Martín Carlos Mottura

Development of microsatellites in *Prosopis* spp. and their application to study the reproduction system

Institute of Forest Genetics and Forest Tree Breeding
Faculty of Forest Sciences and Forest Ecology
Georg-August University of Göttingen
Germany

Development of microsatellites in *Prosopis* spp. and their application to study the reproduction system

Dissertation
submitted in partial fulfillment of
the requirements for the degree of
Doctor of Forestry Science
at the Faculty of Forest Sciences and Forest Ecology,
Georg-August University of Göttingen

by

Martín Carlos Mottura

born in San Guillermo, Santa Fe, Argentina

Göttingen 2006

D 7

Supervisors: Prof. Dr. Reiner Finkeldey

Dr. Aníbal R. Verga

Referees for the dissertation: Prof. Dr. Reiner Finkeldey

Prof. Dr. Ursula Kües

Dr. Aníbal R. Verga

Date of oral examination: 12th October 2006

This dissertation has been published through the library of Lower Saxony State and Georg-August University of Göttingen, Platz der Göttingen Sieben, 37073, Göttingen under the internet address:

http://resolver.sub.uni-goettingen.de/purl/?webdoc-1578

To my family and to Ana

ACKNOWLEDGEMENTS

Though the following dissertation is an individual work, it would not have been possible without the support, guidance and efforts of numerous people. These pages reflect not only the culmination of years of study, but also the relationships with many generous and inspiring people that accompanied me in this adventure. It is a pleasant aspect that I have now the opportunity to express my gratitude to all of them.

I am extremely grateful to Prof. Dr. Reiner Finkeldey for his continuous support, his excellent supervision, patience, and encouragement throughout this work. I would like to thank him for his help not only in my studies, but also in my daily life.

I am deeply grateful to my mentor, Dr. Aníbal R. Verga, for providing me the opportunity to work in this project, for his great guidance, his continuous support, and for encouraged me to develop independent thinking. My thanks go to him and his family for their help and friendship.

I am also very thankful to Prof. Dr. Ursula Kües for consenting to be one of the referees of this dissertation.

I express my deepest thanks to Dr. Oliver Gailing for his guidance in laboratory work, his generous time, and his commitment in this work. I am also very grateful for having the opportunity to work with an exceptional group of people, and wish to thank Prof. Dr. Hans-Rolf Gregorius, Prof Dr. Hans H. Hattemer, Prof. Dr. Martin Ziehe, Dr. Elizabeth Gillet, Dr. Dirk Kownatzki, Dr. Ludger Leinemann and Dr. Barbara Vornam for their support, constructive comments and discussions, help, and guidance during my studies.

Special thanks are given also to Olga Artes, August Capelle, Gerold Dinkel, Oleksandra Dolynska and Thomas Seliger for laboratory technical assistance. My thanks as well to our secretary Marita Schwahn, for her invaluable help in secretarial work.

I would like to express my warm gratitude to my colleagues at the Institute of Forest Genetics: Cui-Ping, Lucian, Natalia, Aki, Sapto, Jörg, Truong, Madhav, Akin, Taye, Abay, Nicolas, Marius, Sylvia, Nag, Hani, Yanti, Marcus and Amaryllis, for their friendship and many interesting academic and non-academic discussions. My thanks go also to my colleagues in Argentina: Gustavo, Cristian, Tomás, Jacke, Diego, and many others who could not be mentioned here, for their friendship and support. Special thanks to Ana Córdoba for providing me data, for sampling, and for her friendship.

I acknowledge the Government of Lower Saxony, Germany, for providing my financial support for my stay in Germany and for my studies.

From the deepest of my hard I want to thank my parents, Imelda and Juan Carlos, my brother Franco, and my sisters Kelly and Laura for their unconditional support and love. They give me strength and help me in countless ways. I also want to express my deepest and genuine gratitude to Ana, my couple, partner and friend, for her support, her patience, her company and her love during all this time.

Finally, my thanks go to many others who could not be mentioned here, but who had contributed to this work. Thanks to them, too.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFLPs:	Amplified Fragment Length Polymorphisms
bp:	Base Pairs
ca.:	Approximately
cm:	Centimeter
DNA:	Deoxyribonucleic Acid
e.g.:	For example
EMBL:	European Molecular Biology Laboratory
FIASCO:	Fast Isolation by AFLP of Sequence Containing Repeats
i.e.:	That is
IAM:	Infinite Allele Model
ISA:	Inter-SSR Amplification
ISSR:	Inter-SSR Amplification
ISSRs:	Inter Simple Sequence Repeats
KAM:	K-Allele Model
kb:	kilo-bases
m.a.r.	Mean Annual Rainfall
MAS:	Marked-Aid Selection
mm:	Millimeters
MP-PCR:	Microsatellite-Primed PCR
mRNA:	Messenger Ribonucleic Acid
ng:	Nanograms
PCR:	Polymerase Chain Reaction
pers. comm.:	Personal Communication
QTLs:	Quantitative Trait Loci
RAMPs:	Random Amplified Microsatellite Polymorphism
RAPDs:	Random Amplified Polymorphic DNA
RFLPs:	Restriction Fragment Length Polymorphisms
SMM:	Stepwise Mutation Model
SNPs:	Single Nucleotide Polymorphisms
SSLP:	Simple Sequence Length Polymorphism
SSM:	Slipped Strand Mispairing
SSR:	Simple Sequence Repeats
STMS:	Sequence Tagged Microsatellite Sites
STMS:	Sequence Tagged Microsatellite Sites
STR:	Short Tandem Repeats
TPM:	Two Phase Model
UCO:	Unequal Crossing-Over
VNTR:	Variable Number Tandem Repeats

1. INTRODUCTION

1.1 Background

The genus *Prosopis* (family Fabaceae) occurs worldwide in arid and semi-arid regions. It includes 44 species (Burkart, 1976; see **Appendix 1**), involving trees and shrubs, found in the Near East, North and Central Africa, North and South America, and the Caribbean. *Prosopis* species vary widely in their productivity and their use and utilization by humans (Pasiecznik et al., 2001). They constitute a very important natural resource for dry zones due to their multi-purpose nature, with the potential to provide a wide range of products, and their ability to grow on the poorest soils where few other useful species can survive. *Prosopis* trees are harvested for pods, fuel or timber wood, and for many other products, such as medicinal extracts or foliage for animal fodder. In addition, they stabilize the soil and prevent erosion, and via biological nitrogen fixation they increase the fertility of soils. *Prosopis* species are seen to fulfill very important roles in both production and protection in many arid and semi-arid regions of the world. Argentina constitutes the main center of diversity of the genus *Prosopis*, with approximately 28 species, 13 of them endemic.

In Argentina, the phytogeographic provinces of the Chaco, Monte, and Espinal (Cabrera, 1976) (**Figure 1**), cover over one million square kilometers, representing approximately one third of the total country area. Within these phytogeographic provinces, *Prosopis chilensis* (Molina) Stuntz and *Prosopis flexuosa* DC., two hardwood arboreal species, constitute important natural resources due to their multipurpose nature (Karlin and Díaz, 1984; Cony, 2000). These species grow in arid and semi-arid regions of the country, which mainly correspond to poor areas from an economical standpoint. These trees play an important ecological role in combating desertification, and controlling erosion. *P. chilensis and P. flexuosa* exhibit the ability to improve soils condition via biological nitrogen fixation, leaf litter addition and incorporation, nutrient pumping, and changes in soil structure and in soil fauna. Besides, these woody plants are of significant economical value because all of their biomass is potentially exploitable (wood, charcoal, pods, etc.). These species have been noted as producers of large quantities of raw materials required by local populations, and the value of *Prosopis chilensis* and *P. flexuosa* for providing much needed

resources for the poor is generally accepted. They offer shade, food for humans, forage for wildlife and livestock, and flowers for honey production. They provide high quality wood for fences, vineyards, furniture, etc., as well as firewood and charcoal. In addition, both species have been reported as promissory tree species for reforestation programs in degraded ecosystems. *Prosopis chilensis* and *P. flexuosa* play a vital role in the ecology and the economy of many arid and semi-arid zones in Argentina, being most of the time the most important resource of the local populations in the arid regions.

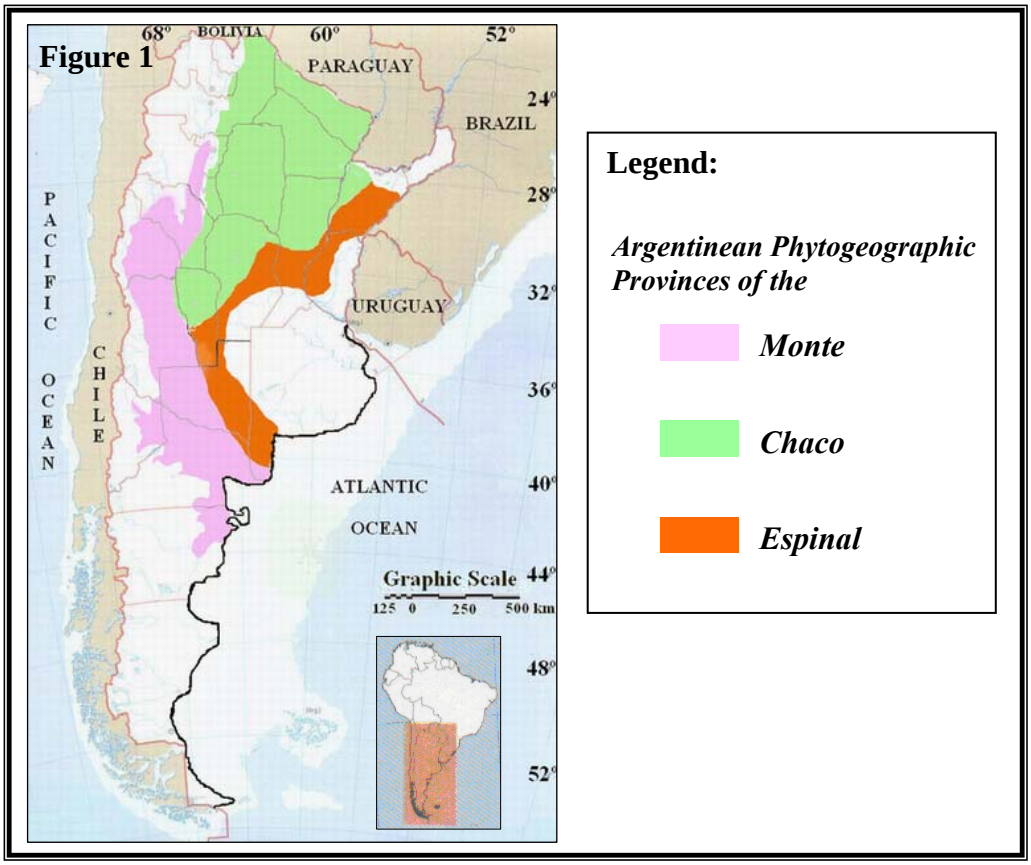

Figure 1: Argentinean phytogeographic provinces of the Monte, Chaco and Espinal according to Cabrera (1976).

In spite of their great value, *Prosopis* forests have been over-exploited without much attention being paid to their conservation. In Argentina, large scale clearance of *Prosopis* forests began after colonization. Prior to colonization, management was minimal but exploitation was also at a very low level, with the large *Prosopis* forests able to support the local populations on sustainable basis (Pasiecznik et al., 2001). It was the advent of major land clearance and industrialization that led to over-exploitation

of native forests. In Argentina, deforestation has been due primarily to land clearance for agriculture or ranching, and logging for a variety of wood uses. Also, where *Prosopis* forests were not completely destroyed, many have suffered from continuous selective felling and genetic depletion with the best trees preferentially removed over time. In Argentina, it has been estimated that the natural coverage of *Prosopis* forests was reduced to between one quarter and one half of its original area between 1500 and 1975, due to the activities of man (D'Antoni and Solbrig, 1977). The irrational exploitation of the *Prosopis* resources and the absence of sustainable productive alternatives for the arid regions have so far caused degradation of the natural resources (vegetation and land), loss of productivity, desertification, and in many areas the total extinction of the native forest (Cony, 2000; Karlin and Díaz, 1984; Verga et al., 2000).

Rational exploitation of promissory species, such as *Prosopis chilensis* and *P. flexuosa*, as well as the development of sustainable management programs and protection strategies for these species, and reforestation plans to control desertification and recover highly degraded areas, are essential in order to preserve the *Prosopis* resources. This, in turn, requires a great in-depth knowledge of the biological characteristics, adaptative strategies, existing variability, genetic structure, and evolutionary relationships of both species. Previous studies of genetic resources of *Prosopis* in Argentina have been performed by means of morphological characters (Cony, 1996; Mantovan, 2004), isozymes (Solbrig and Bawa, 1975; Saidman, 1985, 1986, 1990, 1993; Saidman and Vilardi, 1987, 1993; Verga, 1995; Saidman et al. 1997, 1998b; Bessega et al., 2000a, b; Joseau, 2006) or RAPD markers (Saidman et al. 1998a; Bessega et al., 2000c). Nevertheless, actual understanding of genetic resources in *Prosopis* forests in Argentina is still rudimentary. Therefore, highly polymorphic molecular markers are indispensable for the study of the genetic resources of *Prosopis chilensis* and *P. flexuosa*.

Microsatellite markers, also known as SSRs (Simple Sequence Repeats), have proven to be an extremely valuable molecular tool for a wide range of genetic studies in many organisms. In the last years, these markers have become one of the most popular molecular markers used with application in many different fields, such as population genetics, conservation and management of biological resources, genome mapping and also forensic studies. Their high polymorphism, co-dominant mode of expression and the relative easy of scoring represent the major features that make microsatellites of

large interest for many genetic studies. The major drawback of SSRs is that they need to be isolated *de novo* for species that are being examined for the first time.

1.2 Objectives of the study

The present study is focused on two main topics: (*i*) development of SSR markers for the genus *Prosopis*; (*ii*) the study of the reproduction system of a hybrid swarm between *Prosopis chilensis* and *P. flexuosa* based on the developed SSR markers.

1.2.1. Development of SSR markers for *Prosopis spp.*

An aim of the present study is to develop novel and variable microsatellite gene markers for the genus *Prosopis* with the purpose of providing new molecular tools for the study of the genetic resources of this genus. In chapter **2** of this work, an introduction into microsatellite markers (SSR markers) is given in order to explain and clarify different topics concerning these DNA markers. Chapter **3** introduces the genus *Prosopis* and its importance as natural resource. Chapter **4**, and the manuscripts **I** and **II**, describe the procedure used for the development of microsatellite markers in *Prosopis*, and the results obtained.

1.2.2. Reproduction system study

Studies on the reproductive system of the species are crucial in order to predict the evolutionary future of populations. The study of hybrid complexes can provide excellent models to approach evolutionary problems. Speciation mechanisms, and the relationship between the degree of genetic differentiation and taxonomic status constitute some of the main questions that may be addressed using these models. Furthermore, they allow to study interspecific gene flow and the consequences of hybridization. Thus, a better understanding on the genetic structure and reproduction system in hybrid swarms between *Prosopis chilensis* and *P. flexuosa* will improve our knowledge on the biological characteristics, adaptative strategies, existing variability, genetic structure, and evolutionary relationships of both species. Consequently, the objectives of the study on the reproduction system are:

- to classify morphologically the individuals that are part of a hybrid swarm between *Prosopis chilensis* and *P. flexuosa*, located in the Argentinean Arid Chaco;

- to assess and compare within and among the obtained morphological groups, the genetic structure and variation inferred from individual multilocus genotypes based on the previously developed SSR markers;

- to analyzed and describe aspects of the reproduction system (e.g., pollen movement, gene flow, degree of reproductive compatibility between groups) within the studied hybrid swarm based on paternity analysis applying the developed SSRs;

- to estimate for the obtained groups mating system parameters such as: outcrossing rate, biparental inbreeding and correlated mating.

The second topic of this thesis, the study of the reproduction system of a hybrid swarm based on SSR markers, is an example of application of these molecular markers in genetic studies. Chapter **4** and the manuscript **III**, describe the methodologies and procedures used in the analysis of the mating system. The obtained results of the mating system analysis of a hybrid swarm between *Prosopis chilensis* and *P. flexuosa* based on SSR markers are presented and discussed in chapters **5** to **7**, and in the manuscript **III**.

2. MICROSATELLITES AND GENETIC MARKERS

2.1. Definition and evolution of microsatellites

In different studies conducted at the beginning of the 1980's, simple sequence repeats were found widespread distributed in many eukaryotic genomes (Hamada et al., 1982; Tautz and Renz, 1984). These ***Simple Sequence Repeats*** (***SSRs***), later also called ***microsatellites*** (Litt and Luty, 1989), are sequences of DNA (nucleotides: adenine - A, thymine - T, guanine - G, cytosine - C) made up of tandemly repeated motifs, from one to six bases in length, which are arranged head-to-tail generally without interruption (Hancock, 1999) (**Figures 2 and 3**). They are often flanked by conserved regions, and seem to be widespread and more or less evenly distributed throughout the genome. In

the literature, microsatellite or SSR regions are also known as SSLP (Simple Sequence Length Polymorphisms), STMS (Sequence Tagged Microsatellite Sites), STR (Short Tandem Repeats), or VNTR (Variable Number Tandem Repeats).

