez quelques cultivars élites de *P. lunatus* L. et de *P. vulgaris* L., *Bulletin des Recherches Agronomiques de Gembloux* 16 (4): 309-326

Maquet, M., Vekemans, X. & Baudoin, J.-P. (1999). Phylogenetic study on wild allies of Lima bean, *Phaseolus lunatus (Fabaceae)*, and implications on its origin, *Plant Systems Evolution* 218: 43-54

Maquet, A. & Baudoin, J.-P. (1997). Aperçu de la distribution néotropicale de *Phaseolus lunatus*, *Belgium Journal of Botany* 130 (1): 93-116

Maquet, A., Zoro, Bi I., Delvaux, M., Wathelet, B. & Baudoin, J.-P. (1997). Genetic structure of a Lima Bean base collection using allozyme markers, *Theoretical & Applied Genetics* 95: 980-991

Maquet, A., Zoro, Bi.I., Rocha, O.J. & Baudoin, J.-P. (1996). Case studies on breeding systems and its consequences for germplasm conservation, 1. Isoenzyme diversity in wild Lima bean populations in Central Costa Rica, *Genetic Resource and Crop Evolution* 43: 309-318

Maquet, A. (1995). *Etude de la diversité génétique de la légumineuse alimentaire P. lunatus L. par l'analyse de caractères morphophysiologiques et des marqueurs protéiques*, Dissertation, FUSA de Gembloux, Belgique, pp. 187

Maquet, A., Gutiereza, A. & Debouck, D.G. (1993). Further Biochemical Evidence for the Existence of two Gene Pools in Lima Beans, *Annual Reports of Beans Improvement Cooperative* 33: 128-129

Maréchal, R., Mascherpa, J.-M. & Stainier, F. (1978). *Etude taxonomique d'un groupe complexe d'espèces des genres Phaseolus et Vigna (Papilionaceae) sur la base de données morphologiques et polliniques, traitées par l'analyse informatique*, C.J.B.G., Genève, Suisse, 28: pp. 273

Marshall, D. & Brown, A.H.D. (1975). Optimum Sampling in Genetic Conservation, *in* Franckel O.M., Hawkes (eds.), *Crop Genetic Resources for Today and Tomorrow*, International Biology Cal Programmes; Cambridge Un. Press, England, pp. 53-83

May, B. (1992). Starch-gel electrophoresis of allozymes, *in* Hoelzel AR (ed), *Molecular genetic analysis of populations – a practical approach*, Oxford University Press, Oxford, England, 1-27

McClean, P., Lee, R. & Miklas, P. (2004). Sequence diversity analysis of dihydroflavonol 4-reductase intron 1 in common bean, *Genome* 47: 266–280

Meurrens, F., Baudoin, J-P., Degreef, J. & Rocha, O.J. (2009). Demographic study in microconservation sites to maintain in situ wild Lima beans (*Phaseolus lunatus* L.) in the Central Valley of Costa Rica, *PGR Newsletter, FAO Biodiversity* 128: 45-50

Motta-Aldana, J.R., Serrano-Serrano, M., Hernández-Torres, J., Castillo-Villamizar, G., Debouck, D. & Chacyn, M. (2010). Multiple Origins of Lima Bean Landraces in the Americas: Evidence from Chloroplast and Nuclear DNA Polymorphisms, *Crop Science* 50: 1773–1787

Motten, A.F. & Antonovics, J. (1992). Determinants of outcrossing rate in a predominantly self-fertilizing weed, *Datura stramonium* (Solanaceae), *American Journal of Botany* 79: 419–427

Moritz, C. & Hillis, D. (1990). *Molecular Systematics*, Hillis & Moritz (eds.), Sinauer Associates, Inc., Sunderland, Massachusetts, U.S.A., 1-10

Murphy, R., Sites, J., Both, D. & Haufler, C. (1990). Proteins, Isozyme Electrophoresis, *in* D. M. Hillis and C. Moritz [eds.], *Molecular Systematics*, Sinauer Associates, Sunderland, Massachusetts, U.S.A. pp. 45-126

Murphy, R., Sites, J., Buth, D. & Haufler, C. (1990). Proteins I: isozyme electrophoresis, *in* D.M. Hillis & C. Moritz (eds.), *Molecular Systematics*, Sinauer Associates Inc, Sunderland, Massachusetts, New York, U.S.A. 1-126

Murphy, R., Sites, J., Buth, D. & Haufler, C. (1990). Proteins I: isozyme electrophoresis. In: *Molecular Systematics*, D.M. Hillis & C. Moritz (eds.), Sin Ass., Inc., New York, U.S.A. 1-10

Murphy, R., Sites J., Buth, D. & Garber, C. (1967a). Genetic studies of variant enzymes. I. An electrophoretic survey of esterases and leucine aminopeptidasein the genus *Phaseolus*, *Canadian Journal of Genetics and Cytology* 9: 646-655

Neale, D. & Adams, W. (1981). Inheritance of isozyme variants in seed tissues of balsam fir (*Abies balsamea*), *Canadian Journal of Botany* 59: 1285-1291

Nienhuis, J., Tivang, J., Skroch, P. & Santos, B. Dos (1995). Genetic relationships among cultivars and landraces of Lima bean (*Phaseolus lunatus* L.) a measured by RAPD markers, *Journal of American Society of Horticultural Science* 120: 300-306

Nikolova, V., Poryazov, I. & Rodeva, V. (1986). Hybridization between *P. acutifolius* Gray and *P. vulgaris* L., *Annual Reports of the Bean Improvement Cooperative* 29: 124-125

Nikolova, V., Petkova, V., Poryazov, I. (2003). Possibilities for selection of garden bean (*Phaseolus vulgaris* L.) genotypes tolerant to high temperature. II. Variation of pollen viability, *Annual report of Bean Improvement Cooperative* 46: 83-84

Nodari, R.O., Koinange, E.M.K. & Gepts, P. (1992). Linkage mapping in common bean, *Annual Reports of the Bean Improvement Cooperative* 35: 27-28

Novak, S., Mack, R. & Soltis, D. (1991). Genetic variation in *Bromus tectorum* (Poaceae): population differentiation in its North American ranges, *American Journal of Botany* 78: 1150-1161

Ocampo, C., Martin, J., Ortiz, J., Sanchez-Yelamo, M., Toro, O. & Debouck, D. (2002). Possible origins of common bean (*Phaseolus vulgaris* L.) cultivated in Spain in relation to the wild genetic pools of Americas, *Annual Reports of the Bean Improvement Cooperative* 45: 236–237

Ollitrault, P. (1987). *Evaluation génétique des sorghos cultivés (Sorghum bicolor L. Moench) par l'analyse conjointe des diversités enzymatique et morphophysiologique – relations avec les sorghos sauvages*, Ph.D. thésis, Université de Paris-Sud, Centre d'Orsey, Paris, France

Oka, H.I. (1975). Consideration on the population size necessary for conservation of crop germplasm, *in* Matsuo T. (ed.), *JIBP Synthesis* pp. 57-63

Osborn, T. (1988). Genetic control of bean seed protein, *CRC Critical Reviews in Plant Sciences* 7(2): 93-116

Pachico, D. (1989). Trends in world common bean production, *in* H.F. Schwartz & M.A. Pastor-Corrales (eds.), *Bean Production Problems in the Tropics*, C.I.A.T., Cali, Colombia, pp. 72

Papa, R. & Gepts, P. (2003). Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (*Phaseolus vulgaris* L.) from Mesoamerica, *Theoretical & Applied Genetics* 106: 239–250

Park, S.O., Coyne, D.P., Jung G., Skroch, P.W., Arnaud-Santana, E., Steadman, J.R., Ariyarathne, H.M. & Nienhuis, J. (2000). Mapping of QTL for Seed Size and Shape Traits in Common Bean, *Journal of American Society of Horticultural Science* 125 (4): 466–475

Pasteur, N., Bonhomme, G., Catalan, F. & Britton-Davidian, J. (1987). *Manuel technique de génétique par électrophorèse des protéines*, Lavoisier (ed.) TEC&DOC, Paris, France, pp. 217

Petkova, V., Denev, I. & Stefanov, D. (2010). Resistance to high temperature stress in various bean (*Phaseolus vulgaris* L.) cultivars and lines, *General and Applied Plant Physiol*ogy 35 (3-4): 117-121

Pinero, D. & Eguiarte, L. (1988). The origin and biosystematic status of *P. coccineus ssp. polyanthus*: electrophoretic evidence, *Euphytica* 37: 199-203

Poryazov, I., Zagorcheva, L., Nikolova, V. & Rodeva, V. (1988). Results of the hybridization between *P. vulgaris* L. and *P. coccineus* L. or *P. acutifolius* A. Gray, *Proceedings of the Symposium on Interspecific Hybridisation and Methods for Transfer of Alien Genetic Material*, Sofia 29-30: 296-300

Poryazov, I. (1990). *Studies in green bean selection*, Habilitation paper, Vegetable Crops Research Institute, Plovdiv, pp. 272

Rodiño, A., Santalla, M., De Ron, A. & Singh, S. (2003). A core collection of common bean from the Iberian peninsula, *Euphytica* 131: 165–175

Rodiño, A., Santalla, M., Montero, I., Casquero, P. & De Ron, A. (2001). Diversity in common bean (*Phaseolus vulgaris* L.) germplasm from Portugal, *Genetic Resources Crop Evolution* 48: 409–417

Romero-Andreas, J. & Bliss, F. (1985). Heritable variation in the phaseolin protein of nondomesticated common bean, *P. vulgaris* L., *Theoretical & Applied Genetics* 71: 478-480

Roux, L. (1987). Utilisation des isozymes comme marqueurs génétiques, *Le Sélectionneur français* 39: 31-39

Rossi, M., Bitocchi, E., Bellucci, E., Nanni, L., Rau, D., Attene, G. & Papa, R. (2009). Linkage disequilibrium and population structure in wild and domesticated populations of *Phaseolus vulgaris* L., *Evolutionary applications*, Published Online July 2009. doi:10.1111/j.1752-4571.2009.00082.x

Serrano-Serrano, M., Hernбndez-Torres, J., Castillo-Villamizar, G., Debouck, D. & Chacyn, M. (2010). Gene pools in wild Lima bean (*Phaseolus lunatus* L.) from the Americas: Evidences for an Andean origin and past migrations, *Molecular & Phylogenetic Evolution* 54: 76–87

Singh, S. (1991). Phylogenetic relations in section *Arachis* based on seed protein profile, *Theoretical & Applied Genetics* 82: 593-597

Scandalios, J. & Espiritu, L. (1969). Mutant aminopeptidases of *Pisum sativum*, I. Development genetics and chemical characteristics, *Molecular Genetics* 105: 101-112

Schaal, B., Leverich, W. & Rogstad, S. (1991). A Comparison of Methods for Assessing Genetic Variation in Plant Coservation Biology, *in* Falk, I.E. Holsinger (eds.) *Genetic and Conservation of Rare &Plants,* Oxford University Press, New York

Schinkel, C. & Gepts, P. (1989). Allozyme variability in the tepary bean *Phaseolus acutifolius* A. Gray, *Plant Breeding* 102: 182-195

Singh, S., Gepts, P. & Debouck, D. (1991a). Races of common bean (*P. vulgaris, Fabaceae*), *Economical Botany* 45 (3): 379-396

Singh, S., Nodari, R. & Gepts, P. (1991b). Genetic diversity in cultivated bean germplasm: I. Allozymes, *Journal of Heredity* 76: 447-450

Shaal, B., Leverich, W. & Rogstad, S. (1991). A Comparison of Methods for Assessing Genetic Variation in Plant Coservation Biology, *in* Falk, I.E. Holsinger (eds.), *Genetic and Conservation of Rare Plants,* Oxford University Press, New York

Slightom, J., Drong, R., Klassy, R. & Hoffman, L. (1985). Nucleotide sequences from phaseolin cDNA clones: the major storage proteins from *P. vulgaris* are encoded by two unique gene families, *Nucleic Acids Research* 13: 6483-6498

Sofkova, S. & Petkova, V. (2007). Evaluation of plastid pigments content in garden bean (*Phaseolus vulgaris* L.) pods, *Annual Report of the Bean Improvement Cooperative* 50: 61-62

Sofkova, S. & Yankova, V. (2008). Vegetable bean recombinant inbred lines (*Phaseolus vulgaris* L.), *Annual Report of the Bean Improvement Cooperative* 51: 118-119

Sofkova, S. & Yankova, V. (2011). Biological activity of some products for plant protection against common bean weevil (*Acanthoscelides obtectus* Say) on Bulgarian snap bean (*Phaseolus vulgaris* L.) cultivars, *Annual Report of the Bean Improvement Cooperative* 54: 92-93

Sozinov, A. (1995). *Protein polymorphism and its importance in genetics and selection*, Nauka, pp. 215

Stoyanova, M. (1989). *Study on nitrogen-containing compounds and protein complex and its use for varietal characteristics of beans*, Dissertation, pp. 176

Stuber, C.W. (1989). Isozymes as markers for studying and manipulating quantitative traits, *in* Soltis D. & Soltis P. (eds), *Isozymes in plant biology,* Dioscorides Press, Oregon, Portland, U.S.A., 206-220

Sturm, A., Van Kuik, J., Vliegenthart, J. & Chrispeels, M. (1987). Structure, position, and cDNA biosynthesis of the high-mannose and the complex oligosaccharide side chains of the bean storage protein phaseolin, *Journal of Biology & Chemistry* 262: 13392-13403

Svetleva, D., Pereira, G., Carlier, J., Cabrita, L., Leitão, J. & Genchev, D. (2006). Molecular characterization of *Phaseolus vulgaris* L. genotypes included in Bulgarian collection by ISSR and AFLPTM analyses, *Science Horticulturae* 109: 198-206

Svetleva, D., Velcheva, M. & Bhowmik, G. (2003). Biotechnology as a useful tool in common bean (*Phaseolus vulgaris* L.) improvement, *Euphytica* 131: 189-200

Talbot, D., Adang, M., Slightom, J. & Holl, T. (1984). Size and organization of the multigene family encoding phaseolin, the major seed storage protein of *P. vulgaris* L., *Molecular Genetics* 198: 42-49

Tanksley, S.D. (1983). *Isozymes in Plant Genetics and Breeding, Part A,* SD Tanksley and TJ Orton (Eds), Elsevier Science Publishers BV, Amsterdam, pp. 109-138

Tanksley, S. & Jones, R. (1981). Effects of O_2 stress on tomato alcohol dehydrogenase activity: description of a seconde ADH coding gene, *Biochemistry & Genetics* 67: 209-216

Tohme, J., Toro, O.Ch., Vargas, J. & Debouck, D. (1995). Variability in Andean Nuña Common Beans *(P. vulgaris, Fabaceae)*, *Economical Botany* 49 (1): 78-95

Tomlekova, N.B. (2010) Induced mutagenesis for crop improvement in Bulgaria. *Plant Mutation Report* 2 (2): 1-32

Tomlekova, N., Poryazov, I., Krasteva, L. & Baudoin, J.-P. (2009). Genetic diversity of the Bulgarian *Phaseolus vulgaris* L. germplasm collection through phaseolin marker, *Acta Horticulturae* I: 213-220

Tomlekova, N., Krasteva L., Poryazov I. & Ivanov P. (2001). Conformity of the electrophoretic families with morphological characters in samples of Bulgarian collection of *Phaseolus vulgaris* L. III. Productivity, *Bulgarian Journal of Agricultural Science* 7 (4-5): 403-407

Tomlekova, N., Cholakov, T., Poryazov, I. & Ivanov, P. (1999). Distribution of Phaseolin Types of Common Beans *Phaseolus vulgaris* L. Depending on Physical and Geographical Conditions of the Territory of Bulgaria, *Bulgarian Journal of Agricultural Science* 5 (6): 713-718

Tomlekova, N. (1999). *Application of protein patterns for establishment of the genetic diversity in common bean (Phaseolus vulgaris L.) in Bulgaria,* Dissertation, pp. 150

Tomlekova, N. (1998). Mechanism of resistance against bruchid beetle (*A. obtectus*)(*Say*) in common bean (*P. vulgaris* L.), *Agricultural Science* 5: 19-21

Toro, O., Tohme, J. & Debouck, D. (1990). Wild bean (*P. vulgaris* L.): Desciption and Distribution. *Working Document №181,* I.B.P.G.R. & C.I.A.T., Cali, Colombia, pp. 106

Torres, R.I.G., Villalobos, R.A., Solís, E.G. & Debouck, D.G. (2004). Wild common bean in the Central Valley of Costa Rica: Ecological distribution and molecular characterization, *Agronomía Mesoamericana* 15 (2): 145-153

Torres, A. (1989). Isozyme analysis of tree fruits, *in* Soltis, D. & Soltis, P. (eds), *Isozymes in plant biology*, Dioscorides Press, Oregon, Portland, U.S.A., pp. 192-205

Torres, A. (1974). Genetics of sunflower alcohol dehydrogenase: Adh$_2$, nonlinkage to Adh$_1$ and Adh1 early alleles, *Biochemistry & Genetics* 12 (5): 385-392

Tuwafe, S., Kalher, A., Boe, A. & Ferguson, M. (1988). Inheritance and geographical distribution of allozyme polymorphisms in chickpea (*Ciser arietinum* L.), *Journal of Heredity* 79: 203-207

Vanderborght, T. (1988). A centralized database for the common bean and its use indiversity analysis, *in Genetic resources of Phaseolus beans; their maintenance, domestication, evolution, and utilization,* Gepts P. (ed.), Kluwer Academic Publishers, Dordrecht, Holland 51-65

van Slotten, D. (1987). The role of curators, breeders and other users of germplasm in characterization and evaluation of crop genetic resources, *Tenth anniversary of*

IBPGR/SEAN June 1987, Special issue on characterization and preliminary evaluation of crop genetic resources, *in* N. Chomchalow (ed.), *IBPGR Newsletter* 54: 18-25

Vavilov, N. (1987). Origin and geography of cultivated plants. L., *Nauka*, pp. 440

Wall, J. (1968). Leucin aminopeptidase polymorphism in *Phaseolus* and differencial elimination on the donor parent genotype in interspecific backcrosses, *Biochemistry & Genetics* 2: 109-118

Wall, J. & Wall, S. (1975). Isozyme polymorphism in the study of evolution in the *P. vulgaris-P. coccineus* complex of Mexicokak, *Isozymes, in* C.J. Markert (ed.), Academic Press, New York U.S.A., 287-305

Wang, R.-L., Wendel, J. & Dekker, J. (1995a). Weedy adaptation in *Setaria ssp*. I. Isozyme analysis of genetic diversity and population genetic structure in *Setaria viridis, American Journal of Botany* 82: 308-317

Wang, R.-L. Wendel, J. & Dekker, J. (1995b). Weedy adaptation in *Setaria* ssp. II. Genetic diversity and population genetic structure in *S. glauca, S. geniculata,* and *S. faberii* (Poaceae), *American Journal of Botany* 82: 1031-1039

Weeden, N. & Wendel, J. (1990). Genetics of plant isozymes, *Isozymes in Plant Biology*, *in* Soltis D. & Soltis, P. (eds.), Chapman and Hall, London, England pp. 5-45

Weeden, N.F. (1983). Evolution of plant isozymes, *in* SD Tanksley and TJ Orton (Eds), *Isozymes in Plant Genetics and Breeding*, Elsevier Science Publishers BV, Amserdam, pp. 139-156

Wendel, J. & Weeden, N. (1990). Visualisation and Interpretation of Plant Isozymes, *in* Soltis D. & Soltis P. (Eds), *Isozymes in Plant Biology*, Chapman and Hall, London, England, ISBN 0-412-36500-6, pp. 5-45

Wendel, J.F. & Parks, C.R. (1984). Distorted segregation and linkage of alcohol dehydrogenase genes in *Camellia japonica L.* (Theaceae), *Biochemistry & Genetics* 22 (7/8): 739-748

West, N. & Garber, E. (1996). Genetic studies of variant enzymes. I. An electrophoretic survey of esterases and leucine aminopeptidase in the genus *Phaseolus*, *Canadian Journal of Genetics and Cytology* 9: 646-655

West, N. & Garber, E. (1967). Genetic studies of variant enzymes. II. The inheritance of esterases and leucine aminopeptidases in *P. vulgaris* x *P. coccineus*, *Canadian Journal of Genetics and Cytology* 9: 646-655

Working materials (2001). Molecular characterization of mutated genes controlling important traits for seed crops, IAEA, Vienna, Austria, pp. 187

Wyatt, R., Evans, E. & Sorenson, J. (1992). The evolution of self-pollination in granite outcrop species of *Arenaria* (Caryophyllaceae). VI. Electrophoretically detectable genetic variation, *Systematic Botany* 17: 201-209

Yankova, V. (2010). Bean weevil (*Acanthoscelides obtectus* Say, Coleoptera; Bruchidae): bionomy, relationships with the host *Phaseolus vulgaris* L. and possibilities for control, Dissertation, Agricultural University, pp. 146

Yankova, V., Boev, B. & Sofkova, S. (2011). Biological control of *Thrips tabaci* Lind. with some phytopesticides in beans (*Phaseolus vulgaris* L.), *Annual Report of the Bean Improvement Cooperative* 54: 90-91

Yankova, V., Tomlekova N., Ilieva A., Penchev E. & Genchev D. (2004). Been weevil (*Acanthoscelides obtectus* Say) infestation in common bean (*Phaseolus vulgaris* L.)

varieties and lines. Proceedings of the International Conference on Horticulture, "Post-Graduate Study System and Conditions in Europe", 16th–19th November 2004, Lednice, Czech Republic, 265–272

Yonezawa, K. & Ichibashi, H. (1989). Sample size for collecting germplasms from natural populations in view of the genotypic multiplicity of seed embryos borne on a single plant, *Euphytica* 41: 91-97

Zagorcheva, L., Poryazov, I. & Nikolova, V. (1983a). Meiosis in pollen mother cells of hybrids *P. vulgaris* L. *x P. coccineus* L. in F$_1$, *Genetics and Selection* 3: 215-225

Zagorcheva, L., Poryazov, I., Rodeva, V. & Nikolova, V. (1983b). Interspecific hybrids in *P. vulgaris* x *P. coccineus*. I. Development and fertility of the hybrids in F$_1$ and F$_2$ generation, *Gradinarska i Lozarska nauka* 4: 44-51

Zohary, D. & Hopf, M. (2000). Domestication of plants in the Old World. Third ed. Oxford University Press, New York

Zoro, Bi.I., Maquet, A. & Baudoin, J-P. (2003). Population genetic structure of wild *Phaseolus lunatus* (Fabaceae), with special reference to population sizes, *American Journal of Botany* 90 (6): 897–904

Zoro, Bi.I., Maquet, A. & Baudoin, J.-P. (1999). Genetic control of iso-zymes in the gene pool *Phaseolus lunatus* L., *Biotechnology, Agronomy, Society and Environment* 13: 10-27

Zoro, Bi.I., Maquet, A., Degreef, J., Wathelet, B. & Baudoin, J.-P. (1998). Sample size for collecting seeds in germplasm conservation: case of the Lima bean *(Phaseolus lunatus* L.), *Theoretical and Applied Genetics* 97: 187-194

Zoro, Bi.I., Maquet, A., Wathelet, B. & Baudoin, J.-P. (1997). Genetic control of alcohol dehydrogenase, malate dehydrogenase, and phosphoglucomutase isozymes in Lima beans (*Phaseolus lunatus L.*), *Plant Breeding* 116 (2): 181-185

Zoro, Bi.I., Maquet, A. & Baudoin, J.-P. (1997). Spatial patterns of allozyme variants within three wild populations of *Phaseolus lunatus* L. from the central valley of Costa Rica, *Belgian Journal of Botany* 129: 149-155

Zoro, Bi.I. (1999). *Variabilité génétique des populations sauvages de Phaseolus lunatus L. dans la vallée centrale du Costa Rica et ses implications dans la mise au point d'une stratégie de conservation in situ.* Ph.D. thesis, Faculté Universitaire des Sciences Agronomiques, Gembloux, Belgique, pp. 194

Zoro, Bi.I. (1994). *Etude par les marqueurs enzymatiques, de la diversité et la structure génétique de populations sauvages du haricot de Lima (P. lunatus L.) originaires de la vallée centrale de Costa Rica*, C.E.A. en Sci. Agron., U.E.R. Phytotechnie des Regions Intertropicales, FUSA de Gembloux, Belgique, pp. 96

Part 2

Diversity of Microorganisms Involved in Plant Life

Use of Soil Microorganisms to Improve Plant Growth and Ecosystem Sustainability

Inés E. García de Salamone

Cátedra de Microbiología, Facultad de Agronomía, Universidad de Buenos Aires
Argentina

1. Introduction

The productive capacity of soils is directly associated with their organic matter content which is the main reservoir of organic carbon and main source of nutrients for plants (Grandy et al., 2006; Studdert et al. 1997; Sisti et al., 2004; Urquiaga et al., 2004). A sustained increase in the concentration of organic carbon is needed to establish a land management system where the degradation of soil organic matter occurs slowly and consequently its quality increases. Despite being underestimated in traditional farming, the cycle of organic carbon decomposition, governs various agronomic processes that occur on the soil and appear to affect productivity (Scow, 1997).

Microorganisms are an essential and large component of the living biomass of the soil (Whitman et al., 1998). They play a key role in the biogeochemical cycles and have a great potential in both agriculture use and environmental protection (Doran & Zeiss, 2000; Grandy et al., 2006). For that reason, the operation of any terrestrial ecosystem depends largely on soil microbial activity (Barea, 2004). It is known that the systems of crop management can modify the structure and biodiversity of soil microbial communities.

2. Microorganisms and sustainable agriculture

Soil microorganisms control the mineralization of organic carbon and thus regulate nutrient recycling in it. Crops are the basis of the trophic pyramid because they represent the source of energy and available nutrients to the heterotrophic microorganisms which contribute largely to the decomposition of plant residues. Soil microorganisms can act not only as a destination of the available nutrients to produce their microbial biomass but also as a source of them for the plant (Andrén et al., 1993). It is known that the quality of plant residues, as defined by the C/N ratio and lignin content, determines their rate of decomposition and the mineralization-immobilization cycle. The understanding of the impact of interactions between degradation and mineralization of carbon, nitrogen and phosphorus occurring in the soil within any particular crop sequence is a relevant goal of the ecological and agronomic research (Rorig et al., 2004a; Triplett & Dick, 2008). Changes in the rate of carbon and mineral nutrient movement in the soil as a result of the interactions between plants and other organisms, apparently, would involve changes in the structure and physiology of soil biotic communities (Rorig et al., 2004b; Garcia de Salamone et al., 2006a; Grandy et al., 2006). The prompt response of the

microbial processes and community structure when physical, chemical and biological changes occur is necessary and beneficial to maintain soil quality. Changes in the structure of microbial communities in perturbed systems are generally associated with emissions of greenhouse gases such as, CO_2, NO or N_2O and nitrate leaching loss (Jackson et al., 2003).

The use of no-tillage in agricultural systems can contribute positively to the sustainability of agro-ecosystems (Doran & Zeiss, 2000). This is due to the occurrence of changes in the pore space structure and water dynamics, increasing infiltration (macropores) and storage of water and in consequence reducing soil loss by erosion. In addition, there is a marked stratification of organic matter in the top layer which corresponds to a better structure and greater mesoporosity (Michelena et al., (2001). However, it is necessary to study in depth the dynamics of microbial diversity and activity in order to obtain biotic indicators to diagnose soil health (Garcia de Salamone et al., 2004, 2005). Based on nitrogen balance studies in the soil-plant system, it is known that the soybean crops leave a negative nitrogen balance so the excessive number of cycles of this crop in the sequence may contribute to the impoverishment of the soil (Zotarelli et al., 2002, Cheng et al., 2003). By contrast, cereal crops provide the most important contribution of plant residues in continuous cropping systems (Grandy et al., 2006). Several experiments were performed with the aim to find variables related to the activity of soil microorganisms that may show the influence of crop sequences on the biological properties of no-tillage managed soils. García de Salamone et al., (2006b) have shown that different crops have great influence on the mycorrhiza occurrence and root length (Figure 1). The amount of root biomass indicates that the introduction of maize or soybean in the crop sequence could have a great influence on nutrient cycling and both physiological and structural microbial diversity.
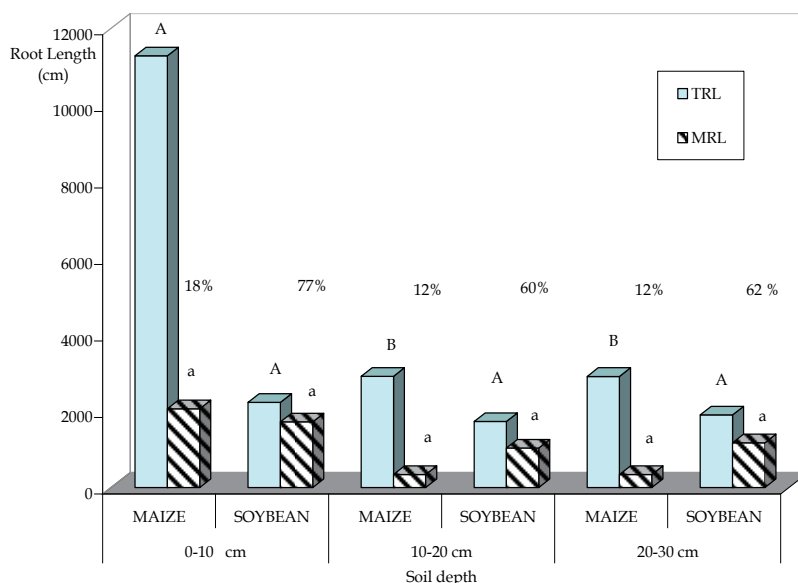
Fig. 1. Total root length (TRL) and Mycorrhiza root length (MRL) of maize and soybean in three soil depth layers at field conditions, (Adapted from García de Salamone et al., 2006b). Bars with the same letter are similar as indicated the Tukey´s test ($p<0.05$).The comparisons between soil depth layers of TRL and MRL are in uppercase and lowercase, respectively. Numbers on each bar indicated percentages of the total root length with mycorrhiza.

3. Numerical dominance and diversity of soil microorganisms

Whitman et al., (1998) could show the numerical superiority of soil microorganisms in different ecosystems. Their estimation of the number of soil microbial cells in the Earth is approximately 10^{30}. The estimated values are probably lower that the real numbers because in many cases the samples did not cover the entire depth of the soil profile. The samples were drawn on average from the first meter of depth and the authors applied methodologies to count cultivable microorganisms. It is known that on their best estimates, these methods cover less than 14% of the microbial community present in a soil sample. Despite these considerations, this figure is elevated when it is compared with the orders of magnitude of other living organisms such as humans (10^9), termites (10^{17}) or other magnitudes of the planet as ocean volume (10^{18} m^3) and Earth's mass (10^{27} g) or the mean distance from Sun to Earth (10^{11} m) and the radio of the universe (10^{26} m). Based on these data, soil microbial community can be named as "the invisible most." Its ecological significance is very strong as well in terms of macro reservoirs, because soil microorganisms contain 350-550 Pg of carbon, 85-130 Pg of nitrogen and 9-14 Pg of phosphorus; while plants only contain 560 Pg of carbon and 10 Pg, 1 Pg of nitrogen and phosphorus, respectively. The analysis of biogeochemical cycles (Schlesinger, 1991) indicates that the global carbon cycle is driven by CO_2 fixation in plant photosynthesis but the soil compartment is more relevant than the atmosphere due to the intensity of flows, amounts and changes. The global nitrogen cycle is driven entirely by microbial reactions (Paul & Clark, 1996). The largest nitrogen compartment is inert and is represented by the atmospheric N_2 (3.9 x 10^{21}g or 78% of gas composition). Microbial processes dominate the exchange between atmosphere and land-sea compartment of that element. Microbial diversity associated with each of the processes involved is clearly evident. The balance between nitrogen fixation and denitrification regulates inputs and outputs of soil nitrogen. Nitrification is the only process of conversion of ammonium to nitrate and is carried out mainly by autotrophic bacteria with fairly tight aeration and pH requirements. In summary, soil microorganisms are an enormous reservoir of genetic diversity, a significant reservoir of nutrients and a primary catalyst for the transformation of nutrients. A number of abiotic and biotic factors modify the genetic and physiological diversity of soil microbial communities (Maddonni et al., 2004).

Various estimates of the total number of bacterial species (Table 1) indicate that they are highly variable and the techniques used probably are the source of this variability. A simplistic analysis would seem to indicate that if there are between 1000 and 10000 species per gram of soil, it would not be necessary to worry about losing a certain amount of critical functions. But even if one accepts that a significant functional redundancy exists in most of the microbial communities (Naeem et al., 1995), we must take into account that: i. The occurrence of certain microbial processes is limited to very specific environments. For example, nitrogen-fixing trees in a rainforest and ii. There are interactions between diversity and invasion in a new habitat. For example, interactions between native microbes and introduced pathogens or biocontrol agents into a developing root. Anderson, (2003) established categories of organisms based on their sensitivity to changes in diversity. It can be found groups of microorganisms with different degrees of sensitivity for

Source	Methodology	Estimated Values species per gram of soil
Bergey's Manual of Systematic Bacteriology (Vol. 2)	Phenotypic Numerical Taxonomy	~3100 spp.
Vitek Database (Biomerieux)	Phenotypic Numerical Taxonomy	~1000 spp.
Torsvik et al. 1990 Torsvik et al. 2002	Hybridization DNA:DNA	~4000-10,000

Table 1. Various estimates of the total number of bacterial species

the same biogeochemical cycle. An example of this is related to the nitrifiers and N_2 fixing microorganisms which are located in the high and low sensitivity groups, respectively.

4. Functionality and microbial diversity

Problems related to the study of microbial diversity and functionality are linked to two main aspects that are characteristic of soil microbial communities. The first is "hyperdensity" since it can be found between 10^7 and 10^9 cells per gram of soil. It can be accepted that it is impossible to obtain a complete catalog of all types present in any particular soil. The second drawback is the "hyperdiversity". As it is pointed out, it is possible that there are 10^3 - 10^4 types per gram of soil. These figures indicated that there is no method that can encompass all of them. On the other hand, the existence of low-density microbial types complicates the possibility of studying by traditional methods of cultivation. In this sense, the frequent problem is related to perform manipulative experiments to study microbial diversity when only a small percentage of soil microorganisms can be recovered by cultivation. This is because when we tried to cultivate microorganisms from a soil sample, even when various culture media and physicochemical conditions of incubation are used, the numbers of organisms that we can count is very low and because of that they are called "the cultivated minority." However, the evaluation of a microbial subgroup should be sufficient to detect changes in the dynamics of the community, if there are interactions between the measured subgroup and the other members of the community.

Techniques for studying microbial diversity can be classified into those that analyze the physiological diversity and those that address structural diversity. In the latter case, it is necessary to apply molecular techniques and the information obtained is based on comparisons with patterns of gene library. In the first case, most of the work, instead of analyzing communities and species abundance matrices as in ecology, physiological profiles are analyzed based on the use of substrates or abundance-based structural profiles of biomarkers such as, acids nucleic membrane lipids. The analysis of the physiological profiles of microbial communities can be achieved through a rapid assessment of multiple physiological properties such as, the use of carbon sources based on various tools such as: i. The reduction of redox dye and measuring the change in absorbance microplate wells after a short period of incubation (Garland, 1996, Di Salvo, García de Salamone 2008), ii. O_2 consumption on a sensitive fluorometer microplate cell (Garland et al., 2003) and iii. The production of carbon dioxide in the bottle headspace at the top of each cell (CS multi-

induced respiration). On the other hand, the analysis of fatty acid esters with phospholipids union (Haack et al., 1997) can be based on: i. The detection of the presence of lipids associated with viable cells (prokaryotes and eukaryotes), ii. Its rapid transformation from polar to neutral lipids (diglycerides) and iii. Cell viability assays based on membrane integrity. This information can be used to evaluate the total biomass of viable cells. Moreover certain lipids are specific and they can be used as biomarkers to detect both the physiological state and composition of the microbial community present in the sample. The counts of cultivable microorganisms belonging to different physiological and taxonomic groups can provide information about the dynamics of the microbial communities along a crop sequence and/or a crop cycle.

5. Crops, sequences and microbial diversity

Regarding the impact of cereal crops on the structure and physiology of the microbial communities, the author's research group has focused on finding variables related to the activity of soil microorganisms that indicate the influence of crop sequence in different soils of the Argentinean Pampas which are managed under no-tillage system. In this sense, it was found that the amount of total bacteria, fungi and actinomycetes in the soil could be used as biological indicators to detect seasonal variations in the wheat-soybean sequence at two Argentinean locations: Ramirez (Entre Rios) and Bengolea (Córdoba) (Figure 2). Both locations are sited at the north and northeast of Buenos Aires (Rorig et al., 2004b, Garcia de Salamone et al., 2006a). These authors also showed that count data for the bacterial genus *Pseudomonas* had a significant variation along the crop sequence which would be indicating versatility. The genus includes pathogenic and non-pathogenic strains. The latter ones have shown great diversity because some of them can detoxify the environment, others are plant growth promoters and/or biological control organisms (Garcia de Salamone et al., 2005).

On the other hand, the potentially mineralizable nitrogen (PMN) has also shown seasonal variations in the wheat-soybean sequence in Typical Argiudolls soils (Zubillaga et al., 2007). It could be also considered that these oscillations are related to changes in the food chains of microorganisms in the soil. The amount of nitrifying microorganisms and chemical availability of nitrate showed significant differences between sampling times in the crop cycle of wheat but not between production environments within a particular field plot (D'Auria et al., 2011). The potential nitrification did not vary between environments and sampling times, indicating that field conditions determine the activity of this group of soil microorganisms (Table 2). Regression analysis between these two variables indicated that the amount of ammonium oxidizing microorganisms increases in anthesis while soil nitrate concentration is reduced between the stages of stem elongation and anthesis of the wheat canopy (Figure 3). The chemical and biological variables analyze in a comprehensive manner the dynamic of nitrogen mineralization in field plots and the use of available nitrogen by the wheat crop.

With respect to carbon and nitrogen mineralization, the counts of cellulolytic and nitrifier microbial communities vary with the conditions imposed by the crop sequence (Rorig et al., 2004b). These authors found that in two different situations of soil and climate, the cultivation of wheat favors the activity of these two physiological groups of microorganisms in contrast to soybean which reduces their numbers. In addition, the

amounts of nitrifiers are significantly enhanced by the presence of plants and reduced in absence of them. The set of microbiological variables determined at different times of the wheat-soybean sequence showed that the conditions imposed cause measurable changes in soil microbiota. Although it is known that these techniques quantify a very small sample of the microbial community, detection of changes within culturable communities can be accepted if functional alterations of them can be estimated (Garcia de Salamone et al. 2006a). Previously, Garcia de Salamone et al., (2004) have observed differences in the physiological profiles of carbon use in the soil microbial communities when several levels of soil compaction were applied to simulate increasing intensities of machinery use (Figure 4). These different patterns of carbon use are indicating that the functional diversity of the microbial communities is changing because of the compaction treatments. The observed changes can be temporary or permanent. The management applied to the soil could finally define each situation. Using the same technique, as described below, it was possible to establish that inoculation with certain soil bacteria also alter temporarily microbial communities associated with wheat (Naiman et al., 2009) and rice (Garcia de Salamone et al., 2010) crops.

	Potential Mineralizable Nitrogen (mg kg^{-1})	Potential Nitrification (NO$_3$ g^{-1} dry soil)	Most Probable Number of Nitrifiers Log g^{-1} dry soil	Soil Nitrate concentration (mg kg^{-1})
Stem Elongation				
H	9.3a	170.7a	5.3a	22.0 bc
M	11.3a	153.1a	5.0a	28.1 c
Anthesis				
H	9.7a	744.3 b	7.7 b	10.0ab
M	16.4a	353.4ab	7.6 b	7.8a

Table 2. Biological and chemical variables of the nitrogen cycle estimated at two different stages of the wheat crop grown at field conditions on two historic levels of crop production. H: High, M: Medium, respectively. For each taxonomic group of microorganisms, means followed by the same letters are similar as indicated by the Tukey´s test (p=0.05).

A significant number of soil microorganisms are involved in mineralization of organic phosphorus. This must be converted to inorganic phosphorus to be used by plants. The mineralization of organic phosphorus is mostly carried out by phosphatase type enzymes, with significant levels of soil microbial phosphatases (García, 1999). Fixation and precipitation of phosphorus in soils is a phenomenon highly dependent on both pH and soil type. An important mechanism of solubilization of mineral phosphates is through the production and release of organic acids synthesized by soil microorganisms (Paul & Clark 1996). This causes an acidification of the medium and inorganic phosphorus release. There would be other ways to solubilize mineral phosphates, such as the production of inorganic acids and chelating substances. The ability of microorganisms to solubilize mineral phosphates depend on the nitrogen and/or carbon sources which are environmentally available. Related to the use of phosphorus by plants is of paramount importance to study arbuscular mycorrhiza. These symbiotic associations increase the area explored by the roots by promoting a better use of available resources (Chiocchio et al., 2002) and soil aggregation

Log cfu/ g dry soil
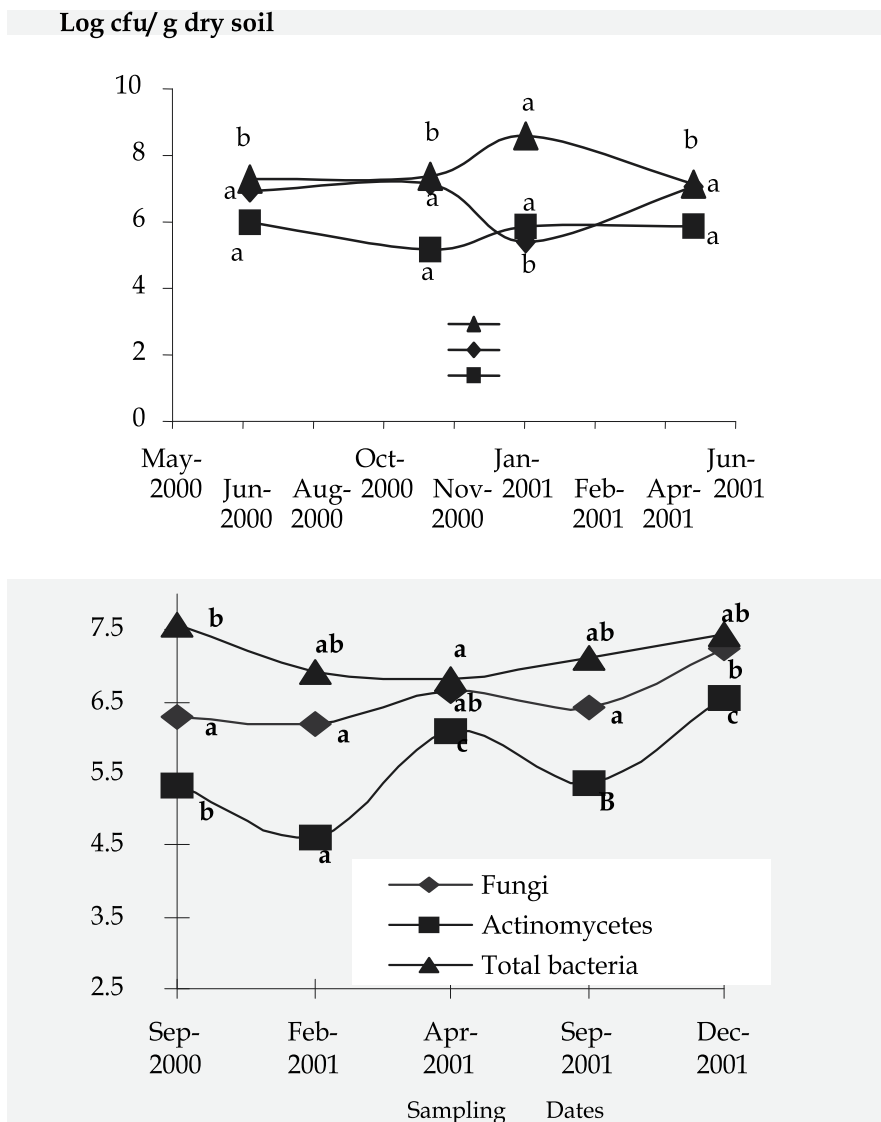
Fig. 2. Dynamics of fungi, actinomycetes and total bacteria during wheat/soybean sequence at two Argentinean locations. Top part: Typic Argiudoll soil in Ramírez (Entre Ríos). Bottom part: Entic Haplustoll soil in Bengolea (Córdoba). Means of each variable followed by the same letters are similar as indicated by the Tukey´s test (*p*=0.05).

(Rillig et al., 2002). The magnitude of natural mycorrhiza is considered a bioindicator of soil quality that responds to management practices (Garcia de Salamone et al., 2006a; Schalamuk et al., 2003) and can be promoted with the application of certain microorganisms (Naiman et al., 2008). The arbuscular mycorrhiza fungi are biotrophic symbionts and they associate with about 80% of the plant species. It has been observed in agroecosystems where the monocroping is applied, that the conventional tillage and the applications of both fertilizers and high soluble pesticides reduced more than 50% the

occurrence of natural mycorrhiza in comparison to the respective non-disturbed systems (Sieverding, 1991). Thus, it is also essential to consider that the maintenance of natural mycorrhiza for each crop in a given edaphoclimatic situation is highly relevant to maintain the sustainability (García de Salamone et al., 2006b; Schalamuk et al., 2006). In this sense,
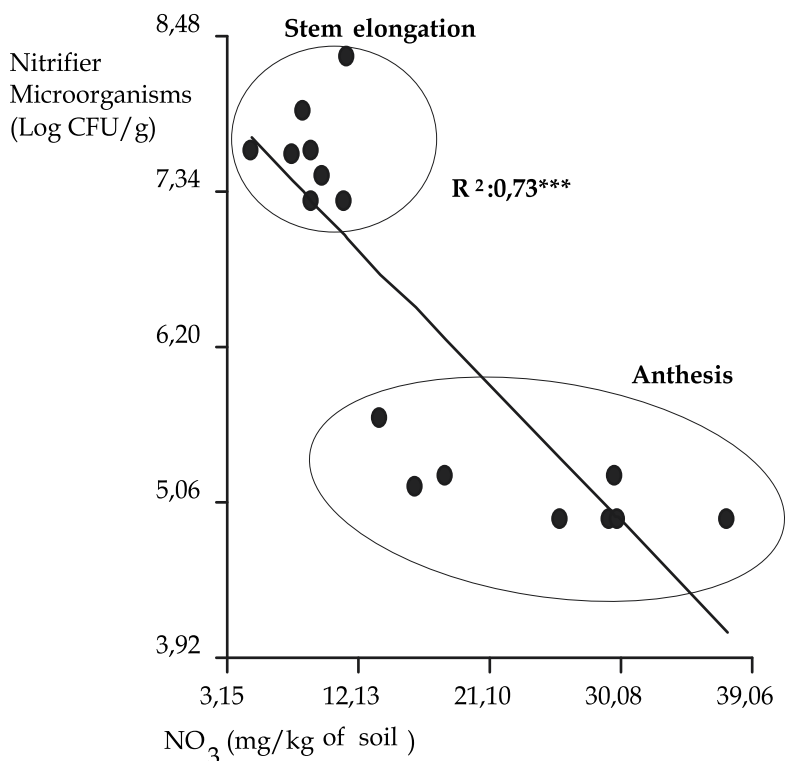
Fig. 3. Regression analysis of soil nitrate concentration and amount of cultivable nitrifiers at two stage of a wheat crop grown at field conditions on a Typic Argiudoll soil in Junin (Buenos Aires.

several studies were performed to evaluate soybean mycorrhiza associations under controlled conditions with different types and doses of pesticides (Martinez et al., 1998; Venedekian et al., 1999). Besides, trials have also been conducted with corn, wheat and soybeans (Garcia de Salamone et al., 2006b), which have shown that natural mycorrhiza is highly dependent on the historical and current management of the soil. Thus, both predecessor crop and phosphorus fertilizer addition exert significant effects on it. For these reasons, it is important to know the microbial dynamics of cereal crops included in the crop

sequence (Chiocchio et al., 2000; Menéndez et al., 1999; Studdert et al., 1997) and the effects of the introduction of microorganisms in the soil-plant system to provide a better understanding of these interactions which are known to help to agroecosystem sustainability (Naiman et al., 2009). It was observed that inoculation of wheat with a phosphorus solubilizer bacterium such as, *Pseudomonas fluorescens* can modify the mycorrhiza percentage and certain fungi structures during the growth cycle (Table 3). This PGPR is changing the population dynamic of the native mycorrhiza associated with the wheat crop and this could probably change the diversity of these microbial communities.

6. From soil microorganisms to microbial inoculants

Several soil microorganisms have natural properties which make them suitable to be used as biotechnological tools. However, any particular soil microorganism should show at least one beneficial characteristic in order to be considered for commercial propagation and consequent field application. There are different alternatives or markets for soil microbial products. In general, they can be included in some of the following types: nitrogen-fixing microorganisms, pest control agents, plant growth promoting rhizobacteria, usually named PGPR, and microorganisms for bioremediation. Any particular isolate should move through several steps before becoming a microbial inoculant. First of all, the development of a new soil microbial product for agriculture or bioremediation begins with the discovery of a useful naturally occurring organism. This can only be obtained establishing specific screening objectives which should be tested
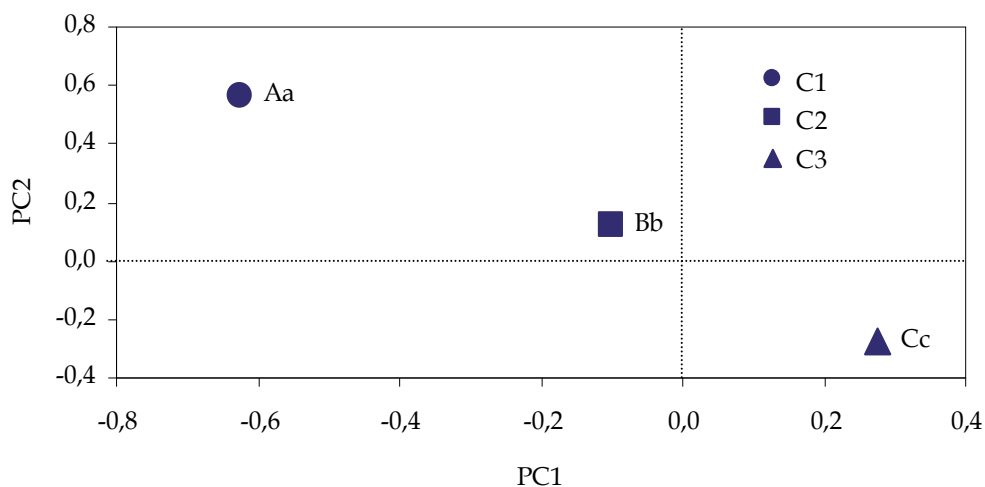
Fig. 4. Principal component analysis of physiological profiles of carbon use of the microbial communities of an Entic Haplustol soil, Bengolea (Córdoba) when three levels of compaction were applied to simulate increasing intensities of machinery use.
Values are means of each compaction treatments, C1: control or field conditions, C2: 100 Kpa and C3: 400 Kpa. C2 and C3 were applied on the soil contained in cylinders with a diameter of 7.1 cm and a height of 6.1 cm. Data of absorbance of 72 h reading were used. PC1 and PC2 scores with the same uppercase and lowercases, respectively, indicate that they were not significantly different as determined by Tukey test at the rejection level of $P<0.05$.

Inoculation treatments	Mycorrhiza percentage at two stages of growth	
	Tillering	*Grain filling*
Control	34 a	36 a
Pseudomonas fluorescens	48 b	51 b
	Arbuscules (%)	
Control	18 a	20 a
Psedomonas fluorescens	25 b	24 b

Table 3. Impact of the inoculation with *Pseudomonas fluorescens* on mycorrhiza and arbuscules percentages found in wheat roots at two stages of growth. Means of each growth stage followed by the same letter are similar as indicated the Tukey's test ($p<0.05$), (Adapted from Naiman et al., 2008).

in every isolate. A sequence of laboratory, greenhouse and field experiments are usually performed in order to reduce the initial huge number of isolates to a small group of selected strains which have a combination of characteristics with commercial potential. Screening programs are usually carried out by companies or university laboratories. Most of the promising soil isolates have been obtained in university laboratories and then they were licensed to private companies to develop commercial formulations. Microbial inoculants for agriculture should be sold in a way that the farmers can easily apply. Liquid inoculants are the most accepted because they can be easily applied over the seeds before sowing. In this chapter, special emphasis is dedicated to a particular group of soil microorganisms. These are the PGPR. They are associated with plants of agricultural interest and in many cases direct beneficial effects on growth and both nitrogen and phosphorous nutrition are observed. Because of that, they constitute an ecological and economical alternative to increase food production (Bashan et al., 2004; Caballero Mellado, 2004; Díaz-Zorita & Canigia Fernandez, 2008; Ferraris & Courerot, 2004; Naiman et al., 2009; Reed & Glick, 2004). However, it should be noted that the impact of interactions between crops and soil microbes on the circulation of nutrients can be considerable. In this respect, plant biomass increases due to inoculation with PGPR can affect nutrient mineralization, solubilization of phosphorus, biological nitrogen fixation (BNF) and other biological processes associated with nutrient dynamics in the productive system. Therefore, it is required a detailed knowledge of the interrelationships between microorganisms added to the system (inoculants) and the native ones in both soil and associated plants. However, we know very little about the influence of management practices, such as PGPR inoculation and fertilization of cereal crops on the structure and functioning of the microbial communities. According to Wardle (2002) the generation of evidence-based information capable of connecting processes occurring in the air portion of the system with processes that take

place in the underground portion is one of the current challenges facing agroecological research.

There is abundant evidence in the literature indicating that the use of PGPR can have a significant role in agroecosystem sustainability (Antoun & Prévost 2006). PGPR microbial inoculants represent an emerging technology designed to improve the productivity of agricultural systems in the long run. However, the current use of microorganisms in agriculture remains at a low level despite the significant investment and scientific work (Catroux, 2007). They can be seen as a technology aligned with principles of sustainable agriculture, as opposed to the increasing use of pesticides and fertilizers. Several microorganisms are used in the normal agricultural practice, and others have potential for future use (Cassan & García de Salamone 2008; Maddonni et al., 2004b]. Most of them have the ability to colonize and establish an ongoing relationship with plants producing increases in biomass, root growth and commercial yield (Glick, 1995). At this regard, in the early '80s, the association between the diazotrophic bacteria *Azospirillum* and cereal crops was already considered a phenomenon of considerable scientific and economic value. Experimentation has also shown that a large number of factors influence the inoculation response in field conditions. Among PGPR, bacteria of the *Azospirillum* genus are the most used. The potential of this bacterium-plant association to increase crop production was first reported several years ago (Baldani et al., 1987). Positive impacts on plant growth through several mechanisms include enhancement of root development, production of growth regulators and nitrogen fixation (García de Salamone et al., 1996; Okon, 1994).

The content of nitrogen, phosphorus, potassium and various micronutrients is higher in plants inoculated with *Azospirillum* (Caballero Mellado et al., 1992; García de Salamone et al., 1996). Significant effects have been observed in wheat (Boddey et al., 1986; Caballero Mellado et al., 1992; Naiman et al., 2009), maize (Caballero Mellado, 2004; García de Salamone & Döbereiner, 1996), soybean (Bashan et al., 1990) and rice (Baldani & Baldani 2005; García de Salamone et al., 2010) among other species including more than a hundred crops and environmentally important plant species (Bashan et al., 2004). A recent publication compiled data from different countries showing the state-of-the-art in the *Azospirillum* inoculation responses (Cassan & García de Salamone, 2008).

The sustainability and profitability analysis requires a detailed knowledge of the interrelationships that exist between microorganisms commonly added as inoculants and those within the natural system (García de Salamone et al., 2006b; Naiman et al., 2009). Several authors have reported the impact of *Azospirillum* in root development with the consequent advantage on the absorption of water and nutrients (Bashan & Levanony 1990; García de Salamone et al., 1996; Okon, 1994). However, most of this information corresponds to experiments performed under controlled conditions. Thus, it is necessary to analyze the effects of commercial and new experimental inoculants in order to obtain better crops and a better use of the environmental resources. To date, several PGPB including *Pseudomonas* strains have been characterized as phosphorus solubilizers with the ability to produce organic acids (such as oxalic acid, fumaric acid and citric acid) and phosphatases that facilitate the solubilization of phosphorus and other nutrients (de Freitas et al., 1997; Rodriguez et al., 2006). In addition, it has been demonstrated that some strains can produce and supply several cytokinins in the rhizosphere of wheat and radish (García de Salamone et al., 2001, 2005). However, the largest amount of information about the activity of

Pseudomonas strains is associated with the indirect effects, through the control of pathogenic microorganisms. This can reduce the incidence of diseases through a number of mechanisms, including increased competitive ability by available nutrients, production of antibiotics, siderophores and induction of systemic resistance (Dowling & O´Gara, 1994; Kloepper, 1993).

The information analyzed in this chapter will show that the inoculation with PGPR such as *Azospirillum brasilense* and *Pseudomonas fluorescens* can modify cultivable microbial rhizosphere communities and plant growth at field conditions. The inoculation with these two rhizosphere microorganisms contributes to the emergence, development and grain production of crops such as wheat (Table 4), rice (Table 5), and maize (Table 6) (Baldani et al., 2008; Garcia de Salamone et al., 2010; Naiman et al., 2009). There is some evidence in controlled conditions for some crops that the application of certain PGPR does not change the structure of microbial communities (Herschkovitz et al., 2005). It was observed that there are some indications of the effect of PGPR inoculation have certain effects on the Shannon-Weaver diversity index of the microbial communities associated to the rhizosphere of wheat (Table 7). However the changes during the crop cycle are stronger that those produce by the PGPR inoculation. However, it is still necessary to understand better the microbial interactions that occur in soil-plant system under field conditions to assess the overall impact of this inoculation technology on the agroecosystem and to achieve the maximum efficiency (García de Salamone & Monzón de Asconegui 2008).

The average of two field experiments has showed that several commercial inoculants have different impact on grain yield of wheat crop in the west semiarid region of the Buenos Aires province (Figure 5). Although, all these commercial inoculants had the same strain of *Azospirillum brasilense*, these observations indicate that the interaction between the crop

Soil Type [a]	Number of plant cultivars	Inoculants		Range of Yield Increase due to Inoculation [b]
		Number	Type	(%)
Humic Hapludoll	2	3	Experimental	18-34
Humic Hapludoll	2	3	Experimental	0-2
Typic Arguidoll	2	3	Experimental	0-5
Typic Arguidoll	1	2	Experimental	7-9
Humic Hapludoll	1	3	Experimental	18-35
Entic Hapludoll	1	4	Commercial	2-32
Typic Arguidoll	1	2	Commercial	1-8
Typic Arguidoll	6	2	Experimental and commercial	9-30
Typic Arguidoll	1	2	Commercial	1-24
Aquatic Arguidoll	1	3	Experimental	0-10

Table 4. Characteristics of wheat field inoculation experiments and PGPR inoculation responses.
[a]All soil are located at the Province of Buenos Aires, Argentina.
[b]Yield increase with respect to the control plants without inoculation

and the microbial inoculant do not allow to confirm an average response with a particular PGPR strain when different formulations are used. Certain unknown ingredients other than the included PGPR strain can change the quality of the inoculants and consequently their efficiency. Standardization of quality control methodologies is still required.

Soil Type [a]	Number of plant genotypes	Inoculants		Range of Yield Increase due to Inoculation [b]
		Number	Type	(%)
Vertic Arguidoll	1	2	Experimental	4-8
Vertic Arguidoll	1	2	Experimental	17 – 20
Vertic Arguidoll	3	1	Commercial	- 7 – 20

Table 5. Characteristics of rice field inoculation experiments and PGPR inoculation response.
[a]All soil are located at the Province of Entre Rios, Argentina.
[b]Yield increase with respect to the control plants without inoculation

The combined analysis of the experiments described in Table 4 shows that inoculation with *Azospirillum* always increases the crop yield of wheat. In addition, the partition spikes and total biomass production can be increased through the practice of inoculation with this PGPR. An example of this kind of inoculation response is reported by Naiman et al (2009) who have observed that, *Azospirillum* inoculation increased aerial biomass by 12%, root biomass by 40% and grain yield by 16% (Table 8). Grain yield increases represent important earnings for the farmer and the increment of both aerial and root biomass may help to obtain a greater sustainability of the agroecosystems because the plant residues added to the system could help to maintain soil organic matter.

Rice is the third largest global food crop. Because of that its cultivation should be part of sustainable agriculture programs. Three experiments performed during the crop seasons 2006-2007, 2008-2009 and 2009-2010 in the rice belt in the province of Entre Rios, NE region of Argentina allowed the evaluation of the response of PGPR inoculation treatments (Table 5). In the first two cases, we studied the effect of *Azospirillum brasilense* on biomass production and grain yield (Garcia de Salamone et al., 2010, Gatica et al., 2009). The control biomass was 7256 and 15183 kg ha^{-1} at tillering and grain filling, respectively. Treatment with *A. brasilense* significantly increased these values by 15 and 35% for tillering and 28 and 50% for grain filling. The control yielded 8370 kg ha^{-1} and inoculation increased that value by 7.5%. In these studies we analyzed the rhizosphere microflora with potential to fix nitrogen using various methodologies. Also we observed increases in the most probable number (MPN) of cultivable rhizosphere diazotrophs when the seeds were inoculated with a mix of two *Azospirillum brasilense* strains (Table 9). However this MPN decreased between tillering and grain-filling stages. In the same experimental work it was observed that the ability to use four carbon sources by the microbial communities associated with rice rhizosphere was different between treatments at the grain-filling stage (Table 9). This approach utilizes the absorbance values on several carbon sources and the carbon level physiological profiles of the microbial communities are obtained through principal components analysis. In this case, malic acid, mannitol, ammonium oxalate and maltose

Soil Type [a]	Number of plant genotypes	Inoculants		Range of Yield Increase due to Inoculation [b]
		Number	Type	(%)
Typic Arguidoll	3	3	Experimental	0 – 4
Aeric Arguidoll	15	8	Experimental	-32 – 94
Aeric Arguidoll	7	8	Experimental	-39 – 42
Humic Hapludoll	4	6	Experimental	-6 – 77
Aeric Arguidoll	6	1	Experimental	2 – 5
Vertic Arguidoll	1	1	Experimental	0 – 5
Humic Arguidoll	1	1	Experimental	22 – 46
Entic Arguidoll	1	1	Experimental	3 – 13
Typic Arguidoll	3	2	Commercial	2 – 7

Table 6. Characteristics of maize field inoculation experiments and PGPR inoculation response
[a]All soil are located at the Province of Buenos Aires, Argentina.
[b]Yield increase with respect to the control plants without inoculation

Treatments	Shanon Weaver Diversity index [1]	
	Tillering	Grain filling
Without N-addition		
Control	3.0	2.9
CI 1	2.9	2.8
CI 2	3.0	2.9
CI 3	3.0	2.9
With N addition [2]		
Control	3.0	2.9
CI 1	2.9	2.9
CI 2	3.0	2.9
CI 3	3.0	2.9
P	0.56	0.29
Growth Stage Mean	2.9 a	2.8 b

Table 7. Effect of three commercial inoculants of PGPB and two levels of nitrogen (N) addition on the Shanon Weaver diversity index of the microbial communities during the wheat growth cycle.
Different letters indicate significant differences with Tukey's test ($p<0.05$). CI: Commercial Inoculant.
[1]Calculated from absorbance values obtained from a set of 23 carbon sources as described Naiman et al., (2009) using the methodology described by Gómez et al., (2004).
[2]45 kg ha^{-1} of N applied as urea at sowing.

showed the highest Pearson correlation coefficients and their averages are in Table 9. Due to these carbon sources are usually frequent in the rhizosphere, the changes observed in their utilization can be related to changes in the physiological diversity of rice associated microbial communities after PGPR inoculation. In the case of the last crop season (Table 5), it was analyzed the effect of combined inoculation of two PGPR. These were *Azospirillum brasilense* and *Pseudomonas fluorescens*. In this experiment, it was analyzed the response to inoculation of three rice genotypes which were coming from different breeding programs. The response rates were significantly different among them (Table 5). This demonstrates the significant interaction between plant genotypes and applied soil bacteria. The analysis of both physiological and structural diversity of microbial communities indicated that they were also associated with the genotypes of the plant and the inoculation did not change those (Garcia de Salamone et al 2011 in press). The results of this work are indicating the high potential of the inoculation practice with PGPR for this crop and the possibility of increasing its efficiency through breeding programs that consider plant-bacteria association during the process. At this regard, Garcia de Salamone & Dobereiner (1996), proposed that the successful use of these plant-bacteria associations in agricultural scale can be achieved through plant breeding programs which consider "the ideotype approach" proposed by Donald (1968). Thus, it can help to obtain better combinations of both partners for any particular environment. Based on the results of several trials, it is possible to conclude about the relevance of the interaction between both maize genotypes and PGPR strains (Table 6). Garcia de Salamone et al., (1993) analyzed the available information about the response to *Azospirillum* inoculation in
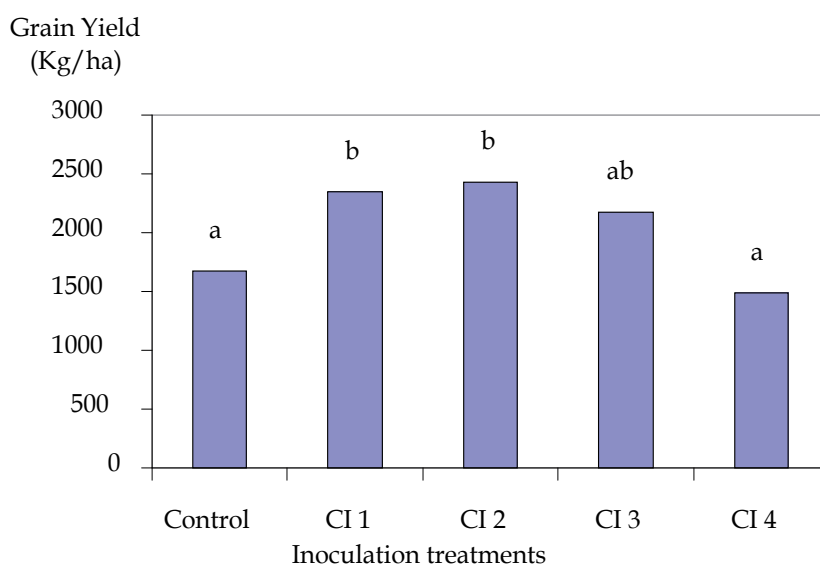

Fig. 5. Impact of *Azospirillum brasilense* inoculation treatments on grain yield of wheat grown at semiarid field conditions on an Entisol soil located in Saliquelo, West of the Buenos Aires, Argentina. CI: Commercial Inoculant.
Different letters indicate significant differences with Tukey's test ($p<0.05$).

association with maize and they observed a great consistency of the results in different crop seasons and environmental conditions. The analysis of consistency revealed significant grain yield increases ($p< 0.05$) due to the inoculation with *Azospirillum* for all the evaluated experiments (Garcia de Salamone & Monzón de Asconegui 2008).

Inoculation Treatments	Grain yield	Biomass				
		Aerial			Roots	
	PM	T [1]	GF [2]	PM [3]	T	GF
		(Kg ha-1)				
Control	2600a	3129a	4911a	8248a	12501a	22969a
CI-1	3400b	3130[a]	5153a	8799a	23992b	25622a
CI-2	2900ab	3189a	5765a	9846a	16737ab	24624a
CI-3	3200b	3168a	5298a	8868a	18367ab	23467a
p	0.02	0,92	0,48	0,41	0,005	0,83

Table 8. Effect of several commercial inoculants of PGPB on grain yield, aerial and root biomass production during the wheat growth cycle.
Different letters indicate significant differences with Tukey's test ($p<0.05$). CI: Commercial Inoculant.
Tillering, GF: Grain-filling, PM: Physiological maturity.

Treatments	Diazotrophic microearophilic bacteria (Log MPN g-1 of root)		Average values of absorbance at 590 nm [1]
	Tillering [2]	Grain filling [3]	
Control	8.5 a	6.2 a	0.184 b
Experimental Inoculants (EI) [4]			
EI 1	9.3 b	6.4 ab	0.137 a
EI 2	8.9 ab	6.7 b	0.128 a

Table 9. Most Probable Number (MPN) of microaerophilic nitrogen-fixing bacteria and average of absorbance values shown by cultivable microorganisms in the rhizosphere of rice plants inoculated at sowing with *Azospirillum brasilense* (Compiled data from García de Salamone et al., 2010).
[1] Absorbance values for four carbon sources showing the highest Pearson correlation coefficients in the principal components analysis of carbon level physiological profiles.
[2] 35 and [3] 117 days after sowing. [4] Experimental inoculants 1 and 2 contained the same two *Azospirillum brasilense* strains but they were differently applied on the rice seeds.

7. Biological nitrogen fixation (BNF)

Biological nitrogen fixation becomes relevant and it can be incorporated by certain cereal-PGPB associations as a relevant source of nitrogen for the agroecosystem (Garcia de Salamone et al., 1996; Urquiaga et al., 2004). It has been shown that inoculation with certain bacteria-plant combinations, such as *Azospirillum*-maize can provide nitrogen from BNF at levels equivalent to 100 kg ha-1 of this element (Garcia de Salamone et al., 1996). The contribution of nitrogen could help to the sustainability of the agroecosystem because it can

improve the quality of the plant residues and enable a reduction in the use of nitrogen fertilizers (Figure 6) Moreover, it was observed that inoculation with *A. brasilense* modify the $^{15}N/^{14}N$ ratio of both rice and wheat plants in the grain filling stage (Garcia de Salamone et al., 2009). In the case of rice, it could be concluded that the amount of nitrogen derived from BNF increased when inoculation with certain *Azospirillum* strains was applied (Table 10). Three weeds plants growing contemporarily with the wheat crop were used as non-fixing controls to estimate the occurrence of nitrogen fixation through the ^{15}N dilution technique (Boddey, 1987). Wheat plants showed significant differences among genotypes in the values of $^{15}N/^{14}N$ ratios (Table 11). The great variability is evident in the $^{15}N/^{14}N$ relationships of uninoculated wheat plants and it would be indicating the possibility of using these values to calculate the amount of nitrogen derived from BNF in inoculated wheat plants. The mean $^{15}N/^{14}N$ ratio of the three weed plants in the experiment of wheat was 8.64 and it was used to estimate the BNF rates. They showed variability between the genotypes of the plant as reported earlier with maize genotypes (Garcia de Salamone et al., 1996) in association with PGPR. The relationship between the $^{15}N/^{14}N$ ratios in both control and inoculated wheat plants has shown that, the average percentage range of nitrogen derived from BNF was 13-55%.

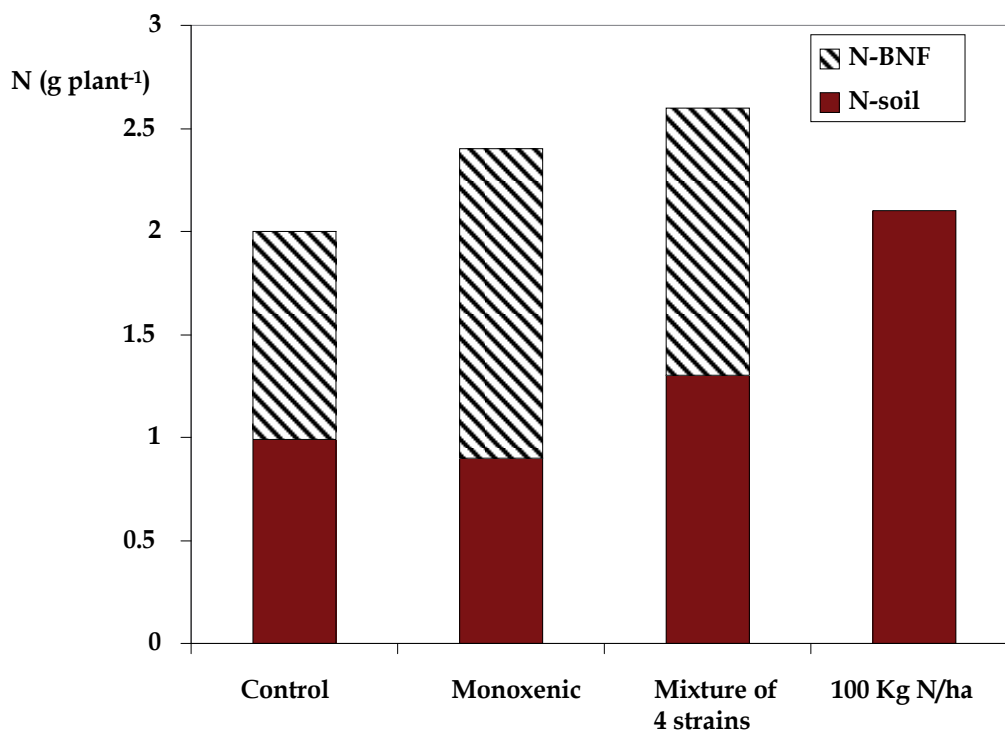Fig. 6. Nitrogen content derived from soil and biological fixation in aerial plant tissues of maize plants fertilized with nitrogen and inoculated with *Azospirillum* strains. Data compiled from García de Salamone, et al., (1996).

These data indicate that inoculation with *Azospirillum* can produce changes in plant physiology which can be related to the occurrence of BNF. To increase accuracy, it is

necessary to obtain better estimates of soil $^{15}N/^{14}N$ values (Alves et al., 2008). The most probable number counts of ammonium oxidizing microorganisms increased when the seeds of rice and wheat crops were inoculated with *Azospirillum brasilense*. However, potentially mineralizable nitrogen (PMN) showed different responses to inoculation for these two crops (D´Auria et al., 2011). PMN values for the rice plants were significantly lower in the soil associated with the inoculated plants. However, the soil PMN associated with uninoculated wheat plants was 5.3 mg $N-NH_4$ kg^{-1} soil, which was significantly ($p<0.05$) lower to 10.3 mg $N-NH_4$ kg^{-1} soil for the soil observed associated with the inoculated plants. From these data, we can infer that changes of plant physiology of rice and wheat plants produced by *Azospirillum* inoculation could favor BNF of these cereal crops and changed activity and numbers of the microorganisms associated to N-cycling.

Inoculation Treatments (1)	Delta $^{15}N/^{14}N$				% N-BNF (3)	PMN $N-NH_4$ (mg Kg^{-1} dry soil)	Log MPN Nitrifiers
	Rice plants	Weed Plants (2)					
		E.p	*Cyp*	*P.v*			
C	16.31	16.9	15.48	18.52	3.9	8.4	5,9
ExIn 1	11.68	13.16	1052	16.25	12.2	3.1	6,0
ExIn 2	11.31	12.22	10.61	16.64	14.1	Wd	6,3

Table 10. Values of Delta $^{15}N/^{14}N$ ratios for rice and associated weed plants, percentages of nitrogen derived from biological fixation (N-BNF), potential mineralizable nitrogen (PMN) and most probable number (MPN) of nitrifiers in the rhizosphere of field-grown rice at grain-filling stage.

$^{15}N/^{14}N$ ratio of weed plants in the experiment of wheat was 8.64. Estimates of BNF rates showed variability between (1) C: control without inoculation; ExIn 1 and 2: Inoculation treatments including the same mixture of two *Azospirillum brasilense* strains but using two different application techniques. (2) Weed plants sampled at the same moment of the crop: *E.p:E. polystachia, Cyp:Cyperus sp., P.v:P. viscose.* (3) % N-FBN: Percentage of N derived from BNF calculated using for calculations the average $^{15}N/^{14}N$ ratio of control plants for each treatment as the value of control. wd: without data.

8. Summary and outlook

The results included in this review show the ability of certain PGPR to modify the ecophysiology of crops such as wheat, corn and rice in field conditions. However, the information available would indicate that the plant-strain-environment interaction is relevant to the results of inoculation that can be obtained (Garcia de Salamone, Monzón de Asconegui 2008). It was observed that the inoculation response is variable and the introduced bacteria can colonize and remain in the rhizosphere. Yield increases and biomass production have ecological relevance and they should be studied from the point of view of microbial ecology. Besides, these aspects should be studied using isolates with ability to fix nitrogen in association with the plant. This could increase the response level and improve the efficiency of use of the available resources. The introduction of PGPR could cause changes in the microbial activity in the rhizosphere and they have to be studied in detail

Genotype of wheat	Inoculation Treatments (1)	Delta $^{15}N/^{14}N$ (2)	% N-FBN (3)	% N-FBN (4)
Buck 75	C	10.08	-16.67	-
	ExIn	8.70	-0.69	13.66
Buck Chacarero	C	7.05	18.34	29.41
	ExIn	8.18	5.38	19.05
Buck Sureño	C	7.72	10.65	22.98
	ExIn	6.06	29.86	39.58
Klein Zorro	C	4.16	51.85	58.57
	ExIn	5.44	37.03	45.99
Klein Castor	C	4.48	48.21	55.27
	ExIn	5.74	33.56	42.83
Austaliano	C	4.97	42.53	50.55
	ExIn	5.89	31.89	41.35

Table 11. Impact of *Azospirillum brasilense* inoculation on values of Delta $^{15}N/^{14}N$ ratios and percentages of nitrogen derived from biological fixation for six genotypes of wheat at grain-filling stage grown at field conditions.
(1) C: control without inoculation, ExIn: Experimental Inoculant formulation of *Azospirillum brasilense*. **(2)** Delta $^{15}N/^{14}N$ of test plants, average of four replicates. **%N-FBN**: Percentage of N derived from BNF calculated using as the value of control plants in **[F1]**: **(3)** average $^{15}N/^{14}N$ ratio of weeds; **(4)**: $^{15}N/^{14}N$ ratio of Buck 75.

(Garcia de Salamone & Cassan 2010). It is known that microbial diversity can be used as an index of soil quality and the management conditions can modify it (Garcia de Salamone et al., 2006, 2010). In connection with this and because the practice of PGPR inoculation is being used by a growing number of farmers in various agricultural areas of the world, it is necessary to provide knowledge about the microbial ecology of the rhizosphere of crops under field conditions when PGPR are applied repeatedly crop after crop.

The results compiled in this review are also consistent and they can be extended with those reviewed by Reed & Glick (2004), Bashan et al., (2004), Cassan & Garcia de Salamone, (2008), Garcia de Salamone & Cassan (2010). All available information indicates that inoculation with PGPR should be favored. However, the variability in both PGPR and plant capabilities should be adjusted and enhanced in order to include alternative mechanisms such as BNF, production of plant growth regulators, biological control of pathogens, among others in order to improve response rates to inoculation in the field.

In summary, the direct effects of the use of PGPR in cereal crops exert a significant role in fundamental processes of the ecosystem such as carbon and nitrogen recycling. The approach to these problems with field experiments provides a set of evidence that has involved the author's interaction in several research teams. It is expected that this review results in a contribution to connect processes occurring in both aerial and soil portions of the ecosystem.

9. Acknowledgment

10. References

Alves, B.J.R., Oliveira, O.C., Boddey, R.M. & Urquiaga, S. 2008. Métodos isotópicos. In: Gabriel de A. Santos; Leandro S. da Silva; Luciano P. Canellas; Flavio A.O. Camargo. (Org.). *Fundamentos da matéria orgânica dio solo: ecossistemas tropicais e subtropicais*. pp. (229-242). 2 ed. : Editora Metrópole, Porto Alegre.

Anderson, T.H. 2003. Microbial eco-physiological indicators to asses soil quality. *Agriculture, Ecosystem and Environment* 98: 285-293

Andrén, O., Hansson, A., Végh, C. & Barley, K. 1993. Nutrient up-take, root growth and depth distribution in two soil types in a rhizotron with vertical and horizontal minirhizotrons. *Swed. J. Agric. Res.,* 23: 115–126.

Antoun, H. & Prevost, D. 2006. Ecology of plant growth promoting rhizobacteria. In: Z.A. Siddiqui, Editor, *PGPR: Biocontrol and Biofertilization*, pp. (1–38) Springer, Dordrecht,.

Baldani, J.I. & Baldani, V.L. 2005. History on the biological nitrogen fixation research in graminaceous plants: Special emphasis on the Brazilian experience, *Anais Academia Brasileira Ciencias*, 77, 3: 549-579.

Baldani, V.L., Baldani, J.I. & Döbereiner, J. 1987. Inoculation of field-grown wheat (*Triticum aestivum*) whith Azospirillum *spp*. In Brazil. *Biol & Fertil Soil* 4:37-40.

Baldani, D.V.L.; Silva Ferreira, J.; dos Santos Teixeira, K.R.; Baldani, J.I. & Massena Reis, V. 2008. Inoculants base don nitrogen-fixing bactéria *Azospirillum* spp. and their application in tropical agriculture. En: *Azospirillum sp.: cell physiology, plant interactions and agronomic research in Argentina.* Cassan, F D; García de Salamone, IE. Eds. pp. (227-237) Asociación Argentina de Microbiología, B.A. ISBN: 978-987-98475-8-9.

Barea, J.M. 2004. Impacto de las micorrizas en la calidad del suelo y la productividad vegetal en sistemas agrícolas y espacios naturales. En: *Biología del suelo. Transformaciones de la materia orgánica, usos y biodiversidad de los organismos edáficos.* Monzón de Asconegui, M.A., García de Salamone, I.E., Miyazaki, S.S. (Eds.) pp. (7-11). Editorial FAUBA. Universidad de Buenos Aires. Argentina.

Bashan, Y. & Levanony, H. 1990. Current Status of *Azospirillum* inoculation technology: *Azospirillum* as a challenge for agriculture. *Canadian Journal of Microbiology*, 36: 591-600.

Bashan, Y., Harrison, S.K. & Whitmoyer, R.E.. 1990. Enhanced growth of wheat and soybean plants inoculated with *Azospirillum* brasilense is not necessarily due to general enhancement of mineral uptake. *Appl. Environ. Microbiol*. 56: 769-775.

Bashan, Y., Holguin, G. & de-Bashan, LE. 2004. *Azospirillum*-plant relationships: physiological, molecular, agricultural and environmental advances (1997-2003). *Canadian Journal of Microbiology*. 50: 521-577.

Boddey, R.M., V.L. Baldani, J.I. Baldani & J. Döbereiner. 1986. Effect of inoculation of *Azospirillum Spp*. on nitrogen accumulation by field grown. *Plant & Soil*. 90: 265-292.

Boddey, RM. 1987.Methods for quantification of nitrogen fixation associated with gramineae. *CRC Critical Review Plant Science*, Boca Raton, v.6, pp. (209-266).

Caballero Mellado, J. 2004. Uso de *Azospirillum* como alternativa tecnológica viable para cultivos de cereales. En: *Biología del Suelo. Transformación de la materia orgánica. Usos y biodiversidad de los organismos edáficos*. de Asconegui, M.A., García de Salamone, I.E., Miyazaki, S.S. (Eds.) pp. (45-49). FAUBA. Universidad de Buenos Aires. Argentina.

Caballero-Mellado, J., M. Carcaño-Montiel & M.A. Mascarúa-Esparza. Field inoculation of wheat (*Triticum aestivum*) with *Azospirillum brasilense* under temperate climate. *Symbiosis* 13 (1992) 243-253.

Cassan, F.D. & García de Salamone, I.E. 2008. *Azospirillum sp.: Cell physiology, plant interactions and agronomic research in Argentina. Asociación Argentina de Microbiología*. Buenos.Aires. ISBN: 978-987-98475-8-9.

Catroux, G. 2007. Uso de microorganismos en agricultura: sueño y realidad. In: *XV Congreso de AAPRESID*, Rosario, Argentina.

Cheng, W., Jonson, D.W. & Shenglei, F. 2003. Rhizosphere effects on decomposition: Control of plant species, phenology and fertilization. *Soil Sci. Soc. Am. J.* 67: 1418-1427.

Chiocchio, V.M., Venedikian, N., Martinez, A.E., Menéndez, A.B., Ocampo, J.A. & Godeas, A. 2000. Effect of the fungicide benomyl on spore germination and hyphal lenght of the arbuscular mycorrhizal fungus *Glomus mossea*e. Revista *International Microbiology* Vol. 3, pp. 163 – 172.

Chiocchio, V.M., Rodríguez, M.A., Venedikian, N., Martinez, A.E., Menéndez, A.B., Scervino, J.M. & Godeas, A. 2002. Una relación con beneficios mutuos: Las raíces de las plantas y los hongos micorrícicos arbusculares. *Bol. Soc. Arg. Bot.* 37 (3-4): 265-275.

D´Auria, F., Escobar Ortega, J., López de Sabando, M., García de Salamone, I. E., Pérez, M. & Zubillaga, M.M. 2011. Mineralización de nitrógeno y microorganismos asociados en un cultivo de trigo de Buenos Aires, Argentina. *VIII Reunión Nacional Científico-Técnica de Biología de Suelos*. 6-8 Julio. Salta, Argentina.

de Freitas, J.R., Banerjee, M.R. & Germida, J.J. 1997. Phosphate solubilizing rhizobacteria enhancer the growth and yield but not phosphorus uptake of canola (*Brassica napus* L.), *Biol & Fertil Soil*. 24: 358-364.

Di Salvo, L.P. & García de Salamone, I.E. 2008. Estudios de comunidades microbianas mediante uso de fuentes carbonadas: comparación entre microplacas comerciales y preparadas en laboratorio.(139) XXI *Congreso Argentino de la Ciencia del Suelo.* Potrero de Funes, San Luis. ISBN 978-987-21419-9-8.

Díaz-Zorita, M. & Fernandez Canigia, M.V. 2008. Análisis de la producción de cereales inoculados con *Azospirillum brasilense* en la República Argentina. 155-166. *En:*

Azospirillum sp.: Cell physiology, plant interactions and agronomic research in Argentina. Cassan, F. & Garcia de Salamone, I.E. (Eds.). Asociación Argentina de Microbiologia, Buenos Aires, Argentina.

Doran, J.W. & Zeiss, M.R. 2000. Soil health and sustaintability: managing the biotic component of soil quality. *Applied Soil Ecology* 15: 3-11.

Ferraris, G. & Courerot, L. 2004. Evaluación de la utilización de bacterias solubilizadores de fósforo (*Pseudomonas* spp) en trigo. *Revista de Tecnología Agropecuaria* –INTA Pergamino. 9:37-39.

García, F. 1999. Aspectos principales de siembra directa y los cultivos de soja y maíz en Argentina. In: *IV Conferencia Anual da Revista Plantio Direto..*21-32.

García de Salamone, I.E. & Döbereiner, J. 1996. Maize genotype effects on the response to *Azospirillum* inoculation. *Biol & Fertil Soil* 21: 193-196.

García de Salamone, I.E. & Monzón de Asconegui; M.A. 2008. Ecofisiología de la respuesta a la inoculación con *Azospirillum* en cultivos de cereales. pp. 209-226. *En: Azospirillum sp.: cell physiology, plant interactions and agronomic research in Argentina.* Cassan, F. & Garcia de Salamone, I.E. (Eds). Asociación Argentina de Microbiología, Buenos Aires, Argentina.

García de Salamone, I.E. & Cassan, F. D. 2010. *Taller Internacional sobre Rizosfera, Biodiversidad y Agricultura Sustentable.* Libro de Resúmenes. ISBN 978-987-9260-78-4

García de Salamone, I.E., Döbereiner, J., S. Urquiaga & Boddey, R.M. 1996. Biological Nitrogen Fixation in *Azospirillum* strain-maize genotype associations as evaluated by ^{15}N isotope dilution technique. *Biol & Fertil Soil,* 23 : 249-256.

García de Salamone, I.E., Hynes, R.K. & Nelson, L.M. 2001. Cytokinin production by plant growth promoting rhizobacteria and selected mutants. *Canadian Journal of Microbiology*, 47:404-411.

García de Salamone, IE; Rorig, M; Bordato, F. & Michelena, R. 2004. Actividad microbiana luego de la aplicación de compactación sobre un suelo Haplustol éntico bajo siembra directa. En: *XVI y XII Congresos latinoamericano y colombiano de la ciencia del suelo.* Cartagena de Indias. Colombia. (pp.27/9-1/10).

García de Salamone, I.E., Hynes, R.K. & Nelson, L.M.. 2005. Role of cytokinins in plant growth promotion by rhizosphere bacteria. En: *PGPR: Biocontrol and Biofertilization*, Editor Z. A. Siddiqui . Springer, Dordrecht, The Netherlands. 173-195.

García de Salamone, I.E.; Bordato, F. & Michelena, R. 2005. Indicadores microbianos de calidad de suelo luego de la aplicación de compactación sobre un suelo Haplustol éntico bajo siembra directa. *V Reunión Nacional Científico-Técnica de Biología de Suelos y V Encuentro sobre Fijación Biológica de Nitrógeno*, San Salvador de Jujuy, Argentina. 6-8 de Julio.

García de Salamone, I.E.; Rorig, M.; Di Salvo, L. & Michelena, R. 2006a. Comunidades microbianas en capas superficiales de un suelo Haplustol éntico bajo siembra directa. *XX Congreso Argentino de la Ciencia del Suelo*. Salta, Argentina. 19-22 de Setiembre.

García de Salamone, I.E., Michelena, R; Rodríguez, A; Montemitoli, I; Gatti, S. & Rorig, M. 2006b. Ocurrencia de micorrizas vesículo arbusculares en plantas de maíz, soja y trigo en sistemas de siembra directa. *Revista de la Facultad de Agronomía*, Universidad de Buenos Aires. 26 (1): 67-72.

García de Salamone, I.E., Di Salvo, L.P., Escobar Ortega, J.S., Boa Sorte, M.P., Urquiaga, S. & Dos Santos Teixeira, K.R. 2010. Field response of rice paddy crop to inoculation with *Azospirillum*: physiology of rhizosphere bacterial communities and the genetic diversity of endophytic bacteria in different parts of the plants. *Plant &Soil* 336:351–362

García de Salamone, I.E., Funes, J.M., Di Salvo, L.P., Escobar-Ortega, J.S., D'Auria, F., Ferrando, L. & Fernandez-Scavino, A. 2011. Impact of plant genotypes on rhizosphere microbial communities and field crop production. *Applied Soil Ecology*. in press.

Garland, J.L. 1996. Patterns of potential C source utilization by rhizosphere communities. *Soil Biol. Biochem*. Vol. 28: 2, 223-230.

Garland, J.L., Roberts, M., Lanfang, S., Levine, H., & Mills, A. L. 2003. Community-Level Physiological Profiling Performed with an Oxygen-Sensitive Fluorophore in a Microtiter Plate. *Applied and Environmental Microbiology* 69, 5:2994–2998.

Gatica, S.M.; Anzovini, M.B.; Di Salvo, L.P.; Escobar Ortega, J.S. & García de Salamone I.E. 2009. Inoculación de arroz con *Azospirillum brasilense*: Incremento en rendimiento con impacto reversible sobre las comunidades microbianas nativas. *VII Reunión Nacional Científico Técnica de Biología del Suelo y Fijación Biológica del Nitrógeno*. 1-3 de Julio. San Miguel de Tucumán, Tucumán,

Glick, B.R. 1995. The enhancement of plant growth by free living bacteria. *Canadian Journal of Microbiology*, 41: 109-117.

Gómez, E., Garland, J., Conti, M. 2004. Reproducibility in the response of soil bacterial community-level physiological profiles from a land use intensification gradient. Applied Soil Ecology 26:21-30.

Grandy, A.S., Robertson, G.P. & Thelen, K.D. 2006. Do productivity and environmental trade-offs justify periodically cultivating No-till cropping systems? *Agronomy Journal*. 98:1377-1383.

Haack, S. K., Garchow, H., Odelson, D. A., Forney, L. J. & Klug, M. J. 1994. Accuracy, Reproducibility, and Interpretation of Fatty Acid Methyl Ester Profiles of Model Bacterial Communities. *Applied Environmental Microbiology.* 60(7): 2483–2493.

Herschkovitz, Y; Lerner, A; Davidov, Y; Rothballer, M; Hartmann, A; Okon, Y. & Jurkevitch, E. 2005. Inoculation with the Plant-Growth-Promoting Rhizobacterium *Azospirillum brasilense* causes little disturbance in the rhizosphere and rhizoplane on maize (*Zea mays*). *Microbial Ecology* 50: 277-288.

Jackson, L E., Calderonb, F. J., Steenwertha, K, L., Scow, K.M. & Rolston, D.E. 2003. Responses of soil microbial processes and community structure to tillage events and implications for soil quality. *Geoderma*, 114, (3-4): 305-317.

Kloepper, J.W.E. 1993. Plant growth-promoting rhizobacteria as biological agents. In: FB Metting, Jr. (ed.), *Soil Microbial Ecology. Applications in agricultural and environmental management.* pp. (255-274). Marcel Dekker.N.Y. U.S.A.

Maddonni, G.A., Ruiz, R.A., Vilariño, P. & García de Salamone, I.E. 2004a. Fertilización en los cultivos para grano. 499-557. En: *Producción de Granos, Bases funcionales para su manejo*. Satorre, E, Benech Arnold, R, Slafer, G.A., de la Fuente, E.B., Miralles,

D.J.; Otegui, M.E. & Savin, R. (Eds.). Editorial FAUBA, Universidad de Buenos Aires.

Maddonni, G.A., Vilariño, P. & García de Salamone, I.E. 2004b Dinámica de los nutrientes en el sistema suelo-planta. En: *Producción de Granos, Bases funcionales para su manejo.* (441-477). Satorre, E, Benech Arnold, R, Slafer, GA, de la Fuente, EB, Miralles, DJ; Otegui, ME; Savin, R. (Eds.). Editorial FAUBA, Universidad de Buenos Aires.

Martinez, A., Menendez, A.B., Chiocchio, V.M.; Venedikian, N. & Godeas, A. 1998. Influencia del insecticida Triflumuron sobre la micorrización por parte de los hongos arbusculares y el crecimiento de las plantas de soja (*Glycine max* var sojae). *Revista de la Facultad de Agronomía de Universidad de La Plata* 103 (2), pp. 135 – 140.

Menéndez, A., Martinez, A.E., Chiocchio, V.M., Venedikian, N., Ocampo, JA. & Godeas, AM. 1999. Influence of the insecticide Dimetoathe on arbuscular mycorrhizal colonization and growth of soybean plants. *Revista International Microbiology.* 2, (1): 43 - 45.

Michelena, R.; Rivero, E.; Rorig, M. & E. Schung. 2001. Stuble evaluation in no-tillage and nutrient availability in soil plant system. *12 th World Fertilizer Congress. International Scientific Centre* – The Chinese Academy of Science. Beijing, China.

Naeem, S., Thompson, L., Lawler S., Lawton, J.H. & Woodfin R.M. 1995. Emprirical evidence that declining species diversity may alter the performance of terrestrial ecosystems. *Phil. Trans. R. Soc. Lond. B.* 345: 249-262

Naiman, A.D.; Latrónico, A.E. & García de Salamone, I.E. 2008. Efecto de la inoculación de Trigo con Pseudomonas *fluorescens* sobre su micorrización natural y sus comunidades de hongos filamentosos en condiciones agrícolas. VI *Congreso Latinoamericano de Micología (VI CLAM).* 10-13 noviembre. Mar del Plata, Argentina.

Naiman, A.D.; Latrónico, A.E. & García de Salamone, I.E. 2009. Inoculation of Wheat with *Azospirillum brasilense* and *Pseudomonas fluorescens*: impact on the production and rhizospheric microflora. *European Journal of Soil Biology* 45:44-51.

Okon, Y. 1994. *Azospirillum*/Plant association. Y. Okon (Ed.). CRC Press, Boca Raton, Florida, USA.

Paul E. A. & F. E. *Clark. 1996*. Soil Microbiology and Biochemistry. Academic Press, Inc., San Diego, CA. 340 pp.

Reed, M.L.E. & Glick, B.R. 2004. Applications of free living plant growth-promoting rhizobacteria. *Antonie van Leeuwenhoek* 86: 1-25.

Rillig, M; Wright, S. & Eviner, V. 2002.The role of arbuscular mycorrhizal fungi and glomalin in soil aggregation: comparing effects of five plant species. *Plant & Soil* 238 (2): 325-333.

Rodriguez, H., R. Fraga, T. Gonzalez &. Y. Bashan. 2006. Genetics of phosphate solubilization and its potencial applications for improving plant growth-promoting bacteria. *Plant & Soil.* 287: 15-21.

Rorig, M., Rivero, E. & Michelena, R. 2004a. Descomposición de rastrojos de trigo y soja en un Argiudol vértico. XIX Congreso Argentino de la Ciencia del suelo. *Paraná, Argentina.*

Rorig, M., Alderuccio, S., Malcolm V., Olsen, D., Michelena, R. & García de Salamone, I.E. 2004b. Estimaciones del número de microorganismos, producción de nitratos y actividad de fosfatasa alcalina en un suelo Argiudol vértico de la localidad de Ramírez, Entre Rios, bajo siembra directa y durante la rotación trigo-soja. En: *Biología del Suelo. Transformación de la materia orgánica. Usos y biodiversidad de los organismos edáficos.* Monzón de Asconegui , M; García de Salamone, I.E. ; Miyazaki, S. Eds. FAUBA. Buenos Aires.

Schalamuk, S., Velazquez, S., Chidichimo, H. & Cabello, M. 2003. Efecto de la siembra directa y labranza convencional sobre la colonización y esporulación en trigo. *Boletín Micológico* 18: 15-19.

Schalamuk, S., Velazquez, S., Chidichimo, H. & Cabello, M. 2006. Fungal spore diversity of arbuscular mycorrhizal associated with spring wheat: effect of tillage. *Mycologia* 98, (1): 16-22.

Schlesinger, W.H. 1991. *Biogeochemistry: An Analysis of Global Change.* Academic Press, San Diego, CA.

Scow, K.M. 1997. Soil Microbial Communities and Carbon Flow in Agroecosystems. In: Ecology in Agriculture, Louise E. Jackson (Ed.). *Physiological Ecology Series, Agricultural and Biological Sciences*, Academic Press. pp. 367-413.

Sieverding, E. 1991. *Vesicular-arbuscular mycorrhiza management in tropical agrosystems.* Bremer Verlag, Berlin.

Sisti, C.P.J., Santos, H.P., Kochhann, R., Alves, B.J.R., Urquiaga, S. & Boddey, R.S. 2004. Change in carbon and nitrogen stocks in soil Ander 13 years of conventional and zero tillage in southern Brazil. *Soil Tillage Research.* 76:39-58.

Studdert , G.A.; Echeverria, H.E. & Casanovas, E.M. 1997. Crop pasture rotation for sustaining the quality and productivity of a Typic Argiudol. *Soil Sci. Soc. Am. J.* 61:1466-1472.

Torsvik, V., Goksøyr, J. & Daae, F. L. 1990. High diversity in DNA of soil bacteria. *Appl Environ Microbiol.* March; 56(3): 782-787.

Torsvik,V., Øvreås, L. & Thingstad, T.F. 2002. Prokaryotic diversity: magnitude, dynamics, and controlling factors. *Science* 296: 1064-1066.

Triplett, Jr., G.B. & Dick, W.A. 2008. No-Tillage Crop Production: A Revolution in Agriculture!. *Agronmy Journal* 100: 153-165.

Urquiaga, S., Jantalia, C.P., Alves, B.J.R. & Boddey, R.M. 2004. Importancia de la FBN en el secuestro de carbono en el suelo y en la sustentabilidad agrícola. *En: Biología del Suelo. Transformación de la materia orgánica. Usos y biodiversidad de los organismos edáficos.* Monzón de Asconegui, M.A., García de Salamone, I.E.; Miyazaki, S.S. (Eds). pp. (1-6). Editoral FAUBA. Universidad de Buenos Aires.

Venedikian, N., Chiocchio, V.M., Martinez, A.E., Menendez, A.B., Ocampo, J.A. & Godeas, A. 1999. Influence of the fungicides Carbendazim and Chlorotalonil on spore germination, arbuscular mycorrhizal colonization and growth of soybean plants. *Revista Agrochimica.* Vol. XLIII – N° 3-4, pp. (105 – 109).

Wardle, D. 2002. Communities and ecosystems: Linking the aboveground and belowground components. Princeton Univ. Press, Princeton, N.J.

Whitman, W. B., Coleman, D. C. & Wiebe, W. J. 1998. Prokaryotes: The unseen majority. *Proc. Natl. Acad. Sci.* USA. 95 6578-6583.

Zotarelli, L., Torres, E., Boddey, R.M., Urquiaga, S. & Alves, B.J.R. 2002. Role of legumes in the N economy of cereal production in crop rotation under conventional and no-tillage. In: World Congress of soil Science. Proceeding *of the 17th World Congress of Soil Science*. Bangkok. Tailandia.

Zubillaga, M.M., Cabrera, M.L., Kissel, D. & Rema, J. 2007. Modelling field scale N mineralization in Coastal Plain soils. *Ecological Modelling* 207: 243-250.

Case Study: *Sclerotinia sclerotiorum*: Genetic Diversity and Disease Control

Silvana Petrofeza[1] and Luiz Carlos Bhering Nasser[2]
[1]Universidade Federal de Goiás, Goiânia – GO
[2]Embrapa Cerrados, Planaltina – DF
Brazil

1. Introduction

Mutation, selection, gene flow between populations, and genetic recombination within them are the primary factors that determine the genetic structure and dynamics of populations. Genetic isolation of populations and drift can lead to genetic divergence and, eventually, speciation. In the Ascomycetes, the capacity for sexual recombination often coexists with enormous asexual reproductive capacity, and may be complemented by parasexual recombination. Some populations within a species have completely lost sexual competence, and the fully asexual species within the Fungi Imperfecti may have originated as such populations diverged from their ancestors. The balance between sexual and asexual reproduction, the contribution of recombination to fitness and pathotypic diversity, the trends toward exclusive asexuality, and the genetic isolation are particularly important in pathogenic fungi, and can have telling consequences for disease management (Taylor et al., 1999).

Sclerotinia sclerotiorum, an important fungal pathogen, offers a pertinent case study of the multidisciplinary approach required to clarify the significance of recombination in agricultural pathosystems. This homothallic ascomycete has a wide host range, sometimes causing extensive necrotic lesions on aerial structures of approximately 400 species of plants worldwide, including important crops and numerous weeds (Boland and Hall, 1994). *S. sclerotiorum* poses a threat to dicotyledonous crops such as sunflower, soybean, oilseed rape, edible dry bean, chickpea, peanut, dry pea, lentils, and various vegetables, as well as monocotyledonous species such as onion and tulip (Boland and Hall, 1994). Annual yield losses due to *Sclerotinia* diseases exceed millions of dollars each year all over the world. Extensive crop damage, lack of high levels of host resistance, and the general difficulty of managing diseases caused by *Sclerotinia* have been the impetus for sustainable research on this pathogen.

2. *Sclerotinia sclerotiorum* reproduction and disease cycle

Sclerotinia sclerotiorum (Lib.) de Bary: kingdom Fungi, phylum Ascomycota, class Discomycetes, order Helotiales, family Sclerotiniaceae, genus Sclerotinia. As a homothallic organism, *S. sclerotiorum* reproduces sexually by self-fertilization, resulting in the formation of apothecia and airborne ascospores, and asexually by formation of sclerotia.

The basic disease cycle of *Sclerotinia* diseases begins with the sclerotium, a pigmented, multi-hyphal structure that can remain viable over long periods of time under unfavorable conditions for growth. Sclerotia can germinate to produce mycelia (myceliogenic germination) or apothecia (carpogenic germination), depending on environmental conditions (Steadman, 1983).