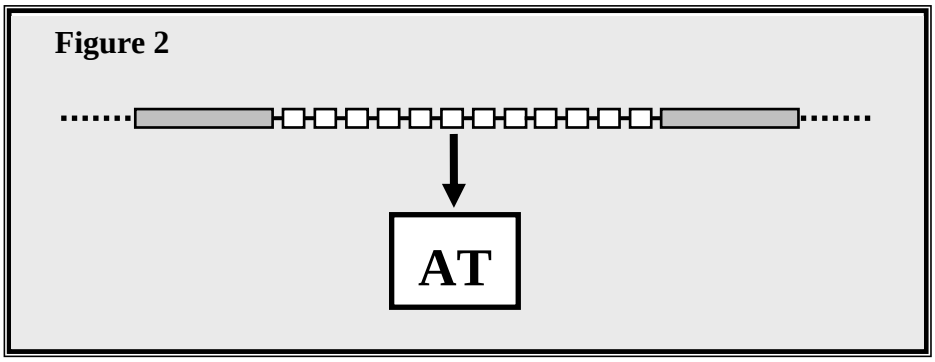

Figure 2: Diagrammatic representation of a microsatellite sequence with a motif ***Adenine-Thymine*** repeated 12 times, i.e., $(AT)_{12}$. The graphic represents a single-stranded DNA molecule. Microsatellite repeats are represented by the small white boxes. The gray boxes at each side of the microsatellite sequence represent the flanking regions.

Microsatellites have been detected within the genomes of every eukaryotic organism so far studied, and also in prokaryotes (Field and Wills, 1998). In plants, SSRs have been detected as well in the chloroplast genome (Valle, 1993; Vendramin et al., 1996; Weising and Gardner, 1999). SSRs are often found at much higher frequencies than they are expected by chance (Hancock, 1999). Edwards et al. (1991) investigated microsatellite loci in the human genome, they observed that at least one SSR was present every 300 to 500 kilo-bases (kb). In plants, a search in sequence databanks revealed that nuclear microsatellites are widely distributed with a frequency of one each 50 kb. Microsatellites show high levels of polymorphism (Litt and Luty, 1989; Weber and May, 1989; Tautz, 1989). The most polymorphic, and therefore the most useful for many purposes, are uninterrupted arrays (Hancock, 1999).

Microsatellites are rare within coding regions of the DNA, although there is evidence of SSR sequences located within protein-coding domains (Tóth et al., 2000; Gentles and Karlin, 2001; Katti et al., 2001). Triplets (i.e., individual codons) comprise by far the most common motif length for SSRs within protein sequences. In plant genomes microsatellites frequency seems to be higher in transcribed regions than in other regions of the genome (Morgante et al., 2002)

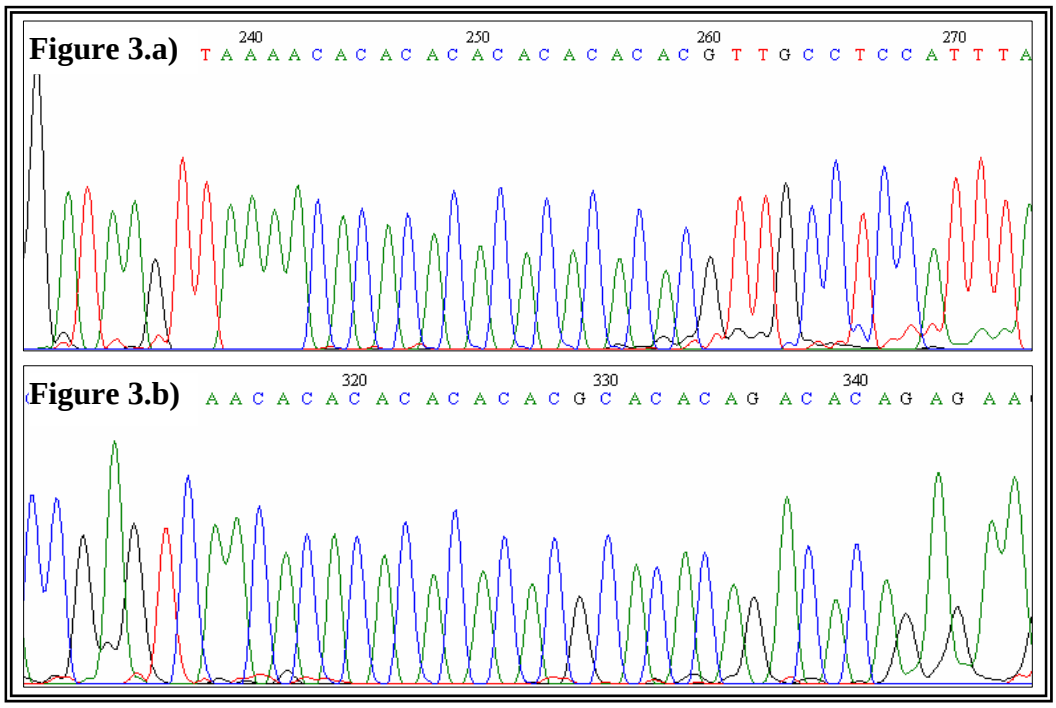

Figure 3: Microsatellite sequences in *Prosopis chilensis* **a)** with the dinucleotide repeat $(AC)_{n}$., and **b)** with the dinucleotide repeat $(AC)_{n}$ interrupted by (G).

Depending on the number of nucleotides per motif, microsatellites are divided into mononucleotide repeats (motif with a single nucleotide, typical for cp SSRs), dinucleotide repeats (motif with two nucleotides), trinucleotide repeats (motif with three nucleotides), and so on. On the basis of purity, microsatellites are divided into three types according to Weber (1990): a) ***perfects*** (i.e., SSR sequences without interruptions in the runs of the repeats) (**Figure 3.a**); b) ***imperfects*** (i.e., microsatellites with one or more interruptions in the run of repeats) (**Figure 3.b**); and c) ***compounds*** (i.e., microsatellite sequences with adjacent tandem simple repeats of a different sequence).

Microsatellite sequences poly (A)/(T) are the most common tandemly repeat sequences in the human genome (Hancock, 1999). In mammalians, (CA)/(GT) repeats are the most common dinucleotide repeats, occurring about twice as frequently as (AT)/(TA) repeats, and three times as often as (AG)/(TC) repeats (Hancock, 1999). Among trinucleotide SSRs, poly (CAG)/(GTC) and poly (AAT)/(TTA) appear to be the most common repeats in mammalian genomes. In *Drosophila melanogaster* (CA)/(GT) repeats form the most common microsatellite found (Schug et al., 1998). In plants, the most common motif repeat is poly (AT)/(TA) (Morgante and Olivieri, 1993), while (CA)/(GT) repeats seems to be the most-scarce.

2.1.1. Nature and origin of polymorphism

A microsatellite region, independently from the repeated motif (CA, TG, ATG, etc.), constitutes in many cases a highly variable multiallelic locus, with a high informative content. The nature of this variation (polymorphism) is due to the fact that the number of repeat motifs in a microsatellite at a particular locus can vary (**Figure 4**). This variation can be produced within species (Gupta et al., 1996) and sometimes even within a single phenotypically homogeneous cultivated variety (Olufowote et al., 1997; Poulsen et al., 1993). The degree of polymorphism in SSRs tends to be proportional to the total microsatellite size. Microsatellite regions smaller than 20 base pairs (bp) in size do not provide enough variation for many applications (Dorocicz, 1999). However, it is not justified to generalize that small microsatellites are always nonvariable, and large SSRs are always polymorphic (Schug et al., 1998; van Treuren et al., 1997). Besides, interruption of microsatellites results in reduced polymorphism and reduced mutation rates, consistent with the greater difficulty of forming slipped intermediates in the presence of sequence interruptions (Hancock, 1999).

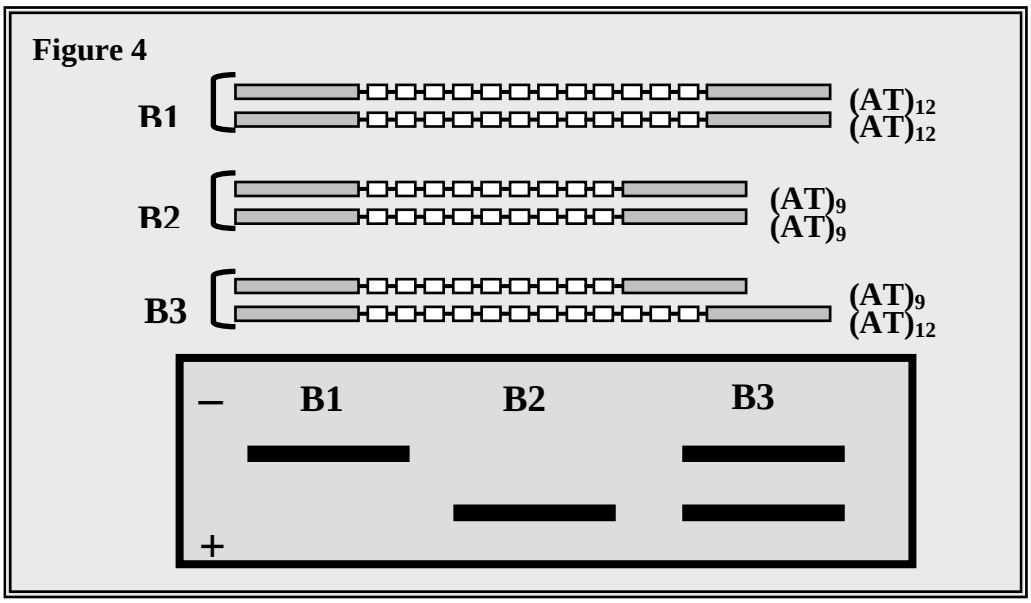

Figure 4: Graphical representation of the nature of polymorphism in microsatellites. Graphics of single-stranded DNA (one strand from each homologous chromosome) containing the microsatellite locus $(AT)_n$ (small white boxes represent SSR repeats, and gray boxes at each side, the flanking regions) for three different individuals (B1, B2 and B3). $(AT)_{12}$ and $(AT)_9$ are the alleles at the $(AT)_n$ locus. The gray box bellow represents the electrophoretic phenotypes of the three individuals at the SSR locus $(AT)_n$.

The reasons for the emergence of SSRs are still discussed. There is considerable theoretical debate about the origins of microsatellite length polymorphism. Variation in microsatellite structure over time reflects the balance between two opposing forces: the mutational forces that generate variation and the cellular error repair systems that remove variation. There are two different models to describe mutation in the number of motifs in microsatellites: ***Slipped Strand Mispairing*** (***SSM***) (**Figure 5**) and ***Unequal Crossing-Over*** (***UCO***).

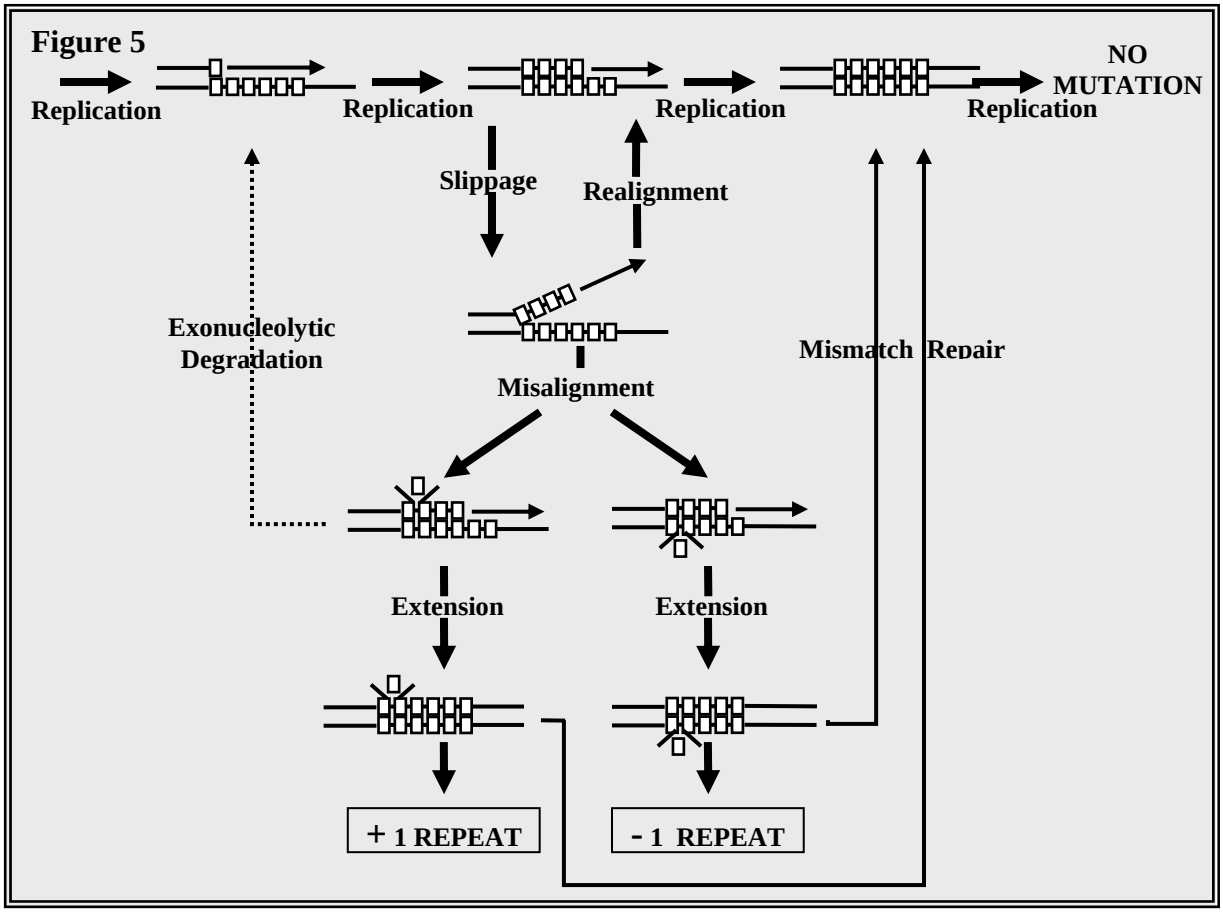

Figure 5: Model of the SSM mutation process at microsatellite loci. Graphics of double-stranded DNA containing an SSR repeat are shown at different stages of the replication and mutation process. DNA strands are represented by thin lines, and microsatellites repeats by small white boxes. Flow arrows point down for steps that lead to mutations, up for steps that prevent mutations, and to the right for steps in the ongoing DNA replication process. The exonucleolytic degradation is shown with dashed line since it has only a limited role in regulating SSR mutations (from: Eisen, 1999).

SSM model (**Figure 5**) involves only a single DNA double helix and slipped strand mispairing (slippage) during DNA replication (Streisinger et al., 1966; Levinson and

Gutman, 1987). This model proposes that polymerase slippage occurs during replication due to the repetitive nature of microsatellites. Slippage during replication can take place when the nascent DNA strand dissociates from the template strand (Levinson and Gutman, 1987). When the new DNA strand is being synthesized, the polymerase can "slip" along the microsatellite repeats, resulting in a mispairing, which can create loops in either the template or the newly synthesized strand (Dorocicz, 1999). When replication continues after such a mispairing, the eventual nascent strand will be longer or shorter than the template, depending whether the loop is created either in the new strand (the resulting strand will be longer) or in the template (the nascent strand will be shorter) (Wierdl et al., 1997).

UCO is believed to occur during recombination at meiosis and/or mitosis (Smith, 1976; Valdes et al., 1993). Recombination could potentially alter the lengths of microsatellites by unequal crossing-over or by gene conversion (Hancock, 1999). UCO results in motif number change when crossing-over occurs between misaligned microsatellites on sister chromatids of homologous chromosomes (Wierdl et al., 1997). Misalignment between sister chromosome strands occurs most easily for long, tandemly repeated sequences where the recombination machinery cannot easily determine the correct register between the two strands (Hancock, 1999). Unequal crossing-over can occur both between chromatids in the same chromosome and between chromosomes.

Regardless as to the cause of microsatellite polymorphism, the SSM hypothesis appears to be the most accepted by the majority of researchers (Schlotterer and Tautz, 1992). Experiments with yeast and *Escherichia coli* have provided strong evidence supporting the SSM model (Wolff et al., 1989). Also, studies of certain human diseases suggested that SSM can explain the relationship between defective DNA repair and the microsatellite instability that is the suspected causative agent of these diseases (Koreth et al., 1996). In order to understand the complete apparatus of the SSM mutation model the cellular mechanisms for error correction should be included in this model (Eisen, 1999). The cellular machinery responsible for removing unimportant or harmful sequences includes two pathways of error correction: exonucleolytic proofreading, and post-replication mismatch repair. If these correction mechanisms are not efficient enough to delete these defective sequences, then they will remain in the genome (Orgel and Crick, 1980). It has been suggested that transcribed DNA regions are more rapidly repaired via error correction mechanisms than non-transcribed DNA, consequently

larger numbers of alleles can accumulate in microsatellites found in non-transcribed regions (Turner et al., 1997).

2.1.2. Theoretical mutation models for microsatellites

Microsatellites display wide variation in their degrees of instability, with reported mutation rates from 10^{-2} to 10^{-10} per locus and per generation (Levinson and Gutman, 1987; Edwards et al., 1992; Schlötterer and Tautz, 1992; Weber and Wong, 1993; Bowcock et al., 1994; Di Rienzo et al., 1994; Forbes et al., 1995; Schug et al., 1997; Hancock, 1999). Different theoretical models have been used in order to understand the evolutionary dynamics of SSRs. These models try to explain the high degree of polymorphism of repeat arrays. Mutation at microsatellite loci usually involves a change in size of one repeat, but this mutation can also involve several repeated units as well (Estoup and Cornuet, 1999). A detailed understanding of the mutational process guiding the evolution of SSRs is necessary in order to optimize the information obtained from these markers (Pandey, 2005). Four theoretical mutation models are briefly described below:

A. Infinite allele model (IAM) (Kimura and Crow, 1964)

This model has been used extensively to study the mechanisms of maintenance of protein polymorphism and has later been applied to SSR markers when they appeared. The model assumes that each mutation can create any new allele randomly. According to this, a mutation involves any number of tandem repeats and always results in an allele state not previously existing in the population. In the IAM, there is no constraint on the potential SSR size, and infinite number of choices (i.e., new alleles) is available and has equal probability of occurring. There are two basic assumptions in this model: 1) all mutations are assumed to be novel; 2) mutation rates are the same for all loci.