Carpogenic germination usually requires the sclerotia to be in wet soil for one to two weeks prior to germination (Figure 1E). At soil depths of up to 2 cm, apothecia can extend from the sclerotia to reach the soil surface. One or several apothecia can emerge from a single sclerotium (Figure 1F). Apothecia are fleshy-colored discs measuring 2 mm to 10 mm in diameter, lined with asci, which are filled with eight hyaline ascospores (110-160 μm x 6-10 μm) (Kohn, 1979). Each apothecium may release up to 10 million ascospores over a period of several days into the environment. Most will fall on susceptible plants in a 100-meter radius from the apothecia, but some can travel up to 3-4 km by wind (Adams and Ayres, 1979; Kohli et al., 1995; Cubeta et al., 1997). Furthermore, viable ascospores were captured at altitudes up to 6000 m (Williams and Stelfox, 1979), indicating a possible dissemination over long distances. Prolonged periods of leaf wetness (16-48 h) with temperatures in the range of 12°C to 24°C promote ascospore germination and infection (Adams and Ayres, 1979; Steadman, 1983).

The pathogen produces oxalic acid and numerous enzymes that break down and degrade plant tissue (Marciano et al., 1983; Godoy et al., 1990), and highly susceptible host tissues, such as flower petals or senescing leaves, are typically the first ones to become infected. With adequate wetness, mycelium moves from the colonized tissue into the vigorous host tissues of stems, leaves, pods, among others (Figure 1B, 1C). Later in the disease process, sclerotia form, either on the plant surface or within stems and other plant parts (Figure 1A, 1B, 1D). As the plant or plant part dies, the sclerotia fall to the soil, where they can survive for multiple years (Figure 1E) (Natti, 1971; Adams and Ayres, 1979; Boland and Hall, 1988).

The other method of germination is myceliogenic, where the sclerotium produces mycelium. White hyphal strands extend from sclerotia, which have been stimulated to germinate by host plant exudates, infecting roots, crowns, and other plant parts that are touching the ground. Infection then spreads to aboveground plant parts (Steadman, 1983).

Ascospores of *S. sclerotiorum* are a major source of inoculum in crops like bean, canola, and soybean, where aerial parts are affected (Steadman, 1983; Boland and Hall, 1988). In other crops, such as sunflower, direct mycelial germination of soilborne sclerotia affects roots and crowns (Huang and Dueck, 1980).

The major adaptation strategies that contribute to the success of *S. sclerotiorum* in establishing a parasitic relationship with crops is the ability of the fungus to continue producing sclerotia on diseased foliar debris even after being detached from the living plant, thus increasing the amount of inoculum in the soil.

3. Genetic diversity, population structure and their implications for disease control

Estimating the genetic diversity of pathogen populations in a region is important to understand the epidemiology of the disease and to establish the strategies for the control,

Fig. 1. Soybean plants showing symptoms of *Sclerotinia sclerotiorum*. Mycelium colonizing tissue into the host tissues of stems (A-B), leaves, and pods (C). Newly produced sclerotia on stem (A-B), diseased pods (D), and soil surface (E). Apothecia from sclerotia on soil surface (F). (Source: Mônica C. Martins, Fundação BA, Barreiras, BA, Brazil, 2008/2009 growing season).

such as the use of fungicides or resistant cultivars. The analysis of *S. sclerotiorum* structure and population dynamics is an essential part of understanding how the underlying mechanisms are involved in the history of this pathogen and its distribution across different geographic areas and wide host range.

3.1 Recombination *vs* clonal reproduction

There are two fundamental means by which fungi and other organisms transmit genes to the next generation – via clonal reproduction or via mating and recombination. Under clonality, each progeny would have one parent, its genome would be an exact mitotic copy of the parental one, and all parts of the genome would have the same evolutionary history. At the other extreme are genetically novel progenies formed by mating and meiotic recombination of genetically different parental nuclei, events that cause different regions of the genome to have different evolutionary histories. However, fungi do not fit neatly into these two categories. Recombination need not be meiotic or sexual because mitotic

recombination via parasexuality can mix parental genomes. Clonality need not be mitotic and asexual since self-fertilizing or homothallic fungi make meiospores with identical parental and progeny genomes. Furthermore, there is the complication that the same fungus may display different reproductive modes in different localities at different times (Taylor et al., 1999).

The methods currently available for studying microbial reproduction in nature can distinguish recombination from its absence, but they cannot reveal the mechanism nor estimate the amount of recombination (Taylor et al., 1999). Therefore, in this chapter, reproduction by recombination is defined as the production of progeny genomes that are mixtures of genetically different parental genomes, and reproduction by clonality is defined as the production of progeny genomes that are identical to the parental genomes.

S. sclerotiorum has a haploid somatic phase in which clonality is the result of both asexual reproduction by means of sclerotia and sexual reproduction by self-fertilization (Kohn, 1995) with the expectation that intraclonal variation is due to mutation (Carbone and Kohn, 2001a; Carbone et al., 1999). Predominant asexual or self-fertilizing modes of reproduction will determine a clonal population.

Prior researches on *S. sclerotiorum* population structure, in areas where canola, sunflower, cabbage, and other crops are grown, have shown that this pathogen populations are mainly clonal (Kohli et al., 1992, 1995; Anderson and Kohn, 1995; Cubeta et al., 1997; Kohli and Kohn, 1998; Carpenter et al., 1999; Hambleton et al., 2002), even though some recombination has been detected (Cubeta et al., 1997; Carbone and Kohn, 2001a), being more evident in subtropical populations than in temperate ones (Carbone et al. 1999; Carbone and Kohn, 2001b; Atallah et al., 2004; Sexton and Howlett, 2004; Sexton et al., 2006; Mert-Türk et al., 2007; Gomes et al., 2011).

Clonality of *S. sclerotiorum* in field populations has been evidenced in several ways: by detecting association of independent markers, such as mycelial compatibility groups (MCGs) and DNA fingerprints; by repeatedly recovering the same genotypes (over a wide geographic area or from year to year); or indirectly by statistical testing for lack of association of markers. Example of this was found by repeated recovery of MCGs in samples of *S. sclerotiorum* from canola made in Ontario in 1989 (Kohn et al., 1991), western Canada in 1990, 1991, and 1992 (Kohli et al., 1992, 1995; Kohli and Kohn, 1998), eastern Ontario and Quebec in 1999, and western Ontario in 2000 (Hambleton et al., 2002). Several clones were dispersed over large geographic areas, with clone 2 repeatedly isolated across 2000 km (recovered from Ontario, Manitoba, Saskatchewan, Alberta, and subsequently from cabbage in New York) over a 4-year period (Kohli et al., 1992; Anderson and Kohn, 1995).

Kohn et al. (1991) demonstrated that MCGs and DNA fingerprints, with a dispersed, repeated element pLK44.20, were linked in clonal populations of *S. sclerotiorum*. Members of different MCGs are incompatible, forming a reaction line where they meet; they also have DNA fingerprints that differ by five or more hybridizing fragments. In haploid organisms that reproduce asexually or by self-fertilization, genetically independent markers, such as MCGs and DNA fingerprints, will remain associated. In an outcrossing sexual fungus, independent markers will be shuffled and disassociated. Based on this concept, a group of isolates sharing the same DNA fingerprint and MCG is interpreted as a clone (Kohn et al. 1991; Kohli and Kohn, 1998; Carbone et al., 1999). Because a clone is a mitotic lineage, some

intraclonal variations due to mutations are expected. This is evident in variant fingerprints among isolates in an MCG (Carbone et al., 1999; Carbone and Kohn 2001b).

Characterizing the genotypic diversity in a population via MCG and fingerprinting constitutes a methodology that allows establishing the heterogeneity in the population and testing for clonality. Several works have deployed MCG typing in population studies of both *S. sclerotiorum* and *Sclerotinia minor* (Kohli et al., 1992, 1995; Cubeta et al., 1997; Carpenter et al., 1999; Carbone and Kohn, 2001b; Hambleton et al., 2002; Phillips et al., 2002; Durman et al., 2003; Hollowell et al., 2003; Kull et al., 2004; Atallah et al., 2004; Sexton and Howlett, 2004). Individual isolates are classified into clonal lineages by the use of two or more independent markers such as MCGs, DNA fingerprinting, or microsatellite markers (Kohn et al., 1991; Carbone et al., 1999; Sirjusingh and Khon, 2001; Hambleton et al., 2002; Auclair et al., 2004). DNA fingerprinting utilizes a probe for a multicopy transposon-like element in Southern blot analyses (Kohn et al., 1991). Microsatellite loci have high mutation rates and are multiallelic in nature, which makes them useful in phylogenetic inference (Sirjusingh and Kohn, 2001).

In an initial description of mycelial interactions in *S. sclerotiorum*, Kohn et al. (1990) reported relatively high mycelial incompatibility, with 21 strains incompatible with all other isolates and the remaining 10 strains tested forming four MCGs of two to three strains each. All self-self pairings were compatible. From each apothecium produced from four strains, all sibling ascospores were mycelially compatible with each other and with the parent homothallic isolate. This is consistent with selfed (homothallic), sexual reproduction, the results of which are indistinguishable from asexual reproduction in a haploid organism such as *S. sclerotiorum*.

Classification of isolates into MCGs is used routinely in many laboratories as a quick marker for genotyping *S. sclerotiorum* within populations. MCG determination is performed using a defined medium, originally developed for vegetative compatibility testing in *S. minor*, namely Patterson's medium (Patterson and Grogan, 1985), amended with red food coloring, and the current standard medium potato dextrose agar (PDA) (Schafer and Kohn, 2006). Compatible pairings formed one confluent colony, whereas incompatible pairings produced a visible reaction in the interaction zone, such as a red line visible on the colony reverse, or a line of fluffy, aerial mycelium or thin mycelium on the colony surface (Figure 2). Microscopically, challenging hyphae in compatible interactions did not necessarily anastomose.

MCG testing is a phenotypic, macroscopic assay of the self recognition system controlled by multiple loci common in fungi (Carbone et al., 1999). A distinction must be maintained between vegetative and heterokaryon compatibility, unless it is known that two strains not only anastomose but also form a stable heterokaryon (Schafer and Kohn, 2006).

In clonal populations, MCGs are transitive and each MCG is associated with one DNA fingerprint; additionally, no fingerprint is associated with more than one MCG. In clonal populations, a few MCGs are frequently sampled with many additional less frequent MCGs or singleton genotypes (Kohli et al., 1995; Hambleton et al., 2002). Even with some evidence of recombination, MCGs can still be identified and each one might be associated with a single fingerprint, or microsatellite, or RAPD genotype (Carpenter et al., 1999; Hambleton et al., 2002; Sexton and Howlett, 2004). If the population is recombining or presents genetic

exchange, markers, MCG, and fingerprint will not be associated. This has been observed in subtropical populations of *S. sclerotiorum* (Cubeta et al., 1997; Carbone and Kohn, 2001b; Meinhardt et al., 2002; Atallah et al., 2004; Sexton et al., 2006; Gomes et al., 2011; Litholdo Júnior et al., 2011). In a highly recombinant population, the expectation is that each isolate sampled, or most of them, is either incompatible with all other isolates or is part of an intransitive MCG, and each isolate either has a unique fingerprint or a fingerprint associated with more than one MCG. In temperate regions, clonal lineages persist over time, with new fingerprints and MCGs arising through mutation and infrequent outbreeding. Consistent with previous studies on *S. sclerotiorum*, there are more fingerprints than MCGs, an indication that transposition leading to new fingerprints occurs more frequently than mutational events resulting in new MCGs (Kohn et al., 1991; Carbone and Kohn, 2001b).
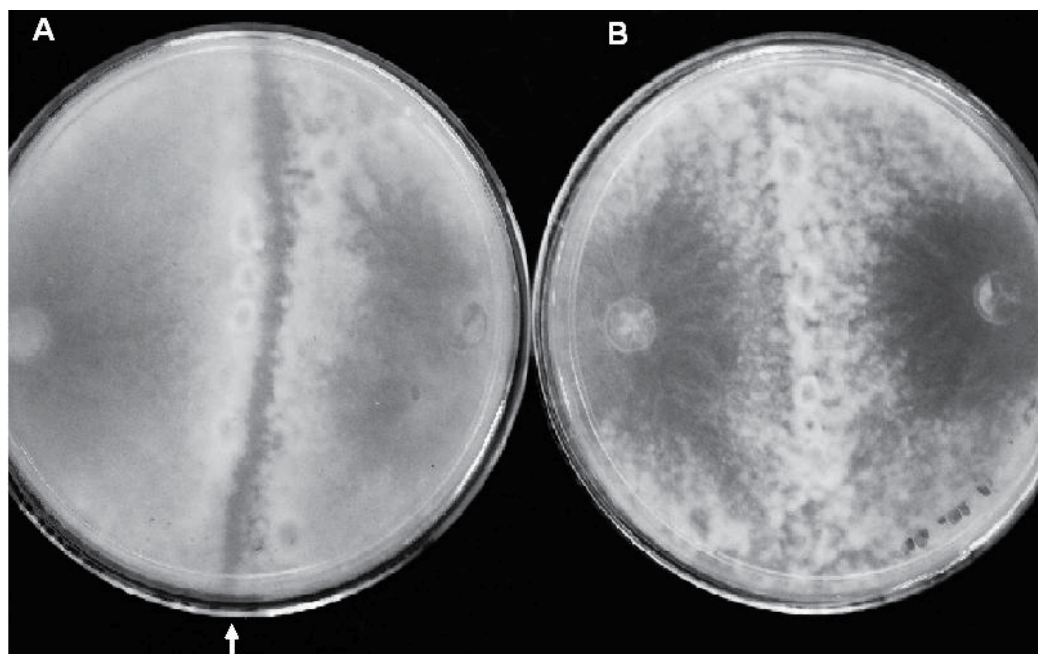
Fig. 2. Representative results from each mycelial compatibility group (MCG). (A) Incompatible reaction. (B) Compatible reaction. MCGs were determined using the method described by Schafer and Kohn (2006).

Factors influencing outcrossing and self-crossing are not well understood, but include weather and agricultural conditions affecting the timing and duration of various stages in the fungal life cycle. For example, if temperature and humidity are more favorable for ascospore production and survival (Clarkson et al., 2003), more simultaneous coinfection occurs, and subsequently more sclerotia might form after fusion of hyphae of two different genotypes, resulting in recombinant genotypes (Sexton et al., 2006).

3.2 Multiple genetic changes associated with gain of virulence

Although *S. sclerotiorum* is considered to exhibit little host specificity (Purdy, 1979), research is still in progress to determine whether genotypes of this fungus vary significantly in

aggressiveness or pathogenicity, or whether host cultivars show differential tolerance to pathogen genotypes. It is important to understand the diversity of this pathogen for the development of effective screening strategies to identify and deploy host resistance.

Pathogen aggressiveness is defined as the relative ability to colonize the host and cause damage, whereas virulence is considered the relative capacity to produce disease (Shurtleff and Averre III, 1997; Kull et al., 2004). Greenhouse inoculation studies on canola suggest that some *S. sclerotiorum* clones are more aggressive in forming lesions (Kohli et al., 1992, 1995).

Studies on *S. sclerotiorum* diversity and pathogenicity have been conducted for different crops. MCGs or microsatellite markers have not been associated with specific virulence characteristics or ecological adaptations of the pathogen, and lack of variation in virulence among isolates from defined geographic areas has been reported in a number of studies on agricultural populations (Kohn et al., 1991; Kohli et al., 1992; Cubeta et al., 1997; Atallah et al., 2004; Auclair et al., 2004; Kull et al., 2004; Sexton and Howlett, 2004; Sun et al., 2005; Malvárez et al., 2007; Litholdo Júnior et al., 2011).

Sexton and Howlett (2004) reported that among 44 isolates tested for virulence, the only two ones that caused relatively small leaf lesions were originally cultured from stem lesions rather than petals. One of these isolates had an identical microsatellite haplotype and MCG to three other stem isolates that caused large lesions, probably indicating that these isolates have genetic differences not reflected by microsatellite and mycelial compatibility markers. Lack of variation in virulence among *S. sclerotiorum* isolates has been noted in other studies. Atallah et al. (2004) found no significant differences in aggressiveness among 35 North American isolates from potato, and Auclair et al. (2004) tested isolates representing four Canadian clonal lineages and did not find any association between genotype and virulence in soybean. In contrast, Kull et al. (2004) reported variation in virulence among isolates from North America (Illinois) and South America (Argentina) collected from soybean. Variation in aggressiveness was highly significant for isolates in MCGs composed of members from different locations. MCGs do not appear to vary in aggressiveness. Although no isolate–cultivar interaction was detected, effective separation of resistant and susceptible cultivars was dependent on isolate selection (Kull et al., 2004).

Differences in virulence may be detected when comparing isolates from widely separate geographic regions. There has been no conclusive evidence to suggest host specialization among isolates of *S. sclerotiorum* (Kull et al., 2004). Diverse isolates of both *S. sclerotiorum* and *Sclerotinia trifoliorum* differed in virulence on alfalfa cultivars, and experiment–cultivar and experiment–isolate interactions were observed, but no isolate–cultivar interaction was detected (Pratt and Rowe, 1995). However, comparisons of *S. sclerotiorum* population on cultivated oilseed rape and on wild perennial host *Rannunculus ficaria* indicated major differences between agricultural and wild populations (Kohn et al., 1988). DNA fingerprint diversity is high in agricultural populations but low in wild populations and there is no evidence of outcrossing in agricultural populations, even though recombination occurs in wild populations.

While differences in pathogenicity or aggressiveness among *S. sclerotiorum* genotypes have been difficult to demonstrate, inconsistent correlations between field and greenhouse disease evaluations (Kim et al., 2000) may stem from incorrect assumptions of genetic

uniformity in pathogen populations. Greenhouse resistance screening programs may be conducted with one isolate of unknown relative aggressiveness and may not represent the range of pathogen variability existing in cultivated field populations. Designing field resistance evaluation experiments to eliminate or compensate for the variability in isolate aggressiveness and MCG profile may not be possible (Kull et al., 2004). It also makes sense to monitor pathogen genotypic diversity in performance trials; new crop varieties with partial resistance can be screened for durability against both established and emerging pathogen genotypes.

Primarily, the production of melanin in *S. sclerotiorum* is intended to protect sclerotia from adverse biological and environmental conditions. An association of melanin with pathogenicity has also been reported in other pathogens, such as *Magnaporthe grisea* and *Colletotrichum lagenarium* (Kubo et al., 2000). However, there are no reports accounting for a relationship between pigmentation and pathogenicity in *S. sclerotiorum*. Furthermore, Garg et al. (2010) found no correlation between pigmentation or colony diameter on PDA with the pathogenicity of different isolates of this pathogen, as measured by diameter of cotyledon lesion on the host genotypes.

Research on the molecular aspects of pathogenicity is mainly concentrated on the contribution of hydrolytic enzyme activity and production of certain metabolites that can act as toxins. This has been the major focus for research on *S. sclerotiorum* as well, with emphasis on the role of cell-wall-degrading enzymes (CWDEs) and oxalic acid. During the interaction with its host, *S. sclerotiorum* secretes a full complement of CWDEs, including pectinases, β-1,3-glucanases, glycosidases, cellulases, xylanases, and cutinases, which can facilitate penetration, macerate tissues, and degrade plant cell-wall components (Riou et al., 1991). Variation in pathogenicity has been associated with the production of pectolytic enzymes (Hancock, 1966; Lumsden, 1976, 1979; Errampalli and Kohn, 1995), cellulase (Lumsden, 1969), hemicellulase, phosphotidase (Maxwell and Lumsden, 1970), and oxalic acid (Noyes and Hancock, 1981; Marciano et al., 1983; Cessna et al., 2000). That oxalic acid is a necessary pathogenicity factor was shown by Godoy et al. (1990), who demonstrated that oxalic acid-deficient mutants were non-pathogenic. Also, at least part of the resistance to *S. sclerotiorum* infection among lines of the scarlet runner bean (*Phaseolus coccineus*) could be attributed to differences in oxalate sensitivity (Chipps et al., 2005). Nonetheless, variation in these factors does not appear to be the primary determinant of pathogenicity as correlations have not always been apparent (Morrall et al., 1972). As an example, in *S. trifoliorum*, a degenerative non-virulent isolate of the pathogen produces more proto-pectinases than a normal string, but only the normal isolate secretes a toxin and is considered virulent (Held, 1955).

In a recent study, our team (Petrofeza et al., data not published) evaluated the genetic interactions for disease response between cultivars of *Glycine max* and isolates of *S. sclerotiorum*. One of the hypotheses tested was the lack of interaction between production of CWDEs and aggressiveness of *S. sclerotiorum* isolates. Aggressiveness of field isolates varied, and this variation was highly significant for isolates in MCGs composed of members from different locations. Significant differences were detected in the activity level of the following CWDEs among the isolates: polygalacturonases, β-1,3-glucanases, cellulases, and xylanases. Furthermore, the level of enzyme activity does not appear to vary in aggressiveness. Nevertheless, the limited number of samples (80 isolates) may not be sufficient to accurately assess the relative contribution of CWDEs production in pathogen aggressiveness.

3.3 Global picture of genetic diversity in *Sclerotinia sclerotiorum* populations worldwide

The major focus of population genetics is to understand the evolutionary processes shaping and maintaining genetic variation within and among populations. Changes in genotype or allele frequencies in populations are considered evolutionary changes, although they often occur on microevolutionary time scales. The evolution of the species in response to selection by deployment of resistant host plants and the evolution of fungicide resistance in response to fungicide applications are perfect examples of population genetics problems in plant pathology. In this sense, plant pathology has provided many examples of how moving a fungus from its normal biogeographic range to a new environment across or between continents can change its reproductive mode and population structure.

The world population of *S. sclerotiorum* is composed of subpopulations that may be associated with phenotypic variability, such as MCGs and aggressiveness, some of which are older or isolated and endemic, whereas others are younger or highly dispersed. And certain crop samples have shown the occurrence of more recombination in subpopulations in subtropical climates than in temperate zones. Unveiling the structure and dynamics of *S. sclerotiorum* populations represents an essential part of understanding the underlying mechanisms involved in the pathogen history and distribution along geographic areas and different hosts. In a brief report, we describe the current knowledge about genetic diversity in *S. sclerotiorum* populations worldwide (Table 1).

Origin	Population/host	Analysis/method	Conclusions	Reference
Canada	Canola (*Brassica napus* L. or *B. rapa*)	MCG DNA fingerprinting Mitochondrial SSrDNA RFLP in Southern probe with mitochondrial 24S RNA from *N. crassa*	Population structure predominantly clonal. Each field infected by several clones. Individuals of some clones can be recovered from different geographic areas and over several years.	Carbone et al. (1999); Kohn et al. (1991); Kohli et al. (1992, 1995)
Norway	*Ranunculus ficaria*	MCG DNA fingerprinting	Genetic markers decoupled. Fingerprinting diversity low and localized. Evidence of some recombination. Spatial substructuring in the distribution of the fingerprints.	Kohn (1995)
USA (North Carolina/ Lousiana)	84 from four cabbage production fields in North Carolina and 16 isolates from an experimental cabbage field plot in Louisiana (*Brassica oleracea* L. var. *capitata*)	MCG DNA fingerprinting	Clonal, but deviated from one-to-one association of markers. Six clones recovered from fields 75 km apart. MCG was associated with more than one DNA fingerprint.	Cubeta et al. (1997)

Canada and USA	Canola (*Brassica napus* L. or *B. rapa*) Cabbage (*B. oleracea* L. var. *capitata*)	DNA fingerprinting	Clonality supported by statistical tests. Clonal fraction higher in canola samples. Some genetic exchange and recombination in cabbage samples.	Kohli and Kohn (1998)
Canada	213 isolates from 10 fields and three performance trials of soybean (*Glycine max*)	MCG DNA fingerprinting	Population structure predominantly clonal. Clones recovered from previous studies. Soybean infected by genotypes residual from other crop or weed hosts.	Hambleton et al. (2002)
USA and Argentina	Diverse hosts/locations, 24 isolates; Argentine set, 21 isolates; soybean set, 254 isolates.	MCG	Within widely dispersed MCGs, isolate aggressiveness varied.	Kull et al. (2004)
USA (Columbia Basin of Washington)	167 isolates from potato	MCG Microsatellite	High haplotypic diversity. Recombinant ascospores in apothecia examined. Microsatellite haplotypes were not correlated with MCGs. Potential for outcrossing.	Atallah et al. (2004)
USA (Alaska)	Vegetable crops	DNA markers	Populations predominantly, if not exclusively, comprised of non-recombining clonal lineages.	Winton et al. (2006)
USA and Canada	294 isolates, three populations: California (CA, lettuce), Washington (WA, pea/lentil), and Ontario (ON, lettuce)	MCG DNA fingerprinting Multilocus DNA sequence	Genetic exchange and recombination could explain the high diversity in California.	Malvárez et al. (2007)
USA (Salinas, San Joaquin, and Santa Maria Valleys in California)	Lettuce drop incidence and structure of *S. minor* and *S. sclerotiorum* populations	MCG	Populations of *S. sclerotiorum* exhibited greater diversity.	Wu and Subbarao (2006)
USA	69 isolates from four natural populations on lentils from eastern Washington	MCG Microsatellite	Clonal lineages dominate each population. 46 MCGs and 53 hyplotypes. Populations differentiated by both location and collection year.	Wang and Chen (2007)
Iran	38 isolates from canola in Iran	rep-PCR primers	High level of genetic diversity. Most isolates from the same regions grouped in the same cluster or in a close cluster.	Karimi et al. (2011)

Iran	48 samples isolated from 29 rapeseed fields in northern Iran	MCG Morphological characteristics	31 MCGs identified. High rates of out crossings as well as evolutionary potentials found within the population.	Barari et al. (2011)
Iran	65 isolates from infected rapeseed, lettuce, bean, tomato, cucumber and wild synapses plants in various fields of northern provinces in Iran	Microsatellite MCG	High level of genetic diversity. 39 MCGs and 26 hyplotypes. High rates of outcrossing as well as evolutionary potential within population of the pathogen in different collecting locations.	Barari et al. (2010)
Iran	276 isolates representing 37 field populations from four provinces in northern Iran	Microsatellite	Moderate levels of gene diversity. Occurrence of outcrossing. Populations randomly mate and have a number of shared haplotypes among regional populations.	Hemmati et al. (2009)
Iran	12 isolates from different provinces of Iran	RAPD	RAPD profiles markedly differed between isolates.	Colagar et al. (2010)
Jordan	25 isolates from vegetable crops in the Jordan valley	MCG RAPD Specific gene amplified	30 MCGs. Occurrence of genetic diversity.	Osofee et al. (2005)
Pakistan	16 isolates associated with the stem rot of chickpea from various localities of Pakistan	MCG Morphological characteristics	Differences in all morphological characteristics observed.	Akram et al. (2008)
Turkey	36 isolates from oilseed rape (*Brassica napus* L.) fields	MCG Microsatellite	High level of genetic diversity within and among population. 19 MCGs identified. Data suggest the possibility of a high rate of outcrossing as well as evolutionary potential within the population of the pathogen in oilseed rape fields.	Mert-Türk et al. (2007)
Turkey	119 isolates from cucumber plants in Kumluca, Finike, and Demre districts of Antalya	MCG Pathogenicity	70 MCGs. Significant differences determined in virulence of isolates within MCGs.	Yanar and Onaran (2011)
China, Canada and the United Kingdom	205 isolates: 5 from England, 33 from Canada, and 167 from China	MCG Pathogenicity	39 MCGs – 64% represented by a single isolate. Isolates from different countries were incompatible. Significant differences were found in radial growth, aggressiveness, production of oxalic acid, and total acids within and among MCGs.	Li et al. (2008)

China, Canada, and the United Kingdom	96 isolates from sunflower	SRAP markers	The five populations isolated from the three countries showed various levels of genetic variability. Population significantly differentiated from others.	Li et al. (2009)
China, Canada, Poland, and Slovakia	Populations from carrot (*Daucus carota* var. *sativa* DC), oilseed rape (*Brassica napus* L.), and soybean plants	RAPD	Highly significant genetic variation among and within populations. Isolates from East Mongolia, in China, and Manitoba, in Canada, were clustered in the same lineage without differentiation.	Sun et al. (2005)
China	76 isolates from different regions and hosts	SRAP markers	Little genetic difference identified among isolates from different host species. Variation among populations differed according to geographic and ecological region.	Chen et al. (2010)
New Zealand	75 isolates from four populations	MCG DNA fingerprinting	High level of variation both within and between populations.	Carpenter et al. (1999)
Australia	154 isolates from four Australian canola fields	MCG Microsatellite	Moderate to high levels of differentiation between populations. Both clonal and sexual reproduction contributed to population structure.	Sexton et al. (2006)
Australia	105 isolates from four Australian canola fields	MCG Microsatellite	Higher genotypic diversity. Occurrence of both clonal reproduction and outcrossing.	Sexton et al. (2006)
Australia	8 isolates from Mount Barker and Walkway regions of Western Australia	Multilocus DNA sequence Pathogenicity Colony characteristics	Significant differences between isolates in relation to pathogenicity. No correlation between pigmentation or colony diameter on PDA with the pathogenicity of different isolates.	Garg et al. (2010)
Australia	250 isolates from sunflower in Queensland and New South Wales	Multicopy Restriction Fragment length polymorphisms MCG RAPD	Individual sclerotia comprised of only one genotype. Temporal studies revealed genetic uniformity was maintained across years.	Ekins et al. (2011)
Brazil	21 isolates from bean (*Phaseolus vulgaris*), one from potato and one from bell peppers	MCG PCR/RFLP of the ITS1-5.8S-ITS2 ribosomal subunit regions Chromosomal telomere sequence Microsatellite	Genetic polymorphisms among isolates within the same MCG.	Meinhardt et al. (2002)

Brazil	isolates from bean (*Phaseolus vulgaris*)	Microsatellite ITS1-5.8S-ITS2 ribosomal subunit regions		Gomes et al. (2011)
Brazil	40 isolates from various fields widely distributed throughout Brazil	MCG RAPD	High level variability. Genetic differences between individuals. The host culture did not have a significant effect.	Litholdo Júnior et al. (2011)
Argentina (Buenos Aires Province)	140 isolates: bean (60 isolates), lettuce (59 isolates), and sunflower (21 isolates) fields	MCG	50 different MCGs. Populations from field crops are made up of various and different MCGs. No differences in aggressiveness between MCGs.	Durman et al. (2003)

Table 1. A list of reported *Sclerotinia sclerotiorum* genetic diversity.

3.3.1 *Sclerotinia sclerotiorum* populations in North America

Over a period of 20 years, population studies on *S. sclerotiorum* carried out in the Unites States and Canada have revealed a predominantly clonal mode of reproduction (Kohn et al., 1991; Kohli et al., 1995; Cubeta et al., 1997; Kohli and Kohn, 1998; Winton et al., 2006), with some evidence of outcrossing contributing to the population structure in a few temperate regions of North America (Atallah et al., 2004; Malvárez et al., 2007), as previously described in item 3.1.

In Canadian oilseed rape (*Brassica rapa* or *B. napus*) fields, clonality dominated in *S. sclerotiorum*, evidenced by the association of each of a relatively large number of DNA fingerprints with MCG. Nevertheless, no evidence of sexual recombination was found within a large population composed of a small number of clones, with a single clone repeatedly isolated across 2000 km over a 4-year period (Kohli et al., 1992; Anderson and Kohn, 1995).

3.3.2 *Sclerotinia sclerotiorum* populations in Asia, Europe and Oceania

A series of intensive analyses of population structure from fields in Australia have been performed. A total of 250 isolates, collected hierarchically from sunflower in Queensland and New South Wales, were proven to belong to one large genetic population of *S. sclerotiorum.* MCGs and molecular markers analyses showed that the majority of sunflower plants were infected by only one genotype and the temporal studies revealed that genetic uniformity had been maintained over the years (Ekins et al., 2011). This corroborates another research that indicated the existence of large populations of *S. sclerotiorum* in canola fields in Australia (Sexton and Howlett, 2004). *S. sclerotiorum* from canola fields in New South Wales were shown to be genetically divergent but with no corresponding difference in virulence from canola-derived isolates collected in Victoria. Genotypic disequilibrium measures were consistent with the occurrence of both clonal reproduction and outcrossing (Sexton et al., 2006).

Several reports concerning the population structure of *S. sclerotiorum* in Asia are available: Iran (Hemmati et al., 2009; Barari et al., 2010, 2011; Colagar et al., 2010; Karimi et al., 2011), Jordan (Colagar et al., 2010); Pakistan (Akram et al., 2008), Turkey (Mert-Türk et al., 2007; Yanar and Onaran, 2011) and China (Sun et al., 2005; Li et al., 2008, 2009; Chen et al., 2010). Different vegetable crop fields were analyzed in these studies and most authors came to the general conclusion that populations exist somewhere in the spectrum between sexual outcrossing and clonal.

In Norway, located in Northern Europe, comparisons between wild and agricultural populations of *S. sclerotiorum* exhibited tendencies for genotypic uniformity of isolates from potato and canola. Also, great genetic diversity was found among isolates from wild perennial host *Ranunculus ficaria,* evidenced by very low fingerprint and MCG diversity accompanied by disassociation of fingerprint with MCG on a very local spatial scale (Kohn, 1995).

Shared haplotypes of *S. sclerotiorum* have been found between countries in different studies. Sun et al. (2005) reported that the relationship between Canadian and Polish populations of *S. sclerotiorum* was very close, although the geographic distance between them was large. Also, Li et al. (2009) analyzed five populations of *S. sclerotiorum* isolated from sunflower in three countries, Mongolia, in China, Canada, and the United Kingdom. Despite the fact that various levels of genetic variability were found, the four populations from West, Middle, and East Mongolia and Canada were clustered into one subgroup, in which the isolates from West and Middle Mongolia belonged to one population and those from East Mongolia and Canada essentially belonged to another. Both East Mongolia and Manitoba, in Canada, are at a high latitude north (over 52°) and have similar climates, which supports the hypothesis that similar ecotypes in two sunflower production areas might produce genetically similar populations.

Comparison of allelic patterns of 22 Australian isolates (Sexton et al., 2006) with Iranian populations (Hemmati et al., 2009) revealed that 16 out of the 30 alleles are common between Australian and Iranian isolates, although no shared haplotypes were found (Hemmati et al., 2009). Furthermore, Australia appears to harbor populations of the fungus that are distinct from North American populations. In a study of southern Australian isolates, Sexton and Howlett (2004) found distinct alleles and allele frequencies from those detected in North America by Atallah et al. (2004). Also, in a study of northern Australian *S. sclerotiorum* isolates from sunflower, using DNA fingerprinting multicopy probe, all Australian genotypes were distinct from Canadian and American genotypes (Ekins et al., 2005).

3.3.3 *Sclerotinia sclerotiorum* populations in South America

In South America, information relating to the population biology of *S. sclerotiorum* is still emerging, with few reports coming from Argentina and Brazil.

In Argentina, *S. sclerotiorum* is one of the most important fungal pathogens of some economically important crops, causing diseases such as *Sclerotinia* stem rot of soybean, head rot of sunflower, and stalk rot of sunflower. In a study designed to analyze MCGs of 140 isolates of *S. sclerotiorum* obtained from soybean (60 isolates), lettuce (59 isolates), and sunflower (21 isolates) fields in a 6500-km² area of the Buenos Aires Province, the partition of total diversity showed that 98.4% corresponded to variations of diversity within populations. These populations were made up of 50 different MCGs (Durman et al., 2003).

In Brazil, white mold caused by the fungus *S. sclerotiorum* occurs primarily in many economically important crops, such as common bean, soybean, cotton, several vegetables, among others, and can cause serious yield losses. This pathogen was firstly reported in the country in the 1920s and spread to the main agricultural regions probably due to infected seeds. White mold is one of the major problems of highly mechanized irrigated winter production of common beans (*Phaseolus vulgaris*) and other vegetable crops in tropical regions in the states of São Paulo, Minas Gerais, and Goiás. During the dry winter months in these states, crops must be grown under center-pivot irrigation, which along with lower ambient temperatures create a microclimate that is suitable for infection and growth of *S. sclerotiorum*. In an initial work, 21 isolates of *S. sclerotiorum* were obtained from a center-pivot-irrigated field and MCG analyses revealed the presence of only two MCGs. PCR/RFLP analyses of the ITS1-5.8S-ITS2 ribosomal subunit regions of these field isolates failed to show any genetic differences between MCGs. DNA amplification with a chromosomal telomere sequence-based primer and one microsatellite primer revealed genetic polymorphisms among isolates within the same MCGs. Isolates taken from beans and two other crops from another region of Brazil showed the same two MCGs and had identical banding patterns for the telomere and microsatellite primers (Meinhardt et al., 2002). In a subsequent work, Gomes et al. (2011) analyzed 79 isolates of *S. sclerotiorum* collected from four common bean center-pivot-irrigated fields and using microsatellite markers observed high genetic diversity within and among populations. The authors suggested that the occurrence of exclusive alleles in the populations analyzed may indicate the introduction of new genotypes in these areas at different moments through contaminated seeds or agricultural implements, in addition to the occurrence of sexual outcrossing and clonal reproduction.

Some specific questions that should be addressed with respect to genetic diversity in *S. sclerotiorum* are:

1. Is *S. sclerotiorum* clonal and/or recombining in nature?
2. What is the relative contribution of both modes of reproduction to genetic diversity, in different locations, and at different times?
3. How variable are virulence factors among populations within the putative recombination belt in different regions?
4. What are the environmental limits to outcrossing sexual reproduction? Do agricultural practices in different regions affect the population biology of the pathogen?

4. Phylogenetics: Inferring evolutionary development

Plant pathologists have led the way in *S. sclerotiorum* population genetics, making fungi in genera such as *Sclerotinia* among the best understood in terms of how they reproduce in nature, the limits of their genetically differentiated or isolated groups, and their spread through space and time. However, a question that has driven much of the recent studies on population genetics of this pathogen has been, "*S. sclerotiorum is a pathogen of virtually all dicots worldwide. Is there a single population with a common origin or are there several populations associated with different geographic areas? Or associated with different hosts?*".

Phylogeography is the study of spatial distributions of genealogical lineages (phylogenies), especially lineages within species and among closely related species

(Taylor et al., 1999). Phylogenetic inference from molecular sequences (loci) is becoming an increasingly popular tool to trace the patterns of pathogen dispersal. Reconstructing both the evolutionary history and spatial process from these sequences provides fundamental understanding of the evolutionary dynamics. Single or multi-locus gene genealogies derived by phylogenetic, coalescent, or Bayesian approaches can be explored to estimate the contribution of the key drivers to evolution in populations: mutations, selection, changes in population size, gene flow, genetic exchange, and recombination (Carbone and Kohn, 2001a, b).

The evolutionary history of population haplotypes indicates that isolates of *S. sclerotiorum* may be grouped into subtropical, temperate, wild, and two relatively recently evolved temperate subtropical populations. Carbone and Kohn (2001b) reconstructed the evolutionary history of haplotypes, on both population and species scales, using a combination of parsimony, maximum likelihood, and coalescent methods. For the population level, samples from Canada, the United States, Norway, and New Zealand isolated from different hosts were analyzed. Samples from extensive crops, like canola (*Brassica napus* or *B. rapa*), sunflower (*Helianthus annuus*), tobacco (*Nicotiana tabacum*), as well as from vegetable crops, like cabbage (*Brassica oleracea*), and smaller samples from other species (*Ranunculus ficaria*, *Apios americana*, *Geranium* sp, *Cannabis sativa*, *Raphanus* sp) provided a broad base for the population analysis. Haplotyping was based on DNA sequence data for seven loci in 385 isolates: the intergenic spacer (IGS) region of the nuclear ribosomal RNA gene repeat unit (IGS; 4,000 bp), an anonymous nuclear region (44.11; 700 bp), portions of the genes encoding translation elongation factor 1α (EF-1α; 350 bp), calmodulin (CAL; 500 bp), chitin synthase (CHS; 300 bp), actin (ACT; 300 bp), and ras protein (RAS; 350 bp). In their study, Carbone and Kohn (2001b) identified five populations from samples of diverse crops and hosts. The most recently diverged populations, 3-1 and 3-2, are dispersed in both temperate and subtropical areas, and diverge via dispersal/isolation by distance. Population 3-5 was limited to temperate areas, whereas 3-4 was a metapopulation restricted to wild buttercup (*Ranunculus ficaria*) in Norway (Carbone and Kohn, 2001a, b). The oldest population, numbered 3-3, appears to be endemic in the subtropical region of the southeastern United States (Carbone and Kohn, 2001a; Phillips et al., 2002).

In a subsequent study, Malvárez et al. (2007) applied this standard to extend a multiple gene genealogical approach aiming to detect geographic differentiation and identify putative biological species among different populations of *S. sclerotiorum*. Isolates of *S. sclerotiorum* were determined to represent three genetically differentiated populations: California (CA, lettuce), Washington (WA, pea/lentil), and Ontario (ON, lettuce). The California population from lettuce was highly diverse, highly divergent from the other populations, and old compared with Washington pea/lentil and Ontario lettuce samples, representing two other populations. The California sample does not represent a new species, despite genetic isolation relative to other populations. As demonstrated in the multilocus phylogeny, the California population was recently derived from a common ancestor in *S. sclerotiorum*, along with representatives of populations 3-3 (southeastern United States), 3-2 (north-central United States and adjacent Canada), and 3-4 (Norwegian wild buttercup), as well as isolates from Uruguay and Argentina, while the representatives of other named and cryptic *Sclerotinia* species are derived from a more distant, more ancient, common ancestor.

This geographic clustering infers that ecological conditions may contribute to adaptations associated with growth temperature range, light intensity, or sclerotial vernalization requirements (Taylor et al., 1999; Kull et al., 2004; Wu and Subbarao, 2006). Nonetheless, outcrossing appears to occur in agricultural populations of *S. sclerotiorum* at a rate that is important in generating new genotypes, which makes it likely that long-term evolution of asexual lineages does not contribute significantly to population genetic structure.

Although pathogens from all hosts are now considered to be one species (Carbone et al., 1999; Carbone and Kohn, 2001b), the extent of gene flow among pathogens of different hosts remains unclear.

5. Disease management and future perspectives

Management practices such as tillage, crop sequence, variety selection, seed treatment, plant population, row width, fertilizer, lime, water (in irrigated crops), herbicide, and fungicide application should be used as integrated management strategies in order to minimize crop losses all over the world.

Management of white mold in any crop requires adoption of an integrated package of practices for disease control. Management practices interact with the white mold pathogen at every stage of the fungal life cycle. No single management practice or variety can be relied on to control white mold. Instead, control of white mold requires application of a combination of management practices to be most effective. The effectiveness of combinations of variety selection and management practices has been tested in Brazil in research plots and on-farm studies examining crop rotation, tillage, variety, and irrigation system selection as tactics for limiting losses with success and substantially reducing white mold incidence.

Establishment and spread of *S. sclerotiorum* among geographic areas is achieved by several mechanisms and disseminative propagules, such as wind dispersal of airborne ascospores and transportation of sclerotia through farming operations, including tillage and irrigation, which are important means for long distance distribution of the pathogen. Sclerotia formed on diseased plant tissues may be dislodged and persist in soil for several years. Sclerotia-contaminated seeds and seeds infected with mycelium are also potential forms of pathogen introduction. Factors such as the different dispersal types for *S. sclerotiorum*, wide host range, genotypic structure with outcrossing potential, and absence of pathogenic races constitute a great challenge for producing resistant crops in the near future.

However, based on the rapid growth of knowledge of molecular genetics and *S. sclerotiorum* genome sequencing, project genome is in progress (Broad Institute through a project sponsored by the USDA Microbial Genomics Program) and advances in the following areas are expected to implement the control of this pathogen:

- Development of methods to assess the amount of recombination as part of larger integrated studies of their epidemiology, evolution, and population biology.
- Development of phylogeography studies, perhaps with multilocus sequence analyses to assess worldwide movement and diversification of genotypes. A better understanding of the evolutionary potential of the pathogen will provide insight into the most suitable breeding strategies for durable resistance.

- Detailed studies are required to understand the mechanism of plant-pathogen interactions, such as virulence factors and aggressiveness, which vary across populations.
- Development of biological control products should be very effective, environmentally safe, and economically acceptable.
- Development of transgenic crops with superior resistance to *Sclerotinia.*

6. Acknowledgments

7. References

Adams, PB, Ayres, WA. (1979). Ecology of *Sclerotinia* species. Phytopathology, 69: 896–899.

Akram, A, Iqbal, SHM, Ahmed, N, Iqbal, U, Ghafoor, A. (2008). Morphological variability and mycelial compatibility among the isolates of S*clerotinia sclerotiorum* associated with stem rot of chickpea. Pakistan Journal of Botany, 40: 2663–2668.

Anderson, JB, Kohn, LM. (1995). Clonality in soilborne, plant-pathogenic fungi. Annual Review of Phytopathology 33: 369–391.

Atallah, ZK, Larget, B, Chen, X, Johnson, DA. (2004). High genetic diversity, phenotypic uniformity, and evidence of outcrossing in *Sclerotinia sclerotiorum* in the Columbia Basin of Washington State. Phytopathology, 94: 737–742.

Auclair, J, Boland, GJ, Kohn, LM, Rajcan, I. (2004). Genetic interactions between *Glycine max* and *Sclerotinia sclerotiorum* using a straw inoculation method. Plant Disease, 88: 891–895.

Barari, H, Alavi, V, Badalyan, SM. (2010). Genetic and morphological diversities in *Sclerotinia sclerotiorum* isolates in northern parts of Iran. World Applied Sciences Journal, 8: 326–333.

Barari H, Alavi V, Yasari E, Badalyan SM. (2011). Study of genetic variations based on the morphological characteristics, within the population of *Sclerotinia sclerotiorum* from the major oilseed planting areas in Iran. International Journal of Biology, 3: 61–66.

Boland, GJ, Hall, R. (1988). Epidemiology of sclerotinia stem rot of soybean in Ontario. Phytopathology, 78: 1241–1245.

Boland, GJ, Hall R. (1994). Index of plant hosts of *Sclerotinia sclerotiorum*. Canadian Journal of Plant Pathology, 16: 93–108.

Carbone, I, Anderson, JB, Kohn, LM. (1999). Patterns of descent in clonal lineages and their multilocus fingerprints are resolved with combined gene genealogies. Evolution, 53: 11–21.

Carbone, I, Kohn LM. (2001a). A microbial population-species interface: nested cladistic and coalescent inference with multilocus data. Molecular Ecology, 10: 947 –964.

Carbone, I, Kohn, LM. (2001b). Multilocus nested haplotype networks extended with DNA fingerprints show common origin and fine-scale, ongoing genetic divergence in a wild microbial metapopulation. Molecular Ecology, 10: 2409–2422.

Carpenter, MA, Frampton, C, Stewart, A. (1999). Genetic variation in New Zealand populations of the plant pathogen *Sclerotinia sclerotiorum*. New Zealand Journal of Crop and Horticultural Science, 27: 13–21.

Cessna, SG, Sears, VE, Dickman, MB, Low, PS. (2000). Oxalic acid, a pathogenicity factor for *Sclerotinia sclerotiorum*, suppresses the oxidative burst of the host plant. Plant Cell, 12: 2191–2200.

Chen, B, Hu, Q, Dixelius, C, Li, G, Wu, X. (2010). Genetic diversity in *Sclerotinia sclerotiorum* assessed with SRAP markers. Biodiversity Science, 18: 509–515.

Chipps, TJ, Gilmore, B, Myers, JR, Stotz, HU. (2005). Relationship between oxalate, oxalate oxidase activity, oxalate sensitivity, and white mold susceptibility in *Phaseolus coccineus*. Phytopathology, 95: 292–299.

Clarkson, JP, Staveley, J, Phelps, K, Young, CS, Whipps, JM. (2003). Ascospore release and survival in *Sclerotinia sclerotiorum*. Mycological Research, 107: 213–222.

Colagar, AH, Saadati, M, Zarea, M, Talei, SA. (2010). Genetic variation of the Iranian *Sclerotinia sclerotiorum* isolates by standardizing DNA polymorphic fragments. Biotechnology, 9: 67–72.

Cubeta, MA, Cody, BR, Kohli, Y, Kohn, LM. (1997). Clonality in *Sclerotinia sclerotiorum* on infected cabbage in eastern North Carolina. Phytopathology, 87: 1000–1004.

Durman, SB, Menéndez, AB, Godeas, AM. (2003). Mycelial compatibility groups in Buenos Aires field populations of *Sclerotinia sclerotiorum* (Sclerotiniaceae). Australian Journal of Botany, 5: 421–427.

Ekins, MG, Aitken, EAB, Goulter, KC. (2005). Identification of *Sclerotinia* species. Australasian Plant Pathology, 34: 549–555.

Ekins, MG, Hayden, HL, Aitken, EAB, Goulter, KC. (2011). Population structure of *Sclerotinia sclerotiorum* on sunflower in Australia. Australasian Plant Pathology, 40: 99–108.

Errampalli, D, Kohn, LM. (1995). Comparison of pectic zymograms produced by different clones of *Sclerotinia sclerotiorum* in culture. Phytopathology, 85: 292–298.

Garg, H, Kohn, LM, Andrew, M, Li H, Sivasithamparam, K, Barbetti, MJ. (2010). Pathogenicity of morphologically different isolates of *Sclerotinia sclerotiorum* with *Brassica napus* and *B. juncea* genotypes. European Journal of Plant Pathololy, 126: 305–315.

Godoy, G, Steadman, JR, Dickman, MB, Dam, R. (1990). Use of mutants to demonstrate the role of oxalic acid in pathogenicity of *Sclerotinia sclerotiorum* on *Phaseolus vulgaris*. Physiological and Molecular Plant Pathology, 37: 179–191.

Gomes, EV, Nascimento, LB, Freitas, MA, Nasser, LCB, Petrofeza, S. (2011). Microsatellite markers reveal genetic variation within *Sclerotinia sclerotiorum* populations in irrigated dry bean crops in Brazil. Journal of Phytopathology, 159: 94–99.

Hambleton, S, Walker, C, Kohn, LM. (2002). Clonal lineages of *Sclerotinia sclerotiorum* previously known from other crops predominate in 1999–2000 samples from Ontario and Quebec soybean. Canadian Journal of Plant Pathology, 24: 309–315.

Hancock, JC. (1966). Degradation of pectic substances associated with pathogenesis by *Sclerotinia sclerotiorum* in sunflower and tomato stems. Phytopathology, 56: 975–979.

Held, VM. (1955). Physiological differences between a normal and a degenerate strain of *Sclerotinia trifoliorum*. Phytopathology, 45: 39–42.

Hemmati, R, Javan-Nikkhah, M, Linde, CC. (2009). Population genetic structure of *Sclerotinia sclerotiorum* on canola in Iran. European Journal of Plant Pathology, 125: 617–628.

Hollowell, JE, Shew, BB, Cubeta, MA, Wilcut, JW. (2003). Weed species as hosts of *Sclerotinia minor* in peanut fields. Plant Disease, 87: 197–199.

Huang, HC, Dueck, J. (1980). Wilt of sunflower from infection by mycelial-germinating sclerotia of *Sclerotinia sclerotiorum*. Canadian Journal of Plant Pathology, 2: 47–52.

Karimi, E, Safaie, N, Shams-Bakhsh, M. (2011). Assessment of genetic diversity among *Sclerotinia sclerotiorum* populations in canola fields by REP-PCR. Trakia Journal of Sciences, 9: 62–68.

Kim, HS Hartman, GL, Manandhar, JB, Graef, GL, Steadman, JR, Diers, BW. (2000). Reaction of soybean cultivars to *Sclerotinia* stem rot in field, greenhouse, and laboratory evaluations. Crop Science, 40: 665–669.

Kohli, Y, Brunner, LJ, Yoell, H, Milgroom, MG, Anderson, JB, Morrall, RAA, Kohn, LM. (1995). Clonal dispersal and spatial mixing in populations of the plant pathogenic fungus, *Sclerotinia sclerotiorum*. Molecular Ecology, 4: 69–77.

Kohli, Y, Kohn, LM. (1998). Random association among alleles in clonal populations of *Sclerotinia sclerotiorum*. Fungal Genetics and Biology, 23: 139–149.

Kohli, Y, Morrall, RAA, Anderson, JB, Kohn, LM. (1992). Local and trans-Canadian clonal distribution of *Sclerotinia sclerotiorum* on canola. Phytopathology, 82: 875–880.

Kohn, LM. (1979). A monographic revision of the genus Sclerotinia. Mycotaxon, 9, 365–444.

Kohn LM. (1995). The clonal dynamic in wild and agricultural plant pathogen populations. Canadian Journal of Botany, 73: S1231–S1240.

Kohn, LM, Carbone, I, Anderson, JB. (1990). Mycelial interactions in *Sclerotinia sclerotiorum*. Experimental Mycology, 14: 255–267.

Kohn, LM, Petsche, DM, Bailey, SR, Novak, LA, Anderson, JB. (1988). Restriction fragment length polymorphisms in nuclear and mitochondrial DNA of *Sclerotinia* species. Phytopathology, 78: 1047–1051.

Kohn, LM, Stasovski, E, Carbone, I, Royer, J, Anderson, JB. (1991). Mycelial incompatibility and molecular markers identify genetic variability in field populations of *Sclerotinia sclerotiorum*. Phytopathology, 81: 480–485.

Kubo, Y, Takano, Y, Tsuji, G, Horino, O, Furusawa, I. (2000). Regulation of melanin biosynthesis genes during appresorium formation of *Colletotrichum lagenarium*. In Freeman, S, Prusky, D, Dickman, M. (Ed.). Host specificity, pathogenicity, and host pathogen interactions of *Colletotrichum* (pp. 99–113). St Paul: American Phytopathological Society.

Kull, LS, Pederson, WL, Palmquist, D, Hartman, GL. (2004). Mycelial compatibility grouping and aggressiveness of *Sclerotinia sclerotiorum*. Plant Disease, 88: 325–332.

Li, Z, Wang, Y, Chen, Y, Zhang, J, Fernando, WGD. (2009). Genetic diversity and differentiation of *Sclerotinia sclerotiorum* populations in sunflower. Phytoparasitica, 37: 77–85.

Li, Z, Zhang, M, Wang, Y, Li, R, Fernando, WGD. (2008). Mycelial compatibility group and pathogenicity variation of *Sclerotinia sclerotiorum* population in sunflower from China, Canada and England. Plant Pathology Journal, 7: 131–139.

Litholdo Júnior, CG, Gomes, EV, Lobo Júnior, M, Nasser, LCB, Petrofeza, S. (2011). Genetic diversity and mycelial compatibility groups of the plant-pathogenic fungus *Sclerotinia sclerotiorum* in Brazil. Genetics and Molecular Research, 10: 868–877.

Lumsden, RD. (1969). *Sclerotinia sclerotiorum* infection of bean and the production of cellulase. Phytopathology, 59: 653–657.

Lumsden, RD. (1976). Pectolytic enzymes of *Sclerotinia sclerotiorum* and their localization in infected bean. Canadian Journal of Botany, 54: 2630–2641.

Lumsden, RD. (1979). Histology and physiology of pathogenesis in plant diseases caused by *Sclerotinia* species. Phytopathology, 69: 890–896.

Malvárez, G, Carbone, I, Grünwald, NJ, Subbarao, KV, Schafer, MR, Kohn, LM. (2007). New populations of *Sclerotinia sclerotiorum* from lettuce in California and peas and lentils in Washington. Phytopathology, 97: 470–483.

Marciano, P, Di Lenna, P, Magro, P. (1983). Oxalic acid, cell wall degrading enzymes and pH in pathogenesis and their significance in the virulence of two *Sclerotinia sclerotiorum* isolates on sunflower. Physiological Plant Pathology, 22: 339–345.

Maxwell, DP, Lumsden, RD. (1970). Oxalic acid production by *Sclerotinia sclerotiorum* in infected bean and in culture. Phytopathology, 60: 1395–1398.

Meinhardt, LW, Wulff, NA, Bellato, CM, Tsai, SM. (2002). Telomere and microsatellite primers reveal diversity among *Sclerotinia sclerotiorum* isolates from Brazil. Fitopatologia Brasileira, 27: 211–215.

Mert-Türk F, Ipek, M, Mermer, D, Nicholson, P. (2007). Microsatellite and morphological markers reveal genetic variation within a population of *Sclerotinia sclerotiorum* from oilseed rape in the Çanakkale Province of Turkey. Journal of Phytopathology, 155: 182–187.

Morrall, RAA, Duczek, LJ, Sheard, JW. (1972). Variation and correlations within and between morphology, pathogenicity, and pectolytic enzyme activity in *Sclerotinia* from Saskatchewan. Canadian Journal of Botany, 50: 767–786.

Natti, JJ. (1971). Epidemiology and control of bean white mold. Phytopathology, 61: 669–674.

Noyes, RD, Hancock, JG. (1981). Role of oxalic acid in the *Sclerotinia* wilt of sunflower. Physiological Plant Pathology, 18: 123–132.

Osofee, H, Hameed, KM, Mahasneh, A. (2005). Relatedness among indiginous membrers of *Sclerotinia sclerotiorum* by mycelial compatibility and RAPD analysis in the Jordan Valley. Plant Pathology Journal 21: 106–110.

Patterson, CL, Grogan, RG. (1985). Differences in epidemiology and control of lettuce drop caused by *Sclerotinia minor* and *Sclerotinia sclerotiorum*. Plant Disease, 69: 766–770.

Phillips, DV, Carbone, I, Gold, SE, Kohn, LM. (2002). Phylogeography and genotype-symptom associations in early and late season infections of canola by *Sclerotinia sclerotiorum*. Phytopathology 92: 785–793.

Pratt, RG, Rowe, DE. (1995). Comparative pathogenicity of isolates of *Sclerotinia trifoliorum* and *S. sclerotiorum* on alfalfa cultivars. Plant Disease, 79: 474–477.

Purdy, LH. (1979). *Sclerotinia sclerotiorum*: history, diseases and symptomatology, host range, geographic distribution, and impact. Phytopathology, 69: 875–880.

Riou, C, Freyssinet, G, Fèvre, M. (1991). Production of cell wall-degrading enzymes by the phytopathogenic fungus *Sclerotinia sclerotiorum*. Applied and Environmental Microbiology, 57: 1478–1484.

Schafer, MR, Kohn, LM. (2006). An optimized method for mycelial compatibility testing in *Sclerotinia sclerotiorum*. Mycologia, 98: 593–597.

Sexton, AC, Howlett, BJ. (2004). Microsatellite markers reveal genetic differentiation among populations of *Sclerotinia sclerotiorum* from Australian canola fields. Current Genetics, 46: 357–365.

Sexton, AC, Whitten, AR, Howlett, BJ. (2006). Population structure of *Sclerotinia sclerotiorum* in an Australian canola field at flowering and stem-infection stages of the disease cycle. Genome, 49: 1408–1415.

Shurtleff, MC, Averre III, CW. (1997). Glossary of plant-pathological terms. St. Paul: American Phytopathological Society.

Sirjusingh, C, Kohn, LM. (2001). Characterization of microsatellites in the fungal plant pathogen, *Sclerotinia sclerotiorum*. Molecular Ecology Notes, 1: 267–269.

Steadman, JR. (1983). White mold – a serious yield-limiting disease of bean. Plant Disease, 67: 346–350.

Sun, JM, Irzykowski, W, Jedryczka, M, Han, FX. (2005). Analysis of the genetic structure of *Sclerotinia sclerotiorum* (Lib.) de Bary populations from different regions and host plants by random amplified polymorphic DNA markers. Journal of Integrative Plant Biology, 47: 385–395.

Taylor JW, Jacobson DJ, Fisher MC. (1999). The evolution of asexual fungi: reproduction, speciation and classification. Annual Review of Phytopathology, 37: 197–246.

Wang, X, Chen, W. (2007). Genetic diversity and population differentiation of natural populations of *Sclerotinia sclerotiorum* on lentils in eastern Washington. Phytopathology,

Williams, JR, Stelfox, D. (1979). Dispersal of ascospores of *Sclerotinia sclerotiorum* in relation to sclerotinia stem rot of rapeseed. Plant Disease Reporter, 63: 395–399.

Winton, LM, Leiner, RH, Krohn, AL. (2006). Genetic diversity of *Sclerotinia* species from Alaskan vegetable crops. Canadian Journal of Plant Pathology, 28: 426–434.

Wu, BM, Subbarao, KV. (2006). Analyses of lettuce drop incidence and population structure of *Sclerotinia sclerotiorum* and *S. minor*. Phytopathology, 96: 1322–1329.

Yanar, Y, Onaran, A. (2011). Mycelial compatibility groups and pathogenicity of *Sclerotinia sclerotiorum* (Lib.) De Bary causal agent of white mold disease of greenhouse grown cucumber in Antalya-Turkey. African Journal of Biotechnology, 10: 3739–3746.

Pyrenophora tritici-repentis, Causal Agent of Tan Spot: A Review of Intraspecific Genetic Diversity

M. V. Moreno[1,2,3], S. A. Stenglein[1,2,3] and A. E. Perelló[1,4]
[1]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
[2]Laboratorio de Biología Funcional y Biotecnología
(BIOLAB)-CEBB-CONICET, Facultad de Agronomía de Azul
Universidad Nacional del Centro de la Provincia de Buenos Aires (UNCPBA)
[3]Microbiología Agrícola, Facultad de Agronomía, UNCPBA
[4]Centro de Investigaciones de Fitopatología (CIDEFI)
Facultad de Ciencias Agrarias y Forestales, Universidad Nacional de La Plata
Argentina

1. Introduction

1.1 Currents status of tan spot

In some countries where the wheat is cultivated, the biological adversities are led by foliar disease. These diseases have emerged as a serious problem in many areas where the wheat is the principal crop. In the last few years, minimum tillage has been considered advantageous to soil conservation, but it leads to a loss of available nutrients and a potential increase in necrotic pathogens whose saprophytic stage lives in the straw of the crop (Annone, 1985). Establishment of the crop under this management can be affected by pathogens of this type. Leaf spotting diseases can be caused by one or a combination of leaf spotting pathogens (Table 1). Leaf spotting diseases affect wheat grown reduce the photosynthetic area of leaves resulting in reduced grain filling and lower yields; particularly when the top two leaves (penultimate and flag leaves) are severely infected. The most of these diseases are similar in host symptomatology, disease cycle, life cycles of pathogens and types of damage induced.

Whitin of these pathogens the Ascomycota fungus, *Pyrenophora tritici-repentis* (Ptr) (Died.) Drechs. It is a facultative pathogen whose asexual stage is *Drechslera tritici-repentis* (Dtr) (Died.). This pathogen is the causal agent of tan spot of wheat.

The effects of tan spot on wheat yields have been reported from 1974 by Hosford & Busch. Several authors had observed severe attacks of *P. tritici-repentis* (Adee & Pfender, 1989; Annone 1994; Fernandez et al., 1998; Galich & Galich 1994; Kohli, 1992; Metha & Gaudencio, 1991; Rees et al., 1982; Rees and Platz, 1983; Sharp et al., 1976). The tan spot has induced losses that reached from 20 % to 70%of yield and it is very destructive on durum, winter and spring wheat (Lamari & Bernier 1989a, b; Misra & Singh 1972). Tan spot reduces total yield,

kernel weight (Shabeer & Bockus 1988; Schilder & Bergstrom 1990), number of grains per head (Schilder & Bergstrom 1990), total biomass (Kremer & Hoffmann 1992), and/or grain quality because of red-smudge symptoms (Fernandez et al., 1994).

Teleomorph	Anamorph	Disease
Pyrenophora tritici-repentis (Died)	*Drechslera tritici-repentis* (Died)	Tan Spot
Mycosphaerella graminicola (Fuckel) J. Schrot. in Cohn	*Septoria tritici* (Roberge in Desmaz.)	Leaf Blotch
Phaeosphaeria nodorum (E. Müller)	*Stagonospora nodorum* [Berk.] Castellani & E.G. Germano)	Stagonospora Nodorum Blotch
Cochliobolus sativus (S. Ito & Kurib.) Drechsler ex Dastur	*Bipolaris sorokiniana*	Spot Blotch

Table 1. Leaf spotting complex of wheat

The increasing inoculum of this pathogen has been attributed to use of conservation tillage systems, shorter crop rotations, continuous wheat cultivation and use of susceptible cultivars. The sexual phase survives in wheat stubble (Rees and Platz, 1980). Currently, we consider the tan spot disease has shown the highest growth in crops under zero-tillage. Interestingly, collateral hosts of *P. tritici-repentis* could play an important role as a source of primary inoculum between growing seasons, as a source of genetic variation, and as a reservoir of a fungal population genetically different from that prevalent on wheat (de Wolf et al., 1998).

Tan spot has been found in several countries worldwide like Australia, Argentina, Asia, Belgium, Brazil, Bolivia, Canada, Colombia, Czech Republic, Denmark, Ecuador, France, Hungary, Kenya; Pakistan, Paraguay, Peru, Poland, United States of America, Ucrania and Uruguay (Ali & Francl, 2001[a]; Annone, 1985, Conners, 1939; Dubin, 1983; Duveiller et al., 1996; Gilchrist et al., 1984; Hosford, 1981 Kohli et al., 1992; Lamari et al., 2005; Linhares & da Luz, 1994; Loughman et al., 1998; Maraite et al., 2006; Mehta et al., 2004; Moreno et al., 2008; Postnifova & Khasanov, 1998; Sarova et al., 2002; Tekauz, 1976; Sim, and Willis, 1982; Watkins, et al., 1978; Zhanarbekova et al., 2007).

The control of this disease is complex and forms part of the concept of integrated disease management, that is, a combination of chemical , cultural, genetic, and sometimes biological controls (Annone, 2005, 2006; Bockus et al., 1992; Carmona, 2003; Duckez & Jones-Flory, 1994; Hosford & Busch, 1974; Li & Sutton, 1995; Loughman et al., 1998; Luz et al., 1998; Perelló et al., 2003, 2006; Pfender et al., 1989; Simón, 2006; Simón et al 2011; Stover et al., 1996; Sutton & Roke, 1996; Zadoks & Schein, 1979).

In the last 30 years, some investigations in the complex wheat-*P. tritici-repentis* have been deeply developed. In 1989a,b, Lamari and collaborators suggested the presence of races of *P. tritici-repentis* based in the interaction with different cultivars and lines of wheat. Before, some researches had already mentioned the existence of physiological specialization of this pathogen. The fungus produces two distinct symptoms, necrotic (nec) and chlorosis (chl), on susceptible wheat genotypes. These symptoms are the result of the production of multiple host-selective toxins (HSTs).

1.2 The pathogen

Pyrenophora tritici-repentis (Died.) Drechs., is the causal agent of tan spot of wheat leaf. It is a facultative parasite fungus (saprotroph) of Phylum Ascomycota, Class Loculoascomycetes, whose anamorph corresponds to *Drechslera tritici-repentis* (Died.) Shoeman., imperfect fungus of the class Hyphomycetes.

It was first isolated from *Agropyron repens* in Germany and determined as *Pleospora trichostoma* by Diecke. In 1928 was isolated from wheat Nisikado, who called *Helminthosporium tritici-repentis* (= *Drechslera tritici-repentis*) (Hosford, 1981).

1.2.1 Taxonomy

Teleomorph:

Pyrenophora tritici-repentis (Died.) Drechs, J. Agr. Research, 24:667. 1923.

_ *Pleospora tritici-repentis* (Died.), Centr. Bakt. Parasitenk, Abt. 2, 11:56. 1903.

_ *Pleospora trichostoma* (Fr.) Wint. f. *sp tritici-repentis* Died., Centr. Bakt. Parasitenk., Abt. 2, 9:329. 1902.

= *Pyrenophora tritici-vulgaris* Dickson, Diseases of Field Crop, ed. 2, Mc Graw-Hill, New York, 1956. P. 247. nom. nud.

Anamorph:

Drechslera tritici-repentis (Died.) Shoem. Canadian Journal of Botany, 37:880, 1959.

_ *Helminthosporium tritici-repentis* (Died.) Died.; Centr. Bakt, Parasitenk, Abt. 2, 11:56.1903.

_ *Helminthosporium graminearum.* Rob. Ex Schlecht. f. sp. t*ritici-repentis* Died, Centr. Bakt. Parasitenk, Abt. 2, 9:329. 1902.

= *Drechslera tritici-vulgaris* (Nisikado), Ito, Proc. Imp. Acad., Tokio, 6-355, 1930.

_ *Helminthosporium tritici-vulgaris* Nisikado, Ann. Phytopath. Soc. Japan, 2:96. 1928.

1.2.2 Morphological characteristics

The ascocarp (protoseudothecia) of *P. tritici-repentis* is dark brown, spherical to subspherical, smooth, from 400 to 500 μ of diameter. The beak bears a number of dark-brown sterile setae near the ostiole. These setaes has from 80 to 120 μ of lenght and 9 μ of wide. The asci measure usually from 170 to 215 by 43 to 50 μ. The ascospores are brownish colored and measuring usually 47-65 μ of length x 20-26 μ of wide. The ascospores are biseriate and uniseriate in the middle towards the apex respectively (Figure 1a) (Ellis & Waller, 1976).

The conidiophores arising singly or in groups of 2-3, emerging through stomata or between epidermal cells, erect, simple, straight or flexuous, sometimes geniculate, cylindrical or slightly tapered, often swollen at the base, mid-pale to brown, smooth, usually up to 250 μ long but occasionally one or a few rather inconspicuous conidial scars. Conidia solitary, straight or slightly curved, cylindrical, rounded at the apex, the basal segment distinctly and characteristically conical or shape of a snake's head, typically subhyaline or rather pale

straw-coloured, smooth, thin-walled, with 1-9 (usually 5-7) pseudosepta, old conidia often constricted at the pseudosepta, 80-250 (117) µ lon, 14-20 (17.7) µ thick in the broadest part, 2-4 (3) µ wide at the base (Figure 1b). (Ellis & Waller, 1976).
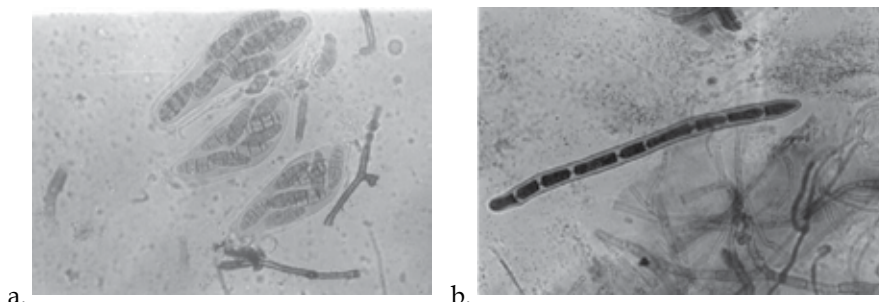

Fig. 1. a: Asci with ascospores b: conidia of *P. tritici-repentis* (=*D. tritici-repentis*) (Moreno 2007)

1.3 Disease cycle

The dispersal and the infection development by *Pyrenophora tritici-repentis* are produce among 10° to 30° C and moisture among 6hs to 48hs (Hosford et al., 1987; Larez et al., 1986; Sah, 1994). These conditions are favorable for the yearly occurrence of tan spot and to distinguish it from other disease, white head that depended of the environmental conditions (Carmona, 2003).

The disease cycle of tan spot (Figure 2) provides a convenient framework to explain the disease progress. The rate of progression through the disease cycle depends upon host, temporal and environmental components of the pathosystem (De Wolf et al., 1998).

The seeds, the straw and collateral hosts are the principal source of inoculum of tan spot. The primary inoculum wade through the wheat areas for long distances and it is introduced in new areas by the seeds. In the seed the pathogen lives in the pericarpio as mycelium and the transmission to plant rest is not systemic (Schilder & Bergstron, 1994). Another source of primary inoculum is the straw of wheat. Several authors considered the straw as the principal source of inoculum (Rees & Platz, 1980). Infested residue usually results in significant disease severity at flag leaf emergence and later growth stages due to secondary infections (McFadden, 1991; McFadden & Harding, 1989; Wright & Sutton, 1990). Collateral hosts could be inoculum source between growing seasons (de Wolf *et al.*, 1998). The tan spot fungus has been reported on many grass species from different parts of the world among these *Agropyron* sp., *Alopecurus arundinaceus*, *Andropogon gerardi*, *Avena fatua*, *A. sativa*, *Bromus inermis*, *Dactilys glomerata*, *Echinochloa* sp., *Elymus innovatus*, *Lolium perenne*, *Phalaris arundinaceae*, *Poa* sp. y *Secale cereale* (Ali & Francl, 2002[b]; Andersen, 1955; Conners, 1939; Dennis and Wakefield, 1946; Dickson, 1956; Diedicke, 1902; Drechsler, 1923; Farr et al., 1989; Hosford, 1971; Howard & Morral, 1975; Krupinsky, 1992[c]; Shoemaker, 1962; Sprague, 1950).

1.4 Host-parasite interactions: Symptomatology of tan spot

On susceptible wheat leaves, *P. tritici-repentis* produces characteristic oval to diamond-shaped lesions. However, newly formed tan spot lesions cannot be separated reliably from those caused by other necrotrophic pathogens. Later, lesions elongate and develop a tan

color with a chlorotic halo and a small dark brown infection site. Chlorotic areas tend to coalesce on heavily infected leaves, especially on young plants, a symptoms which led to the disease name "yellow leaf spot" (Figure 3). On resistant and partially resistant wheat, lesions size is reduced and chlorosis and necrosis may be absent (de Wolf et al., 1998). Lamari & Bernier (1989b) identified two different types of symptoms produced by the pathogen, tan necrosis and extensive chlorosis.
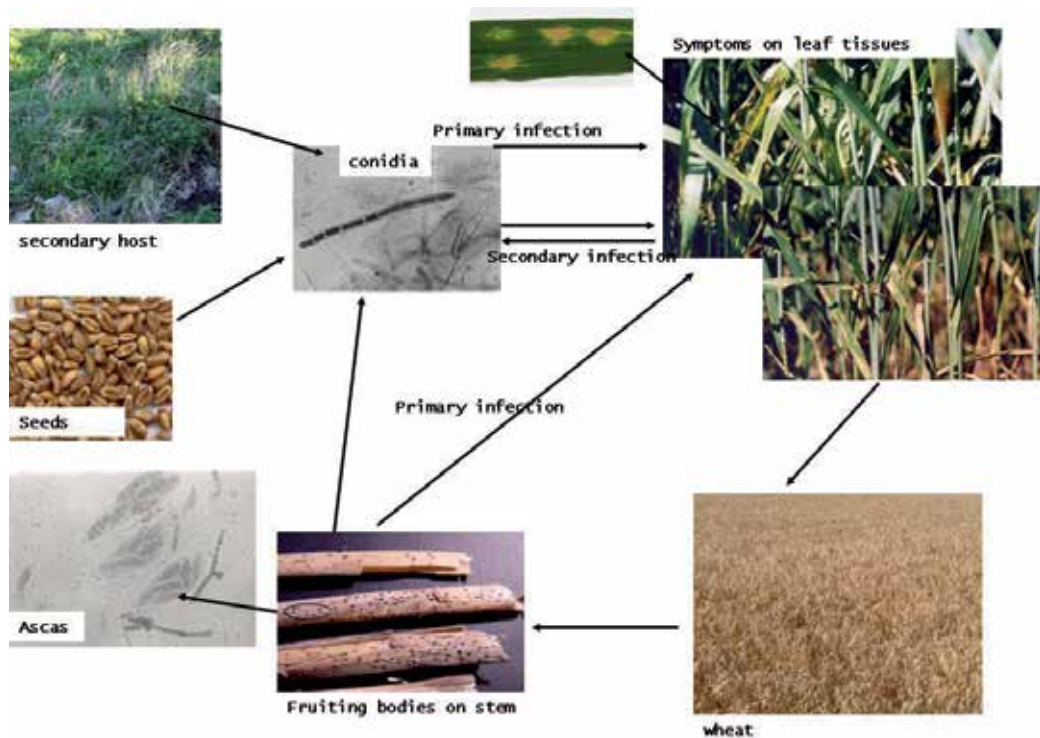

Fig. 2. Disease cycle of tan spot (Moreno & Perelló 2010)

P. tritici-repentis can also infect wheat seed during the grain filling period (Schilder & Bergstrom, 1994). This disorder is called red smudge because infected seed has a reddish discoloration (Valder, 1954).

Fig. 3. Symptoms of tan spot (Moreno & Perelló 2010)

2. The importance of host-selective toxins (HSTs) produced by *Pyrenophora tritici-repentis*

It is known that in the relation gene-for-gene, a pathogen carries a dominant Avr gene and its action result in the induction of the resistance response in a plant that carries the cognate resistance gene (Flor, 1971). Some plant pathogenic fungi produce agents of compatibility, namely host-selective toxins (HSTs). "The genetic of the *P. tritici-repentis*-wheat interaction is reminiscent of gene-for-gene type interactions" (Manning et al., 2004). In the case of *P. tritici-repentis*-wheat the interaction between Tox A and its corresponding host genes results not in resistance but susceptibility (Lamari et al., 2003).

2.1 The host selective toxin produced by *Pyrenophora tritici-repentis*

In recent decades due a great interest on the tan spot of wheat, several research groups worked about composition and identification of host-selective toxin produce by *P. tritici-repentis.* Why is important the characterization of these toxins?. The HSTs are responsible for disease symptoms on susceptible wheat cultivars and lines. At the present, three toxins have been described (Ballance et al., 1989; Ciuffetti et al., 1998; Effertz et al 2002; Martinez et al 2001; 2004; Tomas et al., 1990; Touri et al., 1995; Strelkov et al 1999, 2006). It is commonly accepted that the names of toxin are Ptr Tox A, Ptr Tox B and Ptr Tox C (Ciuffetti et al 1998). The Tox A has been described by several laboratories (Ballance et al., 1989; Ciuffetti et al., 1998; Tomas et al., 1990; Touri et al., 1995). Tomás and Bockus (1987), cited for first time the possible presence of a necrosis-inducing toxin. Five research groups have worked on the isolation and characterization of the Tox A using isolates obtained from Kansas, United State of America (Pt1c) (Ciuffetti et al., 1997; Tomás et al., 1990; Touri et al., 1995) and Canadá (82-124) (Ballance et al., 1989; Zhang et al., 1997). Only three groups have isolated and cloned the gene encoding for Tox A. Ciuffetti and colleagues (1997) used the isolate Pt1c; Zhang et al., (1997) and Ballance et al., (1989) worked about the isolate 82-124. The comparison of these sequences shows differences only for one base. Different studies corroborated the association between toxin production and necrotic symptoms, Lamari and Bernier (1989b), enumerated some points in a) toxin sensitivity correlated with pathogen-induced necrosis in test involving different cultivars, b) sensitivity to the toxin and development of necrosis in response to fungal infection were controlled by a single dominant gene, in all crosses made, c) the isolates nec+ induced necrosis on susceptibles cultivars or lines of wheat, d) those isolates that non produce necrotic symptoms on susceptibles plants did not produce Ptr Tox A *in vitro*. Additionally, Ciuffetti and colleagues (1997), demonstrated that Ptr Tox A functions are the factor of pathogenicity in the interaction wheat- *P. tritici-repentis*, through the transformation of non-pathogenic isolates of *P. tritici-repentis* with the *Tox* A gen. The toxin was shown to be a protein with a weight of a 13.2 kDa and contains a N-terminal pyroglutamate (Ballance et al., 1989; Tomas et al., 1990; Tuori et al., 1995, 2000, Zhang et al., 1997). This toxin is produced by race 1 and 2 (Lamari & Bernier 1989b). Lamari et al., (2003) showed that the isolates classified as race 7 and 8 also synthetized Ptr Tox A.

The pathogen *P. tritici-repentis* is also capable to produce one toxin that induced chlorosis in sensitive lines. This toxin is named Ptr Tox B (Ciuffetti et al., 1998), and was initially identified in culture filtrate by Orolaza et al., (1995). Previouslly, Brown and Hunger (1993), reported a phytotoxin that induced chlorosis in susceptible wheat genotypes. In the same work was

Pyrenophora tritici-repentis, Causal Agent
of Tan Spot: A Review of Intraspecific Genetic Diversity
287

determined that this toxin to has a molecular mass between 800 and 1800 kDa. Orolaza et al., (1995) reported isolates obtained from Algerian that produced chlorosis and these isolates were classified as race 5 (Alg. 3-24). Strelkov et al., (1999) purified and characterized this toxin as a small protein (6.61kDa). Ptr Tox B is encoded by a multiple-copy gene, *Tox B*, comprising a 261-bp open reading frame (ORF) (Lamari et al., 2003; Martinez et al., 2001; Strelkov et al., 2002, Strelkov & Lamari, 2003). The races 7 and 8 produced Ptr Tox B inducing chlorosis on the cultivar Katepwa (Lamari et al., 2003; Strelkov et al., 2002).

Several studies predicted the existence of one third toxin (Gamba et al., 1998; Lamari & Bernier, 1991b), named Tox C. This toxin also produced chlorosis but in different cultivars than Ptr Tox B did. Effertz et al., (2002), proved its production in culture from the isolate of race 1 and it was partially purified. The production of this toxin by isolates classified as races 3, 6 and 8 was proved by several researches (Gamba et al., 1998; Lamari & Bernier, 1991; Lamari et al., 2003; Strelkov et al., 2002;). This toxin is not proteinaceous in the nature, but rather seems to be a nonionic, polar, low molecular mass molecule (Effertz et al., 2002).

At the present, Pandelova and Ciuffetti (2005) suggested that the necrotic symptoms produced by the isolates SO3 are caused by a previously undescribed proteinaceous. They have a tentative name for this toxin Ptr Tox D, in agree with the standard nomenclature published by Ciuffetti et al., (1998).

In summary, the HSTs produced by different isolates of *P. tritici-repentis* are Ptr Tox A, Ptr Tox B (and its homologs), Ptr Tox C, and a potential Ptr Tox D. At the present the studies that combined phenotypic and genotypic determination of races of *P. tritici-repentis* suggested the presence of new races of *P. tritici-repentis* that potentially produce new toxic components (Ali et al., 2010; Andrie et al., 2007; Moreno et al., (unpublished data)).

2.2 Physiological races of *Pyrenophora tritici-repentis*

The ability of an organism to produce disease can be evaluated as pathogenicity or virulence. These terms can differenciate because the pathogenicity is a general attribute of species and virulence is an attribute reserved for a particular isolate of a pathogen in relation of particular host genotype (Day, 1960). Other option according to Van der Plank (1978, 1984), considers that specificity in host-pathogen relationships is often indicated by significant isolate x cultivar interaction in the analysis of variance of an experiment where a number of pathogen isolates are tested in all possible combinations on a set of host genotypes. Non-specificity is identified by a lack of such interaction. Variation in virulence in the population of this pathogen is essential in understanding the interaction of the genomes involved in tan spot. Studies of the diversity of virulence within a pathogen population should help in the development of a successful disease management program, particularly resistant cultivars.

Basing in the host-pathogen interaction and in the reaction type on cultivars and lines of wheat, the complex wheat- *P. tritici-repentis* has been classified as a complex system for the determination of physiological specialization of the causal agent of tan spot (Lamari & Bernier, 1989a,b). Before the proposed of Lamari and Bernier (1989a), several works showed differences on virulence for *P. tritici-repentis* isolates (Cox & Hosford, 1987; Diaz de Ackermann 1987; Gilchrist et al., 1984; Krupinsky 1987, 1992 a,b; Luz & Hosford 1980; Misra & Singh 1972; Nagle et al., 1982;). In 1971, Hosford observed differences between the

reaction types on wheat cultivars produced by isolates of *P.tritici-repentis*. Misra and Singh (1972) tested isolates originated from India and detected significant differences in virulence, based on lesion size. Similar results were observed by Glichrist et al., (1984) when they tested isolates collected from Mexico on the wheat cultivar Morocco. Luz and Hosford (1980) grouped the isolates tested into 12 races based on statistical mean separation. However, Díaz de Ackermann et al., (1988) did not find differences in virulence among the isolates tested by Luz and Hosford (1980). Hunger and Brown (1987) tested nine isolates originated from USA these isolates showed significant differences on the susceptible cultivar TAM 105. Krupinsky (1987) showed differences in length lesion and percentage of severity, among isolates of *P. tritici-repentis* obtained from *Bromus inermis*.

Basing in the reaction type produced by 92 isolates of *P. tritici-repentis* obtained from Canada on 11 wheat cultivars, Lamari and Bernier (1989a) reported specific interactions between cultivars and isolates, indicating that these isolates differed in virulence. In this work they proposed the presence of three pathotypes. Hosford et al., (1990) suggested the presence of extensive chlorosis that could fit the chlorotic symptoms reported by Lamari and Bernier (1989a). Schilder and Bergstrom (1990) tested 70 isolates obtained from Canada on 12 wheat cultivars and detected significant differences among the interaction isolate x cultivar. Similar results were reported by Sah and Fehrmann (1992) for isolates origin from Brazil, Germany, India, Nepal and USA. However, Ali and Buchneau (1992) observed physiological specialization based in the reaction type for isolates obtained from USA. Mehta et al., (2004) tested 40 isolates obtained from Parana (Brazil) on six cultivars of wheat, and observed low interaction for isolate x cultivar. In 2007, Moreno detected significant differences isolate x cultivar for isolates of *P. tritici-repentis* obtained from different grown wheat areas in Argentina.

Actually, the worldwide knowledge indicates the existence of 10 races for *P. tritici-repentis* evaluated on a set international of wheat cultivars (Ali & Francl, 2001$_{a,b}$, 2002$_b$; Lamari & Bernier, 1989$_{a,b}$; Lamari & Gilbert 1998; Lamari et al., 1995, 1998, 2003, 2005).

2.3 Determination and identification of *Pyrenophora tritici-repentis* races

Lamari and Bernier (1989b), suggested that individual isolates of *P.tritici-repentis* induced individual symptoms on specific cultivar/lines of wheat. In this work, they made the following rating system base on lesion type recognized previously by Hosford (1971) and Gilchrist et al., (1984). The rating systems consisted in: 1= Small dark brown to black spots without any surrounding chlorosis or tan necrosis (resistant); 2= Small dark brown to black spots with very little chlorosis or tan necrosis (moderately resistant); 3= Small dark brown to black spots completely surrounded by a distinct chlorotic or tan necrosic ring, lesion generally not coalescing (moderately resistant to moderately susceptible); 4= Small dark brown or blacks spot completely surrounded with chlorotic or tan necrotic zones, some of the lesions coalescing (moderately susceptible); 5= The dark brown or black centres may or may not be distinguishable, most lesions consist of coalescing chlorotic or tan necrotic zones (susceptible). In other work (Lamari and Bernier 1989a), they taken the same rating system and suggested the presence of four pathotypes based in the results obtained according with specific interactions between wheat cultivars and individual isolates of *P. tritici-repentis* as follow: Pathotype 1=tan necrosis or extensive chlorosis on differential cultivars (nec+ chl+); pathotype 2= only tan necrosis (nec+ chl-); pathotype 3= only extensive chlorosis (nec- chl+)

and pathotype 4= avirulent (nec- chl-). Lamari et al., (1995) suggested the incorporation of a new pathotype that corresponding with the isolates obtained from Algerian. In this work they proposed the pathotype/race classification of *P. tritici-repentis*.

The history of determination of races of *P. tritici-repentis* has been well documented (Ali et al., 1990; Ali & Francl, 2002a,b, 2003; Lamari & Bernier 1989a,b; Lamari et al., 1995, 1998; Strelkov et al., 2002). It is commonly accepted the existence of eight races that it have been characterized on three effective host differential of wheat (Lamari et al., 2003). Although, in some studies the authors proposed the existence of two races more (Ali & Francl, 2002a,b).

In summary, the races 1, 2, 3, 4 and 5 of *P. tritici-repentis* corresponds with those determined by Lamari et al., (1995). The races 1 and 2 are predominant in North America (Ali & Francl, 2003). The greater part of isolates identified as race 5, are originated from North Africa, North America and Algerian (Ali et al., 1990; Ali & Francl, 2003; Lamari et al., 1995, 1998; Strelkov et al., 2002). The races 6, 7 and 8 were identified from collections originated from Azerbaïdjan (Caucasus), Syria, Turkey and South America (Ali & Francl, 2002a; Lamari et al., 2003; Strelkov et al., 2002). By other hand, the races 9 and 10 were identified from isolates originated from South America (Ali & Francl, 2002a,b). The results of these studies suggested that the variation in the virulence of population of *P. tritici-repentis*, can be detected through quantitative and/or qualitative rating scales.

The race characterization based only on phenotypic features could carry some mistakes. Andrie et al., (2007), proposed minimized these errors using a combination of phenotypic and genotipic characterization for *P. tritici-repentis* race identification.

At the present, the genotypic characterization of *P. tritici-repentis* population is wide due the incorporation of molecular tools and massive use of Polymerasa Chain Reaction (PCR). Moreover, the availability of sequences of *Tox* A, *Tox* B and *tox* b gene sequences, allowed by many researchers as Andrie et al., (2007) designed a multiplex PCR races of *P. tritici-repentis*. Its known that the virulence patterns of *P. tritici-repentis* are based on a toxin production and each compatible interaction between an isolate and its corresponding susceptible host cultivar. In this way, we considered the race structure proposed by Andrie et al., (2007), which was based in different studies (Lamari et al., 1995; 2003; Strelkov et al., 2002) (Table 2).

Races	Glenlea	Katepwa	6B662	6B365	Salomouni	M3
1	N (Tox A)	N (Tox A)	R	Cl (Tox C)	R	R
2	N (Tox A)	N (Tox A)	R	R	R	R
3	R	R	R	Cl (Tox C)	R	R
4	R	R	R	R	R	R
5	R	Cl (Tox B)	Cl (Tox B)	R	R	R
6	R	Cl (Tox B)	Cl (Tox B)	Cl (Tox C)	R	R
7	N (Tox A)	N (Tox A) Cl (Tox B)	Cl (Tox B)	R	R	R
8	N (Tox A)	N (Tox A) Cl (Tox B)	Cl (Tox B)	Cl (Tox C)	R	R

References: N= necrosis. C= chlorosis, R= resistance, Tox A= presence fragment of *Tox A* gene, Tox B= presence fragment of *Tox B* gene, Tox C= presence of Tox C as evidenced by inoculation.

Table 2. *Pyrenophora tritici-repentis* pathotype/race classification according Andrie et al., (2007).

3. Genetic variability of plant pathogenic fungi

During the last 35 years the genetic of fungal pathogen population of different crops have been extensively analysed (Bayraktar et al., 2007; Clay, 1995; Dinolfo et al., 2010; Goodwing et al., 1995; Stenglein & Balatti 2006; Stenglein et al., 2010; Mc Dermot & Mc Donald, 1993; Moreno et al., 2008, 2009; Tóth et al., 2004). The genetic variability is frequent among plant pathogen populations and its understanding can aid to diseases management decisions (Brown & Wolfe, 1990; Clay, 1995; Milgroon et al., 1992). The plant pathogenic fungi included a heterogeneous group of organisms occuping positions of great importance in both agriculture and plant communities. The range of pathogens found on different hosts also shows considerable diversity that may be associated with the evolutionary history of their host (Clay, 1995) or ecological criteria like the host's architectural complexity or the extent of the pathogen's natural range (Strong & Levin, 1979). The plant pathogenic fungi rely on the processes of mutation and recombination as the ultimate source of genetically based variation (Burdon & Silk, 1997). Burdon and Silk (1997), suggested the different sources and patterns of diversity in plant pathogenic fungi (mutation, migration, gene flow and recombination). The frequency of mutation is highly variable and dependent of rate mutation and ploidy level of pathogen, the size of pathogen population and the selective advantage conferred by mutant phenotype (Burdon, 1997). For example, the origin of the *Puccinia striiformis* in Australia (Welling & Mc Intosh, 1990) is product of mutation process; also, *Phytophtora infestans* in United States (Goodwing et al., 1995). Another mechanism by which one pathogen can be differentiate one pathotype is gene flow, one example of this case was the introduccion of *Puccinia graminis* f. sp. *tritici* in Australia (Burdon & Silk, 1997). The installation of new areas crops can produce fluctuations in the genetic structure of populations of hosts as well as from pathogens (Mc Dermot & Mc Donald, 1993). The level of genetic diversity in fungi is usually the highest in isolates obtained from the center of origin. The genetic variability is a present condition in wild pathosystem that is incrementing in the agroecosystems. Different views (asexual and sexual) that have many of the fungal phytopathogen, provide variability and may affect the genetic structure of a population (Brown & Wolfe, 1990; Milgroon et al., 1992). Those fungi that have the ability to reproduce sexually can produce genetically more variable populations than those that only make it asexually (Goodwin et al., 1992).

Why is important to study the genetic structure of pathogen fungal population of plants? The population genetic structure has been show to be important in the management of a fungal plant disease. Determination of the genetic variability in organisms such as fungal pathogens should be an ongoing task for successful breeding programs. The loss of resistance in commercial cultivars, produced by the emergence of new unknown pathotypes has been documented (Araya & Cárdenas 1999; Young & Kelly 1997). It is important to know the existence of races for different pathogens and the relationship of each one with the host, in order to detect genes for resistance. Genetic variation among isolates of a fungus has been examined through a variety of techniques (Arenal et al., 2002; Dorrance et al., 1999; Moreno et al., 2008a,b; Moriwaki et al., 2003; Pioli et al., 2003). Methods commonly used to assess variation in fungal plant pathogen population include virulence, vegetative compatibility, biochemical and molecular analysis.

3.1 Current status of *Pyrenophora tritici-repentis*

It is recognized that *P. tritici-repentis* presents great variability in its genome and pathogenicity (de Wolf et al., 1998). Several studies included tools for detection variability

using of morphocultural, pathogenic, bioquemical and molecular variability. Several studies to detect variability include different tools, the most of them used pathogenic and genetic characterization (PCR techniques) (Ali et al., 2010; Andrie et al., 2007; Friesen et al., 2005; Leisová et al., 2008; Lepoint et al., 2010; Litcher et al., 2002; Mehta et al., 2004; Moreno et al., 2009; Pujol Vieira dos Santos et al 2002; Singh & Hughes 2006). A lesser extent of research focuses on characterization based on morphological and biochemical aspects (Diaz de Ackermann 1987; Gilchrist et al., 1984; Hosford, 1971; Hunger & Brown 1987, Moreno et al., 2008a; Pujol Vieira dos Santos et al., 2002). It has been noted in these studies that isolates of *P. tritici-repentis* vary greatly.

There are numerous works that refer to the population structure of *P. tritici-repentis* in races (Ali & Francl 2002, 2003; Ali et al., 2002; 2010; Andrie et al., 2007; Lepoint et al., 2010; Manning et al., 2002; Strelkov et al., 2002; Lamari et al., 2003; 2005). Some of these papers involve aspects as morpholocultural and virulence variability (Ali et al., 2010). In general in the mostly combined with molecular markers and PCR (Ali & Francl 2002, 2003; Ali et al., 2010; Andrie et al., 2007; Lepoint et al., 2010; Manning et al., 2002; Strelkov et al., 2002; Lamari et al., 2003; 2005).

Several molecular methods have been used to analyze the diversity of plant pathogens at the genome level, such as random amplified polymorphic DNA (RAPD), amplified fragments length polymorphism (AFLP), enterobacterial intergenic consensus (ERIC), repetitive extragenic palindromic (REP), simple sequence repeats (SSRs) and inter simple sequence repeats (ISSR) (Mehta 2001; Mehta et al., 2004; Pujol Vieira dos Santos et al., 2002; Stenglein & Balatti, 2006).

A preliminary study with Brazilian isolates of *P. tritici-repentis* demonstrated high levels of polymorphisms, but no correlation between RAPD, geographic origin and/or pathogenic data (Pujol Vieira dos Santos et al., 2002). The pathogenic and molecular analyses (ISSR) showed intraspecific variability within *P. tritici-repentis* isolates and it was not possible to establish a relationship between this variability and the geographical regions and/ or wheat cultivars from where *P. tritici-repentis* isolates were obtained (Moreno et al., 2008a). Friesen et al., (2005) using AFLP markers determined a high level of variability among isolates of *P. tritici-repentis* and these results shown no genetic grouping of pathotypes or grouping for geographic location. According to Pujol Vieira dos Santos et al., (2002), there are many factors that could affect polymorphism analysis, e. g. the intra-specific variants of a pathogen, the number of samples selected for analysis, genetic flow between populations, environmental adaptation and adaptation to a new host and selective pressure and migration. On the other hand, *P. tritici-repentis* is a homothallic fungus that readily produces the sexual stage on field stubble, giving this fungal population the opportunity for adaptation by sexual recombination. The occurrence of sexual recombination in nature is likely the reason for the high level of genetic variability among isolates of *P. tritici-repentis* (Singh & Hughes, 2006). Some works have studied the homothallic nature of *P. tritici-repentis* using the specific molecular markers for detection of the MAT genes (Lepoint et al., 2010).Under favourable conditions conidiospores can travel 10-200 km (de Wolf et al., 1998). The tan spot fungus is also seed-borne and thus, long distance travel of fungal inoculum can occur by seed transmission (Singh & Hughes, 2006). Hence, we can assume that the occurrence the sexual state and long distance dispersal of inoculum could have contributed to pathogenic and genetic variability independent of geographic origin.

Recently, primers to detect the gene that codified for specific toxins Ptr ToxA, Ptr ToxB and Ptr toxb have been developed and used for many studies, and several works have enunciated that the classification in races of *P. tritici-repentis* based only on phenotypic symptoms is in disagreement with the obtained results of genotypic analysis (Ali et al., 2010; Andrie et al., 2007).

3.2 Isozymes

Isozyme electrophoretic analysis is a powerful biochemical technique used by mycologist and plant pathologist to asses the genetic variation within and among fungal populations, to trace the geographic origins of pathogens, and to determine ploidy levels in fungi (Brammer, 2000; Bonde et al., 1993; Castro & Bach, 1993; Mc Donald & Mc Dermott, 1993; Micales et al., 1998, 1986; Reynolds et al., 1983). Isozyme analysis has proved to be a successful tool in studying the evolution of races in other fungal species (Bocchese et al., 2003; Burdon & Roelfs, 1985; Fry et al., 1991; Koch & Kölher, 1990). In several of these studies, isozyme markers were of particular helpfulness, because no other genetic markers were available or because the only available markers, viz, virulence genes, were not selectively neutral (Welz et al., 1994).

With isozymes, a large numbers of staining systems can be used for the comparison of the numerous genetic loci coding for enzymes from many metabolic pathways (Bonde et al., 1993). This has the advantage of allow us to draw conclusions about the genetic variability existing within and between species under investigation (Bonde et al., 1993). The selection of enzymes to study is an important part of the enzyme analysis. Some strains, such esterases (EST), phosphatases (ACP) and peroxidases (POX), are not substrate specific and detect entire groups of enzymes (Bonde et al., 1993). These enzyme systems have been used by others authors in previous studies (Chen, 1992; Elias & Schneider, 1992; Petrunak & Christ, 1992).

The analysis of isozyme data showed different resolution for the three enzyme systems used among the 155 *P. tritici-repentis* isolates (Moreno et al., 2008a). Poor resolution was observed for the POX enzyme, while EST and ACP enzymes showed strong and acceptable resolution. *P. tritici-repentis* isolates variability was shown by the analysis of the different isozyme patterns. The highest indications of variability were found with the esterases and acid-phosphatase systems. Peroxidase was less useful to distinguish differences between isolates. The variability within a worldwide collection of isolates *P. tritici-repentis* was studied using pathogenicity test (Ali & Francl, 2001; Lamari & Bernier, 1989a, b; Lamari et al., 1998, 1995; Luz da & Hosford, 1980). According to Ali and Francl (2002) *P. tritici-repentis* has a diverse population in South America. In this study, our collection of *P. tritici-repentis* isolates from Argentina showed a considerable isozyme variation for two of the three systems assayed. This fungus is the predominant necrotrophic pathogen of wheat crops in Argentina and an increase in occurrence of this disease is reported since the least decade (Carmona et al., 1999). Isozyme analysis has been recommended as a possible assistance in the systematic and genetic of fungi (Garber & Behara, 1966; Faure-Raynaud et al., 1987). This technique has also been assayed to characterize interespecific and intraspecific polymorphisms of fungi (Anné-Peberdy, 1981; Boshoff, 1996; Bridge et al., 1986). The use of isozymes as markers for estimating the amount of variability within fungal populations is well documented (Mc Donald & Mc Dermott, 1993; Micales et al., 1998). The lack of genetic

variation within fungal pathogens has previously been correlated with various factors. Newton (1987) reported that a specialized obligate parasite should exhibit little variation due to the uniformity of its substrate whereas facultative pathogens, which are more exposed to diverse substrates and environment will be variable. *P. tritici-repentis* has been reported from or observed on several hosts (Ali & Francl, 2002; Hosford, 1971; Krupinsky, 1992). The isozyme assay revealed considerable diversity among the *P. tritici-repentis* isolates assayed than would be expected from a homothallic fungus (Linde et al., 1997). Matsumura (1991) showed variability among isolates of *Bipolaris sorokiniana* from the same wheat cultivar and from different cultivars. Valim-Labres (1997) observed that the host determined the genotype pathogen. In *Uromyces appendiculatus*, *Gliocladium* sp. and *Pythium* sp. populations, isozyme polymorphisms were observed (Chen et al., 1992; McCain & Groth, 1992). In other pathogens as *Colletotrichum acutatum* and *C. fragariae*, the variability observed was low but high in *C. gloesosporioides* (Bonde et al., 1991).

In several studies, this heterogenicity was associated with the parasitic type exerced by the pathogen (Chen et al., 1992; Elias & Schneider, 1992; Petrunak & Chrits, 1992). Those pathogens with a saprophytic stage have a high diversity of enzymes for use on different substrates (Petrunak & Chrits, 1992). Dorrance (1999) observed that the obligate pathogens showed low levels of isozyme polymorphisms. This observation is attributed because these pathogens live on more uniform substrates than facultative ones. *P. tritici-repentis* is a facultative pathogen whose pathogenic stage can be observed on diverse environments and broather hosts range, like different gramineous species. These observations could explain the results observed in this study for *P. tritici-repentis* for EST and PHOSP systems. The same considerations were proposed by Pujol Vieira dos Santos et al., (2004) for *Bipolaris sorokiniana*. Moreno et al (2008a), showed for *P. tritici-repentis* no association among bands and origin of the isolates. This result was observed by others authors for other pathogens (Chen et al., 1992; Hellmann & Christ, 1991; Pujol Vieira dos Santos et al., 2004). Hellmann and Christ (1991) observed that *Ustilago hordei* isolates from USA and Etiopia, showed the same isozyme pattern. By other hand, Chen et al., (1992) observed groups of *Pythium graminicola* isolates originated from the same localities, but also groups of *P. arrhenomanes* isolates originated from different localities in USA. Bocchese et al., (2003) did not observed relation among isozyme patterns and origin of *Pyrenophora avenae* isolates. Moreno et al., (2008a) suggested further studies of esterases for *P. tritici-repentis*, because within this isozymes, the carboxylesterases, have relation with resistance mechanisms to insecticides, cupper fungicides, plaguicides and fosforate compounds (Bisset, 2002). Furthermore, it would be of interest to study this system in detail due the high polymorphisms observed for argentinian *P. tritici-repentis* isolates. Moreover, Nicholson et al., (1972) showed that microorganisms as *Xanthomonas axonopodis* pv. *manihotis*; *Botrytis cinerea* and *Venturia inaequalis* presented estearase activity near penetration site. Pascholati et al., (1992) observed correlation between presence of estearase and conidia of *Erysiphe graminis*. Bocchese et al., (2003) observed correlation between EST isozyme patterns and virulence of *Pyrenophora chaetomioides*, causal agent of foliar disease on oat. In Argentina, only a few reports are available about electromorphic types of *P. tritici-repentis* based in isozyme analysis (Moreno et al., 2008a).