B. Stepwise mutation model (SMM) (Kimura and Ohta, 1978)

The SMM was developed in the days of isozyme electrophoresis. This model, in contrast to the IAM, describes mutation of microsatellite alleles by the loss or gain of a single tandem repeat, with the same probability in both directions (deletion or addition). Therefore, alleles may possibly mutate towards allele states already present in the population. In this model the underlying assumption is that the most frequent mutational

events only alter microsatellites by one "step" (a single tandem repeat) at a time. Besides, alleles of similar size are expected to be more closely related to each other than alleles of completely different size.

C. K-allele model (KAM) (Crow and Kimura, 1970)

The KAM assumes that there are *K* possible allelic states at the locus in question. Any of these possible alleles has a constant probability of mutating towards any of the *K – 1* other allelic states. Under this model, the number of alleles observed represents all possible alleles, and mutation is equally likely among any pair of alleles.

D. Two phase model (TPM) (Di Rienzo et al., 1994)

Under this model, the allelic variation at a microsatellite locus is assumed to be produced primarily by single-step changes in allele size, but also by rare but important events of large magnitude. In the TPM model, mutations introduce a gain/loss of X repeats. With probability *p*, X is equal to one (this corresponds to the SMM) and with probability, *1-p*, X follows a geometric distribution.

There are other models for microsatellite evolution which have been proposed besides the four abovementioned. Much focus has centered on the development of realistic models for SSR mutation. Which theoretical model should be applied to microsatellites is a main question, because population genetic inferences are sensitive and dependent on the assumed model. In addition, the models discussed above assume that allelic differences are entirely due to changes in the number of the basic repeated unit. However, other forms of mutational changes can occur, such as insertions and/or deletions in the flanking sequences (Curtu et al., 2004). This is particularly observed when the same SSR locus is compared between different species, while it is more rarely observed within species (Estoup and Cornuet, 1999).

Several lines of evidence suggest that the mutation behavior of microsatellites is a complex phenomenon involving many different factors that influence the mutation rate at a given locus. For example, the mutation behavior is dependent on the internal structure of a SSR locus. Mutation of repeat arrays has a complicated dependence on allele size, composition of the repeat unit and purity, among other factors. It has been proved in different studies that independently from the repeat type, polymorphism of

SSR increases with increasing number of repeats. Furthermore, dinucleotide loci appear to evolve at a higher rate than tetranucleotide loci (Estoup and Corneut, 1999), showing a clear relationship between repeat composition and mutation rate. Purity also influences the mutation rate, imperfect SSRs show lower levels of polymorphism than perfect ones with the same repeat motif. Differences in mutability among microsatellite loci were also observed in relation to the location of the SSR in the genome, and among species. Another factor that may influence the variability of a microsatellite locus is selection. The microsatellite itself may have an important function in the genome (see section **2.1.3.**), and the role of selection is to maintain this function by constraining SSR variability. Alternatively, selection may not act on the microsatellite itself but on the DNA region closely linked to the SSR and therefore, influence the microsatellite region.

2.1.3. Possible roles of microsatellites in the genome

SSRs are commonly regarded as "junk" DNA elements (i.e., with no significant roles as genomic information) which are nothing more than parasitic, selfish DNA that persists because cellular repair machinery cannot efficiently remove them (Orgel and Crick, 1980). Although, accumulated evidence support that these motifs have important roles in the cells, and that they could play a positive role in adaptative evolution (Naylor and Clark, 1990; Moore et al., 1991; Li et al., 2004; King et al., 2006). On the other hand, some evidence, such as connection of microsatellites to certain disease phenotypes and to reduced fitness, suggests that repetitive DNA elements can be detrimental (Künzler et al., 1995; Karlin, 2002; Brown and Brown, 2004; Rockman et al., 2004). Alternatively, it has been suggested that only a fraction of the repetitive DNA is functional, while the rest makes a neutral contribution to the organism (Dorocicz, 1999).

In numerous instances microsatellites are found in upstream promoter regions of coding sequences. Several studies documented the conservation of SSRs sequences found in corresponding upstream locations of the same gene in different species. Conservation of sequences across species is often an indication of biological function (Kashi and Soller, 1999). Some of the functional roles attributed to microsatellite sequences are: protein binding sites, sites for nucleosome assembly, enhancer elements that increase the activity of the nearest promoter, and also enhancer *per se* (Ishii et al., 1987; Gilmour et al., 1989; Lue et al., 1989; Suen and Hung, 1990). It has also been proposed that some SSRs may act as genomic tags for the identification of recombination "hot spots"

(Bullock et al., 1986). Some other suggested roles for repetitive DNA include involvement in chromosome pairing, control of gene expression, mRNA processing, participating in DNA replication, and regulation of transcription.

The properties of microsatellites (e.g., extremely high rate of reversible, length-altering mutations; site-specific mutability; diversity; abundance in both coding and noncoding domains) confer virtually ideal "mutator" properties to these sequences. SSRs thus provide a prolific source of quantitative and qualitative variation. It has been suggested that microsatellites, due to their special mutational and functional qualities, play a major role in generating the genetic variation underlying adaptive evolution (Kashi and King, 2006). Recent reports support an evolutionary role for SSRs as important sources of adaptive genetic variation, both within and between species (Sawyer et al., 1997; Fahima et al., 2002; Fondon and Garner, 2004; Hammock and Young, 2005; Nevo et al., 2005; Verstrepen et al., 2005). Some of these researches documented not only the quantitative phenotypic effects of repeat-number alleles (i.e., variation supplied by SSRs is in many cases at least partially responsible for phenotypic differences), but also that natural selection acts upon these alleles shaping their frequencies.

Effects of coding SSRs may be surprisingly sophisticated, but microsatellite effects are not limited to coding sequences. Repeat variation commonly exerts the functional influence on DNA structure and transcription activity, even when the microsatellites are located in introns or other noncoding sites where they do not affect protein structure directly (Kashi and King, 2006). Whatever role an SSR plays within genes, changing the number of repeats can modulate its genetic function. Any genomic variable that routinely affects genetic function must surely play an evolutionary role as well. Therefore, microsatellites may "provide a ready and virtually inexhaustible supply of new quantitative variation for rapid evolutionary adaptation" (Kashi and King, 2006).

2.2. Genetic markers

Following the definitions given by Gillet (1999), a trait is termed a ***genetic trait***, if any two individuals possessing the same genotype also have the same phenotype, regardless of the environmental conditions in which they exist. If after a successful inheritance

analysis the relationship holds that each phenotype can be unambiguously assigned to a set of genotypes at one or more specified loci, then a genetic trait is qualified as a **genetic marker**. This means that genetic markers are traits controlled by one or a few loci. Different types of genetic markers have been used in plant genetic studies throughout time.

When differences in DNA occur within genes, the differences have the potential to affect the function of the gene and hence the phenotype of the individual. ***Morphological markers*** were the first markers used in plant genetic studies. Mendel (1866) looked at morphological traits in some of his basic experiments. Morphological markers contributed significantly to the development of genetic linkage analysis and to the construction of the first linkage maps, among other studies. These markers are controlled by genes associated to a morphological character and the phenotypes are, generally, easy to identify. Some of the morphological traits used in plant genetics as markers are rare leave forms or dwarfishness. Certain ***color traits*** (e.g., the "*purpurea*" form of some trees) were also used in genetic studies. Secondary products of metabolic pathways, such as ***terpenes*** and other products of the secondary metabolism have also been used in the past as genetic markers, especially in conifer species (Hattemer, 1991). However, there are relatively few such markers and its use is reduced to a restricted number of species.

In the early 70th, the development of the ***isoenzyme markers*** represented a great advance in the field of genetic markers. ***Isozymes*** or ***isoenzymes*** are multiple forms of one enzyme with equal or similar catalytic function arising from genetic control of primary protein structure (Gailing et al., 2005). The "one gene - one polypeptide" principle suggests a close relation between the variation at gene loci coding for isoenzymes and isoenzyme phenotypes, since enzymes are the most important group of polypeptides (Finkeldey and Hattemer, in press.). Because the nucleotide sequence in the DNA codes for the corresponding sequence of amino acids, a change in the DNA sequence possibly results in a change in the respective sequence of amino acids. These changes might lead to new forms of the enzyme by changing its structure. These different forms of an enzyme can be detected electrophoretically due to changes in the overall molecular charge (Cooke, 1984). Isozymes analyses have been widely used over the past several decades as a powerful technique to investigate many aspects of the

genetic system of a large number of plant species (Hamrick and Godt, 1989; Finkeldey and Hattemer, in press.).

Since the advent of modern molecular techniques, the number of genetic markers available has increased enormously. DNA genetic markers allow the observation of differences (polymorphisms) directly in the molecule which contains the genetic information. Initially, direct investigation of DNA was performed by means of restriction enzymes (Grodzicker et al., 1974). This technique was called ***"Restriction Fragment Length Polymorphism"*** (***RFLPs***). In summary, the RFLP procedure is based on the digestion of DNA with restriction enzymes, and subsequent detection of changes in the lengths of the produced DNA bands. This technique requires large amounts of DNA. Later on, the development of the process of Polymerase Chain Reaction (PCR) (Mullis and Fallona, 1987; Saiki et al., 1988), together with novel cloning and sequencing techniques, new DNA markers became available. Some of the commonly used DNA-markers are: ***Amplified Fragment Length Polymorphisms*** (***AFLPs***), ***Random Amplified Polymorphic DNA*** (***RAPDs***), ***Minisatellites***, ***Microsatellites*** or ***Simple Sequence Repeats*** (***SSRs***), ***Inter Simple Sequence Repeats*** (***ISSRs***) and ***Single Nucleotide Polymorphisms*** (***SNPs***) (Weising et al., 2005). Nowadays, DNA-markers give the possibility to work with a virtually infinite number of markers.

2.2.1. Microsatellites as genetic markers

Microsatellite regions constitute useful genetics markers mainly because of their **high level of polymorphism**. Through PCR, which uses the conserved flanking regions on either side of the repeat sequences as primer binding sites, microsatellite regions can be specifically amplified. The flanking regions are critical because they allow the development of **locus-specific** SSR primers. The segments amplified from a single SSR locus present many different alleles due to the high degree of length-polymorphism of the repeat region. An allele at a particular microsatellite locus is defined by the number of repeats there are at the same location (i.e., at that particular locus). The amplified segments can be separated by gel electrophoresis on high resolution agarose gels (**Figure 6.a**), or on polyacrylamide gels (**Figure 6.b**). Automation in SSR analysis can also be achieved if sizing of microsatellite alleles is done on an automatic sequencer. The technique involves labeling of one of the primers of the PCR with a colored fluorescent tag. Separation of the resulting products can be achieved on acrylamide gels

with the help of an automatic DNA sequencer, or by capillary electrophoresis with an automatic genetic analyzer (**Figure 6.c**).

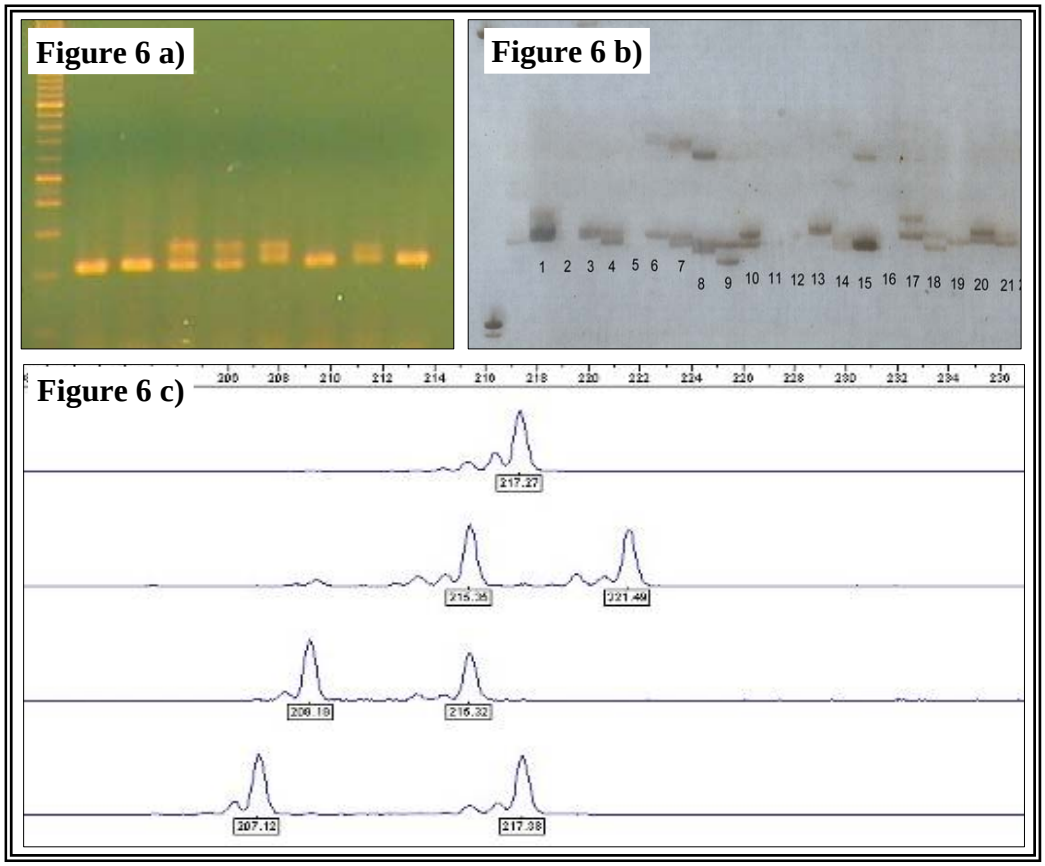

Figure 6: Microsatellite fragments separation **a)** by high resolution agarose gel electrophoresis; **b)** on polyacrylamide gel electrophoresis; and **c)** on the automatic DNA sequencer ABI PRISM® 3100 (Applied Biosystem) by capillary electrophoresis. The figures show different DNA samples isolated from *Prosopis spp*. In the three cases heterozygotes are recognize by the presence of two fragments (bands on the gels, and peaks on the genetic analyzer) of different size. (Photos: M C Mottura)

Besides the abovementioned strategy for SSR detection based on PCR (Tautz, 1989; Weber and May, 1989), other methods have been designed to exploit microsatellite sequences for the study of DNA polymorphism. The first effort made used of in-gel hybridization of digested and electrophoresed genomic DNA with end-labeled oligonucleotides, complementary to microsatellites (Ali et al., 1986). Later, several PCR based approaches were used for developing microsatellite-based markers in a variety of species. The PCR based approach first mentioned, detects SSR polymorphism at individual loci using locus specific primers flanking the microsatellite region (Tautz,

1989; Weber and May, 1989). This technique is also known as ***Sequence Tagged Microsatellite Sites*** (***STMS***), and nowadays is the most popular strategy for SSRs detection. Other PCR approaches detect microsatellite variation by using as primers synthetic oligonucleotides, each complementary to a microsatellite motif randomly distributed throughout the genome. Some of the strategies that use this last mentioned approach are: ***Microsatellite-Primed PCR*** (***MP-PCR***) (Meyer et al., 1993); ***Inter-SSR Amplification*** (***ISA*** or ***ISSR***) (Zietkiewitcz et al., 1994); and ***Random Amplified Microsatellite Polymorphisms*** (***RAMPs***) (Wu et al., 1994).

2.2.2. Advantages of microsatellite (SSR) markers

Microsatellites (SSR) markers have important qualities that make them very desirable and useful molecular markers:

- **Hypervariability:** Most SSRs are highly polymorphic, with extensive allelic variation in repeat number (Litt and Luty, 1989).

- **Locus-specificity**: Refers to amplification of specific alleles, or DNA sequence variants, at the same SSR locus. In SSR markers, the design of long primers (with more than 20 bp) in the flanking region of the microsatellite provides specificity to the marker, and allows the selective amplification of a particular SSR locus in contrast to multi-locus markers such us RAPDs or minisatellite markers (Morgante and Olivieri, 1993; Powell et al., 1996).

- **PCR-based**: Low quantities of template DNA are required (10 to 100 ng per reaction) for sequence tagged SSRs amplifications. Highly degraded or "ancient" DNA can be successfully used as template (Morgante and Olivieri, 1993).

- **Co-dominant Mendelian inheritance**: Heterozygous individuals can be distinguished from homozygous, which is not the case for dominant markers (e.g., RAPDs and AFLPs) (Dorocicz, 1999).

- **Widespread distribution**: SSRs are found in genomes of all species so far examined. They are ubiquitous distributed in the genome and abundant in

both coding and non-coding regions. Many genes are associated with more than one microsatellite (Hokanson et al., 1998; Hancock, 1999).

- **Diversity**: SSRs are based on many different motifs and occur in various functional domains (Hancock, 1999).

- **Versatility**: Microsatellites are very versatile in their application (Rafalski and Tingey, 1993; Gupta et al., 1996; Jarne and Lagoda, 1996). They may be used to detect genetic variability and differentiation in and among populations, identify hybridization between species, determine patterns of pollen and seed dispersal, allow paternity analysis, and evaluate recent genetic history such as population bottlenecks. Useful at a range of scales from individual to fine-scale phylogenies.