3.3 Molecular markers

Since 1990, the molecular biology through the use of the PCR technique facilitated the availability of different tools for identification and characterization with high precision

different microorganisms. The PCR technique was developed by Mullis in 1983 (Mullis, 1990). The DNA fingerprinting consists in the display of a set of DNA fragments from a specific DNA sample (Vos et al 1995). It is used to visualized and identify DNA polymorphisms between individuals. This technique is used to determining the identity of a specific, to identify molecular markers linked to phenotypic traits and/or genetic loci to generate a linkage map. It is known that the PCR, use either specific of arbitrary primers to develop the *in vitro* amplification of template DNA. The different amplification patterns are a function of the primer sequence and the nature of the template DNA. The molecular markers have been used to evaluate genetic diversity in and between species of fungal plant pathogens (Kalkar et al., 2006; Moreno et al., 2009; Saharan & Naef 2008; Stenglein & Balatti, 2006; Stenglein et al., 2010; Yli-Mattila et al., 2004). DNA fingerprint allows visualize and indentify DNA polymorphisms between fungi.

Strelkov and Lamari (2003), published a review where indicated the new advances in the study of wheat- *P. tritici-repentis* interaction from 1999 to 2003. In this review they suggested the existence of eight races of *P. tritici-repentis* determined on four lines of wheat (Glenlea, 6B662, 6B365 and Salomouni). They considered that the molecular basis of virulence has been explained, with the identification of three toxins. The reaction of the host to pathogen is controlled by single dominant and independently inherent genes, suggesting a one-to-one relationship between wheat- *P. tritici-repentis*. In the same work they highlighted the importance that the genes codified for these toxins (*Tox* A, *Tox* B, *tox* b) have been cloned and provided the basis for further works. Also, Singh et al., (2010) published a review with the last advances in the characterization of wheat- *P. tritici-repentis* interactions. None of three reviews proposed a point about the genetic intraspecific variability of *P. tritici-repentis* utilizing the use of molecular markers. However, at present exists different studies where the aim consists in the characterization of genetic variability of *Pyrenophora tritici-repentis* population in different localities, areas and countries. (Ali et al., 2010; Andrie et al., 2007; Friesen et al., 2005; Leisová et al., 2008; Lepoint et al., 2010; Litcher et al., 2002; Mehta et al., 2004; Moreno et al., 2009; Pujol Vieira dos Santos et al 2002; Singh & Hughes, 2006).

During the last ten years, several isolates obtained from the both same and different area of *P. tritici-repentis,* have been study with different molecular markers. Within them Amplified Fragment Length Polymorphism (AFLP) (Friesen et al., 2005; Leisová et al., 2008), Enterobacterial Repetitive Intergenic Consensus (ERIC) and Repetitive Extragenic Palindromic (REP)- PCR technique (Mehta et al., 2004), Internal Transcribed Spacer Regions (ITS) (Lepoint et al., 2010), Mating Type locus (MAT) (Lepoint et al., 2010), Inter Simple Sequence Repeat (ISSR) (Moreno et al., 2009) and Random Amplified Polymorphisms DNA (RAPD) (Pujol Vieira dos Santos et al., 2002; Singh & Hughes, 2006), Restriction Fragment Length Polymorphism (RFLP), Toxin production genes (*Tox* A, *Tox* B and *tox* b genes) (Ali et al., 2010; Andrie et al., 2007; Lepoint et al., 2010).

3.3.1 Restriction Fragment Length Polymorphism (RFLP)

During the past few years also work with the procedure known as Restriction Fragment Length Polymorphism (RFLP) is used for characterization of plant pathogen fungi. Isolates of different species of have been characterized used this methodology (Arabi & Jawhar, 2007; Wu et al., 2003; Zein et al., 2010). This procedure was developed by Botstein et al.,

(1980). The strategy of this technique is the hybridization: consist in the detection of fragment of genomic DNA cut with restriction endonucleases, separated based on their size by electrophoresis and identified by Southern Hybridization with specific labelled probes. This technique offers an alternative means by which genetic diversity can be measure in pathogens (Kohli et al., 1992; Mc Donald & Mc Dermott, 1993; Milgroom et al., 1992; Mueller et al., 1996; Peever & Milgroom, 1994; Zhong & Steffenson, 2001). Wu et al., (2002), showed that every *P. teres* isolates obtained from geographically diverse collection, exhibited a unique RFLP pattern. However, the same isolates showed a broader spectrum and a higher level of virulence on the host differentials (Wu et al., 2003). Recently, Zein et al., (2010), analyzed isolates of *P. graminea*, the causal agent of leaf stripe disease obtained from Syria, using ITS-RFLP. They observed the fingerprints generated from the six restriction digestions of the DNA ITS region showed high levels of intraspecific variation within the *P. graminea* population. Zein et al., (2010), used IRAP (inter-retrotransposon amplified polymorphism) and ITS-RFLP, each marker system could discriminate between all of the isolates in detecting polymorphism. Similar studies with a global collection of *P. tritici-repentis* could be developed to provide information for future selection of isolates to develop durable wheat tan spot resistance.

3.3.2 Amplified Fragment Length Polymorphisms (AFLP)

AFLP, have been used in different works for characterize the population of different *Pyrenophora* species (Friesen et al., 2005; Leisová et al., 2005, 2008; Serenius et al., 2007). This procedure is based on the selective amplification of genomic restriction fragments. The AFLP techniques, generated fingerprint without previous knowledge of the sequence using a limited set of generic primers (Vos et al., 1995). It is a robust and reliable technique, based on the high reproducibility, and on the homology between bands of the same weight. In the case of causal agent of tan spot, only two researches worked with this technique (Bentata et al., 2011; Friesen et al., 2005; Leisová et al., 2008). Friesen et al., (2005), presented a study of genetic diversity of global collection of *P. tritici-repentis*. They analyzed different isolates (97) obtained from Canada, Czech Republic, United States of America (Arkansas, Kansas and Oklahoma) and South America (Argentina, Brazil and Uruguay). They observed that the level of genetic diversity detected by AFLP was "*reasonably*" high, were not able to group isolates of fungal races or geographic location. They suggested that these results were obtained because the fungi have preferentially out crossing in nature and the fungus possibly can be travel in seed movement between continents. Similar study was developed by Leisová et al., (2008). They analyzed fingerprinting resulted of AFLP analyses of 100 *P. tritici-repentis* isolates obtained from Canada, Czech Republic, Russia, Slovak Republic, United States of America and Argentina (South America). They observed that the most variability occurred within local population, and among population this variability is lesser. They attributed these results to sexual reproduction and that *P. tritici-repentis* is a seed pathogen and it can travel long distances. The studies suggested that the AFLP technique is appropriate to detect the intraspecific variability for *P. tritici-repentis* population, and that the genetic variability is most than race structure or/and geographic origin. This type of molecular markers did not reveal a positive correlation among race and geographic origin with identified haplotypes. Similar number of isolates obtained from of different wheat area around the world could be compared with this technique.

3.3.3 Random Amplified Polymorphisms DNA (RAPD)

RAPD-PCR is performed at low annealing temperatures in order to allow primers to anneal to not fully homologous loci. The RAPD technique is quick, effective and produces reliable markers which have been used for the identification of various fungal pathogens. RAPD analyzes the entire genome the fungus, this technique has been extensively used for characterization of plant pathogen fungi (Assigbetse et al., 1994; Bayraktar et al., 2007; Guthrie et al., 1992; Tóth et al., 2004; Pujol Vieira dos Santos et al., 2002; Di Zinno et al., 1998; Singh & Hughes, 2006). In general, it is useful in detecting inter and intra-specific variation. Within the genera *Pyrenophora* the use of RAPD to detect intraspecific variability was used by several researches (Campbell & Crous, 2002; Mehta, 2001; Peever & Milgroom, 1994; Pujol Vieira dos Santos et al., 2002; Singh & Hughes, 2006). Mehta (2001), observed intraspecific variability among isolates of *Dreschlera avenae* obtained from both black spot and leaf spot symptoms. Mehta (2001) suggested that the difference observed between the isolates originating from two types of symptoms is due to intraspecific variants of *D. avenae*. Pujol Vieira dos Santos et al., (2002), analyzed the intraspecific diversity of *P. tritici-repentis* isolates obtained from seeds in Brazil. The RAPD analyses showed intraspecific polymorphisms within the isolates, but it was not possible to establish a relationship between these polymorphisms and the geographical regions where the host seeds were collected. Guthrie et al., (1992) with isolates of *Colletotrichum graminicola* and Assigbetse et al., (1994) with isolates of *Fusarium oxysporum* f. sp. *vasinfectum* were able to detect a relationship between intraspecific polymorphisms and geographic location. Sing & Hughes (2006), published the results obtained from RAPD analyzes of *P. tritici-repentis* isolates. They observed, the level of polymorphisms detected of analysis of molecular variation (AMOVA) was of 96.8%among isolates, but none differentiated isolates of *P. tritici-repentis* on the basis of race classification or geographical origin. Several studies (Crous et al., 1995; Di Zinno et al., 1998; Pujol Vieira dos Santos et al., 2002; Mehta, 2001; Singh & Hughes, 2006) and highlighted the high level of genetic variability among isolates of *P. tritici-repentis* and that this genetic variability was independent of race classification or geographical origin.

3.3.4 Enterobacterial Repetitive Intergenic Consensus (ERIC) and Repetitive extragenic palindromic (rep-PCR)

The ERIC and REP-PCR technique is a rapid and highly reproducible method that involves PCR amplification using primers that correspond to ERIC and REP elements (Versalovic et al., 1991). Identifying families of repetitive DNA sequences has been show also to be an useful and reliable strategy to determine the genetic relationships within groups of microorganisms. Although rep-PCR are used to identify highly conserved, repeated nucleotide sequences in bacterial DNA (Hulton et al., 1991), Gillings and Holley (1997) reported that rep-PCR fingerprinting can also amplify other regions and even non-bacterial sequences of the bacterial genome . Despite this limitation, rep-PCR fingerprinting has been used to analyze the DNA of various fungal species (Edel et al., 1995; Mehta et al 2004; Reynaldi et al., 2003). Mehta et al., (2004) working with 40 isolates of *P. tritici-repentis*, and showed that the ERIC-PCR and rep-POCR patterns for almost all the isolates were homogeneous, suggesting that the genetic variability of *P. tritici-repentis* population in this area in Brazil (State of Paraná) was small. However, Mehta et al., (2002), using ERIC and rep-PCR to analyze de genetic variability among isolates of *Stemphyllum solani*, found clear

differences between the cotton and tomato isolates as well between the tomato isolates from different geographic regions. Redondo et al., (2009) applied these techniques to discriminate *P. glabrum*, *P. adametzioides* and *P. thomii* from other *Penicillium* species. In Argentina, future studies should be order to validate the ability of these techniques to distinguish between races of *P. tritici-repentis*.

3.3.5 Inter Simple Sequence Repeat (ISSR)

ISSR consists of the amplification of DNA sequences between SSR by means of anchored or non-anchored SSR homologous primers (Zietkiewicz et al., 1994). SSR are tandem repeat motifs composed of one to six nucleotides, which are ubiquitous, abundant and highly polymorphic in most eukaryotic genomes (Tautz & Renz 1984). This technique involves the use of one primer complementary to a target microsatellite region in PCR and allows DNA amplification between microsatellite regions. ISSR does not require a previous knowledge of the sequence and generates specific and reproducible patterns due to the high stringent conditions of annealing (Bornet & Branchard, 2001). Several authors has been using this technique to determine the genetic variability in plant pathogenic fungi (Camacho & Liston, 2001; Dinolfo et al., 2010; Moreno et al., 2009; Stenglein & Balatti, 2006; Statkevi´ciute et al., 2010). Within Ascomycota fungi, Stenglein and Balatti (2006), observed a high genetic variability among isolates of *Phaeoisariopsis griseola*. Regarding the genera *Pyrenophora*, Statkevi´ciute et al., (2010) used ISSR to determine the genetic diversity of *P. teres* obtained from Lithuania. They observed that all isolates had a unique ISSR haplotype. However, Moreno et al., (2008b) observed several ISSR haplotypes for *P. tritici-repentis* isolates obtained from Argentina, using five of 28 ISSR primers tested. Dinolfo et al., (2010), detected similar results for other necrotrophic fungus, *Fusarium poae* and Misrha et al., (2004) for *F. graminearum*. In all these works, the conclusion was the high intraspecific variability, but they did not reveal a clear relationship between variability and the host/geographic origin. Moreno et al., (2008b), suggested that the genetic variability at the DNA level among isolates of *P. tritici-repentis* based in ISSR markers was unrelated with pathogenic variability. In agreement with Sicard et al., (1997) and Stenglein and Balatti (2006), we considered that isolates with the same pathogenicity patterns are not necessarily closely related based on DNA analysis.

3.3.6 Internal Transcribed Spacer regions (ITS)

Some genera of fungi are intensively studied and characterized with the use of the ITS region. DNA sequence data from ITS region demonstrated that is the indispensable information source in the taxonomy of fungi, in the diversity of fungi under different scales of time and space (Horton & Bruns, 2001; Hsiang & Wu, 2000; Redecker, 2000; White et al., 1990). It has been used to distinguish between closely related fungal isolates. The ITS region has been used wide because it is a noncoding region and its size and sequence are less conserved, it is used to examine related population of fungi (Pritsch et al., 1997). Green et al., (2004), analyzed different set of primers for ITS region and 18 S rDNA to design specific primers to Pyrenomycetes and to determine the utility of ITS region for genotyping Pyrenomycetes isolates. Phylogenetic studies of genera *Pyrenophora* has been developed by Zhang and Berbee (2001). Arabi and Jawhar (2007) observed high genetic diversity in Syrian population of *P. graminea* with ITS region. Zein et al., (2010) observed fingerprints generated

from the five restriction digestions of the nrDNA ITS region denoted high levels of intraspecific variation within the *P. graminea* population. Bakri et al., (2011) observed high levels of intraspecific variation within the *P. graminea* population. Arabi and Jawhar (2007) and Bakri et al., (2011) suggested the use of other markers to better clarify genetic diversity in Syrian population *P. graminea*. In general, the used of ITS region to study *P. tritici-repentis* has been used in combination with others region as nuclear gene coding region glyceraldehido-3-phosphate dehydrogenase (*gpd*) and mating type locus (*MAT*) (Lepoint et al., 2010). They suggested that with ITS region and *gpd* gene did not detected genetic variability among 20 isolates of *P.tritici-repentis* obtained from different countries. These sequences suggested that *P. bromi* is one of the *P. tritici-repentis*`s closest relatives. In the general use of ITS region and others region of rDNA are use for phylogenetic studies genera level. This technique has been used in the last years for detection and follow sequenciation of specific fragment of DNA fungi and revealed by denaturing gradient gel electrophoresis from environmental samples (Anderson et al., 2003; Green et al., 2004). Since the ITS region is highly conserved intraspecifically but variable between different species it is often used in taxonomy (Bruns et al., 1991; Hillis & Dixon, 1991).

3.3.7 Mating type locus (*MAT*)

The study of genetic diversity through of the mating type genes (*MAT*) allowed to clarify the population genetic structure of some plant pathogens fungi and provided the opportunity to evidence the sexuality reproduction in several fungi (Bennett et al., 2003; Lepoint et al., 2005, 2010; Rau et al., 2005; Serenius et al., 2005).

The mating-type genes of *P. teres* have been studied by different researchs (Rau et al., 2005; Serenius et al., 2005). The *P. teres* isolates obtained from Italy showed two *MAT* genes (1:1), suggesting that the sexual reproduction is the major influence in the population structure. On other hand, Serenius et al., (2005) observed a high variability between regions suggesting that the sexual reproduction occurs but its influence on the population structure is relative, and may be based in environmental differences. The taxonomic status of formae speciales of *P. teres* was proposed for reconsideration by Rau et al., (2005). A similar study of Rau et al., (2005) and Serenius et al., (2005) was developed by Lepoint et al., (2010) with isolates of *P. tritici-repentis* obtained from naturally infected leaves since 1980s in wheat-growing areas worldwide. They showed that all isolates amplified to *MAT* gene and they highlighted some points: a) the organization of *MAT* gene the *MAT1-1* gene is followed by the *MAT1-2* gene, with both ORFs transcribed left to right; b) some strains presented one fragment as insertion between primer PtrPLP5 and PtrPLP6 (Lepoint et al., 2010) in a variable region; c) some strains of *P. tritici-repentis* revealed a nonspecific profile with the primer par PtMat_fw (Rau et al., 2005) and PtrPLP2 (Lepoint et al., 2010). They concluded that *P. tritici-repentis* has a simple *MAT* locus organization containing both idiomorphs tandemly arranged in a single individual. This result corroborated the homothallic condition of *P. tritici-repentis*. They could differentiate two groups, one group more homogeneus including isolates obtained from different areas worldwide; the other one including isolates of *P. tritici-repentis* classified as race 4 and in the other groups these isolates are lack. Also in the second group they observed the presence of others isolates of *P. tritici-repentis* classified as race 1 and 2 (Lepoint et al., 2010). They suggested that the clear-cut is due to the presence of isolates of race 4 in only one group. However, the number of isolates classified as race 4 is low respect

the others races. Lepoint et al., (2010) suggested the new distinction between the strains of *P. tritici-repentis* and they proposed the revision of the intra-infraspecific taxonomic status of the pathogen. They suggested that the MAT is the first gene follow the *HTS* gene that has been able to distinct the different lineages within *P. tritici-repentis*.

3.3.8 Toxin production genes (*Tox* A, *Tox* B and *tox* b)

The technique wider used to the characterization of *P. tritici-repentis* population has been based in the interaction wheat-*P. tritici-repentis* (Ali & Francl, 2001[a,b], 2002[b]; Lamari & Bernier, 1989[a,b]; Lamari & Gilbert, 1998; Lamari et al., 1995, 1998, 2003, 2005) add the phenotypic characterization in races. Due the subjective of this technique add the availability of the sequences of the *Tox* genes (Ballance et al., 1996; Ciuffetti et al., 1997; Martinez et al., 2001; 2004). Different researches has been taken the sequences for construction of different specific primers to detect the presence of Tox genes in the different strains of *P. tritici-repentis* (Ali et al., 2010; Andrie et al., 2007; Martinez et al., 2001; 2004; Moreno et al unpublished). The *ToxA* gene codified to the production of Ptr ToxA, this toxin produces necrotic symptoms in Glenlea and Katepwa and must be present in the strains classified as races 1, 2, 7 and 8. The Ptr Tox B is produced by races 5, 6, 7 and 8; and races 1, 3, 6 and 8 produce Ptr Tox C. The race 4 is nonpathogenic and do not produces any symptoms. The races 5, 6, 7 and 8 produce chlorosis in Katepwa and 6B662. In this case of Ptr Tox B is encoded by the multiple copy of *ToxB* gene (Martinez et al., 2001, 2004; Strelkov et al., 2006). It is know that the races 3 and 4 have present in their genomes singles copies of distinct *ToxB* homolog (Martinez et al., 2004; Strelkov et al., 2006). Therefore found *P. tritici-repentis* strains carrying the *ToxB* gene but without production of Ptr ToxB were found (Andrie et al., 2007; Lamari et al., 1995; Martinez et al., 2004; Strelkov et al., 2006). In this way and considering that the virulence patterns of the various *P. tritici-repentis* isolates are based in that each compatible interaction host-pathogen that it is mediated by HST (Strelkov & Lamari, 2003); Andrie et al., (2007) suggested that the use of PCR multiplex is appropriated to corroborate the phenotypic identification of *P. tritici-repentis* races. In their manuscript their suggested suggests the erroneous identification for two strains of *P. tritici-repentis* (SO3 and PT82), the phenotypic assignation for this strains corresponded with race 2 and 8 respectively. Specifically they observed that SO3 lacks the *Tox A* gene and Ptr Tox A production and PT82 lacks *ToxB* gene and Ptr Tox B production. This point is fundamental to their recommendation "*results suggested that genotypic confirmation shoul accompany, where possible, phenotypic characterization when designating a P. tritici-repentis isolate to race*" (Andrie et al., 2007). According the previous work of Andrie et al., (2007), other publications refers studies of *P. tritici-repentis* population structure in races (Ali et al., 2010; Lepoint et al., 2010; Moreno et al., (unpublished data 2011)). Ali et al., (2010) tested 42 isolates obtained from Arkansas (United State of America) to the presence or absence of the *ToxA* and *ToxB* gene. Their results are agreement with the results of Andrie et al., (2007). Some isolates ToxA gene deficient caused necrosis on Glenlea and others that did not cause necrosis. Then, they suggested that isolates of *P. tritici-repentis* deficient in *ToxA* and T*oxB* gene and that induce necrosis and/or chlorosis may produce a novel toxin(s) on wheat (Ali et al., 2010). In the same year, Lepoint et al., (2010) tested the presence of *HTS* gene and others markers as ITS, *gpd* gene and *MAT* gene to determine the genetic diversity of an extensive collection of *P. tritici-repentis* isolates. In basis of the results obtained they observed discrepancies between both methods, and they highlighted the potential existence of novel(s) toxins and new races.

Similar results were observed by Moreno and colleagues (unpublished) for isolates of *P. tritici-repentis* obtained from Argentina.

4. Discussion and future prospect

Wheat is one of the most important crops in the world. Global wheat supplies for 2011/12 are projected 11.4 million tons higher with higher beginning stocks and a sharp increase in production (font: WASDE – 497 United States Department of Agriculture ISSN: 1554-9089). The future of this production and its participation of the international markets depend on the evolution of proposal of different countries. It is clear that, the start in the field and the exportation of the crop depend on the introduction of new technological develops. Among these developments are the new cultivar composition, management of crops and management of future areas for yield expanding (Ekboir & Morris 2001). The fungi pathogens are become by the combination of these factors (Klein 2001). The management of these diseases to ask the knowledge specific and the increased ability to identify the fungus and the techniques for reduce to minimum the losses of crops (Kohli, 1995). Tan spot disease began to noticeably affect wheat crops region around the world (Annone, 1985; Duveiller et al., 1998; Lamari et al., 2003). It is well known the existence of a fourth HST, additional race that has also been suggested (Andrie et al., 2007; Ali et al 2010; Moreno et al., (unpublished data)). In some studies it has been observed discrepancies between both pathogenic and genetic methods to determine the races (Andrie et al., 2007; Ali et al., 2010; Lepoint et al., 2010; Moreno et al., (unpublished data)). Moreover, several works determined a high level of variability among isolates of *P. tritici-repentis* and these results shown no genetic grouping of pathotypes or grouping for geographic location. These results suggested that the population of *P. tritici-repentis* is complex in reaction types and this characteristic should be considered. The different works cited above have corroborated the intraspecific variability of *P. tritici-repentis* using different techniques. This intrasepecifc variability is not high enough to consider the revision of taxonomic status of *P. tritici-repentis*. However, Lepoint et al., (2010) based on their results with *MAT* genes suggested that could be also be several distinct taxa. Population diversity and genetic structure are two important factors in the management of plant disease. The variability in the population of plant pathogen fungi can be very highly, if we know the degree of variability we could foresee the potential occurrence of new races, speciation process and new isolate fungicide resistance. The pathogenic and genetic analysis of *P. tritici-repentis* showed intraspecific variability. However, in any case was possible to establish a relationship between this variability and the geographical origin or physiological race. In some analyses, it was shown that *P. tritici-repentis* isolates from the same region appeared in the same group. Due the last studies, and the discrepancies showed between methods to determine the population race structure is that we suggested an exhaustive genetic study of a worldwide collection of *P. tritici-repentis* isolates, with similar number of isolates and races for countries in order to obtain a clear genetic mapping of *P. tritici-repentis* population around the world.

5. References

Adee, E.A. & Pfender, W.F. (1989). The effect of primary inoculum level of *Pyrenophora tritici-repentis* on tan spot epidemic development in wheat. *Phytopathology* 79: 873-877. ISSN: 0031-949X.

Agrios, G.N. (2004). Plant Pathology. 5th edition. Elsevier, Academic Press, Burlington, United State of America. ISBN-10:0120445654.

Ali, S. & Buchenau, G. W. (1992), Incidence of toxin-producing pathotypes of *Pyrenophora tritici-repentis* in South Dakota. *Phytopathology*. 82: 1159 (Abst) ISSN: 0031-949X .

Ali. S. & Francl, L.J. (2001). Recovery of *Pyrenophora tritici-repentis* from barley and reaction of 12 cultivars to five races and two host-selective toxins. *Plant Disease* 85: 580-584. ISSN: 0191-2917.

Ali, S. & Francl, L.J. (2002). Race structure of *Pyrenophora tritici-repentis* isolates obtained from wheat in South America. *Plant Protection Science* 38 (Special Issue 2): 302-304. ISSN: 1212-2580

Ali, S. & Francl, L.J. (2003). Population race structure of *Pyrenophora tritici-repentis* prevalent of wheat and non-cereal grasses in the Great Plains. *Plant Disease* 87: 418-422. ISSN: 0191-2917

Ali, S.; Francl, L.J. & de Wolf, E.D. (1990). First report of *Pyrenophora tritici-repentis* race 5 from North America. *Plant Disease* 83: 591. ISSN: 0191-2917

Ali, S.; Ling, H., Meinhardt, S. & Francl, L.J. (2002) A new race of *Pyrenophora tritici-repentis* that produces a putative host –selective toxin. *Phytopathology* 92: S3.ISSN: 0031-949X

Ali, S.; Gurung, S. & Adhikari, T.B. (2010) Identification and characterization of novel isolates of *Pyrenophora tritici-repentis* from Arkansas. *Plant Disease* 94: 229-235. ISSN: 0191-2917

Andersen, H. (1955). Species of *Helminthosporium* on cereals and grasses in Denmark. *Friesia* 5: 80-89. ISSN: 0016-1403.

Anné-Peberdy, J.F. (1981) Characterization of interspecific hybrids between *Penicillium chrysogenum* and *P. roqueforti* by isozyme analysis. *Trans Brit Mycol Soc* 77: 401-408. ISSN: 0007-1536

Annone, J,G. (1985). Presencia de la mancha tostada del trigo (*Helminthosporium tritici-repentis*). Carpeta de Producción Vegetal. TRIGO. TOMO VII. Informe Nº88. Estación Experimental Agropecuaria, Instituto Nacional de Tecnología Agropecuaria, Buenos Aires: Argentina.

Annone, J.G. (1995). Resistencia genética a agentes fitopatógenos: Conceptos sobre hospedante, patógeno, relación hospedante-patógeno y manejo de la resistencia. In: Curso de manejo de enfermedades del Trigo. Kohli, M. M., Annone, J. G. y García, R (eds). Centro Internacional de Capacitación INTA-CIMMYT. Pergamino, Buenos Aires, Argentina. 1995, pp. 237.

Annone, J.G. (2005). El desafío del manejo integrado de enfermedades de los cultivos en sistemas agrícolas conservacionistas: el caso de la sanidad del trigo en siembra directa. 13th Congreso Latinoamericano de Fitopatología, 19–22 April 2004, Villa Carlos Paz, Córdoba, Argentina, pp. 161–164.

Annone, J. (2006). Las principales enfermedades del trigo en Argentina: su importancia relativa en las Regiones productoras Norte y Sur. In: Congreso A Todo Trigo. Conocimiento y Producción. Federación de Centros y Entidades Gremiales de Acopiadores de Cereales, Hotel Sheraton, 18 y 19 de Mayo de 2006. Mar del Plata, Argentina. 53-58.

Annone, J. G.; Botta, G. E. & Ivancovich, A. (1994). Ocurrencia de la mancha bronceada del trigo en el área norte de la provincia de Buenos Aires. In: III Congreso Nacional de

Trigo y I Simposio Nacional de Cereales de siembra Otoño-Invernal. Dpto. de Agronomía, Universidad Nacional del sur, Asociación de ingenieros Agrónomos del Norte de la provincia de Buenos Aires, Bahía Blanca, Buenos Aires. 26-28 de octubre de 1994. p. 205-207.

Andrie, R.M.; Pandelova, I. & Ciuffetti, L.M. (2007). A combination of phenotipic and genotypic characterization *Pyrenophora tritici-repentis* race identification. *Phytopathology* 97: 694-701. ISSN: 0031-949X

Arabi, M.I.E. & Jawhar, M. (2007). Heterogeneity in *Pyrenophora graminea* as revealed by ITS-RFLP. *Journal of Plant Pathology* 89 (3), 391-395. ISSN: 1125-4653

Araya, C.M. & Cárdenas C. (1999). Variabilidad patogénica y fuentes de resistencia a poblaciones de Colletotrichum lindemuthianum en Costa Rica. Actas de Congreso XXXIX Reunión Anual APS-CD. Puerto Rico p. 92.

Arenal, T.; Platas, G., Martin, J., Asenso, F.J., Salazar, O., Collado. J., Vicente, F., Basilico, A., Ruibal, C., Roya, I., de Pedro, N. & Peláez F. (2002). Comparison of genotypic and phenotypic techniques for assessing the variability of the fungus *Epicoccum nigrum*. *Journal of Applied Microbiology* 93: 36-45.ISSN: 1364-5072

Assigbetse, K.B., Fernandez, D., Dabois, M. P. & Geiger, J. P. (1994). Differentiation of *Fusarium oxysporum* f. sp. *vasinfectum* races of cotton by random amplified polimorphic DNA (RAPD) analysis. *Phytopathology* 84:622-626. ISSN: 0031-949X

Bakri, Y.; Arabi, M. I. E.; & Jawhar, M. (2011). Heterogeneity in the ITS of the Ribosomal DNA of *Pyrenophora graminea* Isolates Differing in Xylanase and Amylase Production. *Microbiology* 80: 492–495. ISSN: 1350-0872

Ballance, G.M.; Lamari, L. & Bernier, C.C. (1989). Purification and characterization of a host-selective necrosis toxin from *Pyrenophora tritici-repentis*. *Physiological and Molecular Plant Pathology* 35: 203-213. ISSN: 0085-5765

Ballance, G.M.; Lamari, L., Kowatsch, R. & Bernier, C.C. (1996). Cloning, expression and occurrence of the gene encoding the Ptr necrosis toxin from *Pyrenophora tritici-repentis*. *Physiological and Molecular Plant Pathology*. On-Line http://www.bspp.org.uk/mppol/1996/1209ballance/. ISSN: 0085-5765

Bayraktar, H.; Dolar, F.S. & Maden, S. (2007). Use of RAPD and ISSR Markers in detection of genetic variation and population structure among *Fusarium oxysporum* f sp. *ciceris* isolates on chickpea in Turkey. *Journal Phytopathology*. Doi: 10.1111/j.1439-0434.2007.01319x. ISSN: 0931-1785

Bennett, R.S.; Yun, S.H., Lee, T. Y., Turgeon, B.G., Arseniuk, E., Cunfer, B.M. & Bergstrom, G.C. (2003). Identify and conservation of mating type genes in geographically diverse isolates of *Phaeosphaeria nodorum*. *Fungal Genet. Biol.* 40:25-37. ISSN:1087-1845

Bisset, J.A .(2002). Uso correcto de insecticidas: control de la resistencia. *Rev Cub Medic Trop* 54:202-219. ISSN: 0375-0760

Bocchese, C.A.C.; Martinelli, J.A., Matsumura, A.T.S., Federizzi, L.C. & Prestes, A.M. (2003). Virulencia, atividade enzimatica e padroes de isoesterases de isolados de *Pyrenophora chaetomioides*, agente etiologico da mancha da graos e folhas de aveia. *Fitopatol Bras* 28:11-16. ISSN: 0100-0148.

Bockus, W. W.; Davis, M. A. & Shroyer, J. P. (1992) Effect of foliar fungicide application on seed size of winter wheat. *J. Applied Seed Production* 10: 1-6. ISSN: 8755-8750.

Pyrenophora tritici-repentis, Causal Agent
of Tan Spot: A Review of Intraspecific Genetic Diversity
303

Bonde, M.R.; Micales, J.A. & Peterson, G.L. (1993). The use of isozyme análisis for identification of plant pathogenic fungi. *Plant Disease.* 77:961-968. ISSN: 0191-2917

Bonde, M.R.; Peterson, G.L., & Mass, J.L. (1991). Isozyme comparison for identification of Colletotrichum species pathogenic to strawberry. *Phytopathology* 81:1523-1528. ISSN: 0031-949X

Boshoff, W.H.P.; Swart, W.J., Pretorius, Z.A., Liebenberg, M.M. & Crons, P.W. (1996). Isozyme variability among isolates *Phaeoisariopsis griseola* in southern Africa. *Plant Pathology.* 45:344-349.ISSN: 1365-3059

Bornet, B. & Branchard, M. (2001). Nonanchored inter sequence repeat (ISSR) markers: Reproducible and specific tools for genome fingerprinting. *Plant Mol. Biol. Rep.* 19: 209-215. ISSN: 0735-9640.

Botstein, D.; White, R.L., Skolnick, & Davis, R.W. (1980). Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. *Am. J. Hum. Genet.* 32:314-331. ISSN:0002-9297.

Brammer, S.P. (2000). Mapeamento de genes de resistência parcial á ferrugem da folha do trigo em cultivares brasileiras de trigo (*Triticum aestivum* L. em Thell). Tese de Doutorado. Porto Alegre, Universidade Federal do Rio Grande do Sul, Brasil.

Bridge, P.D.; Hawksworth, D.L., Kozakiewicz, Z., Onions, A.H.S. & Paterson, R.R.M. (1986). Morphological and biochemical variation in single isolates of *Penicillium. Trans. Brit. Mycol. Soc.* 87:389-396. ISSN: 0007-1536.

Brown, D. A. & Hunger, R. M. (1993). Production of a chlorosis-inducing, host-specific, low-molecular weight toxin by isolates of *Pyrenophora tritici-repentis*, cause of tan spot of wheat. *J. Phytopathol.* 137: 221-232. ISSN: 0931-1785

Burdon, J.J.& Roelfs, A.P. (1985). Isozyme and virulence variation in asexually reproducing populations of *Puccinia graminis* and *P. recondita* on wheat. *Phytopathology* 75:907-913. ISSN: 0031-949X

Burdon, J.J. & Silk, J. (1997). Sources and patterns of diversity in plant pathogenic fungi. *Phytopathology* 87: 664-669. ISSN: 0031-949X

Burguess, T.; Malajckuz, N. & Dell, B. (1995). Variation in *Pisolithus* and basidiospore morphology, culture characteristics and analysis of polypeptides using 1D SDS-PAGE. *Mycological Research* 99: 1-13. ISSN: 0953-7562.

Campbell, G.F.; Lucas, J.A. & Crous, P.W. (2002). Evidence of recombination between net and spot-type populations of *Pyrenophora teres* as determined by RAPD analysis. *Mycol Research* 106: 602-608. ISSN: 0953-7562.

Carmona, M.; Reis, E.M. & Cortese, P. (1999). Manchas foliares del trigo. Diagnóstico, epidemiología y nuevos criterios para el manejo. Carmona M, Reis EM, Cortese P (eds). Buenos Aires, Argentina. 32pp.

Carmona, M., Ferrazini, M. & Barreto D.E. (2006). Tan spot of wheat caused by *Drechslera tritici repentis*: Detection transmission and control in wheat seed. *Cereal Research Communications* 34(2-3):1043-1049. ISSN: 0133-3720.

Carmona, M. & Reis, E. M. (1998). Las enfermedades y su manejo en siembra directa. Cuaderno de Actualización Técnica Nº 59 CREA Siembra Directa pp 68-81, 1998.

Carmona M., Sugia V., Schmidt A. & Delfino S. V. (2001). Control potencial de manchas foliares de trigo, cebada y avena por la rotación de cultivos. Congreso Nacional de Trigo y III Simposio Nacional de Cereales de invierno. Villa Carlos Paz, Córdoba, 24-28 de septiembre de 2001.

Castro, T.A.M.G. & Bach, E.E. (1993). Electroforese na agricultura. Zootecnia, Nova Odessa. *SP*. 31:73-88

Chen, W.; Schneider, R.W. & Hoy, J.W. (1992). Taxonomic and phylogenetic analyses of ten *Pythium* species using isozyme polymorphisms. *Phytopathology* 82:1234-1244. ISSN: 0031-949X

Christensen, J.J. & Graham, T.W. (1934). Physiologic specialization and variation in *Helminthosporium gramineum* Rob. University of Minnesota Agricultural Experiment Station, United State of America.

Ciuffetti, L.M.; Francl, L.J., Balance, G.M., Bockus, W.W., Lamari, L., Meinhardt, S.W. & Rasmussen, J.B. (1998). Standarization of toxin nomenclature in the *Pyrenophora tritici-repentis*/wheat interaction. *Canadian Journal of Plant Pathology* 20: 421-424. ISSN: 0706-0661.

Ciuffetti, L.M.; Tuori, R.P. & Gaventa, J.M. (1997). A single gene encodes a selective toxin causal to the development of tan spot of wheat. *Plant Cell* 9: 135-144. ISSN: 1040-4671.

Clay, L. (1995). Correlates of pathogen species richness in the grass family. *Can. Journal Bot.* (Suppl. 1): S42-S49. ISSN: 0008-4026.

Conners I. L. (1939). Yellow leaf botch. *Can. Plant. Dis. Surv*. 19: 12-14. ISSN: 0008-476X.

Cox, D. J. & Hosford, R. M. Jr. (1987). Resistant winter wheat compared at differing growth stages and leaf positions for tan spot severity. *Plant Dis*. 71: 883-886. ISSN: 0191-2917.

Crous, P.W., Janse, B.J.H., Tunbridge, J. & Holz, G. (1995). DNA homology between *Pyrenophora japonica* and *P. teres*. *Mycological Research, 99,* 1098-1102. ISSN: 0953-7562.

Day, P. R. (1960). Variation in phytopathogenic fungi. *Annual Rericw of Microbiology* 14, I - 16. ISSN: 0066-4227.

da Luz , W.C. & Hosford, R.M. Jr (1980). Twelve *Pyrenophora trichostoma* races for virulence to wheatin the central plains of North America. *Phytopathology* 70: 1193-1196. ISSN: 0031-949X.

de Wolf, E.D.; Effertz, R.J., Ali, S. & Francl, L.J. (1998). Vistas of tan spot research. *Canadian Journal of Plant Pathology* 20: 349-370. ISSN: 0706-0661.

Dennis, R.W.G. & Wakefield, E.M. (1946). New or interesting Bitrish fungi. *Trans Brit. Mycol Soc*. 29: 141-166. ISSN: 0007-1536.

Diaz de Ackermann, M. (1987). Resistance among winter wheat to *Pyrenophora tritici-repentis* isolates and variability in fungal growth, color, sporulation and pathogenicity. Th. North Dakota State University Department of Plant Pathology, Fargo, North Dakota, United State of America.

Dickson, J. G. (1956). Diseases of field crops. Mc Graw Hill Book Co. New york, 517 pp.

Diedicke, H. (1902). Uberden zusammenhangwischen *Pleospora* und *Helminthosporium* arten. Centrablatt fur Bakteriologie und Parasitenkunde Jena. Abt. 11: 52-59.

Dinolfo M.I.; Stenglein S.A., Moreno M.V., Nicholson P., Jennings P. & Salerno G. (2010). ISSR markers detect high genetic variation among Fusarium poae isolates from Argentina and England. *European Journal of Plant Pathology* 127: 483-491. ISSN:0929-1873.

Di Zinno, T.; Longree, H. & Maraite, H. (1998). Diversity of *Pyrenophora tritici-repentis* isolatesfrom warm wheat growing areas: pathogenicity, toxin production, and

RAPD analysis. (In E. Duveiller, H.J. Dublin, J. Reeves & A. McNab (Eds.), Helminthosporium Blights in Wheat: Spot Blotch and Tan Spot. (pp. 302-311). CIMMYT: México.

Dorrance, E.A.; Miller, O.K. & Warren, H.L. (1999) Comparison of *Stenocarpella maydis* isolates for isozyme and cultural characeristics. *Plant Dis* 83:675-680. ISSN: 0191-2917.

Drechsler, C. (1923). Some graminicolous species of *Helminthosporium*. *J. Agricul. Res*. 24: 641-740. ISSN: 0718-5839.

Dubin, H. J. (1983). Ocurrence of *Pyrenophora tritici-repentis* in the Andean Countries of South America. *Plant Disease*. 67: 1040. ISSN: 0191-2917.

Duczek, L. J. & Jones-Flory, L. L. (1994). Effect of timing of propiconazole application on foliar disease and yield of irrigated spring wheat in Saskatchewan fro 1991 to 1992. *Can, J. Plant Sci*. 74: 205-207. ISSN: 0008-4220.

Duveiller, E.; Dubin, H.J.; Reeves, J. & MacNab, A. eds. (1998). Helminthosporium Bligths of Wheat: Spot Blotch and Tan Spot. CIMMYT. El Batán México.

Edel, V.; Steinberg, C., Avelange, I., Laguerre, G., Alabouvette, C. (1995). Comparison of 3 molecular methods for the characterization of *Fusarium oxysporum* strains. *Phytopathology* 85: 579-585. ISSN: 0031-949X.

Effertz, R.J.; Meinhardt, S.W., Anderson, J.A., Jordahl, J.G. & Francl LJ. (2002). Identification of chlorosis-inducing toxin from *Pyrenophora tritici-repentis* and the chromosomal location of an insensitivity locus in wheat. *Phytopathology* 92: 527-533. ISSN: 0031-949X.

Eikboir, J. & Morris, M. (2001). Perspectivas del Mercado Internacional de Trigo: Implicancias para los países de América del Sur. In MM Kohli, M Diaz, M Castro (Eds.), Estrategias y Metodologías utilizadas en el mejoramiento de trigo. La Estanzuela, Uruguay. Seminario Internacional CIMMYT-INIA 3-10 pp.

Elias, E. (1987). Evaluation on tan spot *Pyrenophora tritici-repentis* in durum wheat *Triticum aestivum* L. var. Durum. PhD Thesis, North Dakota State University, Fargo, United State of America.

Elias, K.S. & Schneider, W. (1992). Genetic diversity within and among races and vegetative compatibility groups of *Fusarium oxysporum* f. sp. *lycopersici* as determined by isozyme analysis. *Phytopathology* 82:1421-1427. ISSN: 0031-949X.

Ellis, M.B. & Waller, J.M. (1976). *Pyrenophora tritici-repentis* (conidial state: *Drechslera tritici-repentis*). *Commonw. Mycol. Inst. Descriptions Pathogenic Fungi Bacteria* Nº 494.

Farr, D. F.; Bills, G. F., Chamuris, G. P. & Rossman, A. Y. (1989), Fungi on Plants and Plant Products in the United States. APS Press, St Paul MN. 1252 pp.

Fernandez, M.R.; Clarke, J.M. & DePauw, R.M. (1994). Response of durum wheat kernels and leaves at different growth stages to *Pyrenophora tritici-repentis*. *Plant Disease* 78: 597-600. ISSN: 0191-2917.

Faure-Raynaud, M.; Bonnefoy, M.A., Perzadin, Y., Simonet, P. & Moiroud, A. (1987). Protoplast formation from *Frankia* strains. *Microbios* 41:159-166. ISSN: 0026-2633.

Flor, H. H. (1971). Current status of the gene-for-gene concept. *Annu Rev Phytopathol* 9: 275–296. ISSN: 0066-4286.

Frazzon, A.P.G.; Matsumura, A.T.S. & Van der Sand, S.T. (2001). Morphological characterization and genetic analysis of *Drechslera teres* isolates. *Genetic and Molecular Biology* 25: 235-241. ISSN: 1415-4757.

Friesen, T.L., Ali, S., Kleim, K.K. & Rasmussen, J.B. (2005). population genetic analysis of a global collection of *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. *Phytopathology, 95,* 1144-1150. ISSN: 0031-949X .

Fry, W.E.; Drenth, A., Spielman, L.J., Mantel, B.C., Davidse, L.C. & Goodwin, S.B. (1991). Population genetic structure of *Phytophtora infestans* in the Netherlands. *Phytopathology* 81:1330-1336. ISSN: 0031-949X.

Galich, M.T.V. de & Galich, A.N. (1994). Enfermedades de trigo en el área sur de Santa Fe y Córdoba correspondientes a la subregión II Norte. Campaña 1993/94. Informe para Extensión Nº 5. EEA INTA Marcos Juárez, Córdoba, pp 7

Gamba, F.M.; Lamari, L. & Brûlé- Babel, A. (1998). Inheritance of race-specific necrotic and chlorotic reactions induce by *Pyrenophora tritici-repentis* in hexaploid wheats. *Canadian Journal of Plant Pathology* 20:401-407. ISSN: 0706-0661.

Garber, E.D. & Behara, L. (1966). Genetics of phytopathogenic fungi pectinolytic enzymes of virulent and avirulent strains of three phytopathogenic *Penicillia*. *Can J Bot* 44:1645-1650. ISSN: 008-4026.

Gilchrist, L., Fuentes, S.F., Isla de Bauer de la M.L. (1984). Determinación de fuentes de resistencia contra *Helminthosporium tritici-repentis* bajo condiciones de campo y de invernáculo. *Agrociencia* 56: 95-106.

Gillings, M. & Holley, M. (1997). Repetitive element PCR fingerprint (rep-PCR) using enterobacterial repetitive intergenic consensus (ERIC) primers is not necessarily directed at ERIC elements. *Lett Appl. Microbiol.* 25:17-21. ISSN: 1472-765X.

Goodwin, S.B.; Schineider R.E. & Fry, W.E. (1995). Use of cellulose-acetate electrophoresis for rapid identification of allozyme genotypes of *Phytophthora infestans*. *Plant Dis.* 79: 1181-1185. ISSN: 0191-2917.

Guthrie, P.A.I.; Magill, R.A.; Frederiksen, R.A. & Odvody, G.N. (1992). Random amplified polymorphic DNA markers: a system for identifying and differentiating isolates of *Colletotrichum graminicola*. *Phytopathology* 82: 832-835. ISSN: 0031-949X.

Green, S.K.; Freeman, S.; Hadar, Y. & Minz, D. (2004). Molecular tools for isolates and community studies of Pyrenomycete fungi. *Mycologia* 96: 439-451. ISSN: 0027-5514.

Hellmann, R. & Christ, B.J. (1991). Isozyme variation of physiologic races of *Ustilago hordei*. *Phytopahthology* 81:1536-1540. ISSN: 0031-949X.

Hosford, R.M. Jr. (1971). A form of *Pyrenophora trichostoma* pathogenic to wheat and other grasses. *Phytopathology* 61: 28-32. ISSN: 0031-949X.

Hosford, R.M. Jr. (1982). Tan spot. In RM Hosford Jr. (Ed.), Tan Spot of Wheat and Related Diseases Workshop, Fargo, North Dakota State University, USA. (p.116).

Hosford, R. M. Jr. & Busch, R. H. (1974). Losses in wheat caused by *Pyrenophora trichostoma* and *Leptosphaeria avenaria* f. sp. *triticea*. *Phytopathology.* 64: 184-187. ISSN: 0031-949X.

Hosford, R. M.; Jr., Larez, C. R. & Hammond, J. J. (1987). Interaction of wheat period and temperature on *Pyrenophora tritici-repentis* infection and development in wheat of differing resistance. *Phytopathology* 77: 1021-1027. ISSN: 0031-949X.

Hunger, R.M. & Brown, D.A. (1987). Colony color, growth, sporulation, fungicide sensivity and pathogenicity of *Pyrenophora tritici-repentis*. *Plant Disease* 71: 907-910. ISSN: 0191-2917.

Kalkar, O.; Carner, G.R., Scharf, D. & Boucias, D.G. (2006). Characterization of an Indonesian isolate of *Paecilomyces reniformis*. *Mycopathologia* 161:109-118. ISSN: 0301-486X.

Klein, O.A. (2001) Trigos adaptados a siembra directa conocimientos y desconocimientos. In MM. Kohli, M Diaz de Akcermann, M Castro (Eds.), Estrategias y metodologías utilizadas en el mejoramiento de trigo: un enfoque multidisciplinario. La Estancuela, Uruguay. Seminario Internacional, CIMMYT-INIA (pp. 37-42).

Koch, G. & Köhler, W. (1990). Isozyme variation and genetic distances of *Erysiphe graminis* DC. Formae speciales. *J Phytopathol* 129:89-101. ISSN: 0931-1785.

Kohli, M.M. (1995). Conceptos básicos en el manejo de enfermedades de cultivos. In MM. Kohli, JG. Annone, R García (Eds.), Curso de manejo de enfermedades del Trigo. Centro Internacional de Capacitación INTA-CIMMYT. Buenos Aires: Argentina.

Kohli, M.; Mehta, Y. & Ackermann, M. (1992). Spread of tan spot in the Southern Cone Region of South America. In: Francl. L., Krupinsky, J., Mc Mullen, M. eds. Advances in tan Spot. Proceedings of the Second International Tan Spot Workshop North Dakota State University, Fargo, ND, pp. 86-90.

Kremer, M. & Hoffmann, G.M. (1992). Effect to *Drechslera tritici-repentis* as the cause of wheat yellow leaf spot disease on kernel yield and dry matter production. *Z Pflazenkrankh Pflanzenschutz (Journal of Phytopathology)* 99: 509-605. ISSN: 0931-1785.

Krupinsky, J.M. (1987). Pathogenicity on wheat of *Pyrenophora tritici-repentis* isolated from *Bromus inermis*. *Phytopathology* 77: 760-765. doi: 10.1094/Phyto-77-760. ISSN: 0031-949X.

Krupinsky, J.M. (1992a). Aggressiveness of isolates *Pyrenophora trictici-repentis* obtained from wheat grown under different field cultural practices. Advances in tan Spot. Proceedings of the Second International Tan Spot Workshop North Dakota State University, Fargo, United State of America.

Krupinsky, J.M.(1992b). Aggressiveness of isolates *Pyrenophora trictici-repentis* obtained from wheat in the northern Great Plains. *Plant Disease* 76: 87-91. ISSN: 0191-2917.

Lamari, L. & Bernier, C.C. (1989a). Evaluation of wheat lines and cultivars to tan spot (*Pyrenophora tritici-repentis*) based on type lesions. *Canadian Journal of Plant Pathology* 11: 49-56. ISSN: 0706-0661.

Lamari, L. & Bernier, C.C. (1989b). Virulence of isolates of *Pyrenophora tritici-repentis* on 11 wheat cultivars and cytology of the differential host reactions. *Canadian Journal of Plant Pathology* 11: 284-290. ISSN: 0706-0661.

Lamari, L. & Bernier, C.C. (1991). Genetics of tan necrosis and extensive chlorosis in tan spot of wheat caused by *Pyrenophora tritici-repentis*. *Phytopathology* 81: 1092-1095. ISSN: 0031-949X .

Lamari, L. & Gilbert, J. (1998). Toward a coherent model of host-pathogen interactions in tan spot of wheat: An epilogue to the Third International to Workshop on Tan Spot of wheat. *Canadian Journal of Plant Pathology* 20: 440-443. ISSN: 0706-0661.

Lamari, L.; Gilbert, J. & Tekauz, A. (1998). Race differentiation in *Pyrenophora tritici-repentis* and survey of physiologic variation in western Canada. *Canadian Journal of Plant Pathology* 20: 396-400. ISSN: 0706-0661.

Lamari, L.; Sayoud, R., Boulif, M. & Bernier, C.C. (1995) Identification a new race in *Pyrenophora tritici-repentis*: implications for the current pathotype classification system. *Canadian Journal of Plant Pathology* 17: 312-318. ISSN: 0706-0661.

Lamari, L.; Strelkov, S.E., Yahyaoui, A., Orabi, J. & Smith, R.B. (2003). The identification of two new races of *Pyrenophora tritici-repentis* from the host center of diversity

confirms a one-to-one relationship in tan spot of wheat. *Phytopathology* 93 : 391-396. ISSN: 0031-949X .

Larez, C.R., Hosford, Jr R.M. &, Freeman, T.P. (1986). Infection of wheat and oats by *Pyrenophora tritici-repentis* and initial characterization of resistance. *Phytopathology* 76:931-938. ISSN: 0031-949X.

Lebeda, A. & Jendrulek, T. (1987). Application of cluster analysis for establishment of genetic similarity in gene-for-gene host-parasite relationships. *Phytopathology*. 119:131-141. ISSN: 0031-949X .

Leisová, L.; Hanzalová, A. & Kucera, L. (2008). Genetic diversity of *Pyrenophora tritici-repentis* isolates as revealed by AFLP analysis. *Journal of Plant Pathology*. 90: 223-245. ISSN: 1125-4653.

Lepoint , P.; Munaut, F. & Maraite, H. (2005). *Gibberella xylarioides sensu lato* from *Coffea canephora*: A new mating population in the *G. fujikuroi* species complex. *Appl. Environ. Microbiol.* 71: 8466-8471. ISSN:0099-2240.

Lepoint, P.; Renard, M.E., Legrève, A., Duveiller, E., & Maraite, H. (2010). Genetic diversity of the mating type and toxin production genes in *Pyrenophora tritici-repentis*. *Phytopathology* 100: 474-483. ISSN: 0031-949X.

Li, B. & Sutton, J. (1995). Evaluation of leaf associated microorganisms for biocontrol on tan spot in wheat foliage. *Fitopatología Brasileira*. 20: 545-552. ISSN: 0100-4158.

Linde, C.; Drenth, A., Kemp, G.H.J., Wingfied, M.J. & Von Broembsen, S.L. (1997). Population structure of *Phytophtora cinnamomi* in South Africa. *Phytopathology* 87:822-827. ISSN: 0031-949X.

Linhares, A. J. & da Luz, W. C. (1994). Detection of *Drechslera tritici-repentis* on wheat seeds from Mato Grosso do Sul. *Fitopatología Brasileira*. 19 (Suplemento): 298. ISSN: 0100-4158.

Loughman, R.; Wilson, R. E., Roake, J. E., Platz, G. J., Rees, R. G. & Ellison, E. W. (1998). Crop management and breeding for control of *Pyrenophora tritici-repentis*, causing yellow spot of wheat in Australia. In: Heminthosporum Blights of Wheat: Spot Blotch and Tan Spot. Duveiller, E., Dubin, H. J., Reeves, J. y Mc Nab, A. (eds.). CIMMYT, Mexico, D.F. Mexico. 10-17.

Luz da, W.C. & Hosford, R.M. Jr (1980). Twelve *Pyrenophora trichostoma* races for virulence to wheatin the central plains of North America. *Phytopathology* 70: 1193-1196. ISSN: 0031-949X.

Mc Donald, W.C. (1967). Variability and inheritance of morphological mutants in *Pyrenophora teres*. *Phytopathology* 57: 747-755. ISSN: 0031-949X.

Mc Donald, B. A. & McDermott, J. M. (1993). Population genetics of plant pathogenic fungi. *Bioscience* 43:311-319. ISSN:

Mc Cain, J.W. & Groth, J.V. (1992).Inter and intrapopulation isozymes variation in collections from sexually reproducing populations of the bean rust fungus, *Uromyces appendiculatus. Mycologia* 84:329-340. ISSN: 0027-5514.

McFadden, W. (1991) The effect of stubble-borne inoculum of leaf spotting pathogens on disease development and yield loss in winter wheat. *Can. J. Plant Pathol.* 13: 280. (Abstr). ISSN: 0706-0661.

McFadden, W. & Harding, H. (1989) Cereal stubble as a source of primary inoculum of leaf spotting pathogens of winter wheat. *Can j. Plant Pathol.* 11: 195. (Abstr). ISSN: 0706-0661.

Manning, V.A.; Pandelova, I. & Ciuffetti, L.M. (2002). A race for a novel host-selective toxin. *Phytopathology* 92: S51. ISSN: 0031-949X.

Manning, V.A.; Andrie, R.M., Trippe, A.F. & Ciuffetti, L.M. (2004). Ptr ToxA requires multiple motifs for complete activity. *Molecular Plant Microbe Interactions* 17: 491-501. ISSN: 0894-0282.

Maraite, H.; Mercado-Vergnes, D.; Renard, M.E. Zhanarbekova, A. & Duveiller, E. (2006). Relevance of pathogen diversity in management of leaf spot and leaf blight disease of wheat in central Asian. *Agromeridian* 2: 105-114. ISSN: 1814-8417.

Martinez, J.P.; Oesch, N.W. & Ciuffetti, L.M. (2004). Characterization of the multiple- copy host-selective toxin gene, ToxB, in pathogenic and nonpathogenic isolates of *Pyrenophora tritici-repentis. Molecular Plant Microbe Interactions* 17: 467-474. ISSN: 0894-0282.

Martinez, J.P.; Ottum, S.A., Ali, S., Francl, L.J. & Ciuffetti, L.M. (2001). Characterization of the ToxB gene from *Pyrenophora tritici-repentis*. *Molecular Plant Microbe Interactions* 14: 675-677. ISSN: 0894-0282.

Matsumura, A.T.S. (1991). Variabilidade intraespecífica quanto patogenicidade, características de la cultura e padrão isoenzimatico em populacões naturais de *Bipolaris sorokiniana* (*Helminthosporium sativum*). Tese de Doutorado em Genética, Universidade Federal do Río Grande do Sul, Porto Alegre, Brasil.

Mehta, Y.R. (2001). Molecular and pathogenic variability of *Drechslera* isolates from oats. *Fitopatologia Brasileira* 26 (3): 703-709. ISSN: 0100-4158.

Mehta, Y. R. & Gaudencio, R. A. (1991). The effects of tillage practices and crop rotation on the epidemiology of some major wheat diseases. In: D. F. Saunders (ed). Wheat for the nontraditional, Warm Areas. Mexico, D. F. CIMMYT.

Mehta, Y.R.; Mehta, A. & Riede, C.R. (2004). Molecular and pathogenic variability of *Pyrenophora tritici-repentis* isolates causing tan spot of wheat in the state of Paraná, Brazil. *Summa Phytopathologica* 30: 436-444. ISSN: 0100-5405.

Micales, J.A.; Alfenas, A.C. & Bonde, M.R. (1998). Isoenzimas na taxonomia e na genética de fungos. In: Alfenas AC (ed) Electroforese de isoenzimas e proteínas afins. Vicosa: Editora UFV, pp 477-502.

Micales, J.A.; Bonde, M.R. & Peterson, G.L. (1986). The use of isozyme analysis in fungal taxonomy and genetics. *Mycotaxon* 27:405-449. ISSN: 0093-4666.

Milgroom, M. G.; Lipari, M. S. & Powell, W. A. (1992). DNA fingerprinting and analysis of population structure in the chestnut blight fungus, *Cryphonectria parasitica*. *Genetics* 131:297-306. ISSN: 0016-6731.

Misra, A.P. & Singh, R.A. (1972). Pathogenic differences among three isolates of *Helminthosporium tritici-repentis* and the performance of wheat varieties against them. *Indian Phytopathology* 25: 350-353. ISSN: 0367-973X.

Mizar, P.K.; Tewari, J.P., Malear, R. & Turkington T.K. (2004). Molecular genetic variation and geographic structuring in *Fusarium graminearum. Annals Applied Biology.* 145: 299-307. ISSN:0003-4746.

Moreno, M.V. (2007). Caracterización cultural, patogénica, genética y bioquímico-molecular de *Pyrenophora tritici-repentis* en Argentina. Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, La Plata, Argentina. Doctoral Thesis 188 pp.

Moreno, M.V. & Perelló, A.E. (2010) Occurrence of *Pyrenophora tritici-repentis* causing tan spot in Argentina. In A Arya, AE Perelló (Eds.), Management of Fungal Pathogens: Current Trends and Progress (pp. 275-290). UK, CABI Publishers. ISBN-13:978 1 84593 6037.

Moreno, M.V.; Stenglein, S.A., Perelló, A.E. & Balatti, P.A. (2008a) Pathogenic and genetic diversity of isolates of *Pyrenophora tritici-repentis* causing tan spot of wheat in Argentina. *European Journal of Plant Pathology* 122: 239-252. ISSN: 0929-1873.

Moreno, M.V.; Tacalitti, M.S., Castro, A.M. & Perelló, A.E. (2008b). Isozyme polymorphisms within population P*yrenophora tritici-repentis* in Argentina. *World Journal of Microbiology and Biotechnology* 24: 849-860. ISSN: 0959-3993.

Moriwaki, J.; Sato, T. & Tsukiboshi T. (2003). Morphological and molecular characterization of *Colletotrichum baninense* sp. nov from Japan. *Mycoscience* 44: 47-53. ISSN:1340-3540.

Mueller, U. G.; Lipari, S.E. & Milgroom, M.G. (1996). Amplified fragment length polymorphism (AFLP) fingerprinting of symbiotic fungi cultured by the fungus-growing ant *Cyphomymex minutus*. *Molecular Ecology* 5:119-122. ISSN: 0962-1083.

Nagle, B.J., Frohberg, R.C. & Hosford, R.M. Jr (1982). Inheritance of resistance to tan spot of wheat. In: Hosford RM Jr (ed) Tan spot of wheat and related diseases workshop. North Dakota State University, Fargo, p 4045.

Newton, A.C. (1987). Markers in pathogen populations. In: Day PR, Jellis GJ (eds) Genetic and Plant Pathogenesis. Blackwell Scientific Publications, Oxford, pp 187-94

Nicholson, R.L.; Kuc, J. & Williams, E.B. (1972). Histochemical demonstration of transitory esterase activity in *Venturia inaequalis*. *Phytopatholgy* 62:1241-1247. ISSN: 0031-949X.

Orolaza, N.P.; Lamari, L. & Ballance, G.M. (1995). Evidence of a host-specific chlorosis toxin from *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. *Molecular Plant Microbe Interactions* 3: 221-224. ISSN: 0894-0282.

Pandelova, I. & Ciuffetti, L.M. (2005). A proteomics-based approach for identification of the ToxD gen. *Fungal Genet. Newsl.* 52 (Suppl.): 133. ISSN: 0895-1942.

Pascholati, S.F.; Yoshioka, H., Kunoh, H. & Nicholson, R.C. (1992). Preparation of the infection court by *Erysiphe graminis* f. sp. *hordei*: cutinase is a component of the conidial exudates. *Physiol Mol Plant Pathol* 41:53-59. ISSN: 0085-5765.

Peever, T.L. & Milgroom, M.G. (1994). Genetic structure of Pyrenophora teres populations determined with random amplified polymorphic DNA markers. *Canadian Journal of Botany* 72: 915-923. ISSN: 0008-4026.

Perelló, A.E.; Moreno, M.V., Simón, M.R. & Sisterna, M. (2003). Tan spot of wheat (*Triticum aestivum* L.) infection at different stages of crop development and inoculum type. *Crop Protection* 22: 157-169. ISSN: 0261-2194.

Perelló, A.; Mónaco, C., Moreno, M.V., Simón, M.R. & Cordo, C. (2006). *Trichoderma* spp. isolates as potential biological agents for foliar wheat diseases under field conditions in Argentina. *Biocontrol Science and Technology* 16(8): 803-813. ISSN: 0958-3157.

Perelló, A.; Moreno, M. V., Mónaco, C. & Simón, M. R. (2008). Effect of *Thichoderma* spp. Isolates for biological control of tan spot of wheat caused by *Pyrenophora tritici-repentis* under field conditions in Argentina. *BioControl.* 53: 895-904. ISSN: 1386-6141.

Petrunak, D.M. & Chrits, B.J. (1992). Isozyme variability in *Alternaria solani* and *A. alternate*. *Phytopathology* 82:1343-1347. ISSN: 0031-949X .

Pfender, W. F.; Zhang, W. & Nus A. (1989). Field performance and greenhouse assay of fungi for biocontrol of residue borne *Pyrenophora tritici-repentis*. *Phytopathology* 79: 1152. ISSN: 0031-949X .

Pioli, R.N.; Morandi, E.N., Marténez, M.C., Lucca, F., Tozzini, A., Bisaro, V. & Hopp, HE. (2003). Morphological, molecular and pathogenic characterization of *Diaphorte phaseolorum* variability in the core soybean-producing area of Argentina. *Phytopathology* 93: 136-146. ISSN: 0031-949X.

Postnifova, E. N. & Khasanov, B. A. (1998). Tan spot in central Asia. In: In: Heminthosporum Blights of Wheat: Spot Blotch and Tan Spot. Duveiller, E., Dubin, H. J., Reeves, J. y Mc Nab, A. (eds.). CIMMYT, Mexico, D.F. Mexico. 107-113.

Pritsch, K.; Munch, J.C. & Buscot, F. (1997). Characterization and identification of black alder ectomicorrhizas by PCR/RFLP analyses of the rDNA internal transcribed spacer (ITS). *New Phytologist*. 137: 357-369. ISSN: 0028-646X.

Pujol Vieira dos Santos, A.M.; Santos Matsumura, A.T., Van der Sand, S.T. (2002). Intraspecific genetic diversity of *Drechslera tritici-repentis* as detected by a random amplified polymorphic DNA analysis. *Genetic and Molecular Biology* 25: 243-250. ISSN: 1415-4757.

Rau, D.; Maier, F.J. PapaR, Brown A.H.D., Balmas, V., Saba E., Schafer, W. & Attene, G. (2005). Isolation and characterization of the mating type locus of barley pathogen *Pyrenophora teres* and frequencies of mating type idiomorphs within and among fungal population collected from barley landraces. *Genome* 48: 855-869. ISSN: 0831-2796.

Redecker, D. (2000). Specific PCR primers to identify arbuscular mycorrizhal fungi within colonized roots. *Mycorrhiza* 10: 73-80. ISSN: 0940-6360.

Rees, L.G. & Platz, G.J. (1983). Effects of yellow spot of wheat in Southern Queensland. *Australian Journal Agricultural Research* 34: 39-46. ISSN: 0004-9409.

Rees, L. G.; Platz G. J. & Mayer, R. J. (1982). Yield losses in wheat from yellow spot: comparison of estimates derived from single tillers and plots. *Australian Journal Agricultural Research*.33: 899-908. ISSN: 0004-9409.

Redondo, C.; Cubero, J. & Melgarejo, P. (2009). Characterization of Penicillium species by ribosomal DNA Sequencing and BOX, ERIC and REP-PCR Analysis. *Mycopathologia* 168: 11-22. ISSN: 0301-486X.

Reynaldi, F.; Lopez, A.C., Albo, G.N. & Alippi, A.M. (2003). Differentiation of *Ascosphaera apis* isolates by re-PCR fingerprinting and determination of chalkbrood incidence in Argentinean honey samples. *J Apic res*.42: 68-76.

Reynolds, M.; Weinhold, A.R. & Morris, T.J. (1983). Comparison of anastomosis groups of *Rhizoctonia solani* by polyacrilamide gel electrophoresis of soluble proteins. *Phytopathology* 73:903-906. ISSN: 0031-949X.

Sah, D. N. (1994). Effects of leaf wetness duration and inoculum level on resistance of wheat genotypes to *Pyrenophora tritici-repentis*. *J. Phytopathol.* 142: 324-330. ISSN: 0931-1785.

Sah, D.N. & Ferhmann, H. (1992). Virulence patterns of geographically differing isolates of *Pyrenophora tritici-repentis* and sources of resistance in wheat. *Plant Disease* 76: 712-716. doi: 10.1094/PD-76-0712. ISSN: 0191-2917.

Saharan, M.S. & Naef, A. (2008). Detection of genetic variation among Indian wheat head scab pathogens (*Fusarium* spp./isolates) with microsatellite markers. *Crop Protection* 27: 1148-1154. ISSN: 0261-2194.

Sarova, J.; Hanzalova, N. & Barlos, P. (2002). *Pyrenophora tritici-repentis* one the most important leaf spot pathogens in the Czech Republic. 4th International Wheat Tan Spot and Spot Blotch Workshop. July 21-24, 2002. Bemidji, M. N., North Dakota, State University.

Schilder, A.M.C. & Bergstrom, G.C.(1990). Variation in virulence within the population of *Pyrenophora tritici-repentis* in New York. *Phytopathology* 80: 84-90. ISSN: 0031-949X.

Serenius, M.; Manninen, O., Wallwar, K.H. & Williams, K. (2007). Genetic differentiation in *Pyrenophora teres* population measured with AFLP markers. *Mycological Research* 11: 213-223. ISSN: 0953-7562.

Serenius, M.; Micromenko, N. & Manninen, O. (2005). Genetic variation, occurrence of mating types and different forms of *Pyrenophora teres* causing net blotch of barley in Finland. *Mycological Research.* 109: 809-817. ISSN: 0953-7562.

Shabeer, A. & Bockus, W.W. (1988). Tan spot effects on yield and yield components relative to growth stage in winter wheat. *Plant Dis* 72:599–602. ISSN: 0191-2917

Sharp, E.L.; Sally, B.K. & McNeal, F.H. (1976). Effect of *Pyrenophora* wheat leaf blight on the thousand kernel weight of 30 spring wheat cultivars. *Plant Disease* 60, 135–138. ISSN: 0191-2917

Shoemmaker, R.A. (1962). Drechslera Ito. *Canadian Journal of Botany* 40: 809-908. ISSN: 0008-4026.

Sicard, D.; Mickalakis, Y., Drn, M. & Neema, C. (1997). Genetic diversity and pathogenic variation of *Colletrotichum lindemuthianum* in the three centers of diversity of its host, *Phaseolus vulgaris. Phytopathology, 87,* 807-813. ISSN: 0031-949X.

Sim, T. & Willis, W. G. (1982). Kansas wheat disease losses. Kansas State University, Manhattan, USA.

Simón, M.R.; Ayala, F., Golik, S., Terrile, I., Cordo, C., Perelló, A., Moreno, V. & Chidichimo, H. (2011). Influence Integrated Foliar Disease Management to Prevent Yield Loss in Argentinian Wheat Production. *Agronomy Journal.* 103: 1441-1451. ISSN: 0002-1962.

Singh, P.K. & Hughes, G.R. (2006). Genetic similarity among isolates of *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. *J. Phytopathology* 154: 178-184. ISSN: 0931-1785.

Singh, P. K.; Singh, R. P., Duveiller, E., Mergoum, M., Adhikari, T.B. & Elias, E. M. (2010). Genetics of wheat–*Pyrenophora tritici-repentis* interactions. *Euphytica*. 171:1–13. ISSN: 0014-2336.

Sprague, R. (1950). Diseases of Cereals and Grasses in North America. Ronald Press Co., New York. 538 pp.

Statkeviciute, G.; Brazauskas, G., Semaskiene, R., Leistrumaitê, A. & Dabkevicius, Z. (2010). Pyrenophora teres genetic diversity as detected by ISSR analysis. *Agriculture* 97: 91:98.

Stenglein, S.A. & Balatti, P.A. (2006). Genetic diversity of *Phaeoisariopsis griseola* in Argentina as revealed by pathogenic and molecular markers. *Physiological and Molecular Plant Pathology, 68,* 158-167. ISSN: 0085-5765.

Stenglein S.A.; Rodriguero, M.S., Chandler, E., Jennings, P., Salerno, G. & Nicholson, P. (2010). Phylogenetic relationships of Fusarium poae based on nuclear and

Pyrenophora tritici-repentis, Causal Agent
of Tan Spot: A Review of Intraspecific Genetic Diversity
313

mitochondrial sequences. *Fungal Biology* (formerly known as Mycological Research). 114: 96-106. ISSN: 1878-6146.

Stover, R. W.; Francl, L. J. & Jordahl. (1996). Tillage and fungicide management of foliar diseases in a spring wheat monoculture. *J. Prod. Agric.* 9: 261-265. ISSN:

Strelkov, S.E.; Lamari, L., Sayoud, R. & Smith, R.B. (2002). Comparative virulence of chlorosis-inducing races of *Pyrenophora tritici-repentis*. *Canadian Journal of Plant Pathology* 24: 29-35. ISSN: 0706-0661.

Strelkov, S.E. & Lamari, L. (2003) Host-parasite interaction in tan spot *Pyrenophora tritici-repentis* of wheat. *Canadian Journal of Plant Pathology* 25: 339-349. ISSN: 0706-0661.

Strelkov, S.E.; Lamari, L. & Ballance, G.M. (1999). Characterization of a host-specific protein toxin (Ptr Tox B) from *Pyrenohora tritici-repentis*. *Molec. Plant Microb. Interact.* 12:728-732. ISSN: 0894-0282.

Strelkov, S.E.; Kowatsch, R.F., Balance, G.M. & Lamari, L. (2006). Characterization of the ToxB gene from North America and Canadian isolates of *Pyrenophora tritici-repentis*. *Physiological and Molecular Plant Pathology* 67: 164-170. ISSN: 0085-5765.

Strong, D.R. & Levin, D.A. (1979). Species richness of plant parasites and growth from their hosts. *Am. Nat.* 114: 1-22.

Sutton, J. C. & Roke, G. (1986) Interactive effects of foliar diseases and fungicide sprays in cultivars of winter wheat in Ontario. *Can. Plant. Dis. Surv.* 66:37-41. ISSN: 0008-476X.

Tautz, D. & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukariotic genomes. *Nucleic. Acids Res.* 23: 249-255. ISSN: 0305-1048.

Tekauz, A. (1976). Distribution, severity and relative importance of leaf spot diseases wheat in Western Canada en 1974. *Can. Plant. Dis. Surv.* 56: 36-40. ISSN: 0008-476X.

Texeira, H.; Vieira, M.G.G.C. & Machado, J.C. (2004). Marcadores morfofisiológicos e isoenzimáticos na análise da diversidade genética de isolados de *Acremonium strictum*. *Fitopatología Brasileira* 29: 413-418. ISSN: 0100-458.

Tomás, A. & Bockus, W.W. (1987). Cultivar specific toxicity of culture filtrate of *Pyrenophora tritici-repentis*. *Phytopathology* 77: 1337-1366. ISSN: 0031-949X.

Tomás, A.; Feng, G.H., Bockus, W.W. & Leach, J.E. (1990). Purification of a cultivar-specific toxin from *Pyrenophora tritici-repentis*, causal agent of tan spot of wheat. *Molecular Plant Microb. Interact.* 3: 221-224. ISSN: 0894-0282.

Tóth, B; Mesterházy, A., Nicholson, P., Téren, J. & Varga, J. (2004). Mycotoxin production and molecular variability of European an American isolates of *Fusarium culmorum*. *European Journal Plant Pathology* 110: 587-599. ISSN: 0929-1873.

Tuori, R.P.; Wolpert, T.J. & Ciuffetti, L.M. (1995). Purification and immunological characterization of toxic components from cultures of *Pyrenophora tritici-repentis*. *Molec. Plant Microb. Interact.* 8:41-48. ISSN: 0894-0282.

Valder, P. G. (1954), Yellow leaf spot and pink grain in wheat. *Agric. Gaz. N. S. W.* 65: 36-37.

Valim-labres, M.E.; Prestes, A.M., van der Sand, S.T. & Matsumura, A.T.S. (1997). Variacao no aspecto cultural, morfología e virulencia em isolados de *Bipolaris sorokiniana* de trigo. *Fitopatología Brasileira* 22: 483-487. ISSN: 0100-4158.

Vanderplank, J.E. (1978). Genetic and molecular basis of plant pathogenesis. Springer-Verlag, Berlin, Germany.

Vanderplank, J.E. (1984). Disease resistance in plants. 2nd de, Academic Press. Inc. Orlando, FL, United State of America.

Versalovic, J.; Koeuth, T. & Lupski, J.R. (1991). Distribution of repetitive DNA sequences in eubacteria and application to fingerprint of bacterial genoma. *Nucleic Acid Research.* 19: 6823-6831. ISSN: 0305-1048.

Vos, P.; Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Freijters, A., Pot, J., Peleman, J., Kuiper, M. & Zabeau, M. (1995). AFLP: a new concept for DNA fingerprinting. *Nucleic Acids Res* 21, 4407-4414. ISSN: 0305-1048.

Watkins, J. E.; Odvody, G. N., Boosalis, M. G. & Partridge, J. E. (1978). An epidemic of tan spot of wheat in Nebraska. *Plant Dis. Reporter.* 62: 132-134. ISSN: 0191-2917.

Wellings, C. R., & McIntosh, R. A. (1990).*Puccinia striformis* f. sp. *tritici* in Australia: Pathogen Change during the First 10 years. *Plant Pathol.* 39: 316-325. ISSN: 1365-3059.

Welz, H.G.; Köhler, W. & Leonard, K.J. (1994). Isozyme variation within and among pathogenic races of *Cochiobolus carborum* on corn in North Carolina. *Phytopathology* 84:31-38. ISSN: 0031-949X.

White, T.J.; Bruns, T., Lee, S. & Taylor. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Chapter 38 In: Innis et al., PCR Protocols: A Guide to Methods and Applications, San Diego, CA.., Academic Press.p 315-322.

Wiese MV (1987). *Compendium of Wheat Diseases*. (2nd ed.). St. Paul, MN, USA. APS Press.

Wrigth, K. H. & Sutton, J. C. (1990). Inoculation of *Pyrenophora tritici-repentis* in relation to epidemics of tan spot of winter wheat in Ontario. *Can. J. Plant Pathology.* 12: 149-157. ISSN: 0706-0661.

Wu, H.L.; Steffenson, B.J., Li, Y., Oleson, A.E. & Zhong, S. (2003). Genetic variation for virulence and RFLP markers in *Pyrenophora teres. Canadian Journal Plant Pathology* 25: 82-90. ISSN: 0706-0661.

Yli-Mattila, T.; Mach, R.L.; Alekhina, I.A.; Bulat, S.A.; Koskinen, S.; Kullnig-Gradinger, C. M.; kubicek, C. P. & Klemsdal, S.S. (2004). Phylogenetic relationship of *Fusarium langsethiae* to *Fusarium poae* and *Fusarium sporotrichoides* as inferred by IGS, ITS, B-tubulin sequences and UP-PCR hybridization analysis. *International Journal of Food Microbiology.* 267:285. ISSN: 0168-1605.

Young, R.A. & Kelly, J.D. (1997). RAPD markers linked to three major anthracnose resistance genes in common bean. *Crop Science.* 37: 940-946. ISSN: 0011-183X.

Zadoks, J. C. & Schein, R. D. (1979) Epidemiology and plant disease management, New york, Oxford University. Press. 427.

Zhang, H.; Francl, L.J., Jordahl, J.G. & Meinhardt, S.W. (1997). Structural and physical properties of a necrosis-inducing toxin from *Pyrenophora tritici-repentis* . *Phytopathology* 87: 154-160. ISSN: 0031-949X.

Zein, I.; Jawhar, M. & Arabi, M.I.E. (2010). Efficiency of IRAP and ITS-RFLP marker systems in accessing genetic variation of *Pyrenophora graminea. Genetics and Molecular Biology.* 33: 328-332. ISSN: 1415-4757.

Zietkiewicz, E.; Rafalski, A. & Labuda, D. (1994). Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genome* 20: 176-178. ISSN: 0831-2796.

Zhong, S. & Steffenson, B.J. (2001). Virulence and molecular diversity in *Cochliolobus sativus. Phytopathology* 91: 469-476. ISSN: 0031-949X.

Magnaporthe oryzae Genetic Diversity and Its Outcomes on the Search for Durable Resistance

Klaus Konrad Scheuermann[1], Juliana Vieira Raimondi[2],
Rubens Marschalek[1], Alexander de Andrade[1] and Ester Wickert[1]
[1]EPAGRI, Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina
[2]UFSC, Universidade Federal de Santa Catarina
Brazil

1. Introduction

Rice is one of the most important cereal crops, feeding more than 50% of the world population. In some places in Asia, this cereal is responsible for more than half of the total calories intake. Also in Brazil this crop is essential to supply food and is one of the crops responsible for great economic incomes for the Brazilian southern states.

If we consider that the world populations will continue to raise for approximately another 30-40 years (until 2040-2050), so the absolute demand for food will also increase and it would be necessary for the agriculture in all parts of the world to make improvements in order to support the demand.

This required increase in crop production will occur in a context of mounting water scarcity, decreasing area and environmental degradation of arable land (that is partially caused by agriculture), increasing pollution, inevitable emergence of new races and biotypes of pathogens and pests, and possible adverse effects of climate change (Collard & Mackill, 2008).

Plant breeding will play a key role in this coordinated effort for increasing food production. Given the context of current yield trends, predicted population growth and pressure on the environment, traits related to yield stability and sustainability should be a major focus of plant breeding efforts (Collard & Mackill, 2008).

Meanwhile, if this requirement must be attended under environmental conditions of sustainability, this means that this increase in food production must be done under conditions of the best use of environmental goods and services while not damaging these resources (Pretty, 2008), and this represents an unprecedented challenge to scientists and farmers.

In front of this, studies about genetic background are essential for rice culture, to surpass actual productivity levels and mainly for adaptation of modern cultivars to actual environmental conditions.

Rice blast disease is the most important and destructive disease for rice crop, and is caused by ascomycete *Magnaporthe oryzae*. The disease is spread worldwide occurring in almost all

rice-producing areas and can be very destructive when environmental conditions are favorable. Disease occurrence and severity vary by year, location and even within a field depending on environmental conditions and crop management practices. Yield loss estimates from other areas of the world have ranged from 1-50%. It is estimated that each year this disease destroys enough rice to feed more than 60 million people. The economic losses are nearly unmeasurable, but some data suggests values over $ 70 billion of dollars.

The fungus *M. oryzae* presents genetic diversity and is able to recombine by parasexual reproduction, and this causes difficulties on develop resistant cultivars, being also responsible for the short living resistance of released cultivars.

Because of the narrow genetic basis of the elite lineages used for crosses, one of the main troubles in rice breeding is the search for resistance gene sources and how to use them in order to prevent great losses in the crop. Obviously it is necessary to apply a correct culture management in order to use obtained resistance as long as possible.

In this way, we present a shortly state-of-the-art about the pathogen and its epidemiology, what is reported about resistance to the disease, and something about our experience of breeding and management of these disease. We also explain that in our view, the most promising alternative is the use of cultivars with stacking resistance genes allied to a constant follow up on about pathogen genotype frequency by molecular tools. So, emerging of new pathogen races able to overcome resistance could be identified early and some decisions can be made in order to avoid the cultivars to lose their resistance to the disease over the time.

Knowledge about the pathogen population genetic structure could also be useful in order to guide breeding actions related to which and how many resistance genes to exploit for develop the new rice cultivars.

2. The pathogen and the disease

2.1 Rice blast

The first records about rice blast were made in China and date from 1637 (Ou, 1972). Since then, this disease has been reported in almost all world areas with Rice production. It is considered the most important and destructive rice disease.

Rice blast can occur on all above-ground parts of the plant and is observed since the earlier growing stages until the final grain production. Symptoms in leaves start as small brown necrotic lesions that evolve to larger elliptical or spindle-shaped and whitish to gray with darker borders (Figure 1A). On leaf margins, mainly on the flag leaf ligule area, the presence of encircling lesions can cause the leaf to fall. In the presence of favorable conditions, the lesions may enlarge and coalesce to kill the entire leaf and sometimes under severe conditions, the plant can be killed. Symptoms are also observed on culm, culm nodes and panicle neck node. Internodal infected areas are darker blocking the distribution of xylem sap and causing plant lodging. Nodal infection causes the culm to break at the infected node. Infection of panicle neck node is known as blast neck and is a critical factor for final productivity (Figure 1B). If infection occurs near the panicle exertion, grains are not formed and the panicle remains upward. When the panicle is infected at later stages, grains are partially formed, and sometimes the grain weight can cause panicle break at node region.

2.1.1 Etiology

The Rice blast disease is caused by the fungus *Pyricularia oryzae*, firstly described by Cavara in 1891 (cited by Ou, 1985). *P. oryzae* teleomorph is the filamentous ascomycete *Magnaporthe oryzae* originally described as *M. grisea* (Herbert, 1971). During some time, it was treated that isolates that infected Rice would be formally designated as *oryzae* and isolates affecting other species as *grisea*. Because there were not morphological differences that could be used to discriminate the rice isolates from isolates of other species, it was established that all members of the group would be designated as *M. grisea* or *P. grisea* (Rossman et al., 1990). However, Couch & Kohn (2002) results using multilocus gene genealogies demonstrated that isolates infecting *Digitaria* spp. were distinct and it was established that these isolates would be designated as *M. grisea*, whereas isolates affecting rice and the other species would be identified as *M. oryzae*.

The anamorph *P. oryzae* produces piriform shaped conidia with one to two transversal septa, slightly darked or hyaline, linked to conidiophore by its larger bottom. Conidiophores are septated, simple, rarely branched, showing sympodial growth and slightly browned. The teleomorph *M. oryzae* has not been found in nature, but it has been produced after crossing appropriate compatible isolates in laboratory. The teleomorph stage produces hyaline ascospores, typically fusiform shaped with three-septate and envolved by a unitunicate asci.

2.1.2 The disease cycle

During maturation, the conidia, who remains tied to conidiophore, produces a mucilage that in presence of water helps its fixation to the host (Hamer et al., 1988). This mucilage is mannose rich and presents affinity with lectin concanavalin A found in plants (Hamer et al., 1988).

The conidia germination and appressoria formation are controlled by environmental stimulus and germ tube is not formed if conidia were maintained in water, avoiding anticipated germination (Lee & Dean, 1993). The principal stimulus is related to hydrophobic conditions on plant surface. According Hamer et al. (1988), as more wax is disposable on leaf surface where conidia are deposited more effective is the formation of appressoria. The appressoria capacity on penetrate the plant surface is dependent of turgor pressure. The lack of capacity to form this high pressure is related to chemical components or genetic alterations that could interfere on melanin biosynthesis. Melanin is a compound of appressoria cell wall acting as barrier against permeability, allowing increase on some cytoplasmatic solute concentrations (Howard & Ferrari, 1989). In presence of high turgor pressure, the penetration peg formed by appressorium press the epidermic cells until its rupture, allowing fungi penetration. On cell lumen, *M. oryzae* colonizes tissue adopting hemibiotrophic lifestyle. Invasive hyphae are involved by a membrane that separes it from cell host cytoplasm, a characteristic of biotrophic fungi (Kankanala et al., 2007). The fungi progresses to the next cell without cause damages to the cell wall, probably because this happens by plasmodesmatas (Kankanala et al., 2007). Invasion of the neighbor cell coincides with the loss of viability of the previously infected cell, initiating the necrotrophic phase and appearance of lesions, around 74 to 96h after conidia deposition on plant surface. On

presence of high relative humidity (≥93%), conidia are released less than 24h after lesions appearing, with the highest conidia production occurring among 3 to 8 days after sporulation beginning. Each lesion from a susceptible host can give rise to more than 20.000 conidia over several days (Barksdale & Asai, 1961).

The infection and colonization described above occurs in aerial plant tissues such as culms and leaves and is different from the process that takes place on root system. On roots, appressoria are not formed and penetration pegs are produced directly from hyphopodium initiating penetration on rhizodermal cells (Sesma & Osburn, 2004; Marcel et al., 2010). On next step, a differentiation of invasive hyphae occurs and new cells are invaded, but the previously infected cells are not damaged. So, a biotrophic interaction occurs without producing lesions on infected tissues. However, around four weeks after root infection, symptom disease appears on plant stems and leaves, showing the systemic pathogen distribution (Marcel et al., 2010).

2.1.3 Epidemiology

The primary inoculum is originated from infected seeds, crop residues and spores dispersed by wind from neighbor farming areas. Kingsolver et al. (1984) reported that the mycelium present on field remainder rice straw could survive for around six years. Infected seeds can cause blast on initial steps of vegetative development, but rarely causes disease epidemies. More than 50 other *M. oryzae* hosts are known, including cultivated species (Metha & Baier, 1998) and weed species (Mackill & Bonman, 1986), however cases of crossed infections are rare. The main innoculum sources for secondary infections are the sporulating lesions on infected leaves, characterizing a policyclic disease.

The rice plants presents its highest vulnerability period when young plants hold 3 to 4 leaves and during flowering. Blast disease is also favored by temperatures among 20 to 30 °C and relative humidity over 90%, with production of conidia on lesions beginning when humidity surpasses 93%. Frequent occurrence of dew, mist and drizzle are suggestive of ideal conditions to the pathogen. Disease severity is also affected by soil fertility. High levels of organic matter or excessive use of nitrogen can lead to the disease increase when the cultivar resistance is not specific. However, nitrogen deficiency can predispose plants to disease (Prabhu et al., 1996).

Fig. 1. Rice blast symptoms on leaves (A) and on panicle neck node (B).

2.2 *Magnaporthe oryza* genetic diversity

2.2.1 Genetic diversity

In order to know the *M. oryzae* population structure around the world, many studies were done using molecular markers mainly based on MGR 586 repetitive DNA sequence (Hamer et al., 1989), and pathogenicity tests on differential set of rice varieties. Levy et al. (1991) analyzed 42 isolates obtained from an isolates bank collected during 30 years representing the most common lineages infecting USA commercial rice cultivars and verified the occurrence of eight lineages. On the Philippines, 1156 isolates were obtained on 38 cultivars sowed in two growing seasons and after analysis were grouped in only ten lineages (Chen et al., 1995). Simple population structure was also observed in Europe (Roumen et al., 1997) and Africa, when isolates from four different countries were separated in nine lineages (Takan et al., 2011). Results found in the literature reports a limited number of the pathogen lineages, showing a clonally typical population structure (Levy et al., 1993; Xia et al., 1993; Xia et al., 2000). Despite the low genetic diversity presented by *M. oryzae*, a high pathogenic variability was observed in pathogenic populations. In Colombia, 39 *M. oryzae* pathogenic races were identified (Levy et al., 1993). In Brasil, a similar study identified 45 races (Anjos et al., 2009), with a frequent change on *M. oryzae* races composition (Ribeiro & Terres, 1987).

Among the factors affecting the low *M. oryzae* population diversity are the uniformity of rice culture management and adverse environmental conditions to the pathogen. In this way, Kumar et al. (1999) realized a study in a Himalaya area (India northern) belonging to rice diversity Center. At this region, there is a high diversity of rice cultivars that remained isolated in different valleys during thousands of years and cultivated under different growing systems, in an heterogeneous environmental that favored the pathogen. Using 458 isolates collected in 29 valleys, it was possible group them in 56 lineages, showing a considerable higher genetic diversity than that obtained out of rice origin center.

The low genetic diversity with a substantial pathogenic diversity found in some places and a higher genetic and pathogenic diversity showed in other areas suggests that different mechanisms are acting in order to create genetic variability in *M. oryzae*.

2.2.2 Variability mechanisms

2.2.2.1 Sexual recombination

M. oryzae is a heterotallic fungus with sexual recombination controlled by a single *locus* (MAT1), that presents two mating types, MAT1-1 e MAT1-2. Both of them are essential for sexual recombination (Kang et al., 1994). *M. oryzae* isolates can be hermaphrodite, male fertile or sterile, but crossing only occurs between fertile isolates from different mating types, and when at least one isolate is hermaphrodite (Zeigler, 1998). Many studies were conducted in order to evaluate the frequency of the different mating types around the world. In Argentina, analysis of 125 isolates collected between 2000 and 2003 showed that all of them belonged to mating type MAT1-1 (Consolo et al., 2005). Similar results were obtained in Korea, when the analysis of 254 isolates demonstrated that all of them also belonged to MAT1-1 (Park, et al., 2008). In Africa studies revealed that 29% of isolates belongs to MAT1-1 and 71% to MAT2-2, but none of the isolates was hermaphrodite, avoiding crossing among them (Takan et al. 2011). When two *M. orizae*

populations from Himalayas were analyzed, it was found that 22% of the isolates belonged to MAT1-1 and 43% to MAT1-2, and isolates male fertile and hermaphrodite were detected, suggesting a possible occurrence of sexual recombination in this region (Kumar et al., 1999).

There is a consensus that the majority of the *M. oryzae* populations lost its capacity of sexual reproduction. One reason could be the founder effect, because when rice culture was introduced in a new area, probably only one mating type was established (Zeigler, 1998). It is also possible that in some cases the MAT *locus* could be linked to an avirulence gene, with no compatibility with the prevailing Rice cultivar, and this could represents a drawback for fitness (Notteghem & Silue, 1992), being eliminated along time. The frequent chromosomic rearrangement derived from DNA mobile genetic elements (Chuma et al., 2011) can cause the degeneration of sexual behavior without producing a drawback of fitness.

One role of the sexual reproduction is to form resistance structures for safeguarding the fungi on adverse situations. When rice is cultivated on areas where the weather conditions enable the fungi to survive on a mycelia form during many fallow seasons, mutations that remove sexual reproduction could be fixed because of high energy costs of this character. It is also possible that rice was introduced in worldwide areas where weather conditions would be adverse for sexual reproduction occurrence (Zeigler, 1998).

2.2.2.2 Parasexual reproduction

Parasexual recombination in *M. oryzae* can be an alternative to the sexual cycle, because these cycle rarely occurs. Parasexual cycle involves the formation of anastomosis (haploid hiphal fusion) resulting in heterokaryotic cells that can undergo karyogamy and become a diploid cell and then be able to realize mitotic recombination. The return to the haploid stage occurs by successive lost of one chromosome from each pair of homologous chromosomes (aneuploidy) (Crawford et al., 1986, Noguchi, 2011). Parasexual recombination possibility was first suggested by Yamasaki & Niizeki (1965) that observed anastomosis formation and obtained variants matching auxotrophy isolates. Concomitant culture of genetically distinct *M. oryzae* isolates originated higher new haplotype frequency spontaneously than pure cultures (Xia et al., 1993). Zeigler et al. (1997) carried out matching of genetically different isolates and observed growing of mycelium tufts between cocultured isolates. From formed anastomosys new haplotypes were found, indicating the presence of recombination. The fact that isolates from the same or from different populations can easily recombine corroborates the theory that vegetative compatibility is weakly controlled in *M. oryzae* (Crawford et al., 1986). When virulence spectrum and adaptability from parasexual recombinant isolates were evaluated, it was verified that they could infect Rice cultivars that were not infected by its parents (avirulent) (Noguchi et al., 2007). This study also found that the pathogenicity of recombinant isolates is stable and did not change after seven generations, and that the adaptability is intermediate when compared to the parents. But it was also observed that aggressiveness of new *M. oryzae* recombinants can increase after some years of rice cultivation (Fujita & Suzuki, 1982). Parasexual recombination is a process that generates genetic variability, but could also be a mechanism to restore the genome, preventing that the mutation accumulation endangers pathogen fitness, but it remains unproven at this date.

2.2.2.3 Mutation

A common trait observed among *M. oryzae* is the high genetic instability of characters related to morphology, fertility and pathogenicity when isolates are mantained in laboratory (Valent & Chumley, 1991). The Buf1 gene, which is involved in melanine synthesis, an essential role for penetration process, presents a mutation rate that can reach 20% (Chumley & Valent, 1990).

The *M. oryzae* x O. *sativa* interaction belongs to gene-gene system (Jia et al., 2000) and mutations in avirulence genes (*Avr*) can be decisive for plant-pathogen relationship. Some of *M. oryzae* avirulence genes are located in telomeric and sub-telomeric regions, as is the case of *Avr-Pita* (Orbach et al., 2000), *AvrMedNoï-1; AvrKu86-1* (Dioh et al., 2000), *Avr-Piz* (Luo et al., 2002), *Avr-Pii* (Yasuda et al., 2006), *Avr-Pi15* (Ma et al., 2006), and *Avr-Pit* and *Avr-Pia* (Chen et al., 2007). Deletions in these regions could result in *Avr* genes lost, resulting in virulence gain, as demonstrated by Orbach et al. (2000). These authors identified spontaneous mutants that suffered deletions on *Avr-Pita* gene and so they were able to infect cultivars with resistance codified by the correspondent *Pi-ta* gene. With *M. oryzae* complete genome sequencing it was verified that 9.7 % of the genome was constituted by repetitive DNA sequences (Dean et al., 2005). Substantial part of these sequences are transposons, belonging to eight families, five of them of retrotransposons and three of DNA transposons (Dean et al., 2005). Some of these transposons presented more than hundred copies dispersed over the genome (Kachroo et al., 1994). The highest transposon concentration was observed on telomeric regions, hosting around 24% of these elements (Rehmeyer et al., 2006). The presence of multiple transposon copies in regions next to *Avr* genes could result in recombination among copies of the same element, resulting in deletion of one copy, which could take away one *Avr* gene (Rehmeyer et al., 2006). Farmanet et al. (2002) studied the *Avr1-CO39* gene structure on isolates from different origin and verified that these structure present variation and is not functional in most of the cases. Analysis of regions flanking *Avr1-CO39* allows the identification of RETRO5 and REP1 transposons suggesting that recombination among copies of these elements is related to the frequent alterations observed on *Avr1-CO39* gene structure.

Transposons are recognized as hotspots for recombinations, but can also operate actively introducing themselves in codifying genome regions, resulting in inhibition or alteration of gene expression. Kanget et al. (2001) showed that avirulent *M. oryzae* isolates to cultivars whose resistance was mediated by *Pi-ta* gene could infect the cultivar when transposon *Pot3* was inserted on *Avr-Pita* regulatory region. Transposon insertions on avirulence genes were observed also for the gene *Avr1-CO39* (Farman et al., 2002), for genes controlling morphologic traits as *Acr1*, which are involved on conidia morphological control (Nishimura et al., 2000). Recently, it was demonstrated that transposons can move through the genome and can carry *Avr* genes located next to them. Chuma et al. (2011) verified that the *Avr-Pita* gene could be located in different places of the same chromosome, in different chromosomes and also in supernumerary chromosomes, and that it was flanked by other transposons possibly involved with its translocation. This work also suggests that the *Avr-Pia*, *Avr-Pik* and *Avr-Pii* genes also suffered multiple translocation events.

Presence of *Avr* genes on *M. oryzae* supernumerary chromosomes indicates its involvement on avirulence genes dynamics. These chromosomes are known to be easily lost from the

genome, probably because they are unnecessary (Johnson et al., 2001). This property could be adaptatively advantageous because pathogen population can adapt quickly to cultivars carrying correspondent *R* genes by losing supernumerary chromosomes. As these chromosomes containing *Avr* genes could be horizontally transferred (Akagiet al., 2009) by recombination to isolates that had lost briefly its *Avr* genes because of cultivar resistance, these restored *Avr* genes could confer some fitness to the pathogen population.

2.2.3 Resistance to blast

Two basic resistance types commonly recognized can be used to control rice blast disease, partial and complete resistance. The complete resistance is race specific and governed by one or few R genes, dominant or recessive, that recognizes the product of the pathogen *Avr* gene (Jia et al., 2000). In this cases, when resistance is effective, absence of disease symptoms are observed, whereas when resistance is not effective, disease symptoms occurs in large scale. On the other hand, partial resistance is not race specific and is controlled by QTL´s (Quantitative Trait Loci) characterized by reducing the number and size of lesions (Vergne et al., 2010), acting in a similar manner for different pathogen races. But specifically for rice, some exceptions can be considered. Zenbayashi-Sawata et al. (2005) identified an avirulence gene (*Avr-Pi34*) correspondent to the resistance gene *Pi34,* that confers partial resistance. Analyzing the resistance spectrum of a set of rice QTL´s, it was verified that only three of them conferred broad spectrum resistance (Ballini et al. 2008). These results suggest that partial resistance is sometimes specific and does not necessarily have a broader resistance spectrum when compared to complete resistance (Ballini et al., 2008).

A careful analysis of a set of previously described rice blast resistance genes allowed identification of 85 resistance genes and around 350 QTLs (Ballini et al., 2008). Until now, 15 *R* genes have been cloned (Table 1, adapted from Liu et al., 2010) and others are mapped (Ballini et al., 2008), being available for use in breeding programs. If we consider the *M. oryzae* genetic variability and its capacity to adapt to new rice cultivars, the use of blast resistance genes in isolate, alternate or simultaneous manner in different cultivars could cause the overcome of each of them individually. This could allow the pathogen population to form a 'pathogenic memory´, or in other words, to select races able to infect the cultivars carrying each of the resistance genes. To overcome the resistance, pathogen populations just need change races frequency according the cultivars.

The short living effectiveness of cultivars with single gene resistance because of race changes in pathogen populations had lead resistance breeding strategies usually take strategies as the gene stacking or the multiline or cultivar mixture approach. The gene stacking approach incorporates different resistance genes into a plant selection that is then grown as a homogeneous population, whereas the multiline and mixtures strategies disperses resistance genes among different plants that are grown as a heterogeneous population (Wilson et al., 2001).

In front of this, we propose that the highest possible number of genes assembled in the same cultivar, in a process known as pyramiding of genes (Servin et al., 2004) could be the more effective way to face the rice blast disease in our rice culture conditions. But pyramiding genes could enable the emerging of a pathogen ´super-race´ able to overcome resistance in a single event. As it was previously discussed, the overcome of a *R* gene by the pathogen is

possible because of the lost of the *Avr* function. Some works had demonstrated that *Avr* genes are involved with pathogen adaptability acting on processes of host infection and colonization (Leach et al., 2001). So, the higher the number of *Avr* genes that the pathogen need to lost, the lower would be its capacity to cause the disease, until a level that it would fail on cause symptoms. Assuming that this is true, why rice genotypes with stacking genes were not naturally be selected during evolution? Possibly because pyramiding genes represents an additional physiological cost for the plant fitness. Tian et al. (2003) measured the introgression cost of the resistance gene *Pseudomonas syringae RPM1* in *Arabidopsis thaliana*, verifying that isogenic lines that received the gene produced around 9% less seeds when compared to normal plants. This resistance cost is also observed on plants that were submitted to chemical resistance induction. Wheat plants treated with acibenzolar-S-methyl to induce resistance, showed reduction on number of tillers and number of spikelets in each ear (Heil et al., 2000). So, if a rice cultivar with pyramiding genes against blast disease would be developed, possibly it would present the similar results, increasing resistance and decreasing productivity. There are also cases where reistance genes are linked to agronomic undesirable traits, as occurs with *Pi-21* gene, that is linked to a gene codifying for chalky on rice grain (Fukuoka et al., 2009).

R gene	Codified protein	Reference
Pib	NBS-LRR	Wang et al., 1999
Pi-ta	NBS-LRR	Bryan et al., 2000
Pi9	NBS-LRR	Qu et al., 2006
Pi2	NBS-LRR	Zhou et al., 2006
Piz-t	NBS-LRR	Zhou et al., 2006
Pi-d2	Kinase receptor	Chen et al., 2006
Pi36	NBS-LRR	Liu et al., 2007
Pi37	NBS-LRR	Lin et al., 2007
Pikm	NBS-LRR	Ashikawa et al., 2008
Pit	NBS-LRR	Hayashi and Yoshida, 2009
Pi5	NBS-LRR	Lee et al., 2009
Pid3	NBS-LRR	Shang et al., 2009
Pikh	NBS-LRR	Rai et al., 2009
Pikp-1	NBS-LRR	Yuan et al., 2011
Pikp-2	NBS-LRR	Yuan et al., 2011

Table 1. Cloned rice blast resistance genes and the proteins codified by them.

So, the selection of genes for pyramiding must be done carefully, in order to introgress genes with broader resistance spectrum and whose *Avr* genes are not easily mutated, as occurs with *Avr-Pita* gene.

3. Breeding rice for resistance to blast

The *Oryza* L. genera belongs to the Poaceae (Gramineae) family and encompasses around 25 species, of which 23 are considered wild and only two are cultivated, *O. sativa* L. e *O. glaberrima* Steud (Khushi, 1997). *O. sativa* species is cultivated worldwide and *O. glaberrima* only in Africa.