- **Across-species transferability**: Reports are available where STMS primers designed for a particular species could be successfully utilized for a study involving related species. Successful cross-species amplification of SSR primers has been widely described in many of the species studied so far (Kijas et al., 1995; Peakall et al., 1998; Pierantoni et al., 2004; Pandey et al., 2004). However, some problems, such us null-alleles (i.e., alleles where locus specific primers give no PCR products), may appear when microsatellite primers are transferred to other species.

2.2.3. Limitations of microsatellite markers

Despite the numerous advantages of SSR markers, drawbacks exist that limit the utility of them. Some of the problems that present working with sequence tagged microsatellite markers are:

- **Development**: The main disadvantage of STMS technique is the great amount of work needed for the development of the markers (see section **2.3.**). SSR primer design is practically complex, work intensive, time-consuming, and expensive. Isolation of microsatellites is technically demanding and often resulting in low efficiency of microsatellite discovery (Zane et al., 2002).

- **Homoplasy**: Two alleles are homoplasic when they are identical in state but not identical by descent. In SSR markers, microsatellite variation is revealed by means of electrophoresis of PCR products. Allelic classes differ by the length (bp) of the amplified fragments. Two PCR products of the same length may not be copies without mutation of the same ancestral sequence, introducing the possibility of size homoplasy. A substantial amount of size homoplasy is expected at most SSR loci (Estoup and Cournet, 1999). Homoplasy at a particular microsatellite locus may also be detected by looking for variation in the flanking regions, such variation being commonly observed among species and occasionally within species (Curtu et al., 2004). In population studies, homoplasy can lead to underestimates of divergence.

- **Null-alleles**: They refer to alleles where locus specific primers give no PCR products. Null alleles are generally attributed to mutations within the binding site for a DNA primer, preventing the binding and leading to the loss of PCR product. Such mutations are more frequently observed among species and occasionally within species (Nascimento et al., 2005), limiting sometimes, the across-species transferability of the SSR primers (Vornam et al., 2004). Null alleles have been reported in many plant species (Gupta and Varshney, 2000). In wheat, in two different studies, 13% and 10% of SSR loci carried each, up to 25% of null alleles (Plaschke et al., 1995; Prasad et al., 2000). The presence of null alleles will lead to an underestimate of heterozygosity due to the fact that heterozygotes may be misclassified as homozygous when a null allele occurs. This misclassification due to null alleles may also cause problems in other applications of SSR markers such as parentage analysis.

- **Mutation model**: Several theoretical models describe the evolutionary dynamics of SSRs (section **2.1.2.**). However, the underlying mutation model controlling SSR mutation is usually unknown. Which theoretical model should be applied is a main question because population genetic inferences are sensitive and dependent on the assumed model.

- **Non-random distribution of SSRs**: The frequency distribution of SSRs with different motifs varies by functional domain (Morgante et al., 2002). Triplet motifs are more common within coding regions. Besides, different species have different motif frequency distribution.

- **Neutrality**: SSR alleles are not always adaptively neutral even when they are commonly analyzed under the assumption that allele frequencies are determined solely by mutational processes and genetic drift. As it was presented before (section **2.1.3.**), the possibility of adaptively relevant functions of SSR should be explicitly recognized and tested.

- **Stutter bands**: During microsatellite PCR, insertion-deletion mutations produce stutter products differing from the original template by multiples of the repeat unit length. These "stutter" or "shadow" bands that arise during PCR may complicate accurate scoring of SSR polymorphism (Viguera et al., 2001; Shinde et al., 2003).

2.3. Strategies for microsatellite markers development

In recent years, a variety of molecular markers based on microsatellites, especially based on STMS technique have become the markers of choice given their large applicability. The first microsatellite markers were developed in humans. Since then, there has been an extraordinary increase of interest in SSR markers. The popularity of these markers is demonstrated by the growing number of reports describing the isolation of SSR markers in many organisms. A clear example of this can be observed in the journal *Molecular Ecology Notes*. This journal registered during January-March 1999 a total of 20 articles reporting the isolation of microsatellites. Two years later, during the same period, *Molecular Ecology Notes* published almost 100 articles on microsatellite isolation (Zane et al., 2002).

The major disadvantage of SSR markers is that they need to be isolated de *novo* from most species that are being examined for the first time for the designing of the primers matching the conserved flanking regions. As it was previously mentioned, this isolation can be time consuming, expensive, practically complex and work intensive. Nevertheless, nowadays several strategies are available for microsatellite isolation. One of the methods used for the development of STMS primers is searching for microsatellites in the current DNA sequence

databases. The presence of SSR arrays in many plant genomes can be searched from the available DNA sequence databases like EMBL or GenBank™, using an appropriate computer program. Nowadays, due to the growing number of genome projects in many different species, partial or total genome sequences are available for different organisms (e.g., humans, *Drosophila melanogaster*, maize, potato, rice, and wheat). In 2004, the International Populus Genome Consortium finished sequencing the complete genome of *Populus*. This represents the first full draft genome sequence of a forest tree species. In addition, the Eucalyptus Genome Initiative is working to sequence and characterize the *Eucalyptus* tree genome.

The traditional method for microsatellite development has been isolation from partial genomic libraries of the species of interest by screening clones through colony hybridization with repeat-containing probes (Rassmann et al., 1991). In this method, fragmented DNA is size selected, ligated into a vector and transformed into bacteria to generate a partial genomic library. Subsequently screening for positive clones (clones carrying SSR motifs) is generally carried out by means of Southern hybridization using repeat-containing probes. Hybridization probe(s) can be labeled by both, radioactive (^{32}P, ^{33}P) and nonradioactive (digoxigenin) methods. Although this approach is relatively simple, this method can be extremely tedious and with very low efficiency of SSR isolation. According to Zane et al. (2002) the average percentage of positive clones obtained from traditional isolation protocols ranged from 0.4% to 3.1%.

To avoid library construction, some authors proposed methods for SSR isolation based on the observed abundance of repeat regions in RAPD amplicons. Isolation of microsatellite regions is achieved by means of Southern hybridization of RAPD profiles with repeat-containing probes, followed by selective cloning of positive bands (Ender et al., 1996). Other RAPD-based method proposed the cloning of all RAPD products and posterior screening of arrayed clones (Lunt et al., 1999). Other nonlibrary PCR-based methods rely on the use of repeat-anchored primers to isolate and then sequence one (Fisher et al., 1996) or both (Lench et al., 1996; Cooper et al., 1997) flanking regions.

STMS markers can also be developed using enriched genomic libraries. The establishment of microsatellite-enriched libraries is a fast and low-cost method for microsatellite marker development. The genomic library can be enriched for one or more different microsatellite motifs. The efficiency of protocols for SSR isolation based on enriched libraries is estimated

as the number of clones that contains microsatellites motifs. Reported efficiencies range from 10% to 95% (Cordeiro et al., 1999; Jakse and Javornik, 2001). Several enrichment protocols have been reported on the basis of:

1) Capture by streptavidin-coated magnetic beads (e.g.; Kijas et al., 1994; Fischer and Bachman, 1998).

2) Microsatellite probes attached to small nylon membranes (e.g., Karagyzov et al., 1993; Edwards et al., 1996).

3) Other uncommon used procedures such as magnetic capture of phagemid DNA (Paetkau, 1999).

One of the latest methods for SSR isolation is the Fast Isolation by AFLP of Sequences Containing Repeats (FIASCO protocol) proposed by Zane et al. (2002). This method is fast and simple, and relies on the efficient digestion-ligation reaction of the amplified fragment length polymorphism (AFLP) procedure. The FIASCO protocol is an enrichment protocol based on the ability to recover microsatellite DNA by PCR amplification, after selective hybridization. Zane et al. (2002) reported 50% to 90% of clones containing dinucleotide SSR using this protocol.

2.4. Application of microsatellite markers

2.4.1 Application in genotyping, diversity assessment, and breeding

In the last few years microsatellites have become one of the most popular molecular markers used with applications in many different fields. Microsatellites represent single loci molecular markers that combine extensive **hypervariability** with **somatic stability** and **co-dominant Mendelian inheritance**. Besides, SSR regions are **abundant** and **ubiquitously distributed** throughout the genome. Since Tautz (1989), and Weber and May (1989) published the first reports on a PCR based strategy for the detection of microsatellite regions and their use as genetic markers, many reports have been published on the application of SSR markers in animal and plant species. In forest tree species specifically, microsatellites are used for a wide range of applications. The first SSR markers developed for a forest tree species were in *Pinus radiata* (Smith and

Devey, 1994). Some of the main applications of SSR markers are listed below with examples reported in forest tree species:

- **Genome mapping and characterization of QTLs**: STMS analysis has been used for genome mapping of microsatellites and the construction of microsatellite linkage maps for the genomes of different species. Among forest tree species, genome maps including SSR markers have been constructed for some *Pinus* species (Devey et al., 1996; Devey et al., 1999; Echt and Nelson, 1997), for *Eucalyptus grandis* x *E. urophylla* (Brondani et al., 1998), and for *Quercus robur* (Barreneche et al., 1998), among others. In tree species, many linkage maps have been developed to locate quantitative trait loci (QTLs), and to provide a basis for marker-aided selection (MAS). Selecting economically important trees for breeding programs can be a long and costly process when selection is based upon physical traits. MAS programs have the potential to provide a rapid, reliable and effective selection many years earlier in the growth cycle of commercial tree species. The application of SSR markers in MAS programs for the generation of linkage maps, and for the characterization of QTLs, is in an advanced stage especially in *Eucalyptus*, *Pinus* and *Quercus* species (Groover et al., 1994; Barreneche et al., 1998; Scalfi et al., 2004; Isoda et al., 2006; Marcucci Poltri, 2006).

- **Fingerprinting (Genotyping)**: The hypervariable nature of microsatellites increases the probability that every individual in a population will have a unique genotype, making microsatellites particularly useful for fingerprinting. Multilocus genotyping systems have been developed for *Eucalyptus* species (Kirst et al., 1999; Kirst and Grattapaglia, 1999). In *Pinus*, fingerprinting has been used for clone identification as an important component of breeding programs (Butcher et al., 1999).

- **Studies of population structure and variation**: SSR markers constitute an especially valuable molecular tool for studies on genetic variation within and among populations, for the construction of genetic inventories, for programs of conservation of genetic resources, or for the study of the spatial genetic structure of populations. A wide range of studies have been

reported on population structure and variation based on SSR markers in many different forest tree species. Some examples are *Eucalyptus nitens* (Byrne et al., 1996), *Pseudotsuge menziesii* (Viard et al., 2001), *Carapa guianensis* (Dayanandan et al., 1999); *Shorea curtisii* (Ujino et al., 1998), *Populus tremuloides* (Wyman et al., 2003), *Pinus strobus* (Marquardt and Epperson, 2004), and *Fagus sylvatica* L. (Vornam et al., 2004).

- **Phylogenetic studies**: Microsatellites are useful for fine-scale phylogenies up to the level of closely related species. They also provide data suitable for phylogeographic studies that seek to explain the concordant biogeographic and genetic histories of the floras. However, there are few phylogenetic studies that use microsatellite markers and most of these studies are based on chloroplast or mitochondrial SSRs. Application of SSR markers in phylogenetic studies have been reported in forest species such as *Pinus* (Gugerli et al., 2001; Karhu, 2001; Soranzo et al., 1999), *Prunus* (Xu et al., 2004), and *Dipterocarpaceae* (Indrioko et al., 2006).

2.4.2. Applications to studies of the reproduction system

Studies on mating systems and gene flow have benefited from the development of biochemical markers, such as isozymes in the 1970s. However, with the development of hypervariable DNA markers, studies of mating systems became more precise. Microsatellite markers present important qualities that make them very desirable and useful molecular markers for determining mating system and gene flow. In forest trees, the first microsatellite markers developed were in *Pinus radiata* D. Don (Smith and Devey, 1994). They have since been developed from the nuclear genomes of a wide range of temperate and tropical forest trees (Butcher et al., 1999).

SSR markers are now the most important molecular tool for genetic studies for many forest tree species. They serve an important role to estimate genetically effective pollen movement among plants, to assess gene flow within and between populations, to study pollen and/or seed dispersal, to estimate mating system parameters (e.g., outcrossing rate, correlated mating, and biparental inbreeding), and for parentage analysis, among other applications. Next, examples of such applications of SSR markers in forest trees are presented.

There are several studies where microsatellite markers were applied to estimate mating system parameters such as outcrossing and selfing rates, correlated mating, biparental inbreeding, or effective number of fathers in forest tree species. Butcher et al. (1999) reported the used of SSR markers for the estimation of outcrossing rates in natural and breeding populations of *Acacia mangium* in New Guinea. Collevatti et al. (2001) investigated the mating system of populations of the endangered tropical tree species *Caryocar brasiliense* using genetic data from ten microsatellite loci to estimate mating system parameters under the mixed mating model. In other tropical tree species from Brazil, *Theobroma grandiflorum*, Alves et al. (2003) studied the mating system using eight microsatellite loci. Estimation of mating system parameters based on microsatellite markers were also reported in *Prunus mahaleb* L (García et al., 2005), in Valley oak (*Quercus lobata* Neé) (Sork et al., 2002), and in *Pinus pinaster* Ait. (González-Martínez et al., 2003), among others.

Microsatellite markers can provide insights into the genetic structure of natural populations and gene flow in species with little or no isozyme variation. Echt et al. (1998) were able to detect variation among populations of *Pinus resinosa* Ait. using chloroplast microsatellite markers. This forest tree species showed little morphological variation, no isozyme variation, and very limited RAPD variation. Dick et al. (2003) studied the outcrossing rates and pollen dispersal of the tropical tree species *Dinizia excelsa* using five microsatellite loci and the TwoGener approach, a novel two-generation (parent-offspring) method for the estimation of pollen movement proposed by Smouse et al., (2001). Chase et al. (1996) examined the impacts of forest fragmentation on genetic diversity and gene flow of *Pithecellobium elegans* using SSR markers.

Parentage analysis allows the detection of the parent or parental couple that could have sired individual seeds or seedlings. It includes the analysis of paternity (i.e. identification of the father) and maternity (i.e. detection of the mother). Hence, the discriminating power of parental analysis approach provides a direct and realistic estimate of gene movement. The high variability of SSR markers increases the probability that every individual in a population will have a unique genotype, making microsatellites particularly useful for fingerprinting and therefore, for parentage analysis. There is an extensive list of studies where microsatellite markers are used

successfully in parentage analyses in order to assess gene flow, pollen dynamics or seed dispersal in forest tree species. Besides, new studies appear each year. Godoy and Jordano (2001) studied seed dispersal by animals in *Prunus mahaleb* based on multilocus paternity analysis at SSR markers. Streiff et al. (1999) studied the pollination dynamics in a mixed oak stand comprising *Quercus robur* L. and *Quercus petraea* (Matt.) Leibl. using six microsatellite primers and paternity assignment. In *Quercus*, as well, Dow and Ashley (1996 and 1998) assessed seed dispersal and pollen flow by means of SSRs and parentage analyses. Microsatellites have also been used successfully in paternity analyses in the tropical species *Gliricidia* to estimate pollen dispersal (Dawson et al., 1997). Tabbener and Cottrell (2003) reported the use of microsatellite markers to assess the incidence of natural hybridization and to provide accurate information regarding the distance traveled by pollen in the genus *Populus*. In their study, they used seven SSR primers to genotype a group of trees belonging to different species of the genus *Populus* and their progenies. Then, by means of paternity analysis they estimate the cross between species, degree of introgression, and the distance over which pollen can travel.

3. THE STUDIED TAXA

3.1. The genus *Prosopis*

3.1.1. Systematics and species classification

The genus *Prosopis* Linnaeus emend. Burkart belongs to the family Fabaceae (Leguminosae), sub-family Mimosoideae. The placing of *Prosopis* in the wider taxonomic classification system is given below (**Table 1**), based on Elias (1981), and Lewis and Elias (1981) (Pasiecznik et al., 2001).

The history of taxonomic confusion within the genus was largely settled with the authoritative monograph of Arturo Burkart (Burkart, 1976), who defined the generic limits and divided the genus in five sections, based on floral characteristics, each also with marked vegetative differences in armature (Pasiecznik et al., 2001). While the generic limits and division into sections defined by Burkart (1976) are generally

accepted, there is continuing debate as to the relative rank of species that he defined. Since the classification proposed by Burkart (1976) is the most recent and complete taxonomical classification of the genus *Prosopis*, it is taken as the basis of *Prosopis* taxonomy in this work.

Table 1: Taxonomic classification of the genus *Prosopis*.

Family:	Fabaceae	around 650 genera and 18000 species
Sub-family:	Mimosoideae	60 to 70 genera, 650 to 725 species
Tribe:	Mimoseae	38 accepted genera
Group:	Prosopis	
Genus:	*Prosopis*	44 species (according to Burkart, 1976)

In his monograph, Burkart (1976) describes ***44 species*** and 27 varieties of *Prosopis* (see **Appendix 1**). The existence of populations of *Prosopis* with distinct characteristics led Burkart (1976) to describe many separate species or varieties, even though several are known to hybridize (Pasiecznik et al., 2001). Based on floral characteristics and vegetative differences, Burkart (1976) divided the genus *Prosopis* into ***five sections*** and defined within two of them ***series*** (see **Appendix 1**).