The commercial exploitation of a reduced genetic base and the prevalence of a small set of landraces in the breeding process has been the general approach for several crops species. Compared with other crop species, the genetic diversity in the world rice germplasm is quite large. Despite the richness of genetic resources, only a small proportion of the world rice germplasm collections have been used in breeding programs. As a consequence a high genetic similarity is found within several commercial rice germplasms around the world. The limited use of the rice genetic diversity available worldwide has been a concern in Latin America since the late 1980s. Although differences in genetic diversity and relatedness are observed within rice germplasms of different countries, the general feature is a very close relationship among cultivars (Cuevas-Perez et al., 1992; Guimarães et al., 1996; Fuentes et al., 1999). For example, a study demonstrated that irrigated rice varieties currently under cultivation in Venezuela were closely related among them and with cultivars from Colombia, Brazil and Ecuador (Cuevas-Perez et al., 1992). In Brazil, studies demonstrated that Brazilian rice cultivars presents a narrow genetic basis, with only 10 parents responsible for 68% of the genetic pools (Rangel et al., 1996). The same study also showed that only Tetep and Tadukan cultivars were used as gene sources for blast resistance to all Brazilian commercial rice cultivars at that time. But there was little progress in this way, and this still a troubling condition if we consider the very large areas that are cultivated with rice in Brazil.

The Brazilian experience has demonstrated that rice cultivars with complete resistance were short lived. As an example, rice cultivars Metica 1 and Aliança were overcome by pathogen only one year after released. Javaé cultivar presented high resistance levels, but was overcome two years after. In 1997, Javaé cultivar was substituted by Formoso cultivar, and the first results were promising. However, in the 1999/2000 growing season, the resistance was overcome, causing great losses for farmers who did not used fungicides.

So, efforts are being done in order to better understand the rice-*Magnaporthe* pathosystem, on aspects like search for resistance genes and establishment of pathogen population structure for development of cultivars with long live resistance. Efforts are also being done about cultural practices that can help farmers to produce rice without damaging environmental and his own health, and so, supply the final consumer with highest quality rice.

3.1 The experience of Epagri´s rice breeding programme

Rice is one of the major crops for Brazil and for the state of Santa Catarina, and the annual production value for this state represents around US$ 300 millions. The state of Santa Catarina has an area of 1.12% of the Brazilian territory, although the area that could be farmed is only 12% of its territory, nevertheless, this state is ranked the fifth food supplier in Brazil. The total state rice area of 149.000 ha in 2010 growing season is cultivated by 8.500 farmers, in 11.000 small farms belonging to 83 counties and the seeding system is water-seeded (Figure 2).

The Santa Catarina State rice yield in 2009/10 was estimated in 7.0 t/ha, whereas in some regions, like High Itajaí Valley, the yield can reach up to 12 t/ha (Figure 3). In fact, yields of 10 t/ha are very common among farmers that use the recommended technologies, including modern varieties. The establishment of paddy rice as a major crop and its economic

feasibility in the small family owned farms in Santa Catarina, is the result of the technologies developed by public research and rural extension efforts.

The consequences, in this sense, are the attainment of very high rice yields, which can, in great part, be attributed to the use of the modern rice varieties developed by Epagri (Santa Catarina State Agricultural Research and Rural Extension Agency). In the late 70´s yields were around 2.3 t/ha. Epagri´s Breeding Program started in 1976 at Itajaí Experiment Station. Since 1980 Epagri has released 17 varieties (Table 2). This was decisive to improve rice yield in Santa Catarina. Epagri´s varieties are spread all over the rice irrigated areas in Brazil , reaching also other countries like Paraguai, Bolivia, Argentina and Venezuela.

The general goals of Epagri´s Rice Breeding Program are developing improved genotypes of long grains, to improve blast resistance and grain quality. Recently, the program is directed to promote genetic variability and also to develop technologies that have lower environmental impact.

Fig. 2. Aspect of rice culture in Santa Catarina State, with rice growing in water-seeded system.

In the beginning of the research program, local grown varieties brought by immigrant farmers from Italy, were replaced by new lines from national and international institutes.

The methods currently used in the Epagri´s breeding program are pedigree, recurrent selection and induced mutations. Annually, around 230-300 different crosses and backcrosses are done. These originate at least 100.000 individual F_2 and F_3 plants which are submitted to field selection to compose the preliminary studies which include 300 F_4 families, and in the next year, 200 F_5 families. The F_1 to F_5 populations are hand transplanted to the field (Figure 4). In the sequence about 30 to 50 F_6 lines form the advanced experiment in broadcast sowed plots. In the final stages, normally not early than in F_7, the advanced lines (around 20) are tested in 5 different locations (on-farm

trials) for three years through the three major rice areas of Santa Catarina. The management of this rice fields follows Epagri´s recommendations. In these experiments, each variety is broadcast sowed in 4 x 15m plots.

Fig. 3. Overview of a rice crop field on High Itajaí Valley.

Fig. 4. Overview of Epagri´s Rice Breeding Program experimental field.

It is unavoidable to conclude that Epagri´s Rice Breeding Program is successful and has contributed, recommending varieties from abroad and, developing new varieties, to increase rice yield in Santa Catarina (only paddy rice considered). The highly quality and yielding varieties released by the Breeding Program of Epagri contributed for the State to hold the highest rice yield in Brazil.

But in recent years, blast disease emerged as a potential destructive disease for rice culture in Santa Catarina State. And new strategies were planned in order to face the problem and improve genetic resistance in Epagri´s rice released cultivars. These strategies include cross elite rice lines with known resistance sources, increase of the genetic base by crossing stabilized lines with other *Oryza* species and developing of cultivars with pyramiding genes.

3.1.2 Crosses with resistance sources

Epagri´s Rice Breeding Program have been using rice cultivars and lineages known and tested for resistance to blast, as Tetep, Oryzica Llanos 5, NP-125, Raminad Str. 3, WC 299 and WC 277. This blast resistance sources were all used in crosses with elite rice cultivars and the progeny was conducted by Pedigree method. Despite all efforts made to improve the efficiency for selecting and developing rice cultivars, blast resistance is continuously being lost in breeding lines after the F3 – F4 or later generations. In our experience, Raminad and Oryzica Llanos 5 produced the higher number of plants with some good agronomic traits and resistance to blast, but they failed in final comparing assays related to productivity and grain cooking quality. So, until now, Epagri´s released cultivars were not directly related to these blast resistance sources.

Cultivar	Releasing year	Yield (t/ha)
Empasc 101	1980	6.79
Empasc 102	1980	6.66
Empasc 103	1981	6.38
Empasc 104	1985	7.41
Empasc 105	1987	7.79
Epagri 106	1992	6.86
Epagri 107	1994	7.27
Epagri 108	1995	7.83
Epagri 109	1996	8.90
SCS BRS 111	2000	7.00
SCS 112	2000	7.29
SCSBRS Tio Taka	2002	8.84
SCS 114 Andosan	2005	9.07
SCS 115 CL	2007	7.90
SCS 116 Satoru	2010	9.40
SCS 117 CL	2011	8.90

Table 2. Epagri´s Rice Breeding Program released cultivars, with respective year or release and yield.

In 2010/2011 growing season, twelve lines originated from crosses with blast resistance sources are being evaluated for agronomic and resistance traits in final assays. Six of these lines were originated from crosses with Oryzica Llanos 5, three from WC 299, two from WC 277 and one from NP 125. On next growing season, this lines would be tested again in field assays order to observe their agronomic and resistance traits. These assays are done for three years in five different counties covering the several rice growing areas in Santa Catarina State.

3.1.3 Crosses with other *Oryza* species

The wild species of *Oryza* contains numerous genes of economic importance and could be used as alternate sources of resistance or tolerance to biotic and abiotic stresses to enrich the cultivated rice gene pool (Jena, 2010).

Efforts were also done related to crosses with other rice species, as *O. rufipogon*, *O. nivara*, *O. latifolia* and *O. glumaepatula*. These four species were introduced in Epagri´s Rice Program in 2006 and since then are used in crosses with elite cultivars (SCS 116 Satoru, SCS 114 Andosan, SCSBRS Tio Taka, Epagri 109, SCS 112, SC 339 and SC 213, among others). The main objective for use this species is the increase of genetic variability of the Epagri´s rice elite lines and the introduction of new genes related to resistance to diseases and pests. This are being done by the backcrossing method, because these species also present undesirable traits for a modern rice cultivar.

These four species are known for its traits related to tolerance/resistance to biotic and abiotic stress (Thanh et al., 2011; Nonomura et al., 2010; Vaughan et al., 2005, Rangel et al., 2006). In the 2010/2011 growing season, when evaluating the F3 generation, segregant populations originated from crosses between *O. glaberrima* and SCS 112 presented good agronomic traits and a productivity of more that 10 t/ha. Despite this promising results, these progenies will be carefully observed in the next segregant generations for its traits related to disease resistance, mainly for blast disease. Also in this growing season, F4 segregating populations originated from the cross *O. glumaepatula*/SC 355//SC213 presented productivities ranging from 8,5 to 10 t/ha, and interesting agronomic traits related to grain quality. The species *O. nivara* and *O. rufipogon* were recently crossed with Epagri´s rice lines, and in the 2010/2011 season were initial segregating generations F2.

Crosses with these species are being done in the last four years, and the resulting progenies are being observed and selected in the four first segregating generations. If the promising initial results becomes true, we hope that in the next years we could release a rice cultivar with desired agronomic traits, quality and productivity allied to traits related to blast resistance.

3.1.4 Breeding for durable resistance

Considering the significance and the potential losses that blast causes, rice breeding programs since early dispended efforts on developing cultivars with resistance to the disease. Therefore, crosses among elite cultivars and lines with resistance genes were extensively done. However, during breeding process, evaluations were done based on resistance phenotypic traits, leading to selection of lines/cultivars with complete resistance. This happen because it is difficult to identify and evaluate the partial-resistance on this segregating lines, and this leads to erosion of this resistance form. As a final result, cultivars with resistance based on one or few *R* genes were obtained and released, causing an overcome of this resistance after two or three years.

With the conclusion of the rice genome sequencing (Goff et al., 2002; Yu et al., 2002) that allowed cloning and mapping several *R* genes and QTLs, a high number of molecular markers were available for MAS (marker assisted selection), mainly to follow traits as resistance gene during segregant generations (Jena & Mackill, 2008). So, it is possible to obtain lines that carry genes codifying for partial and complete resistance. These lines are interesting because if the pathogen overcomes the *R* genes, plants are not completely vulnerable, remaining with partial resistance.

Incorporation of durable disease resistance into susceptible rice backgrounds has often been difficult using conventional breeding methods. Conventional breeding programs have been

limited to monogenic, race-specific resistance genes since they are easily introgressed into susceptible rice elite lines through simple backcrossing techniques. However, single gene resistance has often proven ephemeral and highly vulnerable to dynamic and diverse plant pathogen populations. Therefore, breeders are endeavoring to shift to breeding durable forms of resistance by pyramiding race-specific genes into a single cultivar. Conventional gene pyramiding requires extensive disease screenings with several races of the pathogen due to the race specificity of many of these genes after each cycle of crossing. Further complicating a pyramiding effort is the frequent absence of an effective selection method due to a lack of differentiating races. Testcrossing to susceptible genotypes is required in each cycle to detect the presence (or absence) of the masked genes due to epistatic interactions between many resistance genes. In addition, as highly effective resistance genes that are effective to many races of the pathogen are incorporated into breeding populations, valuable hypostatic genes are easily lost in each backcross cycle. MAS of hypostatic genes would facilitate their maintenance in pyramiding populations. Molecular markers tightly linked to resistance genes have facilitated their pyramiding into single elite cultivars, with the possibility of creating more durable or broad spectrum resistance (Saghai-Maroof et al., 2008).

Breeding for durable resistance against blast disease would be more effective with a previous knowledge of the pathogen populations from the country or place where the future cultivar will be cultivated. This characterization can be done by pathogenicity tests on international set of rice varieties (Atkins et al., 1967) or by molecular markers (George et al., 1998). When pathogen population structure is known, effective resistance genes could be easily identified. At the present time, isolines (Mackill & Bonman, 1992) and lineages holding only one known resistance gene are available, in order to help plant breeder in this first steps to develop a resistant cultivar.

We have done a previous study on *M. oryzae* pathogen population covering Santa Catarina State (unpublished data) using Pot-2 marker. We were able to identify 13 haplotypes that were inoculated on international set of rice varieties (Atkins et al., 1967), allowing us to classify the isolates in six distinct races. The inoculation of isolates representing these six races on isolines developed by Mackill & Bonman (1992) suggests that *Pi-1*, *Pi-11* e *Pi-4b* genes ensure resistance against all tested races, despite the fact that pathogenic variations among races are known and documented. Based on these informations crosses are being done in order to incorporate these three genes on elite lines and cultivars. The next step is to follow up these genes by molecular markers. On each generation the lines are being submitted to anther cultivation in order to stabilize segregation. The regenerated lines would be analyzed by MAS and submitted to pathogenicity tests to allow selection only of lines with homozygous resistance genes only.

A study covering all the country is also being done to evaluate *M. oryzae* population present on main Brazilian Rice production areas. This is being done by Embrapa (Empresa Brasileira de Pesquisa Agropecuária) and has the objective to suggest to plant breeders the more effective resistance genes to be incorporated on new rice cultivars.

4. Future perspectives

During the 20th and 21st centuries, widespread, intensive, and systematic efforts have been made and continue to be made by plant breeders throughout the world to breed plants that

combine the most useful genes for higher yields, better quality, uniform size of plants and fruits, uniform ripening, cold hardiness, and disease resistance. In searching for new useful genes, plant breeders cross existing, local, cultivated varieties with one another, with those of other localities, both here and abroad, and with wild species of crop plants from wherever they can be obtained (Agrios, 2005).

Because of the genetic variability of *M. oryzae* and the possibility of parasexual recombination added to favorable environmental conditions to the disease in the tropics (high temperature and relative humidity), the practice had demonstrated that the resistant cultivars lost its effectiveness in less than five years, causing great yield losses. This also happened to Epagri´s rice released cultivars, which were selected for exhibit resistance. As Santa Catarina State is considered a new production area (culture ´boom´ begins in late 1970), the first Epagri´s cultivars showed resistance for many years. But with intensive use of same culture field year after year, some years ago several problems related to pests and disease began to emerge, among that the rice blast. And today almost all 17 Epagri´s released cultivars are affected by blast in some level. This prompts us to search for strategies to manage the problem. The main of them is to develop cultivars with durable resistance.

Developing durable resistance requires strategies that limit the ability of pathogen populations to adapt to host resistance. The ephemeral effectiveness of resistance to some diseases with great potential has resulted in considerable research into the nature of durable disease resistance, but despite the efforts devoted to understand it, very few successful examples have been confirmed (Wilson et al., 2001). To date, the acknowledged forms to face plant diseases are the use of resistant cultivars, the use of multilines or cultivar mixtures and the use of fungicides. Culture management techniques as crop rotation, control of weed plants and incorporation to soil of crops residues are also recommended for rice culture.

The development of blast resistant rice cultivars by pyramiding genes is a process that is hard-working and time-consuming (at least 8 to 9 years by traditional breeding methods) and the expected result is that the resulting cultivar would show effectiveness and longevity. Theoretical models have been used to estimate the longevity of crop protection afforded by three highly effective resistance genes used under different strategies. In absence of stabilizing selection in the pathogen population, the models indicated few differences if the genes were used sequentially (13 to 14 years), stacked into a single variety (13 to 14 years) or used in multilines (12 to 13 years), considering that composition of multigenic and multilines host population was static year to year (Wilson et al., 2001).

The durability of resistance for a large number of fungal and oomycete pathogens was evaluated by McDonald & Linde (2002). These authors proposed that the risk of failure of genetical control depends not on the nature of the resistance genes but on the evolutionary potential of the pathogen. So, the evolutionary potential of a pathogen population is reflected in its population genetic structure. They identified three factors, population size, rate of gene and genome flow (migration) and the reproduction or mating system (i.e. asexual or sexual), from which to produce a risk factor for loss of resistance. The calculated risk for *P. grisea* was considered high because of the large effective population size, mixed reproduction/mating system and medium mutation rate. So, these authors estimated that the average number of years before the pathogen causes detectable damage on previously resistant cultivars was 1-3 years.

According our understanding, the more effective form to control the disease is the use of cultivars with pyramiding genes for blast resistance. As the development of a cultivar often needs a long time to select and test its agronomic traits, there are few rice released cultivars with pyramiding genes actually available, and none evaluated and ready for cultivation in Santa Catarina State. And while there is not a stacking cultivar available, other strategies need to be implemented.

The use of cultivar mixtures or multilines needed to be carefully considered. Different studies had demonstrated that complex races that can attack various resistance genes have been assumed to be unable to increase in a cultivar mixture or multilines because of the costs associated with a lack of avirulence genes (Zhu et al., 2000; Mundt, 2002; Garret & Mundt, 1999; Wilson et al., 2001). But recently, a study suggested that complex races can overcome multiple host genotypes in a multiline system by producing variants that can infect resist lines through parasexual recombination (Noguchi et al., 2007). The variants had a more complex virulence than the parents and exhibited a level of fitness equal to that of the parents. In particular, in a cultivar mixture, leaf blast caused by the variants was more severe than that caused by either parent. These results suggest that parasexual recombination not only alters pathogenicity but also enhances fitness, such as lesion enlargement and spore production, leading to lack of resistance (Noguchi et al., 2011).

So, a specific strategy to Santa Catarina State could be developed. Considering that the majority of rice is cultivated by little farms, owing areas as 1 to 10 ha, the Epagri´s rice cultivars could be grown in a scheme like patchwork, with each farm cultivating a different rice cultivar, maintaining disease levels in a low and manageable scale. But in presence of favorable climatic conditions to the pathogen, the use of fungicides must be considered in order to reduce economic losses.

Considering that we know that *M. oryzae* presents low genetic diversity in Santa Catarina State, the use of Epagri´s rice cultivars in an organized manner by farmers and allied to recommended cultural management could help to decrease the use of fungicides and to maintain pathogen populations under control until new resistant cultivars will be released.

With the long history of coevolution between pathogens and its hosts, it seems unlikely that we will be able to eliminate plant diseases by engineering cultivars with new genes or receptors or combinations of genes/receptors and putting them into our crops. Pathogens will continue to evolve. However, genetic engineering offers new opportunities to stay few steps ahead of the pathogen, and the technology can be used for helping to create novel pyramids of resistance alleles and transfer them to the cultivars (McDonald & Linde, 2002). The use of new breeding technologies is likely to underpin future gains in crop productivity and to generate new datasets for crop species (Langridge & Fleury, 2011).

Probably the more reasonable and effective strategy to deal with the rice-*P. oryzae* pathosystem is the use of modern molecular tools in order to access pathogen population structure in spatial and time scale allied to the use of the same techniques to identify potential sources of resistance genes in *O. sativa* and its related species.

5. Acknowlegements

6. References

Agrios, G.N. (2005). *Plant Pathology*. Elsevier Academic Presss, ISBN 13: 978-0-12-044565-3, San Diego, California.

Akagi, Y.; Akamatsu, H.; Otani, H. & Kodama, M. (2009). Horizontal chromosome transfer, a mechanism for the evolution and differentiation of a plant-pathogenic fungus. *Eukaryotic Cell*, Vol.8, No.11, (November 2009), pp. (1732–1738), ISSN 1535-9786.

Anjos, L.M.; Santos, G.R.; Dias Neto, J.J.; Oliveira, W.F. & Castro Neto, M.D. (2009). Identificação de raças fisiológicas de *Magnaporthe grisea* em áreas de arroz irrigado no Estado do Tocantins. *Tropical Plant Pathology*, Vol.34, No.3, (May-June 2009), pp. (182-185), ISSN 1982-5676.

Ashikawa, I.; Hayashi, N.; Yamane, H.; Kanamori, H.; Wu, J.; Matsumoto, T.; Ono, K. & Yano, M. (2008). Two adjacent nucleotide-binding site-leucine-rich repeat class genes are required to confer Pikm-specific rice blast resistance. *Genetics*, Vol.180, No.4, (December 2008), pp. (2267-2276), ISSN 1943-2631

Atkins, J.G.; Robert, A.L.; Adair, C.R.; Goto, K.; Kozaka, T.; Yanagita, R.; Yamada, M. & Matsumoto, S. (1967). An International set off rice varieties for differentiating races of *Pyricularia oryzae*. *Phytopathology*, Vol.57, pp. (297-301), ISSN 0031-949X.

Ballini, E.; Morel, J-B.; Droc, G.; Price, A.; Courtois, B.; Notteghem, J-L & Tharreau, D. (2008). A genome-wide meta-analysis of rice blast resistance genes and quantitative trait loci provides new insights into partial and complete resistance. *Molecular Plant-Microbe Interactions*, Vol.21, No.7, (July 2008), pp. (859-868), ISSN 0894-0282.

Barksdale, T. & Asai, G.N. (1961). Diurnal spore release of *Piricularia oryzae* from rice leaves. *Phytopathology*, Vol.51, No.5, (May 1961), pp. (313-317), ISSN 0031-949X.

Bryan, G.T.; Wu, K.; Farrall, L.; Jia, Y.; Hershey, H.P.; McAdams, S.A.; Faulk, K.N.; Donaldson, G.K.; Tarchini, R. & Valent, B. (2000). A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene Pi-ta. *The Plant Cell,* Vol.12, No.11, (November, 2000),pp. (2033-2045), ISSN 1532-298X.

Chen, Q.H.; Wang, Y.C.; Li, A.N.; Zhang, Z.G. & Zheng, X.B. (2007). Molecular mapping of two cultivar-specific avirulence genes in the rice blast fungus *Magnaporthe grisea*. *Molecular Genetics and Genomics*, Vol. 277, No.2, (February 2007), pp. (139–148), ISSN 1617-4623.

Chen, X.; Shang, J.; Chen, D.; Lei, C.; Zou, Y.; Zhai, W.; Liu, G.; Xu, J.; Ling, Z.; Cao, G.; Ma, B.; Wang, Y.; Zhao, X.; Li, S. & Zhu, L. (2006). A B-lectin receptor kinase gene conferring rice blast resistance. *Plant Journal*, Vol.46, No.5, (December 2006), pp. (794-804), ISSN 1365-313X.

Chen, D.; Zeigler, R.S.; Leung, H. & Nelson, R.J. (1995). Population structure of *Pyricularia grisea* at two screening sites in the Philippines. *Phytopathology*, Vol.85, No.9, (September 1995), pp. (1011-1020), ISSN 0031-949X.

Chuma, I.; Isobe, C.; Hotta, Y.; Ibaragi, K.; Futamata, N.; Kusaba, M.; Yoshida, K.; Terauchi, R.; Fujita, Y.; Nakayashiki, H. & Valent, B. (2011). Multiple translocation of the *AVR-Pita* effector gene among chromosomes of the rice blast fungus *Magnaporthe oryzae* and related species. *PLoS Pathogens*, Vol.7, No.7, (July 2011), pp. (1-20), ISSN 1553-7374.

Chumley, F.G. & Valent, B. (1990). Genetic analysis of melanin-deficient, nonpathogenic mutants of *Magnaporthe grisea*. *Molecular Plant Microbe Interactions*, Vol.3, No.3, (May-June 1990), pp. (135-143), ISSN 0894-0282.

Collard, B. C. Y. & Mackill, D. J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophycal Transactions of the Royal Society B*, Vol.363, No.1491 (February, 2008), p. (557-572), ISSN 1471-2970.

Consolo, V.F.; Cordo, C.A. & Salerno, G.L. (2005). Mating-type distribution and fertility status in *Magnaporthe grisea* populations from Argentina. *Mycopathologia*, Vol.160, No.4, (November 2005), pp. (285-290), ISSN 1573-0832.

Couch, B.C. & Kohn, L.M. (2002). A multilocus gene genealogy concordant with host preference indicates segregation of a new species, *Magnaporthe oryzae* from *M. grisea*. *Mycologia* Vol.94, No.4 (July –August 2002), pp. (683-693), ISSN 1557-2536.

Crawford, M.S.; Chunley, F.G. & Weaver, C.G. (1986). Characterization of the heterokaryotic and vegetative diploid phases of *Magnaporthe grisea*. *Genetics*, Vol.114, No.4, (December 1986), pp. (1111-1129), ISSN 1943-2631.

Cuevas-Perez, F. E.; Guimaraes, E. P.; Berrio, L. E. & Gonzales, D. I. (1992). Genetic base of the irrigated rice in Latin America and the Caribbean. *Crop Science*, Vol.32, No.4, (July 1992), pp. (1054-1059), ISSN 1435-0653.

Dean, R.A.; Talbot, N.J.; Ebbole, D.J.; Farman, M.L.; Mitchell, T.K.; Orbach, M.J.; Thon, M.; Kulkarni, R.; Xu, J-R.; Pan, H.; Read, N.D.; Lee, Y-H.; Carbone, I.; Brown, D.; Oh, Y.Y.; Donofrio, N.; Jeong, J.S.; Soanes, D.M.; Djonovic, S.; Kolomiets, E.; Rehmeyer, C.; Li, W.; Harding, M.; Kim, S.; Lebrun, M-H.; Bohnert, H.; Coughlan, S.; Butler, J.; Calvo, S.; Ma, L-J.; Nicol, R.; Purcell, S.; Nusbaum, C.; Galagan, J.E. & Birren, B.W. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature*, Vol.434, No.7036, (April 2005), pp. (980-986), ISSN 0028-0836.

Dioh, W.; Tharreau, D.; Notteghem, .J.L.; Orbach, M.J. & Lebrun, M.-H. (2000). Mapping of avirulence genes in the rice blast fungus, *Magnaporthe grisea*, with RFLP and RAPD markers. *Molecular Plant Microbe Interactions*, Vol.13, No.2, (February 2000), pp. (217–227), ISSN 0894-0282.

Farman, M.L.; Eto, Y.; Nakao, T.; Tosa, Y.; Nakayashiki, H.; Mayama, S. & Leong, S.A. (2002). Analysis of the structure of the AVR1-CO39 avirulence locus in virulent rice-infecting isolates of *Magnaporthe grisea*. *Molecular Plant Microbe Interactions,* Vol.15, No.1, (January 2002), pp. (6-16), ISSN 0894-0282.

Fuentes, J.L.; Escobar, F.; Alvarez, A.; Gallego, G.; Duque, M. C.; Ferrer, M.; Deus, J. E. & Tohme, J. (1999). Analyses of genetic diversity in Cuban rice varieties using isozyme, RAPD and AFLP markers. *Euphytica*, Vol.109, No.2, (September 1999), pp. (107-115), ISSN1573-5060.

Fujita, Y. & Suzuki, H. (1982). Aggressiveness of race 047 of *Pyricularia oryzae* with years after initial occurrence. *Annual Review of Phytopathology*, Vol.48, pp. (290-294), ISSN 0066-4286

Fukuoka, S.; Saka, N.; Koga, H.; Ono, K.; Shimizu, T.; Ebana, K.; Hayashi, N.; Takahashi, A.; Hirochika, H.; Okuno, K. & Yano, M. (2009). Loss of function of a proline-containing protein confers durable disease resistance in rice. *Science*, Vol.325, No.3, (August 2009), pp. (998-1001), ISSN 1091-6490.

George, M.L.C.; Nelson, R.J.; Zeigler, R.S. & Leung,H. (1998). Rapid population analysis of *Magnaporthe grisea* by using rep-PCR and endogenous repetitive DNA sequences. *Phytopathology*, Vol.88, No.3, (March 1998), pp. (223-229), ISSN 0031-949X.

Goff, S.A.; Ricke, D. Lan, T.; Presting, G.; Wang, R.; Dunn, M.; Glazebrook, J.; Sessions, A.; Oeller, P.; Varma, H.; Hadley, D.; Hutchison, D.; Martin, D.; Katagiri, F.; Lange, B.M.; Moughamer, T.; Xia, Y.; Budworth, P.; Zhong, J.; Miguel, T.; Paszkowski, U.; Zhang, S.; Colbert, M.; Sun, W.-L.; Chen, L.; Cooper, B.; Park, S.; Wood, T.C.; Mao, L.; Quail, P.; Wing, R.; Dean, R.; Yu, Y.; Zharkikh, A.; Shen, R.; Sahasrabudhe, S.; Thomas, A.; Cannings, R.; Gutin, A.; Pruss, D.; Reid, J.; Tavtigian, S.; Mitchell, J.; Eldredge, G.; Scholl, T.; Miller, R. M.; Bhatnagar, S.; Adey, N.; Rubano, T.; Tusneem, N.; Robinson, R.; Feldhaus, J.; Macalma, T.; Oliphant, A. & Briggs, S. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science,* Vol.296, No.5, (April 2002), pp. (92-100), ISSN 1091-6490.

Guimaraes, E.P.; Borrero, J. & Ospina-Rey, Y. (1996). Genetic diversity of upland rice germplasm distributed in Latin America. *Pesquisa Agropecuaria Brasileira*, Vol.31, No.3, (Março 1996) pp. (187-194), ISSN 0100-204X.

Hamer, J.E.; Farral, L.; Orbach, M.J.; Valent, B. & Schumley, F.G. (1989). Host species-specific conservation of family of repeat DNA sequences in the genome of a fungal plant pathogen. *Proceedings of the National Academy of Sciences of the United States of America*, Vol.86, No.15, (December 1989), pp. (9981-9985), ISSN 1091-6490.

Hamer, J.E.; Howard, R.J.; Chumley, F.G. & Valent, B. (1988). A mechanism for surface attachment in spores of a plant pathogenic fungus. *Science*, Vol.239, No 15., (January 1988), pp. (288-290), ISSN 1095-9203.

Hayashi, K. & Yoshida, H. (2009). Refunctionalization of the ancient rice blast disease resistance gene Pit by the recruitment of a retrotransposon as a promoter. *Plant Journal*, Vol.57, No.3, (February 2009), pp. (413-425), ISSN 1365-313X.

Hebert, T.T. (1971). The perfect stage of *Pyricularia grisea*. *Phytopathology*, Vol.61, No.1, (January 1971), pp. (83-87), ISSN 0031-949X.

Heil, M.; Hilpert, A.; Kaiser, W. & Linsenmair, K.E. (2000). Reduced growth and seed set following chemical induction of pathogen defence: does systemic acquired resistance (SAR) incur allocation cost?. *Journal of Ecology*, Vol.88, No.4, (August 2000), pp. (645-654), ISSN 1365-2745.

Howard, R.J. & Ferrari, M.A. (1989). Role of melanin in apressorium function. *Experimental Mycology*, Vol.13, No.4, (December 1989), pp. (403-418), ISSN 0147-5975.

Jena, K.K. The species of the genus *Oryza* and transfer of useful genes from wild species into cultivated rice, *O. sativa*. *Breeding Science*, Vol.60, No.5 (December, 2010), pp. (518-523), ISSN 1347-3735.

Jena, K.K. & Mackill, D.J. (2008). Molecular markers and their use in marker-assisted selection in rice. *Crop Science*, Vol.48, No.4, (July-August 2008), pp. (1266-1276), ISSN 1435-0653.

Jia, Y.; McAdams, S.A.; Bryan, G.T.; Hershey, H.P. & Valent, B. (2000). Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO Journal*, Vol.19, No.15, (August 2000), pp. (4004-4014), ISSN 1460-2075.

Johnson, L.J.; Johnson, R.D.; Akamatsu, H.; Salamiah, A.; Otani, H.; Kohmoto, K. & Kodama, M. (2001) Spontaneous loss of a conditionally dispensable chromosome from the *Alternaria alternata* apple pathotype leads to loss of toxin production and pathogenicity. *Current Genetics,* Vol.40, No.1, (August 2001), pp. (65–72), ISSN 1432-0983.

Kachroo, P.; Leong, S.A. & Chattoo, B.B. (1994). *Pot*-2, na inverted repeat transposon from the rice blast fungus *Magnaporthe grisea. Molecular Genetics and Genomic*, Vol.245, No.3, pp. (339-348), ISSN 1617-4623.

Kang, S.; Lebrun, M.H.; Farrall, L. & Valent, B. (2001). Gain of virulence caused by insertion of a Pot3 transposon in a *Magnaporthe grisea* avirulence gene. *Molecular Plant-Microbe Interaction*, Vol.14, No.5, (May 2001), pp. (671-674), ISSN 0894-0282.

Kang, S.; Chumley, F.G.; Valent, B. (1994). Isolation of the mating-type genes of the phytopathogenic fungus *Magnaporthe grisea* using genomic subtraction. *Genetics*, Vol.138, No.2, (October 1994), pp. (289–296), ISSN 1943-2631.

Kankanala, P.; Czymmek, K. & Valent, B. (2007). Roles for rice membrane dynamics and plasmodesmata during biotrofic invasion by the blast fungus. *Plant Cell*, Vol.19, No.2, (February 2007), pp. (706-724), ISSN 1532-298X.

Khushi, G.S. (1997). Origin, dispersal, cultivation and variation of rice. *Plant Molecular Biology*, Vol.35, No.1/2, (September 1997), pp. (25-34), ISSN 1573-5028.

Kingsolver, C.H.; Barksdale, T.H. & Marchetti, M.A. (1984). Rice blast epidemiology. Pennsylvania State University: Pennsylvania, 33p. (Bulletin, 853).

Kumar, J.; Nelson, R.J. & Zeigler, R.S. (1999). Population structure and dynamics of *Magnaporthe grisea* in the Indian Himalayas. *Genetics,* Vol.152, No.3, (July 1999), pp. (971-984), ISSN 1943-2631.

Langridge, P. & Fleury, D. (2011). Making the most of ´omics´ for plant breeding. Trends in Biotechnology, Vol.29, No.1, (January 2011) pp.33-40, ISSN 0167-7799.

Leach, J.E.; Cruz, C.M.V.; Bai, J. & Leung, H. (2001). Pathogen fitness penalty as a predictor of durability of disease resistance genes. *Annual Review of Phytopathology*, Vol.39, pp. (187-224), ISSN 0066-4286.

Lee, S.; Song, M.; Seo, Y.; Kim, H.; Ko, S.; Cao, P.; Suh, J.; Yi, G.; Roh, J.; Lee, S.; An, G.; Hahn, T.; Wang, G.; Ronald, P. & Jeon, J. (2009). Rice Pi5-mediated resistance to *Magnaporthe oryzae* requires the presence of two coiled-coil-nucleotide-binding-leucine-rich repeat genes. *Genetics*, Vol.181, No.4, (April 2009), pp. (1627-1638), ISSN 1943-2631.

Lee, Y.H. & Dean, R.A. (1993). Cyclic AMP regulates infection structure formation by the plant pathogenic fungus *Magnaporthe grisea. Plant Cell,* Vol.5, No.6, (June 1993), pp. (693-700), ISSN 1532-298X.

Levy, M.; Correa-Victoria, F.J.; Zeigler, R.S. Xu, S. & Hamer, J.E. (1993). Genetic diversity of the rice blast fungus in a disease nursery in Colômbia. *Phytophalology*, Vol.83, No.12, (December 1993), pp. (1427-1433), ISSN 0031-949X.

Levy, M.; Romao, J.; Marchetti, M.A. & Hamer, J.E. (1991). DNA fingerprint with a dispersed repeated sequence resolves pathotype diversity in the rice blast fungus. *Plant Cell,* Vol.3, No.1, (January 1991), pp. (95-102), ISSN 1532-298X.

Lin, F.; Chen, S.; Que, Z.; Wang, L.; Liu, X. & Pan, Q. (2007). The blast resistance gene Pi37 encodes a nucleotide binding site-leucine-rich repeat protein and is a member of a resistance gene cluster on rice chromosome 1. *Genetics*, Vol.177, No.3, (November 2007), pp. (1871-1880), ISSN 1943-2631.

Liu, J.; Wang, X.; Mitchell, T.; Hu, Y.; Liu, X.; Dai, L. & Wang, G.L. (2010). Recent progress and understanding of the molecular mechanisms of the rice-*Magnaporthe oryzae* interaction. *Molecular Plant Pathology*, Vol.11, No.3, (May 2010), pp. (419-427), ISSN 1364-3703.

Luo, C.X.; Hanamura, H.; Sezaki, H.; Kusaba, M. & Yaegashi, H. (2002). Relationship between avirulence genes of the same family in rice blast fungus *Magnaporthe grisea*. *Journal of General Plant Pathology,* Vol.68, No.4, (December 2002), pp. (300–306), ISSN 1610-739X.

Ma, J.H.; Wang, L.; Feng, S.J.; Lin, F.; Xiao, Y. & Pan, Q.H. (2006). Identification and fine mapping of *AvrPi15*, a novel avirulence gene of *Magnaporthe grisea*. *Theoretical and Applied Genetics*, Vol.113, No.5, (September 2006), pp. (875-883), ISSN 1432-2242.

Mackill, D.J. & Bonman, J.M. (1992). Inheritance of blast resistance in near-isogenic lines of rice. *Phytopathology*, Vol.82, No.7, (July 1992). pp. (746-749), ISSN 0031-949X.

Mackill, A.O. & Bonman, J.M. (1986). New Hosts of *Pyricularia oryzae*. *Plant Disease*, Vol.70, No.2, (February 1986), pp. (125-127), ISSN 0191-2917.

Marcel, S.; Sawers, R.; Oakeley, E.; Angliker, H. & Paszkowski, U. (2010). Tissue-adapted invasion strategies of the rice blast fungus *Magnaporthe oryzae*. *The Plant Cell*, Vol.22, No.9, (September 2010), pp. (3177-3187), ISSN 1532-298X.

McDonald, B.A. & Linde, C. (2002). Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytophololy*, Vol.40 (September, 2002), pp. (349-380), ISSN 0066-4286.

Metha, Y.R. & Baier, A. (1998). Variação patogênica entre isolados de *Magnaporthe grisea* atacando triticale e trigo no Estado do Paraná. *Summa Phytopathologica*, Vol.24, No.2, (Abril-Junho 1998), pp. (119-125), ISSN 0100-5405.

Mundt, C.C. (2002). Use of multiline cultivars and cultivar mixtures for disease management. *Annual Review of Phytophololy,* Vol.40 (September,2002), pp. (381-410), ISSN 0066-4286.

Nishimura, M.; Hayashi, N.; Jwa, N.-S.; Lau, G.W.; Hamer, J.E. & Hasebe, A. (2000). Insertion of the LINE retrotransposon MGL causes a conidiophore pattern mutation in *Magnaporthe grisea*. *Molecular Plant-Microbe Interactions,* Vol.13, No.8, (August 2000), pp. (892-894), ISSN 0894-0282.

Noguchi, M.T. (2011). Parasexual recombination in *Magnaporthe oryzae*. Japan Agricultural Research Quarterly, Vol.45, No.1, (January 2011), pp. (39-45), ISSN 0021-3551.

Noguchi, M.T.; Yasuda, N. & Fujita, Y. (2007). Fitness characters in parasexual recombinants of the rice blast fungus, *Pyricularia oryzae*. Japan Agricultural Research Quarterly, Vol.41, No.2, (February 2007), pp. (123-131), ISSN 0021-3551.

Nonomura, K.I.; Morishima, H.; Miyabayashi, T.; Yamaki, S.; Eiguchi, M.; Kubo, T. & Kurata, N. (2010). The wild *Oryza* collection in National BioResearch Project (NBRP) of Japan: History, biodiversity and utility. *Breeding Science*, Vol.60, No.5, (December 2010), pp. (502-508), ISSN 1347-3735.

Notteghem, J.L. & Silue, D. (1992). Distribu-tion of the mating type alleles in *Magnaporthe grisea* populations pathogenic on rice. *Phytopathology*, Vol.82, No.4, (April 1992), pp. (421–424), ISSN 0031-949X.

Ou, S.H. (1972). *Rice diseases* (1). Eastern Press, ISBN 0 85198 217 4, London.

Ou, S.H. (1985). *Rice disease (*2). Commonwealth Mycological Institute, ISBN 0 85198 545 9, Kew.

Orbach, M.; Farrall, L.; Sweigard, J.A.; Chumley, F.G. & Valent, B. (2000). A telomeric avirulence gene determines efficacy for the rice blast resistance gene *Pi-ta*. *The Plant Cell*, Vol.12, No.11, (November 2000), pp. (2019-2032), ISSN 1532-298X.

Park, S.Y.; Milgroom, M.G.; Han, S.S.; Kang, S. & Lee, Y.H. (2008). Genetic differentiation of *Magnaporthe oryzae* populations from scouting plots and commercial rice fields in Korea. *Phytopathology*, Vol.98, No.4, (April 2008), pp. (436-442), ISSN 0031-949X.

Prabhu, A.S.; Filippi, M.C. & Zimmermann, J.P. (1996). Genetic control of blast in relation to nitrogen fertilization in upland rice. *Pesquisa Agropecuária Brasileira*, Vol.31, No.5, (May 1996), pp. (339-347), ISSN 0100-204X.

Pretty, J. (2008). Agricultural sustainability: concepts, principles and evidence. *Philosophical Transactions of the Royal Society B,* Vol.363, No.1491 (February, 12 2008), pp. (447-465), ISSN 1471-2970.

Qu, S.; Liu, G.; Zhou, B.; Bellizz, L.M.; Zeng, L.; Dai, L.; Han, B. & Wang, G.L. (2006). The broad-spectrum blast resistance gene Pi9 encodes a nucleotide-binding site-leucine-rich repeat protein and is a member of a multigene family in rice. *Genetics*, Vol.172, No.3, (March 2006), pp. (1901-1914), ISSN 1943-2631.

Rai, A.K.; Kumar, S.; Gautam, N.; Gupta, S.K.; Chand, D.; Sinh, N.K. & Sharma, T.R. (2009). Cloned rice blast resistance gene Pi-kh confers broad spectrum resistance to *Magnaporthe oryzae*. Proceedings of XIV Congress on Molecular Plant-Microbe Interactions, International Society of Molecular Plant-Microbe Interactions, Quebec, Canada, 19-23 July, 2003.

Rangel, P.H.N.; Brondani, C.; Ferreira, M.E.; Rangel, P.N. & Brondani, R.P.V. (2006). Utilização de espécies silvestres *Oryza glumaepatula* no pré-melhoramento de arroz. Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia, (Agosto 2006) pp. (94-98), ISSN 0102-0110 *(Documentos 185)*.

Rangel, P.H.N.; Guimaraes, E.P. & Neves P.C.F. (1996). The genetic base of Brazilian irrigated rice (*Oryza sativa* L) cultivars. *Pesquisa Agropecuaria Brasileira*, Vol. 31, No. 5, (Maio 1996), pp. (349-357), ISSN 0100-204X.

Rehmeyer, C.; Li, W.; Kusaba, M.; Kim, Y.S.; Brown, D.; Staben, C.; Dean, R. & Farman, M. (2006). Organization of chromosome ends in the rice blast fungus, *Magnaporthe oryzae*. *Nucleic Acids Research*, Vol.34, No.17, (October 2006), pp. (4685-4701), ISSN 1362-4962.

Ribeiro, A.S. & Terres, A.L.S. (1987). Variabilidade do fungo *Pyricularia oryzae* e sua relação com cultivares resistentes à brusone. *Fitopatologia Brasileira*, Vol.12, No.4, (December 1987), pp. (316-321), ISSN 0100- 4158.

Rossman A.Y.; Howard, R.J. & Valent, B. (1990). *Pyricularia grisea*, the correct name for the rice blast fungus. *Mycologia,* Vol.82, No.4 (July-August 1990), pp. (509–512), ISSN 1557-2536.

Roumen, E.; Levy, M. & Nottegham, J.L. (1997). Characterization of the European pathogen population of *Magnaporthe grisea* by DNA fingerprinting and pathotype analysis. *European Journal of Plant Pathology*, Vol.103, No.4 (May 1997), pp. (363-371), ISSN 1573-8469.

Saghai Maroof, M.A.; Jeong, S.C.; Gunduz, I.; Tucker, M.D.; Buss, G.R. & Tolin, S.A. (2008). Pyramiding of soybean mosaic virus resistance genes by marker-assisted selection. *Crop Science*, Vol.48, No.2, (March-April 2008), pp. (517-526), ISSN 1435-0653.

Servin, B.; Martin, O.C; Mézard, M. & Hospital, F. (2004). Toward a theory of marker-assisted gene pyramiding. *Genetics*, Vol.168, No.1, (September 2004), pp. (513-523), ISSN 1943-2631.

Sesma, A. & Osbourn, A.E. (2004). The rice leaf blast pathogen undergoes developmental processes typical of root-infecting fungi. *Nature*, Vol.431, No.7008, (September 2004), pp. (582-586), ISSN 0028-0836.

Shang, J.; Tao, Y.; Chen, X.; Zou, Y.; Lei, C.; Wang, J.; Li, X.; Zhao, X.; Zhang, M.; Lu, Z.; Xu, J.; Cheng, Z.; Wan, J. & Zhu, L. (2009). Identification of a new rice blast resistance gene, Pid3, by genome-wide comparison of paired NBS-LRR genes and their pseudogene alleles between the two sequenced rice genomes. *Genetics*, Vol.182, No.4, (August 2009), pp. (1303-1311), ISSN 1943-2631.

Takan, J.P.; Chipili, J.; Muthumeenakshi, S.; Talbot, N.J.; Manyasa, E.O.; Bandyopadhyay, R.; Sere, Y.; Nutsugah, S.K.; Talhinhas, P.; Hossain, M.; Brown, A.E. & Sreenivasaprasad, S. (2011). *Magnaporthe oryzae* populations adapted to finger millet and rice exhibit distinctive patterns of genetic diversity, sexuality and host interaction. *Molecular Biotechnology*, DOI 10.1007/s12033-011-9429-z (Online First), Available from http://cogeme.ex.ac.uk/talbot/pdf/2011_Takan_molbio_moryzae_population.pd, ISSN 1559-0305.

Thanh, P.T.; Phan, P.D.T.; Mori, N.; Ishikawa, R. & Ishii, T. (2011). Development of backcross recombinant inbred lines between *Oryza sativa* Nipponbare and *O. rufipogon* and QTL detection on drought tolerance. *Breeding Science*, Vol.61, No.1, (February 2011), pp. (76-79), ISSN 1347-3735.

Tian, D.; Traw, M.B.; Chen, J.Q.; Kreitman, M. & Bergelson, J. (2003). Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature*, Vol., No.1, (May 2003), pp. (74-77), ISSN 0028-0836.

Valent, B. & Chumley, F.G. (1991). Molecular genetic analysis of the rice blast fungus, *Magnaporthe grisea*. *Annual Review of Phytopathology*, Vol.29, (September 1991), pp. (443-467), ISSN 0066-4286.

Vaughan, D.A.; Kadowaki, K.; Kaga, A. & Tomooka, N. (2005). On the phylogeny an biogeography of the Genus *Oryza*. *Breeding Science*, Vol.55, No.2, (June 2005), pp. (113-122), ISSN 1347-3735.

Vergne, E.; Grand, X.; Ballini, E.; Chalvon, V.; Saindrenan, P.; Tharreau, D.; Nottéghem, J.L. & Morel, J.B. (2010). Preformed expression of defense is a hallmark of partial resistance to rice blast fungal pathogen *Magnaporthe oryzae*. *BMC Plant Biology*, Vol.10 (September 2010), pp. (1-17), ISSN 1471-2229.

Yamasaki, Y. & Niizeki, H. (1965). Studies on variation of the rice blast fungus, *Pyricularia oryzae* Cav. I. Karyological and genetic studies on variation. *Bulletin of the National Institute of Agricultural Sciences (Japan)*, Vol.13, pp. (231–273).

Yasuda, N.; Tsujimoto-Noguchi, M. & Fujita, Y. (2006). Partial mapping of avirulence genes *AVR-Pii* and *AVR-Pia* in the rice blast fungus *Magnaporthe oryzae*. *Canadian Journal of Plant Pathology*, Vol.28, No.4, (October 2006), pp. (494–498), ISSN 1715-2992.

Yu, J.; Hu, S.; Wang, J.; Wong, G.K-S.; Li, S.; Liu, B.; Deng, Y.; Dai, L.; Zhou, Y.; Zhang, X.; Cao, M.; Liu, J.; Sun, J.; Tang, J.; Chen, Y.; Huang, X.; Lin, W.; Ye, C.; Tong, W.; Cong, L.; Geng, J.; Han, Y.; Li, L.; Li, W.; Hu, G.; Huang, X.; Li, W.; Li, J.; Liu, Z.; Li, L.; Liu, J.; Qi, Q.; Liu, J.; Li, L.; Li, T.; Wang, X.; Lu, H.; Wu, T.; Zhu, M.; Ni, P.; Han, .H.; Dong, W.; Ren, X.; Feng, X.; Cui, P.; Li, x.; Wang, H.; Xu, X.; Zhai, W.; Xu, Z.; Zhang, J., He, S.; Zhang, J.; Xu, J.; Zhang, K.; Zheng, X.; Dong, J.; Zeng, W.; Tao, L.; Ye, J.; Tan, J.; Ren, X.; Chen, X.; He, J.; Liu, D.; Tian, W.; Tian, C.; Xia,. H.; Bao, Q.; Li, G.; Gao, H.; Cao, T.; Wang, J.; Zhao, W.; Li, P.; Chen, W.; Wang, X.; Zhang, Y.; Hu, J.; Wang, J.; Liu, S.; Yang, J.; Zhang, G.; Xiong Y.; Li, Z.; Mao, L.; Zhou, C.; Zhu, Z.; Chen, R.; Hao, B.; Zheng, W.; Chen, S.; Guo, W.; Li, G.; Liu, S.; Tao, M.; Wang, J.; Zhu, L.; Yuan, L. & Yang, H. . A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). (2002). *Science*, Vol.296, No.5, (April 2002), pp. (79-91), ISSN 1091-6490.

Yuan, B.; Zhai, C.; Wang, W.; Zeng, X.; Xu, X.; Hu, H.; Lin, F.; Wang, L. & Pan, Q. The P*ik-p* resistance to *Magnaporthe oryzae* in rice is mediated by a pair of closely linked CC-NBS-LRR genes. *Theoretical and Applied Genetics*, Vol.122, No.5, (March 2011), pp. (1017-1028), ISSN 1432-2242.

Xia, J.Q.; Correll, J.C.; Lee, F.N. & Ross, W.J. (2000). Regional population diversity of *Pyricularia grisea* in Arkansas and the influence of host selection. *Plant disease*, Vol.84, No.8, (August 2000), pp. (877-884), ISSN 0191-2917.

Xia, J.Q.; Correll, J.C.; Lee, F.N. Marchetti, M.A. & Rhoads, D.D. (1993). DNA Fingerprinting to examine microgeographic variation in the *Magnaporthe grisea* (*Pyricularia grisea*) population in two rice fields in Arkansas. *Phytopathology*, Vol.83, No.10, (October 1993), pp. (1029-1035), ISSN 0031-949X.

Wang, Z.; Yano, M.; Yamanouchi, U.; Iwamoto, M.; Monna, L.; Hayasaka, H.; Katayose, Y. & Sasaki, T. (1999). The Pib gene for rice blast resistance belongs to the nucleotides binding and leucine-rich repeat class of plant disease resistance genes. *Plant Journal*, Vol.19, No.1, (July 1999), pp. (55-64), ISSN 1365-313X.

Wilson, J.P.; Gates, R.N. & Panwar, M.S. (2001). Dynamic multiline population approach to resistance gene management. *Phytopathology*, Vol.91, No.3, (March 2001), pp. (255-260), ISSN 0031-949X.

Zeigler, R.S. (1998). Recombination in *Magnaporthe grisea*. *Annual Review of Phytopathology*, Vol.36, pp. (249-276), ISSN 0066-4286.

Zeigler, R.S.; Scott, R.P.; Leung, H.; Bordeos, A.A.; Kumar, J. & Nelson, R.J. (1997). Evidence of parasexual exchange of DNA in the rice blast fungus challenges its exclusive clonality. *Phytopathology*, Vol.87, No.3, (March 1997), pp. (284-294), ISSN 0031-949X.

Zenbayashi-Sawata, K.; Ashizawa, T. & Koizumi, S. (2005). *Pi34-AVRPi34*:a new gene-for-gene interaction for partial resistance in rice to blast caused by *Magnaporthe grisea*.