According to the generic description given by Burkart (1976), the *Prosopis* species are all trees or shrubs, rarely sub-shrubs, predominantly xerophilous, aculeate, spiny or rarely unarmed. Leaves are bipinnate, often with few pairs of opposite pinnae. Leaflets are small, rarely large, numerous, mostly opposite, linear, oblong, fusiform, entire and of the same color on both sides. Shoots are dimorphic, with long megablast and brachyblasts or short shoots emerge from multiple axillary bud, from which develop the cauline spines when extant, leaf fascicles and racemes. Flowers are small, actinomorphic, pentamerous, hermaphroditic, and grouped in racemes. Racemes are spike-like, amentiform, axillary, mostly densiflorous, but sometimes globose heads (Burkart, 1976). The fruit is a modified, indehiscent, fleshy legume called a "drupaceous loment" (Burkart, 1976). These are linear, straight, falcate, annular to spirally coiled with fleshy mesocarp. The mesocarp is sugary or fibrous, containing

endocarps, divided into one-seeded coriaceous to bony segments. Seeds are ovoid, compressed, hard, brown, with a mucilaginous endosperm surrounding the embryo.

The flowers of *Prosopis* are pollinated by insects. Seeds are endozoochorous dispersed and germinate freely after passing through the digestive tracts of cattle, sheep, goats, or wild animals. The nutritive legumes, with seeds resistant to the digestive juices of herbivores, demonstrate a superior adaptation of the species of *Prosopis* in their struggle for survival and adaptation (Burkart, 1976). *Prosopis* species are protogynous (Burkart, 1976) and were believed to be obligate outcrosser (Simpson, 1977; Balboa and Parraguez, 1986). However, more recent researches showed levels of outcrossing ranging from 0.6 to 1 in different species of this genus (Keys and Smith, 1994; Bessega et al., 2000a).

Most *Prosopis* species studied so far are diploid, with a somatic number of ***2n = 28***. Diploid and tetraploid (***2n = 56***; ***x = 14***) ecotypes have been described for *P. juliflora* in Haiti, Aruba, Colombia and Venezuela (Galera, 2000). Chromosomes in *Prosopis* are small and uniform, without special distinctive marks. Meiosis seems to be normal even in hybrid populations (Burkart, 1976).

3.1.2. Distribution and origin

The natural distribution of the genus *Prosopis* includes arid and semi-arid regions of Africa, Asia and the Americas (**Figure 7** and **Table 2**). *Prosopis africana* is native to Tropical Africa, in the Soudano-Guinean zone and neighboring areas of Africa, from Senegal in the west to Sudan and Kenya in the east (**Figure 7** and **Table 2**). In Asia there are three native species, the three species included in the section Prosopis (see **Appendix 1**). These species are original to the Middle East, stretching east to India, north to Georgia and Turkmenistan, reaching also the North African coast to west Algeria (**Figure 7** and **Table 2**). The greater part of the species (40) is native to the Americas, and can be approximately divided into three geographic areas (**Figure 7** and **Table 2**). There is the northern Texan/Mexican center, the Argentinean/Paraguayan center including all neighboring countries, and an intermediate Tropical Andean Region which includes the Caribbean and Central America to the extreme north of Chile and northwest of Argentina (Burkart, 1976).

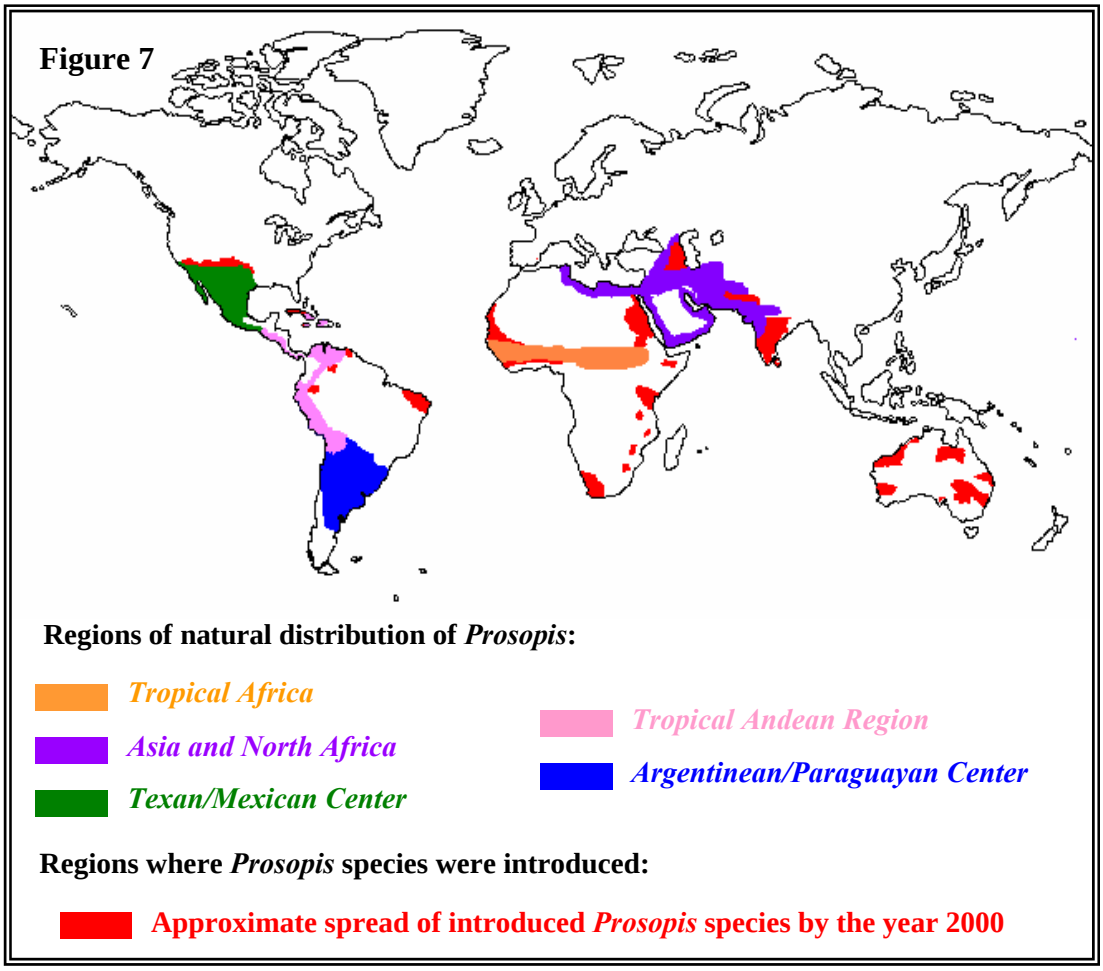

Figure 7: Approximate native and present distribution of the genus *Prosopis* in the world (Pasiecznik et al., 2001). The species occurring in each region of natural distribution are listed in **Table 2** in colors in correspondence with the colors of the regions in this figure.

Prosopis trees have been also introduced widely by man (**Figure 7**). This is because of the perceived value of the trees' products, the multiple products obtained, the high yields under even the poorest conditions, tolerance to drought and poor soils, and their values in agroforestry systems (Pasiecznik et al., 2001). The first records of *Prosopis* introduction are those to West Africa and Pacific islands around the 1820s, to India and Pakistan in the 1870s, and to Australia and South Africa before 1900. There have been, however, many other unrecorded introductions before (Pasiecznik et al., 2004), for example, several species from the Americas with useful attributes may have been introduced to neighboring regions following the arrival of man in prehistory. The majority of intentional introduction reported of *Prosopis* were of four species: *P.*

juliflora, *P. pallida*, *P. glandulosa*, and *P. velutina*. Besides, species such as *P. alba* and *P. chilensis* have proved to be well adapted and are locally common in some regions. Some examples of intentional introductions of *Prosopis* are the introduction of *P. pallida* and *P. juliflora* in Australia, Hawaii, Ethiopia and South Africa; *P. juliflora* from Mexico to India and Pakistan; and *P. juliflora* and *P. glandulosa* into the Middle East.

Table 2: Natural geographic distribution of the genus *Prosopis* (adapted from Burkart, 1976).

Region	Total nº of species	Species
Tropical Africa	1	***P. africana***
Asia and North Africa	3	***P. cineraria*; *P. farcta*; *P. koelziana*.**
Texan/Mexican Center	8	***P. pubescens*; *P. palmeri*; *P. articulata*; *P. tamaulipana*; *P. juliflora*; *P. laevigata*; *P. glandulosa*; *P. velutina*.**
Tropical Andean Region	6	***P. burkartii*; *P. ferox*; *P. tamarugo*; *P. pallida*; *P. chilensis*; *P. juliflora*.**
Argentinean/Paraguayan Center	28	***P. strombulifera*; *P. reptans*; *P. abbreviata*; *P. torquata*; *P. sericantha*; *P. kuntzei*; *P. ruscifolia*; *P. fiebrigii*; *P. vinalillo*; *P. hassleri*; *P. humilis*; *P. rojasiana*; *P. rubriflora*; *P. campestris*; *P. affiants*; *P. elata*; *P. chilensis*; *P. nigra*; *P. caldenia*; *P. flexuosa*; *P. alpataco*; *P. alba*; *P. pugionata*; *P. argentina*; *P. denudans*; *P. ruizleali*; *P. castellanosii*; *P. calingastana*.**

Prosopis species and their hybrids have been reported also as invasive weeds becoming a problem of major importance in many regions worldwide. There are reports from almost every country where they are introduced, of exotic *Prosopis* invading agricultural and pasture land, nature reserves, watercourses, roadsides and wasteland. This is the case of Australia, where *P. pallida*, *P. glandulosa* and *P. velutina* have been recognized as a weed of national significance due to their invasiveness and potential negative impacts. However, the most studied weedy invasions have occurred within the native ranges of *Prosopis* species. *P. glandulosa* in the USA, particularly Texas, and *P. ruscifolia* in northern Argentina and the Chaco region of Paraguay, have dramatically

spread since 1850. Other species have also spread in these regions to a lesser extent, including *P. velutina* in the USA and Mexico, and *P. campestris*, *P. hassleri*, *P. humilis* and *P. kuntzey* in northern Argentina and Paraguay (Pasiecznik et al., 2001).

The genus *Prosopis* is a rather primitive member within the Mimosoideae (Burkart, 1976). This theory is supported on the fact that *Prosopis* species maintain 14 as the base chromosome number (i.e., ***2n = 28***), and pollen is released in single grains and not two, four or more (Burkart and Simpson, 1977). According to Burkart (1976), the genus *Prosopis* possibly originated in tropical Africa, where only *P. africana*, the least specialized species, persists. From this or a similar ancestral stock may have evolved the more specialized groups. The prickly Asian species of section Prosopis developed in the eastern, arid desert zones. The spiny American ones developed in the western hemisphere into two differentiated groups: section Strombocarpa and section Algarobia, both also with desert climate preferences. In-between the two appears the small section Monilicarpa, with only one member, *P. argentina*.

The accepted centre of polymorphism for *Prosopis* is Argentina, with approximately 28 species, 13 of them endemic. The Texan-Mexican region is considered a secondary centre of polymorphism with eight species. Both centers of polymorphism present closely related species.

3.1.3. Uses of *Prosopis* spp.

In many different regions of the world the species of the genus *Prosopis* are considered valuable trees with multiple uses. In the Americas, where the majority of the *Prosopis* species grow naturally, the aborigines knew and exploited the benefits of *Prosopis* as multipurpose species long before the colonizers arrived from the Old World. *Prosopis* species do not only provide wood, but also a wide range of non-wood forest products.

The wood is probably the most important natural resource from *Prosopis* species for use either for structural purpose or as a fuel (Pasiecznik et al., 2001). As timber wood it can be used as poles or roundwood, or cut into boards and cants. The heartwood of *Prosopis* species is strong and durable, hard and heavy. The specific gravity of dry wood of different species ranges from 0.65 to 1.35 g/cm^3 (Alden, 1995). The wood has a relatively high dimensional stability over other timbers, meaning less shrinkage and

cracking, and abundant wood elements give a high tensile strength (Tortorelli, 1956; Weldon, 1986). Although the wood is hard and heavy, it generally works well and takes a good finish. The heartwood is dark red to dark brown in color. As a fuel wood it can be burnt directly or made into charcoal. *Prosopis* species produce a wood which is a very high quality fuel, having a high calorific value of approximately 5000 kcal/kg (Pasiecznik et al., 2001). Although the wood burns better when dry, a great advantage over some other species is an ability to burn well when freshly cut or "green".

The fruit produced by *Prosopis* species are legume pods which have been a historic source of food for human populations where *Prosopis* species are found. However, during the last few centuries pods became less important as a human food and more important as a livestock feed. Pods are high in sugars, carbohydrates and protein. For example, pods from species of section Algarobia contain between 7 to 22 % protein, and 30 to 75% carbohydrates. *Prosopis* species produce fruit every year and can be termed an "unfailing crop" (Simpson, 1977). They produce a greater yield of pods in years of below average rainfall. This characteristic makes them very suitable as source of food and/or fodder. *Prosopis* fruits help sustain human life by providing a valuable source of food for animal species in most arid and semi-arid zones where the production of other food crops, whether wild of cultivated, is highly dependent on the rainfall. Pods can also be processed into human food. Some products for human feeding that can be obtained from processing of pods are: pulp flour, syrup, protein-enriched flour, alcohol by fermentation, additive for dietetic foods, gum, protein concentrate, and coffee substitute.

Leaves of *Prosopis* species are often used as food for animals. The use of leaves varies widely between *Prosopis* species and between the different animals that may consume those (Pasiecznik et al., 2001). A minority of species are known to have leaves which are palatable to livestock. All the Afro-Asiatic species have foliage that is readily consumed by all livestock and are considered main fodder species.

Another very important *Prosopis* resource is the flowers. *Prosopis* flowers are a valuable source of bee forage. The flowers produce copious quantities of pollen and nectar over a relatively long period of time. *Prosopis* honey is light yellow in color, and generally of good quality with a pleasant taste and only a slight aroma.

Exudate gums, tannins, fibers, dyes, and medicines are also potential uses of *Prosopis* species. Exudate gum is produced from natural wounds in the bark of plants as a defense mechanism, but can be stimulated by artificial wounding. *Prosopis* gums are water soluble, liquid, and of a very high quality, similar to the gums produced from *Acacia* species. Tannins can be extracted from bark, wood, and fruit. The tanning content of various plant parts from different *Prosopis* species is 6 to 20% (Pasiecznik et al., 2001). Fibers can be obtained mainly from the inner bark or roots with bark removed. *Prosopis* fibers have been used to make strong ropes or employed in basketry (Felger, 1977). Bark and gums can be used to produced paints, dyes, cosmetics, and hair cleanser (Felger, 1977).

Many medicinal uses have been recorded for extracts from *Prosopis* plant parts. The medical properties of *Prosopis* have been and are still exploited since centuries by humans in areas of the entire native range of the genus. Products used with medicinal purposes are leaf and bark extracts, leaf smoke, flower extracts, gums, buds preparations, and infusions from most plant parts. Examples of ailments treated with these products are mouth and throat infections, parasites, urinary disorders, asthma, liver stones, and venereal disease. Chemical compounds such as alkaloids have been isolated, which contribute to the medical properties of *Prosopis*.

From an environmental and ecological point of view *Prosopis* species constitute a highly valuable resource in many arid and semi arid regions of the world where they grow naturally. In their native range, and occasionally where introduced, *Prosopis* are often the only, or the dominant, tree species of the ecosystem. They can grow on the poorest soils, in hot, arid climates where few other tree species are found. They give shade and shelter for humans and livestock, as well as by a wide range of other animals and insects. Besides, shade and shelter positively affect water balance and the microclimate under the tree canopy. They provide shelter from the wind, reducing soil movement. Soil and sand are fixed by roots giving *Prosopis* an important role in erosion control. *Prosopis* trees also have ameliorating effects on surrounding soil, increasing soil fertility and decreasing salinity and alkalinity. They have the ability to improve soils via biological nitrogen fixation, leaf litter addition and incorporation, nutrient pumping, change in soil structure and in soil fauna and microbial populations. *Prosopis* role in recovering otherwise unproductive sites has been noted (Pasiecznik et al., 2001).

3.2. The studied species *Prosopis chilensis* and *P. flexuosa*

3.2.1. Description

Prosopis chilensis (Molina) Stuntz and *P. flexuosa* DC. are closely related hardwood arboreal species belonging to the Series Chilenses, Section Algarobia of the genus *Prosopis* (see: **Appendix 1**). In Argentina and other areas of natural distribution of these tree species in South America, the most common name given to them is "***algarrobo***". This name was given by the invading Spaniards because of the observed similarity between *Prosopis* trees and the carob (*Ceratonia siliqua*) of the Mediterranean basin, both with similar, sweet edible pods (D'Antoni and Solbrig, 1977; Cruz, 1999). In Argentina *P. chilensis* is commonly named as "***algarrobo blanco***" (white algarrobo), and *P. flexuosa* as "***algarrobo negro***" (black algarrobo).

According to the description given by Burkart (1976), *P. chilensis* and *P. flexuosa* are trees, long-living, 3 to 10 meters tall. The tree top is rounded or hemispherical; branches flexuous and knotty. **Spines** axillary, germinate, uninodal. **Leaves** deciduous, uni- to trijugate. **Leaflets** linears, glabrous or scarcely puberulous, 10 to 29 pairs per pinna. **Flowers** complete, greenish-white to yellowish, grouped in racemes. **Racemes** spiciform, densiflorous (ca. 200 to 250 flowers each), axillary. **Fruit** indehiscent legume, straw-yellow, tinged with violet to nearly black-violet. **Roots** deep. Both species are diploid, with a somatic chromosome number of ***2n = 28***. The main morphological differences between *P. chilensis* and *P. flexuosa* are shown in **Table 3** (see also: **Figures 8** and **9**).

Table 3: Main morphological differences between *P. chilensis* and *P. flexuosa* according to Palacios (1988).