Journal of General Plant Pathology, Vol.71, No.6, (December 2005), pp. (395-401), ISSN 1610-739X.

Zhou, B.; Qu, S.; Liu, G.; Dolan, M.; Sakai, H.; Lu, G.;Bellizi, M. & Wang, G. L. (2006). The eight amino-acid differences within threeleucine-rich repeats between Pi2 and Piz-t resistance protein determine the resistance specificity to Magnaporthe grisea. *Molecular Plant-Microbe Interactions*, Vol.19, No.11 (November 2006), pp. (1216-1228), ISSN 0894-0282.

Zhu, Y.; Chen, H.; Fan, J.; Wang, Y.; Li, Y.; Chen, J.; Fan, J.; Yang, S.; Hu, L.; Leung, H.; Mew, T.W.; Teng, P.S.; Wang, Z. & Mundt, C.C. (2000). Genetic diversity and disease control in rice. *Nature*, Vol.406, No.17, (August 2000), pp. (718-722), ISSN 0028-0836.

The Nutrigenome and Gut Microbiome: Chronic Disease Prevention with Crop Phytochemical Diversity

Erica Daniell and Elizabeth P. Ryan
Colorado State University
Ft. Collins, Colorado
USA

1. Introduction

Lifestyle factors including, but not limited to, dietary intake, have led to the increasing global prevalence and incidence of major chronic diseases (e.g. obesity, diabetes, cardiovascular disease and cancer) in children and adults. Numerous reports have documented the relationships between dietary patterns, weight gain (Mozaffarian, Hao, Rimm, Willett, & Hu, 2011; Romaguera et al., 2011) and chronic disease prevalence (Adebamowo et al., 2005; de Munter, Hu, Spiegelman, Franz, & van Dam, 2007; L. Hooper et al., 2008; Hung et al., 2004; Song, Manson, Buring, Sesso, & Liu, 2005). Few studies have evaluated the potential contribution of decreased food crop genetic diversity in these metabolic and inflammatory disorders (Jew, AbuMweis, & Jones, 2009a; M. D. T. Thompson & Thompson, 2009). Integrating our understanding of nutrigenetics and the gut microbiome with genetic and phytochemical diversity of food crops represents a novel systems level approach for determining the important contributions of plant genetic diversity to human health.

The composition of one's diet is a strong environmental pressure that can influence the gut microenvironment, as well as the nutrigenomic evolution of the human species. The human genome has adapted to changes in the available food supply through a complex set of interactions. Humans have also strongly influenced the evolution of plant genomes, particularly those plant foods domesticated as staple crops (e.g. rice, wheat, beans, potatoes, corn). While staple crop genomes have been widely studied for agronomic traits, such as yield and disease resistance, little is known regarding the impact of genetic selection on staple food crop traits that are of human health importance. Changes in the "genome" of both plants and animals are complex and multi-layered fields of study. Beyond the foundational layer of genetic sequence, plants and animals have modifications to DNA that create a pattern of inheritance known as epigenetics. Moreover, humans have another layer of extensive genetic diversity bestowed by the microbial kingdom that co-habituates within the human body that may be referred to as metagenomics. Each of these layers interacts to influence host gene expression and have evolved over time to affect host metabolism.

2. Definition of terms for understanding human nutrigenetic co-evolution with crops

Definitions for common terms relevant to the synthesis of information on genetic diversity of plant foods for human chronic disease prevention are provided to advance our ability to integrate these systems level concepts.

- **Genotype:** Genetic makeup consisting of nucleic acid sequence and containing a combination of DNA mutations. Genotype is unique to every individual, is determined at conception, cannot be changed throughout one's lifetime, and will be the foundation for determining genetic predisposition for disease (DeBusk, Fogarty, Ordovas, & Kornman, 2005).
- **Nutrigenetics:** The field of study that focuses on the effect of one's genotype on dietary needs as it relates to risk for developing disease (DeBusk et al., 2005; Simopoulos, 2010).
- **Nutrigenomics:** The field of study that focuses on the effect of dietary components on gene expression as it relates to disease processes (DeBusk et al., 2005).
- **Epigenetics:** The pattern of DNA modifications that affect gene expression, but do not involve changing nucleotide sequence (DNA methylation and histone acetylation are examples of epigenetic modifications). Epigenotype is an inherited pattern which varies from cell to cell, is determined at conception, and may also be highly influenced by environmental conditions during development and throughout one's lifetime. Unlike genotype, it is not a static pattern, but a mutable one (Feil, 2006; Fraga et al., 2005).
- **Microbiomics:** The study of all of microorganisms within a single ecosystem. For nutritional contexts, this refers to all microorganisms co-existing within the human gastrointestinal (GI) tract (Dimitrov, 2011).
- **Metagenomics:** The study of a collection of genetic material (genomes) from a mixed community of organisms.For nutritional contexts, this refers to the microbial communities residing within the human GI tract (Hattori, 2009).
- **Phylotype:** The composition of microflora that make up a particular microbiome. It is established after birth and changes in response to environmental influences and disease (Let et al., 2008; Turnbaugh et al., 2009).
- **Phenotype:** The measurable, observable, or experienced expression of a trait. It is the manifestation of genotype as influenced by epigenotype, phylotype, and environment (DeBusk et al., 2005).
- **Cultivar:** A variety of a plant or crop that has been deliberately selected for a specific trait or characteristic (e.g. yield or disease resistance) (Zeven, 1998).
- **Phytochemical:** Any of the chemical substances produced by plants.
- **Primary Metabolites:** Products of plant biosynthesis that are considered essential building blocks for growth and development, e.g. macromolecules such as carbohydrates, proteins, and fats.
- **Secondary Metabolites:** Small molecules synthesized by plants for signaling or defense purposes. There are over 200,000 different secondary metabolite compounds that have been characterized (M.D.T. Thompson & Thompson, 2009).
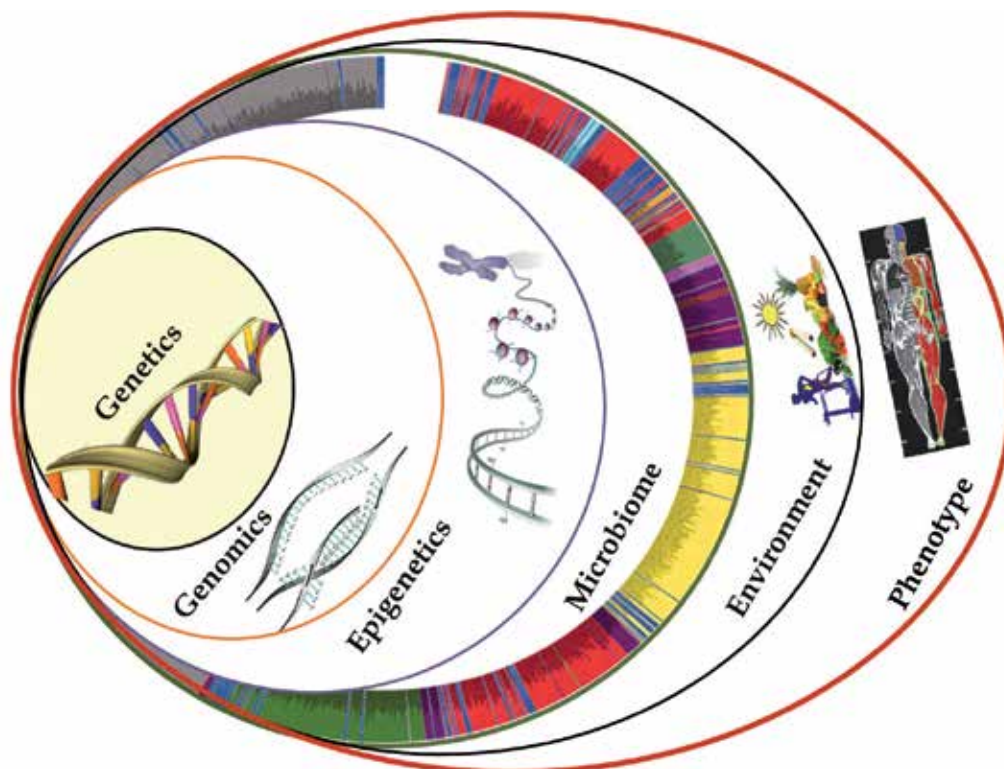
Fig. 1. The Layers of the Nutrigenome: Genetic make-up (genotype), genomic expression, epigenetic modifications (epigenotype), the microbiome (phylotype) and environmental factors contribute to the overall health phenotype of humans. Genotype serves as the foundation, while epigenotype, and phylotype are modifiers of gene expression in response to dietary components and other environmental factors.

3. Gene-nutrient interactions

After a decade into the 21st century, there is concern that today's children may be the first generation in America to have a shortened life expectancy compared to their parents (Olshansky et al., 2005). One major reason for this concern is the growing obesity epidemic and the myriad of health issues associated with obesity including, diabetes, heart disease, non-alcoholic liver disease, and cancer (Bray, 2004). A theory to address this global surge in metabolic related chronic diseases is related to the deviation of human populations from the food supply that their ancestors have co-evolved with (Cordain et al., 2005; Jew, AbuMweis, & Jones, 2009b).

The nutritional and biomedical sciences have developed areas of investigation that delve into possible causal events for metabolic and inflammatory related chronic disorders, while simultaneously searching for effective preventive measures (Garcia-Canas, Simo, Leon, & Cifuentes, 2010; Go, Nguyen, Harris, & Lee, 2005; Goodacre, 2007). These emerging fields of study include Nutrigenetics, Nutrigenomics, Nutri-Epigenetics, Nutri-Proteomics, Nutri-Metabolomics, and Microbiomics (**Figure 1**). Nutrigenetics, Nutrigenomics, and Nutri-

Epigenomics seek to characterize genes that participate in diet-disease relationships, while Proteomics and Metabolomics involve measuring the protein or metabolite end products of gene expression in response to dietary influences, and can be considered critical measures of function or phenotype (Cobiac, 2007; Go et al., 2005). Microbiome studies represent the youngest of the aforementioned high throughput technologies, and reveals how the phylogenetic make up of our GI tract microbiota may influence overall metabolic status and diet-disease relationships (Hattori, 2009), and emerging models support the importance of gut microbial community modifications by diet (Flint, Duncan, Scott, & Louis, 2007; Kau, Ahern, Griffin, Goodman, & Gordon, 2011; Ley et al., 2005).

3.1 Evolutionary nutrigenetics

In parallel with major shifts in human diet composition and dietary patterns, such as the incorporation of animal products like meat or dairy, and the increased consumption of a single staple food during the green revolution (i.e. wheat, rice), the human body has had to adapt metabolically for optimal absorption of essential nutrients (Jew et al., 2009b; Luca, Perry, & Di Rienzo, 2010). These adaptations have been targeted by natural selection, resulting in genetic mutations that may vary between individuals and are an essential part of a person's unique nutrigenetic code. Specific examples of gene mutations that have arisen in concordance with changes in the available food supply have become increasingly apparent as the field of nutrigenetics unfolds. Though there are countless known genetic mutations relevant to nutrition and chronic disease risk, we will highlight three significant examples. Additional examples of gene mutations that have arisen in response to changes in diet composition are listed in Table 1.

3.1.1 Lactase persistence

A token example of human genetic co-evolution with the food supply is lactase persistence in concordance with incorporation of dairy within the diet of some human populations. Lactose intolerance is the "default" phenotype yielded by the ancestral or wild type version of the human LCT (lactase-phlorizin hydrolase) gene and results from diminished expression of the LCT gene post-weaning (Hollox, 2005). Individuals with mutant forms of the LCT gene have certain DNA mutations, called single nucleotide polymorphisms, or SNPs, within the regulatory region of the gene. These mutations allow for continued expression of the LCT gene post-weaning, resulting in the ability to digest and gain energy from lactase in dairy products. These mutations were positively selected for during human evolutionary history in populations with increased incorporation of dairy in the diet. Several different SNPs in the LCT gene have independently evolved in distinct populations of northern Europe and eastern Africa (Jarvela, Torniainen, & Kolho, 2009). This mutation is a classic example of evolutionary nutrigenetics, and exemplifies the genetic co-evolution of populations with dietary intake and availability.

3.1.2 Alcohol dehydrogenase

The human ADH1B gene, which encodes for a subunit of the alcohol dehydrogenase enzyme responsible for catalyzing the breakdown of alcohol into acetaldehyde, is another

example of genetic selection that occurred in response to changes in the food supply. The ancestral genotype for the ADH1B gene encodes for the amino acid arginine at amino acid position 47 while the mutant version encodes for histidine at this location. This mutation has been correlated with increased alcohol metabolism, and decreased propensity for developing alcoholism (Chen, Peng, Wang, Tsao, & Yin, 2009). Interestingly, the ADH1B*47His SNP is most prevalent in the southeast populations of China with the geographic distribution overlapping the areas of origin and expansion of rice domestication (Li et al., 2007; Peng et al., 2010). It is thought that as rice became a Neolithic staple for use in fermented food and beverages, the *47His allele was beneficial for preventing some of the deleterious effects of alcohol consumption (Peng et al., 2010).

3.1.3 Salivary amylase

The human salivary amylase gene (AMY1) provides an example of a gene where the variation observed between populations is not differences in genetic sequence, but differences in gene copy number. The human salivary amylase enzyme serves an important role in the digestion of dietary starch and oligosaccharides by breaking them down into maltose molecules which can then be further broken down by the enzyme maltase to produce glucose (Meisler & Ting, 1993). Geographic populations were found to have extensive variation in the copy numbers of AMY1 they possess, and it has been determined that the number of AMY1 copies in a population correlates to salivary amylase activity and efficiency of starch digestion (Perry et al., 2007). AMY1 copy number has been found to correlate with the starch intake of a population, with populations consuming a high starch diet having more copies of AMY1 than populations that are hunter-gatherer or pastoral in nature and depend on a low starch diet (Coyne & Hoekstra, 2007; Perry et al., 2007). As populations moved toward increased dependence on high starch containing diets (e.g. wheat, rice), increased efficiency in digesting starches became a beneficial trait.

4. Nutritional epigenetics and microbiomics

Beyond genetic traits selected naturally over many generations, it is also possible for changes in diet composition to affect gene expression and manifest as epigenetic modifications ("epigenotype") or changes in microbiome composition ("phylotype"). These epigenetic and microbiome modifications occur continuously as a real-time response to diet and other environmental factors. It is these layers of the human genome that we do have the power to change with dietary and lifestyle modifications.

Genetic diversity of plant food crops may be considered a significant dietary feature, and as such, may play a large role in chronic disease risk by influencing the epigenotype and phylotype of humans. It can be hypothesized that the decrease in the botanical and genetic diversity of plant foods in today's diet may influence the evolution of epigenotype and phylotype to affect disease risk within one's lifetime. To our knowledge, there has been limited investigation of the relationships between staple food crop biodiversity and their dietary effects on human epigenotype and phylotype in relation to risk for developing chronic diseases such as obesity, type II diabetes, heart disease and cancer.

Gene Abbreviation	Protein	Phenotype with selected mutation	Selection Advantage
LCT	Lactase-phlorizin hydrolase	Ability to digest lactose in milk	Allowed for consumption of dairy as staple food(Jarvela et al., 2009).
TAS2R38	Taste receptor, type 2, member 38	Ability to detect certain substances with bitter taste	Allowed for detection and aversion of certain plants that contained substances which could cause iodine deficiencies and thyroid disordes(Wooding et al., 2004).
AMY1	Salivary Amylase Gene	More copies of the AMY1 gene result in increased salivary amylase protein	Populations that consume high starch diets have more copies of the AMY1 gene, allowing for more efficient digestion of starches(Perry et al., 2007).
FOXI1	Forkhead-box transcription factor I1	Sensorineural deafness, distal renal tubular acidosis, male infertility	Possibly allowed for climate adaptation through water-electrolyte homeostasis and prevention of dehydration(Moreno-Estrada et al., 2010).
GIP	Gastric inhibitory peptide (GIP)	GIP is resistant to serum degradation, and exhibits a higher bioactivity	May have been beneficial for individuals who had unconstrained access to the food supply in agricultural societies by preventing hyperglycemia (Chang et al., 2011)
ADH1B	Alcohol Dehydrogenase 1B (class 1), beta polypeptide	Increased activity for ethanol oxidation	Ability to metabolize fermented rice products (Peng et al., 2010)

Table 1. Examples of diverse genetic mutations that may have co-evolved with changes in food supply.

4.1 Dietary influences on epigenetic evolution

Epigenetic modifications are biochemical alterations that are made to DNA or DNA packaging proteins that modify the availability of genes to transcription machinery, thereby affecting gene expression. Epigenetic modifications such as DNA methylation or histone acetylation, for example, are responsible for differential expression allowing for a liver cell to greatly differ in structure and function from a skin cell, even though they share the same genetic make-up (Morgan, Santos, Green, Dean, & Reik, 2005). The "epigenotype," determined at conception, is inherited from the parents, but it can be greatly influenced by gestational conditions (Cooney, Dave, & Wolff, 2002; Nicoletto & Rinaldi, 2011; Wolff, Kodell, Moore, & Cooney, 1998). This is best exemplified by the *Agouti* mouse, where the epigenetic status of the Agouti gene, namely presence or absence of methylation at nine particular cytosines within a regulatory region 5′ of the *Agouti* start site, results in a strikingly visible phenotype, as it causes variation in coat color and body weight as well as proneness to developing Type 2 Diabetes (Dolinoy, 2008). Experiments with these mice have demonstrated that maternal diet lacking in certain nutrients affects the epigenetic status of this gene and the related phenotypes (Cooney et al., 2002; Dolinoy, Huang, & Jirtle, 2007; Wolff et al., 1998).

In addition to the influence of gestational environment, epigenotype can also be influenced by environmental factors post-development and can be thought to evolve throughout one's lifetime. The changes that occur in epigenotype over time are referred to as "Epigenetic Drift" (Fraga et al., 2005; Nicoletto & Rinaldi, 2011). Studies conducted with identical twins that analyzed epigenetic patterns at different ages demonstrated that while twins share highly similar epigenetic patterns at young ages, this similarity decreases over time, and the disparity is likely due to environmental influences including diet composition (Fraga et al., 2005). Prospective intervention studies in twins that consume foods comprising of distinct dietary patterns, and including diverse staple foods may be useful to determine the relative importance of epigenetic drifts to the diet-disease risk relationship.

4.1.1 Effects of dietary folate on epigenotype

DNA methylation is one example of an epigenetic modification, where a methyl group is covalently added to the C5 position of cytosine residues that occur directly adjacent to guanine residues. Many genes contain regulatory regions that are rich in cytosine-guanine dinucleotide repeats ("CG islands") (Hirst & Marra, 2009), and in general, a high degree of methylation in these regulatory regions results in diminished gene expression- a silencing that is often critical to many cellular processes. Methylation is a heritable modification and is maintained with cell division; however, methyl groups can be lost due to enzymatic removal or failure to accurately copy methylation pattern during replication (Oommen, Griffin, Sarath, & Zempleni, 2005). Alterations in methylation patterns have been associated with various disease states including cancer, Type 2 diabetes, and Alzheimer's Disease (Coppieters & Dragunow, 2011; Hirst & Marra, 2009; Ling & Groop, 2009; Martin-Subero & Esteller, 2011). These changes in epigenotype occur as a result of aging and environmental influences, including dietary exposures, and more work needs to be done to fully demonstrate a clear cause and effect relationship between diet induced changes in epigenotype and risk for disease (Jaenisch & Bird, 2003).

Folate is a micronutrient that can directly influence DNA methylation status via its effects on the one carbon metabolism pathway, whereby folate serves as an important co-enzyme

in the production of S-adenosylmethionine, the DNA methyl donor (Kim, Friso, & Choi, 2009). Folate depletion and/or folate supplementation can affect the methylation status of certain genes. Studies beyond the scope of this chapter demonstrate the precise mechanisms by which folate status effects epigenotype and gene expression (Jacob et al., 1998; Rampersaud, Kauwell, Hutson, Cerda, & Bailey, 2000).

Legumes represent a major staple plant food that is naturally rich in folate (e.g. lentils and dry beans) (U.S.D.A, 2009). It is unknown whether the vast genetic diversity of the Leguminosae family and consumption rates of legumes can influence DNA methylation. Populations with staple legume consumption have shown lower chronic disease prevalence that may be due to a number of reported bioactivities. (Adebamowo et al., 2005; Bazzano et al., 2001; Michels et al., 2006; Papanikolaou, 2006; Singh & Fraser, 1998; Villegas et al., 2008). There is potential for legume folate, bioactive compounds, and essential nutrients to affect methylation status and influence epigenotype, yet the relationship to disease risk in humans is currently unknown.

4.1.2 Effects of butyrate on epigenotype

Modifications to the histone proteins, which interact and tightly bind DNA so it can eventually be condensed into chromatin, are another category of epigenetic modifications that make up epigenotype. The tails of the histone proteins are targets for modifications because they protrude from the histone/DNA complex known as the nucleosome. Specific residues in these histone tails may undergo methylation, phosphorylation or acetylation (Cobiac, 2007; Sawan & Herceg, 2010). Modifications to histones affect their DNA binding affinity and may result in more or less condensation of the DNA, and thus influences gene transcription at that location. Histone acetylation is generally associated with less condensation and increased gene transcription. Histone acetyltransferases and deacetyltransferases affect acetylation, and aberrant acetylatione has been associated with disease such as cancer, neurodegenerative disorders, and Type 2 Diabetes (Cobiac, 2007; Sawan & Herceg, 2010; Gray & DeMeyts, 2005; Mattson, 2003).

There have been some accounts of dietary factors affecting histone aetyltransferases. In particular, the short chain fatty acid, butyrate, was reported to inhibit histone acetytransferase (Timmermann, 2003). Short chain fatty acids are produced in the gastrointestinal tract from dietary fiber and resistant starch microflora fermentation. Legumes and whole grains represent food staples that are rich sources of dietary fiber that leads to production of short chain fatty acids, including butyrate. This is another example of how changes in staple food crops may interact with the microbiome and epigenome to affect risk of developing chronic diseases.

4.2 The gut microbiome and dietary evolution of phylotype

The community of microbes within our gut and the genes they harbor is known as the microbiome (Zaneveld et al., 2008). Each individual has a distinct microbiota fingerprint, and the composition is subject to change from acute and chronic environmental influences such as diet, illness and travel (Dethlefsen, McFall-Ngai, & Relman, 2007). The microbiome includes species of bacteria, archaea, fungi, viruses, protozoans, and sometimes multicellular organisms, though bacteria are the predominate population and reach 100 trillion cells in the colon (Lee & Mazmanian, 2010). The microbiome composition can also be referred to as "phylotype". In mammals, phylotype is established after birth, becomes more

established in childhood, and continues to evolve throughout one's lifetime. GI microbes are critical components of the early digestive development process and affect the amount of energy extracted from the diet (Ley et al., 2008). The microbiome acts as a metabolic filter between what is ingested, what is absorbed into the bloodstream, and what small molecules or nutrients are presented to the intestinal tract.

The microbiome provides protective immune and metabolic functions for the human host (Laparra & Sanz, 2010). The metabolic enzymes and pathways provided by the microbiome are vast, allowing for biotransformation of many molecules including lipids, carbohydrates and phytochemicals (Laparra & Sanz, 2010). It is an integral part of the human digestive system, providing essential functions for the host such as biosynthesis of vitamins (e.g. vitamin K) and digestion of otherwise non-digestible carbohydrates (e.g. cellulose, psyllium, and pectin) (L. V. Hooper, Midtvedt, & Gordon, 2002). An oligosaccharide that is non-digestible by the host but can be fermented by the microbiome is known as a "prebiotic", and these molecules will stimulate the growth and/or activity of different species of microorganisms within the microbiome, thus having the potential to change phylotype(Laparra & Sanz, 2010). Emerging evidence suggests that the diet modifiable gut microbiome is a promising area of exploration for chronic disease control and prevention (Cani et al., 2008; Delzenne & Cani, 2011; Kau et al., 2011; Ley et al., 2005) **(Figure 2)**. In addition to essential nutrients, there are a number of non-essential nutrients that differ across genetically diverse varieties of a single plant food, such as rice and dry beans (Heuberger et al., 2010; Mensack et al., 2010). These foods demonstrate promising potential to affect the microbiome via fibers and bioactive phytochemicals (e.g. polyphenolics, triterpenoids etc.), and may reveal a role for phytonutrient/phytochemical teamwork to influence dietary-mediated host protection against chronic disease.

4.3 Nutri-metabolome interactions and chronic disease

Nutrient-based dietary guidelines for overweight and chronic disease prevention may undoubtedly require an assessment of the gut microbiome to establish ideal nutritional phylotypes. A growing body of evidence exists for determining which dietary patterns are associated with improved human health (Adebamowo et al., 2005; Batres-Marquez, Jensen, & Upton, 2009; de Munter et al., 2007; Hung et al., 2004; Lanza et al., 2006; Michels et al., 2006; Mozaffarian et al., 2011; Sofi, Abbate, Gensini, & Casini, 2010). While age, sex, life stage, and other factors contribute to the variation in our nutrient requirements among individuals within a population, the opportunities to explore how plant diversity may benefit human health are astounding. Because human genetic variation confers tolerance/intolerance for certain foods and the genetic contribution to dietary requirements within and among human populations remains to be evaluated rigorously, the potential to continue to co-evolve with our food supply is a promising endeavor to consider.

Brown rice consumption was recently shown to have an inverse association with risk for developing Type 2 Diabetes (Sun et al., 2010), however the importance of the rice genotype for this effect is unknown. Recent evidence for metabolome diversity in cooked brown rice from genetically diverse varieties suggests that rice crop varieties may differ in these health promoting and disease fighting properties (Heuberger et al., 2010; Ryan et al., 2011).
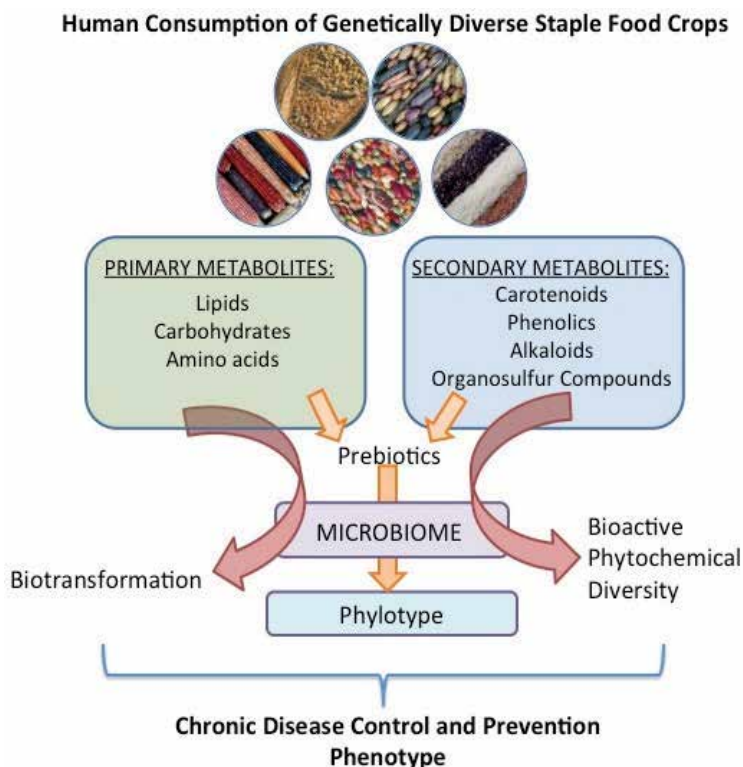
Fig. 2. Interactions between bioactive phytochemicals from genetically diverse staple food crops and the gut microbiome for chronic disease control and prevention.

5. Plant crop genetic diversity

Plants are among the oldest living complex organisms, and have had to face evolutionary pressures from pathogens such as insects, bacteria, and fungi from the very beginning of their evolutionary history. As they have developed defense mechanisms to protect themselves from infectious threats, these organisms have evolved counter-mechanisms of their own, necessitating further adaptations by plants, thus creating a cycle of co-evolution. This co-evolution of plants and the species that attack them has led to the development of an immense array of over 200,000 different phytochemicals present in plant life today (Hartmann, 2007; Macias, Galindo, & Galindo, 2007).

As humans transitioned from hunter gatherer societies to agricultural societies, they began cultivating crops, and both consciously and unconsciously selected for certain agronomic traits that rendered the crop more desirable for production in a particular climate and environment (Ross-Ibarra, Morrell, & Gaut, 2007). These selective pressures resulted in what is known as the "domestication syndrome", whereby certain traits are present in all domesticated crops compared to wild plants. These traits are ones that have made the crop easier to cultivate, including traits such as agricultural production of a larger fruit or grain, a more robust plant, more robust growth of the central stem compared to the side stems, and the loss of natural seed dispersal, which renders the plant dependant on humans for propagation (Doebley, Gaut, & Smith, 2006).

Since this time of early domestication, agricultural developments involve continuous improvements of genetic traits in our staple crop species, and have not considered traits of nutritional importance (Sands, Morris, Dratz, & Pilgeram, 2009). This has also been referred to as the "breeders dilemma". Recently, crop domestication has come with a price, and that price is a loss of genetic diversity that has occurred on two levels. First, there has been a loss of biodiversity in overall diet composition, as nearly 70% of all of the calories consumed by humans are supplied by only 15 crops (Ross-Ibarra et al., 2007), with the majority of these calories not coming from the basic raw or cooked form of the plant food, but from some more processed form. Secondly,there has been a loss of genetic diversity within each individual staple crop species themselves, as it has been estimated that cultivation has resulted in the loss of up to 95% of the genetic variation for many traits (M. D. T. Thompson & Thompson, 2009).

5.1 Loss of dietary phytochemical diversity

Primary metabolites produced by plants are compounds such as proteins, carbohydrates, and lipids that serve structural and functional purposes, and comprise essential nutrients in the human diet. Other phytochemicals produced by the plant that have no recognized role in the maintenance of fundamental life processes in the plants that synthesize them, but do have an important role in the interaction of the plant with its environment are known as secondary metabolites (Oksman-Caldentey & Inze, 2004). Secondary metabolites serve as chemical messengers functioning in the interaction of plants with their abiotic and biotic environment in processes such as communication, reproduction, or defense mechanisms (Hartmann, 2007). There are over 200,00 different plant secondary metabolites, and many have been studied for their bioactivity and enhanced human health importance **(Table 2)**, which has led in large part to their development as dietary supplements (Espin, Garcia-Conesa, & Tomas-Barberan, 2007), though secondary metabolites are typically included in the human diet via consumption of fruits, vegetables, spices, flavouring agents, or beverages (Mandlekar, Hong, & Kong, 2006). While the increased consumption of dietary supplements may have replaced the availability and intake of these chemicals from food in the past couple decades, there is emerging interest in achieving intakes of these "non-essential nutrients" from whole foods (Liu, 2003, 2004). The loss of genetic variation in staple food crops that make up our food supply is a serious concern for a number of reasons. Of particular emphasis in this chapter is the loss of genetic diversity that translates to reduced dietary phytochemical intake from foods that make of the bulk of our caloric intakes.

5.2 Food crop diversity as a feasible dietary solution to chronic disease susceptibility

One of the main tenets of a Paleolithic diet that we have diverged from compared to our present diets is the amount and diversity of plant foods and fiber that is consumed. Diet composition of plant foods is believed to have decreased from about 2/3 of total intake to less than 10% of total intake, while fiber intake is believed to have decreased from about 104 g/day in Paleolithic times to about 15.2 g/day in present times (Jew et al., 2009b). Because there is growing evidence that maintaining intake of diverse phytochemicals is of utmost importance for maintaining health and preventing chronic disease (M. D. T. Thompson &

Secondary Metabolites in whole grains and legumes	Examples (Amarowicz, 2008)	Biologic Activity
Antioxidants and Polyphenols	γ-Oryzanol (Rice) Flavonoids Phenolic acids Procyanidins Anthocyanidins α-lipoic acid	- Antibacterial - Antioxidant - Reduces cholesterol absorption - Anti-Cancer
Vitamin E	Alpha, gamma, delta tocotrienols and tocopherols	- Anti-tumor - Antioxidant - Antibacterial
Phytosterols	β-sitosterol Campesterol Stigmasterol	- Reduces cholesterol absorption - Anti-inflammatory - Antioxidant - Stimulates lymphocyte proliferation

Table 2. Secondary metabolites from plants and bioactive mechanisms of action for chronic disease prevention (Amarowicz, 2008).
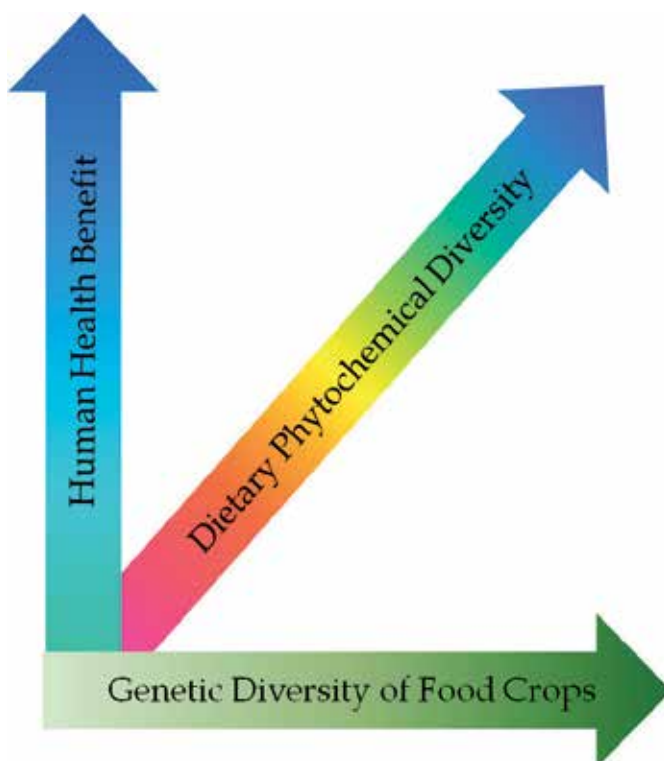

Fig. 3. Identifying relationships between food crop genetic diversity, dietary phytochemical diversity, and the potential opportunities for human health benefit.

Thompson, 2009; Heuberger et al., 2010; Ryan et al., 2011), it stands to reason that including more phytochemicals (both in terms of quantity and variety) in the diet has great potential for benefiting human health and decreasing risk for chronic disease (Figure 3). Recent evidence from the NIH-AARP cohort study that prospectively evaluated the diet-disease relationships of more than a half a million Americans supports that legumes and whole grains were the most significant sources of fiber intake compared to fruit and vegetable consumption for protection against chronic disease risk(Park, Subar, Hollenbeck, & Schatzkin, 2011). These findings provide strong rationale for evaluating **staple** food crops as a powerful vehicle for delivery of health promoting phytochemicals and bioactive food components that comprise a large part of total plant food and caloric intake.

6. Concluding remarks

The shift to the modern Western diet, lacking in fiber and phytodiversity while providing an overabundance of macronutrients, has happened relatively fast in evolutionary terms. Much of human nutrigenetic variation is the result of natural selection for genotypes that allowed for metabolism of diet available at the time. The selection of these gene mutations with the evolving food supply was a very slow process requiring thousands of years and many generations. While it is not feasible to change genotype that is amendable to current changing environmental conditions, or poor lifestyle choices, we have other layers to our genome that are modifiable. The epigenome and the microbiome are by nature short term modulators between our environment and our genes. In our quest to diminish chronic disease, we will need to harness this ability to affect the short-term evolutionary potential of the epigenome and microbiome and determine what dietary patterns have the most optimal effects on epigenotype and phylotype. Maintaining genetic diversity within our food crops is an important concept that can be appreciated across diverse scientific disciplines for providing an extensive array of molecules that, like folate or phenolics, may be beneficial modulators of the epigenome and microbiome. Genetic Diversity in staple food crops, because they are the most widely consumed, will play an especially important role in optimizing the diet-gene-epigenetic-microbiomic-disease relationships.

7. Acknowldegements

8. References

Adebamowo CA, Cho E, Sampson L, et al.: Dietary flavonols and flavonol-rich foods intake and the risk of breast cancer. Int J Cancer 114:628-633, 2005.

Amarowicz RP, R.: Legumes as a source of natural antioxidants. Eur. J. Lipid Sci. Technol. 110:865-878, 2008.

Batres-Marquez SP, Jensen HH, Upton J: Rice consumption in the United States: recent evidence from food consumption surveys. J Am Diet Assoc 109:1719-1727, 2009.

Bazzano LA, He J, Ogden LG, et al.: Legume consumption and risk of coronary heart disease in US men and women: NHANES I Epidemiologic Follow-up Study. Arch Intern Med 161:2573-2578, 2001.

Bray GA: Medical consequences of obesity. J Clin Endocrinol Metab 89:2583-2589, 2004.

Cani PD, Bibiloni R, Knauf C, et al.: Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. Diabetes 57:1470-1481, 2008.

Chang CL, Cai JJ, Lo C, et al.: Adaptive selection of an incretin gene in Eurasian populations. Genome research 21:21-32, 2011.

Chen YC, Peng GS, Wang MF, et al.: Polymorphism of ethanol-metabolism genes and alcoholism: correlation of allelic variations with the pharmacokinetic and pharmacodynamic consequences. Chem Biol Interact 178:2-7, 2009.

Cobiac L: Epigenomics and nutrition. Forum Nutr 60:31-41, 2007.

Cooney CA, Dave AA, Wolff GL: Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. J Nutr 132:2393S-2400S, 2002.

Coppieters N, Dragunow M: Epigenetics in Alzheimer's Disease: A Focus on DNA Modifications. Curr Pharm Des, 2011.

Cordain L, Eaton SB, Sebastian A, et al.: Origins and evolution of the Western diet: health implications for the 21st century. The American journal of clinical nutrition 81:341-354, 2005.

Coyne JA, Hoekstra HE: Evolution of protein expression: new genes for a new diet. Curr Biol 17:R1014-1016, 2007.

de Munter JS, Hu FB, Spiegelman D, et al.: Whole grain, bran, and germ intake and risk of type 2 diabetes: a prospective cohort study and systematic review. PLoS Med 4:e261, 2007.

DeBusk RM, Fogarty CP, Ordovas JM, et al.: Nutritional genomics in practice: where do we begin? J Am Diet Assoc 105:589-598, 2005.

Delzenne NM, Cani PD: Gut microbiota and the pathogenesis of insulin resistance. Curr Diab Rep 11:154-159, 2011.

Dethlefsen L, McFall-Ngai M, Relman DA: An ecological and evolutionary perspective on human-microbe mutualism and disease. Nature 449:811-818, 2007.

Dimitrov DV: The human gutome: nutrigenomics of the host-microbiome interactions. OMICS 15:419-430, 2011.

Doebley JF, Gaut BS, Smith BD: The molecular genetics of crop domestication. Cell 127:1309-1321, 2006.

Dolinoy DC, Huang D, Jirtle RL: Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. Proc Natl Acad Sci U S A 104:13056-13061, 2007.

Dolinoy DC: The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. Nutr Rev 66 Suppl 1:S7-11, 2008.

Espin JC, Garcia-Conesa MT, Tomas-Barberan FA: Nutraceuticals: facts and fiction. Phytochemistry 68:2986-3008, 2007.

Feil R: Environmental and nutritional effects on the epigenetic regulation of genes. Mutat Res 600:46-57, 2006.

Flint HJ, Duncan SH, Scott KP, et al.: Interactions and competition within the microbial community of the human colon: links between diet and health. Environ Microbiol 9:1101-1111, 2007.

Fraga MF, Ballestar E, Paz MF, et al.: Epigenetic differences arise during the lifetime of monozygotic twins. Proceedings of the National Academy of Sciences of the United States of America 102:10604-10609, 2005.

Garcia-Canas V, Simo C, Leon C, et al.: Advances in Nutrigenomics research: novel and future analytical approaches to investigate the biological activity of natural compounds and food functions. J Pharm Biomed Anal 51:290-304, 2010.

Go VL, Nguyen CT, Harris DM, et al.: Nutrient-gene interaction: metabolic genotype-phenotype relationship. J Nutr 135:3016S-3020S, 2005.

Goodacre R: Metabolomics of a superorganism. J Nutr 137:259S-266S, 2007.

Gray SG, De Meyts P. Role of histone and transcription factor acetylation in diabetes pathogenesis. Diabetes/metabolism research and reviews. 2005;21(5):416-33

Hartmann T: From waste products to ecochemicals: fifty years research of plant secondary metabolism. Phytochemistry 68:2831-2846, 2007.

Hattori MT, T.: The Human Intestinal Microbiome: A New Frontier of Human Biology. DNA Research 16:1-12, 2009.

Heuberger AL, Lewis MR, Chen MH, et al.: Metabolomic and functional genomic analyses reveal varietal differences in bioactive compounds of cooked rice. PLoS One 5:e12915, 2010.

Hirst M, Marra MA: Epigenetics and human disease. Int J Biochem Cell Biol 41:136-146, 2009.

Hollox E: Evolutionary genetics: genetics of lactase persistence--fresh lessons in the history of milk drinking. European journal of human genetics : EJHG 13:267-269, 2005.

Hooper L, Kroon PA, Rimm EB, et al.: Flavonoids, flavonoid-rich foods, and cardiovascular risk: a meta-analysis of randomized controlled trials. Am J Clin Nutr 88:38-50, 2008.

Hooper LV, Midtvedt T, Gordon JI: How host-microbial interactions shape the nutrient environment of the mammalian intestine. Annu Rev Nutr 22:283-307, 2002.

Hung HC, Joshipura KJ, Jiang R, et al.: Fruit and vegetable intake and risk of major chronic disease. J Natl Cancer Inst 96:1577-1584, 2004.

Jacob RA, Gretz DM, Taylor PC, et al.: Moderate folate depletion increases plasma homocysteine and decreases lymphocyte DNA methylation in postmenopausal women. J Nutr 128:1204-1212, 1998.

Jaenisch R, Bird A: Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet 33 Suppl:245-254, 2003.

Jarvela I, Torniainen S, Kolho KL: Molecular genetics of human lactase deficiencies. Annals of medicine 41:568-575, 2009.

Jew S, AbuMweis SS, Jones PJ: Evolution of the human diet: linking our ancestral diet to modern functional foods as a means of chronic disease prevention. J Med Food 12:925-934, 2009a.

Jew S, AbuMweis SS, Jones PJ: Evolution of the human diet: linking our ancestral diet to modern functional foods as a means of chronic disease prevention. Journal of medicinal food 12:925-934, 2009b.

Kau AL, Ahern PP, Griffin NW, et al.: Human nutrition, the gut microbiome and the immune system. Nature 474:327-336, 2011.

Kim KC, Friso S, Choi SW: DNA methylation, an epigenetic mechanism connecting folate to healthy embryonic development and aging. J Nutr Biochem 20:917-926, 2009.

Lanza E, Hartman TJ, Albert PS, et al.: High dry bean intake and reduced risk of advanced colorectal adenoma recurrence among participants in the polyp prevention trial. J Nutr 136:1896-1903, 2006.

Laparra JM, Sanz Y: Interactions of gut microbiota with functional food components and nutraceuticals. Pharmacol Res 61:219-225, 2010.

Lee YK, Mazmanian SK: Has the microbiota played a critical role in the evolution of the adaptive immune system? Science 330:1768-1773, 2010.

Ley RE, Backhed F, Turnbaugh P, et al.: Obesity alters gut microbial ecology. Proc Natl Acad Sci U S A 102:11070-11075, 2005.

Ley RE, Hamady M, Lozupone C, et al.: Evolution of mammals and their gut microbes. Science 320:1647-1651, 2008.

Li H, Mukherjee N, Soundararajan U, et al.: Geographically separate increases in the frequency of the derived ADH1B*47His allele in eastern and western Asia. American journal of human genetics 81:842-846, 2007.

Ling C, Groop L: Epigenetics: a molecular link between environmental factors and type 2 diabetes. Diabetes 58:2718-2725, 2009.

Liu RH: Health benefits of fruit and vegetables are from additive and synergistic combinations of phytochemicals. Am J Clin Nutr 78:517S-520S, 2003.

Liu RH: Potential synergy of phytochemicals in cancer prevention: mechanism of action. J Nutr 134:3479S-3485S, 2004.

Luca F, Perry GH, Di Rienzo A: Evolutionary adaptations to dietary changes. Annu Rev Nutr 30:291-314, 2010.

Macias FA, Galindo JL, Galindo JC: Evolution and current status of ecological phytochemistry. Phytochemistry 68:2917-2936, 2007.

Mandlekar S, Hong JL, Kong AN: Modulation of metabolic enzymes by dietary phytochemicals: a review of mechanisms underlying beneficial versus unfavorable effects. Curr Drug Metab 7:661-675, 2006.

Martin-Subero JI, Esteller M: Profiling epigenetic alterations in disease. Adv Exp Med Biol 711:162-177, 2011.

Mattson MP. Methylation and acetylation in nervous system development and neurodegenerative disorders. Ageing research reviews. 2003;2(3):329-42

Meisler MH, Ting CN: The remarkable evolutionary history of the human amylase genes. Crit Rev Oral Biol Med 4:503-509, 1993.

Mensack MM, Fitzgerald VK, Ryan EP, et al.: Evaluation of diversity among common beans (Phaseolus vulgaris L.) from two centers of domestication using 'omics' technologies. BMC Genomics 11:686, 2010.

Michels KB, Giovannucci E, Chan AT, et al.: Fruit and vegetable consumption and colorectal adenomas in the Nurses' Health Study. Cancer Res 66:3942-3953, 2006.

Moreno-Estrada A, Aparicio-Prat E, Sikora M, et al.: African signatures of recent positive selection in human FOXI1. BMC evolutionary biology 10:267, 2010.

Morgan HD, Santos F, Green K, et al.: Epigenetic reprogramming in mammals. Hum Mol Genet 14 Spec No 1:R47-58, 2005.

Mozaffarian D, Hao T, Rimm EB, et al.: Changes in diet and lifestyle and long-term weight gain in women and men. N Engl J Med 364:2392-2404, 2011.

Nicoletto SF, Rinaldi A: In the womb's shadow. The theory of prenatal programming as the fetal origin of various adult diseases is increasingly supported by a wealth of evidence. EMBO reports 12:30-34, 2011.

Oksman-Caldentey KM, Inze D: Plant cell factories in the post-genomic era: new ways to produce designer secondary metabolites. Trends Plant Sci 9:433-440, 2004.

Olshansky SJ, Passaro DJ, Hershow RC, et al.: A potential decline in life expectancy in the United States in the 21st century. N Engl J Med 352:1138-1145, 2005.

Oommen AM, Griffin JB, Sarath G, et al.: Roles for nutrients in epigenetic events. J Nutr Biochem 16:74-77, 2005.

Papanikolaou Y: Bean consumption by adults is associated with a more nutrient dense diet and reduced risk of obesity. FASEB J 20:A575, 2006.

Park Y, Subar AF, Hollenbeck A, et al.: Dietary fiber intake and mortality in the NIH-AARP diet and health study. Arch Intern Med 171:1061-1068, 2011.

Peng Y, Shi H, Qi XB, et al.: The ADH1B Arg47His polymorphism in east Asian populations and expansion of rice domestication in history. BMC evolutionary biology 10:15, 2010.

Perry GH, Dominy NJ, Claw KG, et al.: Diet and the evolution of human amylase gene copy number variation. Nature genetics 39:1256-1260, 2007.

Rampersaud GC, Kauwell GP, Hutson AD, et al.: Genomic DNA methylation decreases in response to moderate folate depletion in elderly women. Am J Clin Nutr 72:998-1003, 2000.

Romaguera D, Angquist L, Du H, et al.: Food composition of the diet in relation to changes in waist circumference adjusted for body mass index. PloS one 6:e23384, 2011.

Ross-Ibarra J, Morrell PL, Gaut BS: Plant domestication, a unique opportunity to identify the genetic basis of adaptation. Proc Natl Acad Sci U S A 104 Suppl 1:8641-8648, 2007.

Ryan EP, Heuberger AL, Weir TL, et al.: Rice bran fermented with saccharomyces boulardii generates novel metabolite profiles with bioactivity. J Agric Food Chem 59:1862-1870, 2011.

Sands DC, Morris CE, Dratz EA, et al.: Elevating optimal human nutrition to a central goal of plant breeding and production of plant-based foods. Plant Sci 117:377-389, 2009.

Sawan C, Herceg Z. Histone modifications and cancer. Advances in genetics. 2010;70:57-85

Simopoulos AP: Nutrigenetics/Nutrigenomics. Annu Rev Public Health 31:53-68, 2010.

Singh PN, Fraser GE: Dietary risk factors for colon cancer in a low-risk population. Am J Epidemiol 148:761-774, 1998.

Sofi F, Abbate R, Gensini GF, et al.: Accruing evidence on benefits of adherence to the Mediterranean diet on health: an updated systematic review and meta-analysis. Am J Clin Nutr 92:1189-1196, 2010.

Song Y, Manson JE, Buring JE, et al.: Associations of dietary flavonoids with risk of type 2 diabetes, and markers of insulin resistance and systemic inflammation in women: a prospective study and cross-sectional analysis. J Am Coll Nutr 24:376-384, 2005.

Sun Q, Spiegelman D, van Dam RM, et al.: White rice, brown rice, and risk of type 2 diabetes in US men and women. Arch Intern Med 170:961-969, 2010.

Thompson MDT, Thompson HJ.: Chapter 1 Biomedical Agriculture: A Systematic Approach to Food Crop Improvement for Chronic Disease Prevention, in Advances in Agronomy, Vol. Volume 102. Edited by Donald LS, Academic Press, 2009, pp. 1-54.

Thompson MDT, Thompson HJ.: Botanical Diversity in Vegetable and Fruit Intake: Potential Health Benefits, in Fruit and Vegetables in Health Promotion. Edited by Watson RRP, V.R. Oxford, Academic Press, 2009, pp. 3-17.

Timmermann S, Lehrmann H, Polesskaya A, Harel-Bellan A. Histone acetylation and disease. Cellular and molecular life sciences : CMLS. 2001;58(5-6):728-36.

Turnbaugh et al, A core gut microbiome in obese and lean twins. Nature. 2009; 457(7228):480-484

U.S.D.A. (2009, 8-13-09). USDA National Nutrient Database for Standard Reference, Release 18. Retrieved 9-12-2011, 2011, from
http://www.ars.usda.gov/Services/docs.htm?docid=9673

Villegas R, Gao YT, Yang G, et al.: Legume and soy food intake and the incidence of type 2 diabetes in the Shanghai Women's Health Study. Am J Clin Nutr 87:162-167, 2008.

Wolff GL, Kodell RL, Moore SR, et al.: Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. FASEB J 12:949-957, 1998.

Wooding S, Kim UK, Bamshad MJ, et al.: Natural selection and molecular evolution in PTC, a bitter-taste receptor gene. American journal of human genetics 74:637-646, 2004.

Zaneveld J, Turnbaugh PJ, Lozupone C, et al.: Host-bacterial coevolution and the search for new drug targets. Current opinion in chemical biology 12:109-114, 2008.

Zeven AC: Landraces: A review of definitions and classifications. Euphytica 104:127-139, 1998.

Edited by Mahmut Çalışkan

The Molecular Basis of Plant Genetic Diversity presents chapters revealing the magnitude of genetic variations existing in plant populations. Natural populations contain a considerable genetic variability which provides a genomic flexibility that can be used as a raw material for adaptation to changing environmental conditions. The analysis of genetic diversity provides information about allelic variation at a given locus. The increasing availability of PCR-based molecular markers allows the detailed analyses and evaluation of genetic diversity in plants and also, the detection of genes influencing economically important traits. The purpose of the book is to provide a glimpse into the dynamic process of genetic variation by presenting the thoughts of scientists who are engaged in the generation of new ideas and techniques employed for the assessment of genetic diversity, often from very different perspectives. The book should prove useful to students, researchers, and experts in the area of conservation biology, genetic diversity, and molecular biology.

IntechOpen