	P. chilensis	***P. flexuosa***
Spines	on strong shoots, up to 6 cm long	commonly strong or absent but sometimes on strong shoots, 3 - 4 cm long
Pinnae	8 – 24.5 cm long	5 – 13 cm long
Leaflets	11 – 54 mm long X 1.1 – 3 mm broad	4 – 15 mm long X 1 – 2 mm broad
Fruits	straw yellow, with parallel margins, compressed	straw yellow tinged with violet to nearly black violet, margins undulate (submoniliform), subcompressed

Figure 8.a)

Figure 8.b)

Figure 8: Photographs of adult trees of **a)** *Prosopis chilensis*, and **b)** *Prosopis flexuosa*. (Photos: Aníbal R. Verga)

P. chilensis and *P. flexuosa* are hermaphroditic species with complete flowers, i.e., both male and female gametes are produced in the same flower. Flowers are produced in masses, mainly on spike-like racemes. The flowers of these species are insect-pollinated, they attract a large number of potential pollinators with the production of copious amounts of pollen. A large number and variety of insects have been observed visiting *P. chilensis* and *P. flexuosa* flowers, but bees are thought to be the most important group of pollinators in these species. Early studies on species of *Prosopis* belonging to Section Algarobia pointed that these species are functional self-incompatible due to the fact that they are protogynous (Burkart, 1952, 1976; Solbrig and Cantilo, 1975; Simpson, 1977; Palacios and Bravo, 1981). This "incompatibility" would

prevent self-fertilization making these species obligate outcrosser (Simpson, 1977; Balboa and Parraguez, 1986). However, more recent researches showed levels of outcrossing ranging from 0.6 to 1 in different species of this Section (Keys and Smith, 1994; Bessega et al., 2000a). The seed distribution of these species is endozoochorous by animals.

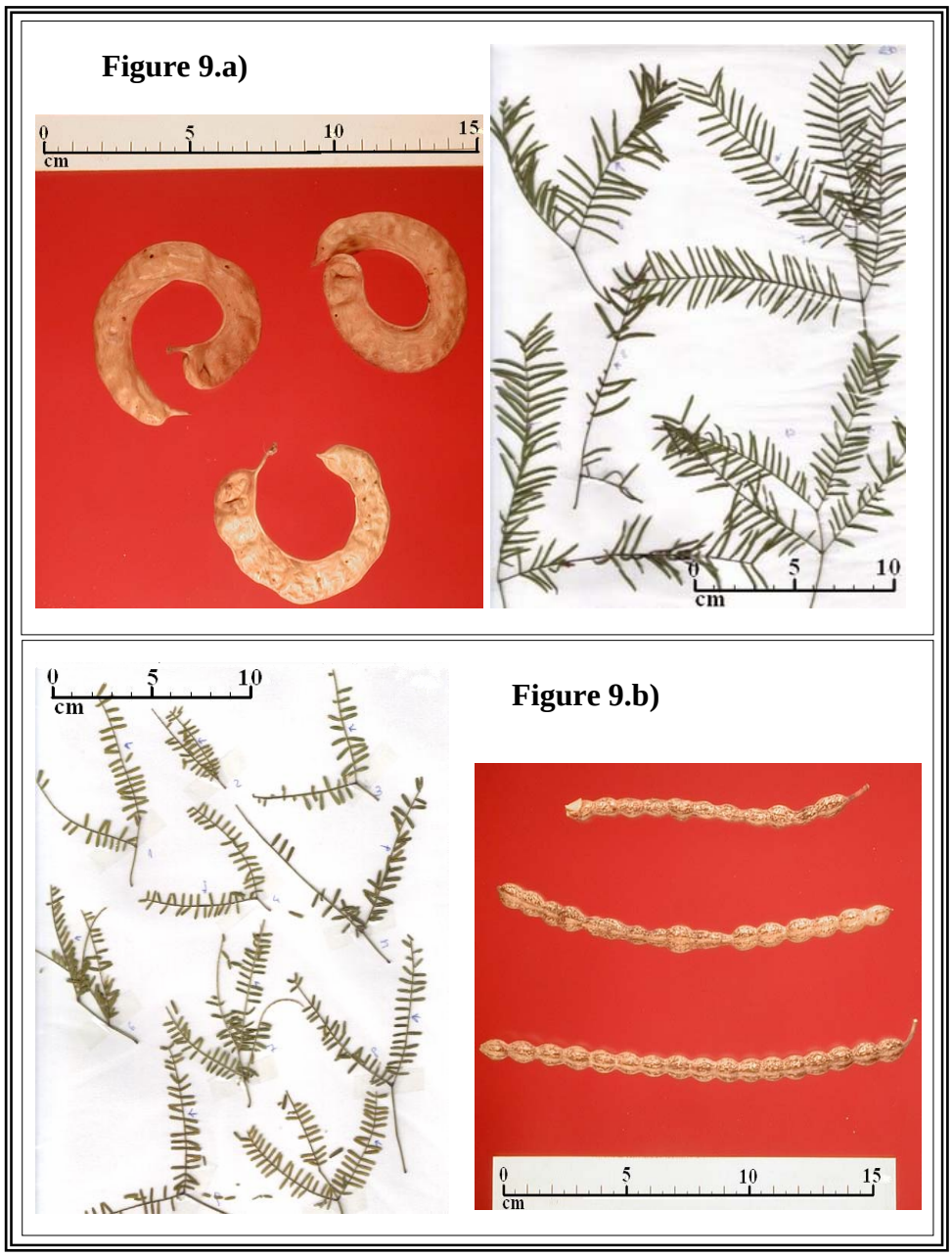

Figure 9: Fruits and leaves of **a)** *Prosopis chilensis*, and **b)** *Prosopis flexuosa*. (Photos: Aníbal R. Verga)

3.2.2. Distribution and ecology

Both species are native to South America. *P. chilensis* is a variable species; its natural area extends from Peru and Bolivia to the parallel 34º S approximately. In Chile this species is commonly found in the North and Central areas in the vicinity of Santiago de Chile. The native range of *P. chilensis* in Argentina is significantly wide, including the phytogeographic provinces of the Monte and Espinal (northern area of these regions), and the driest area of the Chaco (**Figure 10.a**). Its distribution in Argentina includes part of the provinces of Salta, Tucumán, Catamarca, La Rioja, San Juan, Mendoza, San Luis, and Córdoba.

P. flexuosa is naturally found in northern Chile (Atacama and Coquimbo), and is very common in the arid regions of central-western Argentina, from Tucumán and Catamarca, through La Rioja, western Córdoba, San Juan, San Luis, and Mendoza, to La Pampa, Río Negro and western Buenos Aires provinces (Burkart, 1976). Its native range in Argentina includes the complete phytogeographic province of the Monte, and part of the Chaco and Espinal (**Figure 10.b**).

P. chilensis and *P. flexuosa* thrive in a wide range of rainfall zones, from 50 mm mean annual rainfall (m.a.r.), to 500 mm m.a.r. The existence of two root systems, a deep tap root (from 10 up to 20 meters depth) to reach ground water, and a mat of surface lateral roots to make use of infrequent rainfall events, places these species as phreatophytes (Mooney et al., 1977). In arid regions with less than 300 mm m.a.r., both species are obligatory phreatophytes (Galera, 2000). Both species develop in a wide variety of soil types, preferring deep and free draining soils due to the fact that they have no tolerance to flooding. They are able to tolerate saline and alkaline soils. Mean annual air temperature where *P. chilensis* is found, is above 20º C, with an upper limit of 48º C (Galera, 2000). *P. chilensis* is frost sensitive, but tolerating occasionally, minimum temperatures of -3 to -7º C for a short period. *P. flexuosa* can thrives in areas of lower temperatures than *P. chilensis* and exhibits higher frost tolerance, reaching areas of lower latitude in the south.

In Argentina, in the phytogeographic province of the Monte, the distribution of *P. chilensis* and *P. flexuosa* is limited to areas where water from aquifers (phreatic stratums) is available. In this region, both species are isolated one from the other due to

the geo-morphological structure of the region, with relatively closed valleys and mountain chains, which generates clear and defined pure-species populations (Verga, 1995).

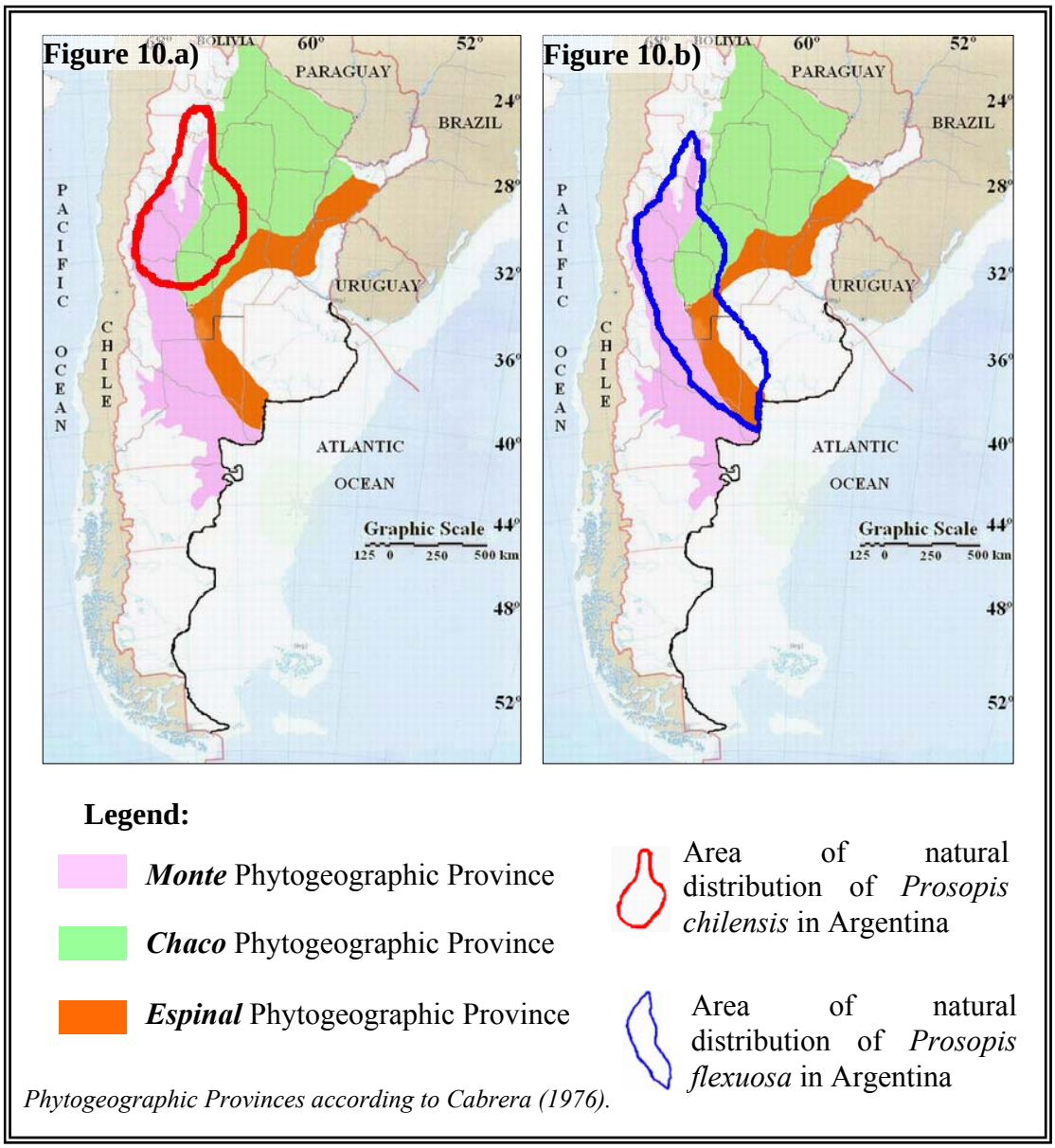

Figure 10: Approximate natural distribution of **a**) *Prosopis chilensis*, and **b**) *Prosopis flexuosa* in Argentina (from Verga, 1995).

In the Chaco phytogeographic province, specifically in the Arid Chaco region, *P. chilensis* and *P. flexuosa* occur in different ecological niches, physically separated. *P. chilensis* thrives along dry watercourses and river beds in the foothills, where there is

extra water contribution supply by ground or superficial water. On the other hand, *P. flexuosa* grows in a wider range, occupying the valleys and flatland areas, where no extra water contribution exists. The relative higher ecological requirements of *P. chilensis* give to their populations a typical structure. In general, *P. chilensis* populations are limited to small isolated areas, and when they grow along watercourses they are linear. In contrast, *P. flexuosa* appears constituting large forests in the flatlands. In the foothills, the borders of the *P. flexuosa* forests come in contact with the *P. chilensis* populations generating hybrid zones (Verga, 1995). Other contact regions of both species are also irrigation areas, as well as areas highly disturbed by man (e.g., areas where native forest was cut down and replaced by agriculture or cattle raising, and later abandoned). In these disturbed areas a secondary forest grows with *Prosopis* species as dominant components. In contact areas where the natural habitat has been disturbed by human impact, interspecific hybrids between *P. chilensis* and *P. flexuosa* can be found together with the pure species.

3.2.3. Genetic variation

The species of the section Algarobia are the most intensively studied genetically among the sections of the genus *Prosopis*. Previous genetic studies on different species of the section Algarobia include studies of genetic variation and population structure (Solbrig and Bawa, 1975; Saidman, 1985, 1986; Saidman and Vilardi, 1987, 1993; Verga, 1995; Cony, 1996; Saidman et al., 1997; Bessega et al., 2000b, c; Mantovan, 2004; Joseau, 2006), studies of interspecific geneflow (Saidman, 1990, 1993; Saidman et al., 1998a), and researches on the reproductive system of the species (Bessega et al., 2000a). These studies were conducted with morphological and physiological characters, isozymes markers, or also by means of different molecular markers such as RAPDs and RFLPs.

Different genetic studies performed in the species of section Algarobia showed that the species of this section exhibited a high genetic variability within populations and that they posses more genetic variation than species of section Strombocarpa (Pasiecznik et al., 2001). Besides, isoenzymatic and molecular studies conducted in different species of section Algarobia indicated a general trend towards significant homozygote excess within populations (Saidman 1986, 1988a, 1990, 1993; Saidman and Vilardi, 1987, 1993; Saidman et al., 1997; Verga, 1995; Bessega et al., 2000 b). One of the most remarkable

characteristics of the section Algarobia is the high genetic similarity among the species so far studied (Saidman et al., 2000). The differentiation among species within this section, measured by means of biochemical or molecular markers, showed relatively low values, despite the important morphological differences that exist between them (Burkart, 1976). Almost all alleles of polymorphic isoenzyme loci are shared by all species. Diagnostic loci are mostly absent and species differ only in allele frequencies.

3.2.4. Hybridization

Another important aspect of the section Algarobia is that interspecific hybridization is very frequent in zones of sympatry. This hybridization creates intermediate phenotypes that render an accurate morphological determination difficult. Within the section Algarobia, a large number hybrids with of two or even three parental species have been postulated and confirmed with a variety of morphological, enzymic, and molecular studies in South America (Palacios and Bravo, 1981; Hunziker et al., 1986; Saidman, 1990; Verga, 1995; Vega and Hernández, 2005) and also in Mexico (Almanza et al., 1992). In Argentina, frequent interspecific hybridization has been confirmed between at least seven species of section Algarobia: *P. alba*, *P. alpataco*, *P. caldenia*, *P. chilensis*, *P. flexuosa*, *P. nigra*, and *P. ruscifolia*. Hybrid swarms composed by two or more of these species are described in zones of natural sympatry and also in areas where the natural habitat has been disturbed by human impact.

The strong evidences of natural hybridization among several species of the section Algarobia, as well as the high genetic similarity between them, has led to the assumption that these species would integrate a ***syngameon*** (Palacios and Bravo, 1981; Saidman, 1985, 1988a, 1993; Saidman and Vilardi, 1987, 1993; Saidman et al., 1998a; Bessega et al. 2000a, b). Grant (1981) defines the syngameon as "the most inclusive unit of interbreeding in a hybridizing species group". Because frequent events of interspecific natural hybridization with fertile hybrid production in areas of sympatry occur, and isolation mechanisms between species seem to be weak or incomplete, the entities members of this syngameon can not be considered species under the classic "biological species concept", which defines species in terms of isolating mechanisms (Mayr, 1963; Dobzhansky, 1970; White, 1978). However, the members (botanical species) of the syngameon constitute real units in terms of morphology, ecology and evolution, and are considered "good taxonomic species". Therefore, the entities

involved in this syngameon might be considered species under the Templeton (1989) cohesive concept that defines species as "the most inclusive group of organisms having the potential for genetic and/or demographic exchangeability".

Natural hybridization between *Prosopis chilensis* and *P. flexuosa* has been confirmed in sympatric areas of the Chaco phytogeographic province in Argentina (Verga, 1995). Both species are highly related genetically, and significant gene flow occurs between them. Natural and fertile hybrids have also been described between these species as well as hybrids swarms in wide areas of sympatry. Morphological analysis performed by Verga (1995) in different populations located in contact areas between *P. chilensis* and *P. flexuosa* in the Chaco region revealed the natural occurrence of individuals with intermediated morphologies between both species. Besides, this morphological evidence for hybridization was confirmed by Verga (1995) by means of genetic studies with isozymes. Using the isozyme ADH (alcohol dehydrogenase), at locus *ADH-A*, which can be considered a diagnostic locus for species differentiation between *P. chilensis* and *P. flexuosa*, Verga (1995) found that morphohybrids, i.e. individuals showing intermediate morphologies, presented also hybrid genotypes at the *ADH-A* locus. Hybridization rates appear to be higher in contact areas where the natural habitat has been disturbed by human impact. It is assumed an introgression process between these two species taking place in sympatric populations (Verga, 1995). In addition, controlled pollination experiments were successfully performed pollinating *P. chilensis* trees with pollen from *P. flexuosa* and the opposite (Córdoba Ana, pers. comm.).

It is unknown for how long the hybridization process between *P. chilensis* and *P. flexuosa* has been taken place, but there is strong evidence that natural hybrids are widespread, fertile, and ancient. The spontaneous crossing between populations of both species is reflected in the Chaco phytogeographic province, Argentina, by the frequent occurrence of "***hybrid swarms***" in wide areas of sympatry (habitat overlap). These hybrid swarms are populations composed of pure individuals of both *Prosopis* species, and also by morphologically distinctive individuals that result from hybridization between the parent species (pure species), backcrossing of the offspring to members of the parent species, and interbreeding among the hybrid individuals, too. Hybrids swarms between *P. chilensis* and *P. flexuosa* are also very common in areas of sympatry

characterized by man-made disturbance, such as irrigation areas, or areas where natural forest was cut down to clear land for agriculture or ranching, and then abandoned.

In spite of hybridization, *P. chilensis* and *P. flexuosa* represent "real" biological units in terms of morphology, ecology, genetics and evolution, and both species have maintained and are maintaining genetic, phenotypic and ecological cohesion within them, and distinction between them, too. They have maintained themselves as distinct evolutionary lineages for million of years. However, the continuous production of hybrids might favor the occurrence of "evolutionary experiments" through the production of novel genetic combinations, and hybridization might have been important in the evolution of this group (Saidman et al., 1998b). Hybrids may display characteristics that allow them to occupy niches different from the parental ones. Frequently, hybrids display special capabilities to exploit open or hybrid habitats (Anderson, 1949). Such specifically adapted populations could be used in forest genetic improvement programs to obtain basic material to restitute the plant cover in degraded areas.

4. MATERIALS AND METHODS

4.1. Materials

4.1.1. The studied plot

4.1.1.1. Location

The study took place in the Provincial Park and Forest Natural Reserve Chancaní (*Parque Provincial y Reserva Forestal Natural Chancaní*). This natural reserve is a protected area of 5000 hectares located in the *Departamento Pocho*, on the western slope of the "*Sierras de Pocho-Guasapampa*", in the west of Córdoba province, Argentina (**Figure 11**). The Chancaní Reserve is located in the *Chaco* phytogeographic province according to Cabrera (1976) (see: **Figure 1**). In terms of the Koppen-Trewartha classification (Koppen, 1931; Trewartha, 1943) the area is classified as Dry Subtropical Arid. The annual average rainfall fluctuates between 400 and 600

millimeters with 400 to 500 millimeters of water deficit during the year (Cabido and Pacha, 2002).

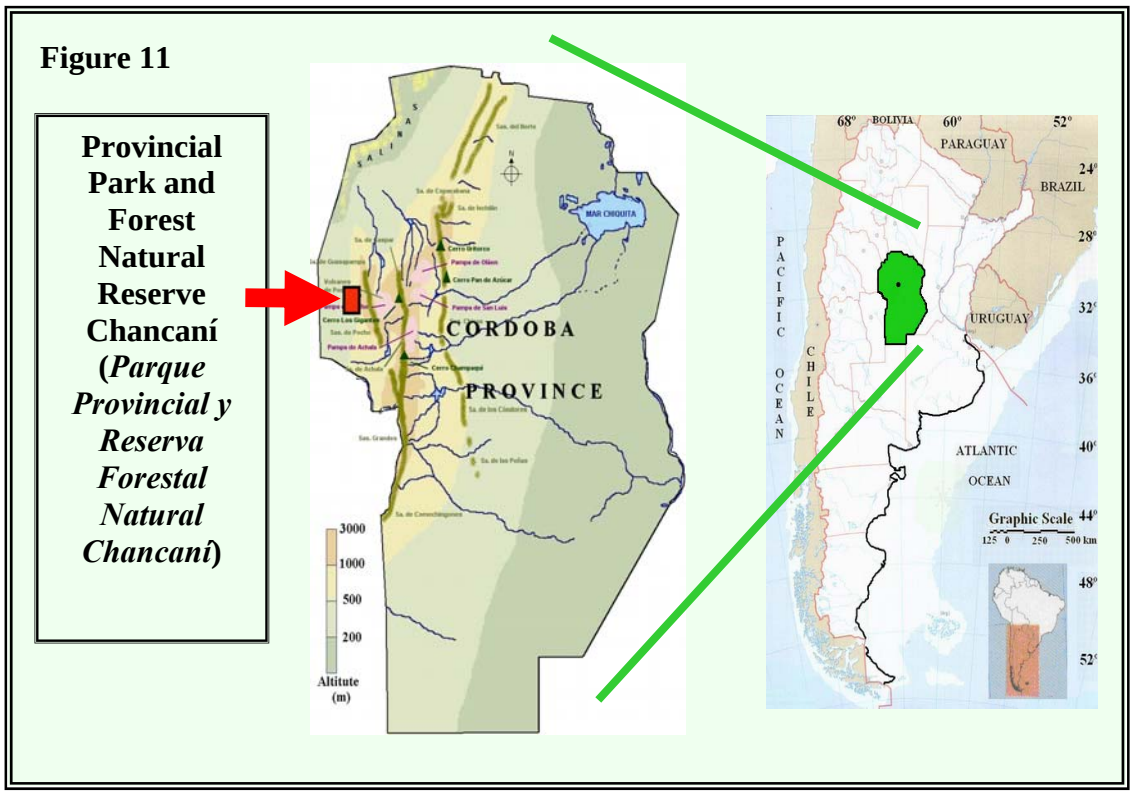

Figure 11: Location of the Provincial Park and Forest Natural Reserve Chancaní (*Parque Provincial y Reserva Forestal Natural Chancaní*) in Córdoba Province, Argentina.

The studied plot has an area of 4700 square meters. It is included in a continuous forest located southeast of the aforesaid reserve (Latitude 31º 23' S, Longitude 65º 27' O, Altitude 360 m). The area belongs to an abandoned farm. Nowadays, it presents around 30 years of natural regeneration of the native forest (Cabido and Pacha, 2002). In this environment, which has been highly modified by man, both species of *Prosopis* (*P. chilensis* and *P. flexuosa*) occur together and many individuals with intermediate morphological characteristics between the two species can be observed. Besides, other tree and shrub species can be found in the native forest, such us: *Aspidosperma quebracho-blanco*, *Acacia aroma*, *Cercidium australe*, *Larrea divaricada*.

4.1.1.2. Field measurements

In the studied plot all *Prosopis* trees, which flowered during the flowering season 2003-2004 were mapped (**Figure 12**). A total of 100 flowering *Prosopis* trees (***adult population***) were identified during that period (see **Appendix 2**). In order to describe the position of each tree in the plot, measurements in the field were performed using polar coordinate system. Each point (tree) was identified by a distance from some fixed feature in space (a fixed point or a neighbor tree), and by one or more subtended angles. Subsequently, to build the map (**Figure 12**) polar coordinates were converted to Cartesian coordinates as follows:

$$x = r\cos\theta \text{ and } y = r\sin\theta$$

where x and y are the standard Cartesian coordinates, and r and θ the standard polar coordinates (distances and angles, respectively).

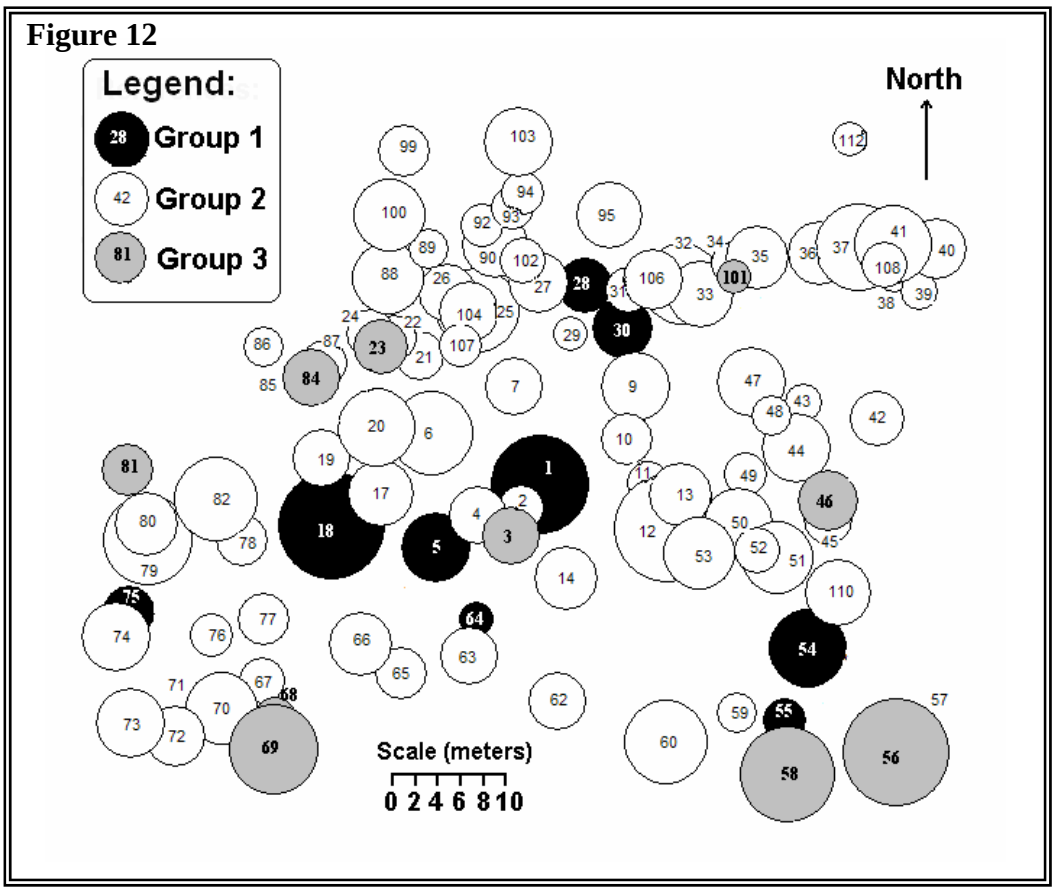

Figure 12: Map of the studied plot with 100 mapped Prosopis trees that are represented by circles. The different colors refer to the morphological groups obtained after morphological analysis (see section **4.2.3.**). The diameter of the circle is proportional to the crown.

For the estimations of the crown size of each tree, the projection of the crown on the ground was measured in two directions, North-South and East-West. Then, the diameter of the crown was estimated as the arithmetic mean between the two values, and crown size was computed considering the crown as a perfect circumference (**Figure 12**).

4.1.1.3. Phenological observations

Observations on flowering phenology were performed during the flowering period 2003-2004. Phenological observations were carried out for the phenological substage "*c*", or completely open and receptive flower (Genise et al., 1990). Phenological events were registered for each individual tree every ten days. Then, with the observations performed the overlapping of flowering between trees was estimated. The results of this study were used in the paternity analysis as complementary data for paternal exclusion (see section **4.2.5.** and manuscript **III**).

4.1.2. Sampling

4.1.2.1. Leaves

From each one of the adult trees ten mature leaves were sampled for herbarium collection and posterior morphological analyses. Besides, samples of fresh leaves from each adult individual were collected for DNA isolation.

4.1.2.2. Fruits

Five mature pods were sampled from each adult tree for herbarium collection and morphological analyses.

4.1.2.3. Seeds

After morphological assignment (see: section **4.2.3.**), twenty three open-pollinated mother trees were selected among the adult tree population for seed collection. The most representative individuals, from a morphological standpoint, of each morphological group, were selected as mother trees. Seeds were harvested from the mother tree during the flowering season 2003-2004. A total of 516 seeds (***seed population***) were collected. **Table 4**

shows the adult trees selected as mother trees among the adult population and the number of seeds harvested for each tree. Sample size per seed tree was from nine to 48 seeds according to the availability of seeds at the time of collection. The seed population was divided into groups according to the morphological classification of the mother trees (see: **Table 1** in publication **III**). Seeds were stored in a freezer at –20°C. All seeds were subjected to microsatellite analyses.

Table 4: Adult trees selected as mother trees among the adult population, and number of seeds harvested for each tree.

Selected Mother Trees	N° of seeds harvested
1	32
3	32
4	32
6	10
10	10
11	20
13	24
14	24
18	9
30	20
31	19
46	20
47	24
49	20
51	24
54	20
58	24
59	32
65	24
69	48
81	20
84	18
101	10
TOTAL	516

4.2. Methods

4.2.1. Development of microsatellite (SSRs) gene markers in *Prosopis*

Microsatellite loci were isolated for the species *Prosopis chilensis* using an enrichment protocol developed by Fischer and Bachmann (1998). This method is a nonradioactive strategy for SSRs isolation based on the establishment of microsatellite-enriched libraries, using microsatellite-oligonucleotide probes, and magnetic capture by streptavidin-coated magnetic beads.

For the development of SSR markers in *Prosopis chilensis*, total DNA was isolated from five-day-old green cotyledons of a single tree using the Dneasy® PlantMiniKit (QIAGEN) following the manufacturer's instruction. The extracted DNA was digested with *RsaI* restriction enzyme. Thereafter, DNA fragments were ligated to adapter-to-primer oligonucleotides. The constructs were hybridized to biotinylated microsatellite oligonucleotide probes $(CA)_{10}$, $(GAA)_8$, $(AAC)_8$. These hybrids were then bound to streptavidin-coated magnetic beads, followed by a magnetic separation of the fragments in order to obtain a microsatellite-enriched product. Adapter-mediated genomic PCR were performed and the products were ligated into a vector. The plasmid vector was then transformed into chemically competent *Escherichia coli* SURE cells (Stratagene), and plated onto LB agar medium in order to obtain a microsatellite-enriched library. Recombinant plasmids (positive colonies) were identified by means of blue-white screening. A detailed protocol of the procedures used for microsatellite isolation in *Prosopis chilensis* is given in **Appendix 3**.

A total of 120 positives colonies were sequenced. PCR primers were designed for those sequences containing microsatellites motifs. PCR protocols for these primers were optimized and an analysis for the identification of alleles was performed. For the developed primers regular segregation was tested in single tree progenies following the method developed by Gillet and Hattemer (1989). Variation was characterized in a natural population of *Prosopis chilensis* (20 trees) and *P. flexuosa* (20 trees) of the Argentinean Arid Chaco. Cross-species amplification tests were performed in seven other *Prosopis* species. Details about the procedures for the development of SSR markers in *Prosopis* are presented in manuscript **II**.

4.2.2. Evaluation of the efficiency of microsatellite enrichment in *Prosopis* using magnetic capture

The results of the enrichment strategy used for the development of SSR markers in *Prosopis,* was evaluated. After the creation of a microsatellite-enriched library previously described (section **4.2.1.**), the obtained level of enrichment was analyzed by sequencing 120 enriched genomic fragments. In the obtained sequences, a search for SSR motifs was performed applying specific search criteria. The search was carried out looking for microsatellites composed of tandemly repeated units of one to six nucleotides, repeated at least three times and of a total size of at least six bases. The efficiency of the protocol was estimated as the number of clones that contained microsatellite motifs with the abovementioned characteristics. Sequences with SSR motifs were compared for homology and similarity with genomic database sequences of other plant species in the European Molecular Biology Laboratory Database (EMBL-Bank) using the program FASTA 3 (Pearson and Lipman, 1988; Pearson, 1990). The number of sequences containing chloroplast and nuclear SSR loci and the number of SSR motifs in transcribed and nontranscribed regions was determined. More details about the method used for the evaluation of the efficiency of microsatellite enrichment in *Prosopis* using magnetic capture are given in manuscript **I**.

4.2.3. Morphological assignment

Qualitative and quantitative morphological characters (**Table 5**) from leaves and fruits were analyzed in the adult population in order to assign individual trees to morphological groups. For each adult tree, morphological characters (**Table 5**) were measured on ten mature leaves and five mature fruits from the herbarium collection. Thus, the following computations were performed over ten repetitions per tree for leaves characters, and five repetitions per tree for fruit characters. The morphological study was performed applying the "Morphological Distance" (d_m) method (Verga, 1995; Verga and Gregorius, in press.) and defined as:

$$d_m(f,g) = \tfrac{1}{2}\int_{-\infty}^{\infty} |f(x) - g(x)| dx$$

where d_m *(f, g)* is the morphological distance between OTUs ("Operational Taxonomic Units"; Sneath and Sokal, 1973) *f* and *g*; and *f(x)* and *g(x)* represent the probability density functions for the character *x* in OTUs *f* and *g*, respectively. This distance measures differences between frequency distributions, and it assumes its maximum value of 1 exactly if two OTUs have no characters in common.

In the present study, each individual adult tree was considered as an OUT. In consequence, it is possible to compute the d_m for each single pair of individuals at each morphological character. Then, the total (multivariate) d_m between each pair of individuals is computed as the arithmetic mean of the distances at each particular character.

Table 5: Alphabetic list of qualitative and quantitative characters used for the morphological analysis.

Abbreviation	Description
AF	leaflet width
AFOL	leaf area
AFR	fruit width
AP_T	division between leaflet-one third superior area and its total area
API	division between leaflet-one third superior area and the rectangle where this area is included
AR	leaflet area
BFR	fruit-margin shape
CFR	fruit color
DIFOL	LPI / NFO
FAL	leaflet falcate
FFR	fruit shape
G_AFR	GFR / AFR
GFR	fruit thickness
L_AF	LF / AF
LF	leaflet length
LFR	fruit length
LPE	petiole length
LPI	pinna length
NFO	number of leaflets per pinna
NPI	number of pinna pairs

The Morphological Distance (d_m) was calculated using the program NTSYS. Then a dendrogram was achieved by UPGMA (Unweighted Pair-Group Method using an Arithmetic Average) with the obtained multivariate morphological distances (total d_m) From the obtained UPGMA dendrogram (see section **5.3.**), ten different grouping alternatives were tested to define the optimal number of groups. The confidence of the different clusters found in each alternative was evaluated by means of Discriminant Analysis. The distribution within each group was assumed to be multivariate normal; consequently a parametric method was used to develop the discriminant function. Within-group covariance matrices were tested for homogeneity in order to define the discriminant function (linear or quadratic). Stepwise-discriminant analysis by stepwise selection was performed in order to find those variables that better discriminate among the obtained classes (groups). The stepwise-discriminant procedure was performed with the program SAS, with a significant level of 0.15.

After groups where obtained, each individual was taxonomically classified in *Prosopis chilensis*, *P. flexuosa* or putative hybrid. This classification was carried out following the key for identification of interspecific hybrids between *Prosopis chilensis* and *P. flexuosa*, based on quantitative characters proposed by Verga (2000). This key uses ten morphological characters (LF, L_AF, AFR, GFR/AFR, LPI, DIFOL; LFR, AF, BFR, and CFR; see **Table 5** for abbreviations) to differentiate between pure species and putative hybrids. For each tree, the aforesaid characters are measured in ten mature leaves and five mature fruits. Then, for every individual the arithmetic mean is computed for each character. According to the obtained value, a table score is given to the character. A total score is achieved for each tree by adding the scores for each character. With the obtained total score the tree is classified based on a given ranking. This ranking was developed on the bases of prior classifications done through numerical taxonomy and identification of hybrids with isozymes.

4.2.4. Characterization of the genetic structure and variation

4.2.4.1. Adult population

The genetic structure of the adult population was studied in order to characterize the levels of genetic variation and differentiation within and among the morphological groups achieved after morphological assignment.

Total DNA was extracted from fresh leaves of each individual tree of the adult population using the DNeasy PlantMiniKit (Qiagen, Hilden, Germany). In the genetic analysis six microsatellite loci (***Mo05***, ***Mo07***, ***Mo08***, ***Mo09***, ***Mo13*** and ***Mo16***) originally developed for *Prosopis chilensis* were used. In **Table 1** in manuscript **II**, the SSR primers used in the genetic analysis are described in details. Details of the PCR amplification reaction used are given in manuscript **II**. SSR fragments were analyzed by means of capillary electrophoresis on an automatic sequencer ABI PRISM® 3100 Genetic Analyzer (Applied Biosystem/ HITACHI). The length of the fragments was determined using the computer programs Genescan 3.7 and Genotyper 3.7 (both from Applied Biosystems). The different alleles founded at each locus were scored as number of base pairs of the amplified fragment.

Population genetic parameters were calculated for each morphological group in the adult population using the software GSED Version 1.1j (Gillet, 2004), Arlequin version 3.01 (Excoffier et al., 2005), and GenAlEx6 (Peakall and Smouse, 2005).

The ***relative frequency*** (p_i) ***of an allele*** was calculated from counting the number of that particular allele (n_i) in a sample of n individuals:

$$p_i = \frac{n_i}{2n}$$

The ***number of observed alleles per locus*** (N_a) was determined by direct count. The ***average number of alleles per locus*** (***A/L***) was computed dividing the total number of alleles observed at all gene loci by the total number of gene loci. A search for ***private alleles*** at group level was performed with the program GenAlEx6 (Peakall and Smouse, 2005). Besides, the presence of individuals with ***rare or unique alleles***, i.e. an allele which is absent in all but one tree in the sampled population, was investigated.

The ***effective number of alleles*** (N_e) was calculated for each single locus according to Brown and Weir (1983). The number of effective alleles across loci was computed as the harmonic mean of single locus N_e.

$$N_e = \frac{1}{\sum p_i^2}$$

where p_i is the frequency of the *i-th* allele.

Genetic diversity was estimated according to Nei (1973). The ***observed heterozygosity*** (***H_o***) is the proportion of all heterozygous genotypes observed at a locus and can be polled based on the average across all loci. The ***expected heterozygosity*** (***H_e***) is equal to the ***total population differentiation*** (***δ_T***; Gregorius, 1988) for a population of effectively infinite size:

$$H_e = 1 - \sum p_i^2 = \delta_T$$

where p_i is the frequency of the *i-th* allele.

The ***fixation index*** (***F***) from Hartl and Clark (1997) was used to measure the excess of homozygosity (or reduction of heterozygosity) in populations relative to the Hardy-Weinberg expectations.

$$F = 1 - \frac{H_o}{H_e}$$

where H_e and H_o are expected and observed heterozygosity, respectively.

Since sample sizes of Groups 1 and 3 (ten individuals each) are very heterogeneous compared to the sample size of Group 2 (80 individuals) (see **Appendix 2**), allelic richness (Petit et al., 1998) per locus and group, and overall groups, was estimated using the program FSTAT Version 2.9.3.2 (Goudet, 2001). The principle is to estimate the expected number of alleles in a sub-sample of *2n* genes, given that *2N* genes have been sampled ($N \geq n$). *n* is fixed as the smallest number of individuals typed for a locus in a group. Allelic Richness is then calculated as:

$$R = \sum_{i=1}^{n} \left[1 - \frac{\binom{2N - N_i}{2n}}{\binom{2N}{2n}} \right]$$

where N_i is the number of alleles of type *i* among the *2N* genes.

In addition, ***N_a***, ***H_o*** and ***H_e*** were recalculated for Group 2 in the adult population considering a sample size equal to the sample sizes of Groups 1 and 3 (ten individuals each). Random sampling of ten individuals was preformed within Group 2 and the aforesaid parameters were estimated. This procedure was repeated 100 times, and the average and standard deviation for the parameters ***N_a***, ***H_o*** and ***H_e*** were computed over 100 replicates.

In order to characterize the genetic variation among morphological groups, the measures proposed by Nei (1972) (***genetic distance D***), and Gregorius (1974) (***genetic distance d_0***) and Gregorius and Roberds (1986) (***genetic differentiation D_j***; ***average differentiation δ***) were computed as follows:

Genetic distance D:

$$D = -\ln I$$

where I is the normalized identity of genes between the two populations

$$I = \frac{J_{XY}}{\sqrt{J_X J_Y}}$$

J_{XY}, J_X and J_Y are calculated as follows:

$$J_{XY} = \sum_{i=1}^{k} p_{iX}\, p_{iY} \quad ; \quad J_X = \sum_{i=1}^{k} p_{iX}^2 \quad ; \quad J_Y = \sum_{i=1}^{k} p_{iY}^2$$

where p_{iX} and p_{iY} are the frequencies of the i-th allele in the populations X and Y, respectively. For multiple loci, J_{XY}, J_X and J_Y are calculated by summing over all loci and alleles, and dividing by the number of loci.

Genetic distance d_0:

$$d_0(x,y) = \tfrac{1}{2}\sum_i \left| p_i(x) - p_i(y) \right|$$

where $d_0\,(x,y)$ = genetic distance d_0 between two populations x and y; and $p_i(x)$ and $p_i(y)$ are respectively the frequencies of the *i-th* allele in population x and y.

Genetic differentiation D_j:

$$D_j = d_0(p_j, \overline{p}_j)$$

where p_j is the frequency distribution of the genetic types in the *j-th* population, $\overline{p}_j$ the frequency distribution of these types in all other populations taken together and d_0 the genetic distance.

Average differentiation δ:

$$\delta = \sum_j c_j D_j$$

where c_j are the weights corresponding to the proportion of genetic elements belonging to the *j-th* population, and D_j is the genetic differentiation.

In the present study, $\boldsymbol{\delta}$ as well as $\boldsymbol{D_j}$, were estimated considering equal sample size for the groups, i.e. c_j=1/*n* (where *n* is the number of groups). Because in the seed population sampling was not performed at random, it can not be assumed that the proportion of sampled individuals in each group reflects reality. Thus, it would not be appropriated to use in the computations for the seed population the weights in proportion to the sample size. In order to be able to compare across populations, in the adult population the same criterion was used in the computations.

Total and pairwise $\boldsymbol{F_{st}}$ was calculated with the following formula:

$$F_{ST} = \frac{f_0 - f_1}{1 - f_1}$$

where f_0 is the probability of identity by descent of two different genes drawn from the same population f_1 is the t probability of identity by descent of two genes drawn from two different populations. The significance of the F-statistic was tested using a non-parametric permutation approach described in Excoffier et al. (1992), consisting in permuting individuals among populations (groups). *P*-values were calculated after 10,000 permutations.

To test the excess of homozygotes or heterozygotes over Hardy-Weinberg expectations in the morphological groups, a ***goodness-of-fit test*** was performed following the model of ***Pearson's χ^2 goodness-of-fit test*** with the statistics (Weir, 1990):

$$x^2 = \sum_{types} \frac{(N_i - E(N_i))^2}{E(N_i)}$$

where N_i and $E(N_i)$ are, respectively, the observed and expected sample counts of each genetic type.

This statistic is asymptotically χ^2 distributed. The number of degrees of freedom (*df*) is defined as:

$$df = (k-1)(n-1)$$

where *k* is the number of populations or groups, and *n* is the number of genetic types.

The theoretical exclusion probabilities for paternity were computed for all loci over all trees of the adult population using the program FAMOZ (Gerber et al., 2003). Details on the formula and procedures for the computation of the probability of identity for codominant markers are given in Jamieson and Taylor (1997).

4.2.4.2. Seed population

Based on the results of the characterization of the genetic structure in the adult population and the theoretical exclusion probabilities, three SSR loci (***Mo08***, ***Mo09*** and ***Mo13***; see: **Table 1** in manuscript **II** for details on the loci) were selected for the genetic analysis of the seed population and subsequent mating system analysis. The three microsatellite primers selected showed high levels of diversity and genetic differentiation among morphological groups in the adult population, and the high values of theoretical exclusion probabilities for paternity.

DNA was isolated from cotyledons of seedlings after germination using the DNeasy 96 Plant Kit (Qiagen, Hilden, Germany). Amplification and detection procedures for the three SSR loci selected were performed as previously described for the adult population.

Levels of genetic variation and differentiation within and among morphological groups at three SSR loci were characterized in the seed population. Population genetic parameters were calculated for each morphological group in the seed population following the methods and measurements described for the adult population. Results were compared with those obtained in the adult population at the same loci.

4.2.5. Mating system analysis

With the genetic data obtained from the seed population, the male gametic frequency of alleles in the effective pollen clouds of individual mother trees was estimated by the Maximum-Likelihood method of Gillet (1997), assuming regular segregation. Detailed

information about the methods and procedures used in the analysis of the mating system are given in manuscript **III**.

The effective pollen cloud for a particular morphological group (*pooled pollen cloud*) was computed using weights in proportion to the number of seeds investigated for each tree. The allele frequencies in the effective pollen clouds of each individual mother tree belonging to a particular group were added in proportion to the number of seeds investigated for each tree. Population genetic parameters were calculated in the pooled pollen cloud and compared with those obtained in the adult population at the same locus. In addition, pairwise genetic distances (D and d_0) were computed for individual pollen clouds. A homogeneity test (G_h-*test*) was used to verify whether the distribution of allelic frequencies was homogeneous among the pollen clouds and morphological groups.

Paternity analysis was applied in order to assess gene flow within and between morphological groups. Paternity assignment was conducted by categorical allocation based on multilocus genotypes. The most likely father for each seed was assigned based on the computed Likelihood ratio (LOD score). The potential father with the maximum LOD score was regarded as the most likely pollen donor for the respective seed. Potential fathers whose flowering period did not overlap with the flowering period of the mother tree were excluded. Considering the mating events detected by paternity assignment, descriptive parameters on the possible mating system prevailing in the hybrid swarm were estimated.

Mating system parameters were estimated for the flowering period 2003-2004 using different methodologies. For each individual mother tree, the proportion of progeny with a "foreign" allele at any gene locus was calculated as a *minimum estimate of the outcrossing rate*. Besides, outcrossing rate for individual trees was estimated based on the results of the paternity assignment. In addition, selfing rate (s) was estimated based on the presence of rare alleles in the adult population. Furthermore, outcrossing rate (t) from single- and multilocus genotypes, correlated mating (r_p), and biparental inbreeding ($t_m - t_s$) were estimated assuming the mixed mating model (Ritland and Jain, 1981). Estimations were computed for each morphological group, and for individual trees within each group.

5. SUMMARY OF THE RESULTS

5.1. Development of microsatellite (SSRs) gene markers in *Prosopis*

Six polymorphic microsatellite markers were developed in *Prosopis chilensis* using the protocol proposed by Fischer and Bachmann (1998). Subsequently, the novel SSR primers were characterized in a natural population of *Prosopis chilensis* (20 trees) and *P. flexuosa* (20 trees) of the Argentinean Arid Chaco. Manuscript **II** describes the development and characterization of the six microsatellite gene markers in *Prosopis*.

From the microsatellite-enriched library, 120 positive colonies were sequenced. The average length of the sequenced fragments was 580 bp (± 96 bp). Out of 120 sequences, 18 were selected for primer development considering those sequences with SSRs with at least five repeats. All primer pairs were tested; nine out of 18 primers amplified products of the expected size. These nine primers were labeled with a fluorescent dye at its 5'-end, and variation was characterized in a natural population of *Prosopis chilensis* (20 trees) and *P. flexuosa* (20 trees) of the Argentinean Arid Chaco. Six out of nine primer pairs amplified reproducible and well-scorable fragments which were polymorphic in both species. Details about the six developed primers are given in **Table 1** in manuscript **II**. The complete sequences of the microsatellite loci, as they appear publicly available in the GenBank® sequence database, are given in **Appendix 4**. The test for regular segregation performed for the six SSR loci in single tree progenies showed regular meiotic segregation and random fertilization of the eggs by pollen at all markers.

The novel SSR markers showed a Polymorphism Information Content (PIC) between 0.14 and 0.70 in *P. chilensis* and, between 0.41 and 0.85 in *P. flexuosa*. The number of alleles (N_a) varied from two to six, and from two to 13 in *P. chilensis* and *P. flexuosa* respectively. Expected heterozygosity (H_e) ranged from 0.14 to 0.73 in *P. chilensis*, and from 0.46 to 0.86 in *P. flexuosa* (**Table 1** in manuscript **II**).

The six primer pairs were tested for cross-species amplification in seven other *Prosopis* species (**Table 2** manuscript **II**). All primers amplified in at least five species, revealing a broad cross-species affinity. The number of alleles found varied from one to five.

5.2. Evaluation of the efficiency of microsatellite enrichment in *Prosopis* using magnetic capture

Manuscript **I** presents the results obtained in the evaluation of the enrichment strategy developed by Fischer and Bachmann (1998), and used for the development of SSR markers in *Prosopis*. From the SSR-enriched library, a total of 120 clones with and average length of 580 bp (± 96 bp) were sequenced.

An analysis of the 120 clones sequenced showed that 95.8% (115 out of 120) had at least one SSR motif matching the search criteria (i.e., microsatellite fragments composed of tandemly repeated units of one to six nucleotides, repeated at least three times and of a total size of at least six bases). When regarding SSR segments of at least five repeated units and a minimum length of ten bp, the level of enrichment was 30.8%. Sixteen percent (84 fragments) of all detected motifs showed specific complementary sequences with the oligoprobes used in the enrichment (11% with the [CA] probe, 4% with [GAA], and 1% with [AAC]).

The comparison of the 115 sequences containing SSR motifs against the EMBL Nucleotide Plant Sequence Database revealed that 7.8% of the clones (9 clones) had high similarity with chloroplast regions. Most of the clones (102 clones) with similarity to nuclear DNA conform to nontranscribed regions of different species. Only 4 clones showed high similarity (>60% identity over >180 bp length with an expectation value <1*e–10) with transcribed regions (**Table 2** in manuscript **I**).

Taking into account those clones with similarity to nuclear sequences (106 clones), a total of 466 SSR motifs were detected including 216 mononucleotide repeats (**Table 1** in manuscript **I**). According to the length of the microsatellite motif, i.e. numbers of repeats, the most frequent SSR segments found were those with three repeated units (**Figure 1** in manuscript **I**).

5.3. Morphological assignment

The dendrogram obtained from UPGMA using a matrix of *Morphological distances* (d_m) derived from fruit and leaf characters measured in the adult population is shown in **Figure 13**. After the UPGMA dendrogram was achieved, Discriminant Analysis was performed in order to evaluate the best grouping alternative. Since the Test of Homogeneity of within-group covariance matrices performed in the Discriminant procedure showed significant chi-square values (significant level 0.05), the within covariance matrices were used in the discriminant function, yielding a quadratic discriminant function. The study based on Discriminant Analysis allowed establishing as adequate criteria of classification the formation of three morphological groups (***Group1***, ***Group 2***, and ***Group 3***). This grouping showed the minimum error rate (no individual was misclassified) for the Discriminant Analysis performed. According to this classification, the first major differentiation occurs at a d_m of approximately 0.80 and it separates the individuals in two large groups (Group 3 from the rest) (see: **Figure 13**). Then, at a d_m of 0.70 appears a second differentiation which distinguishes the individuals belonging to Group 1 from the individuals of Group 2 (see: **Figure 13**). As a result of this grouping, the number of individuals in each group is: Group 1, ten individuals; Group 2, 80; and Group 3, ten individuals.

According to the Stepwise-discriminant Analysis (level of significance: 0.15) the variables that better discriminate among morphological groups are, in decreasing order of importance: LF, BFR, NFO, AF, FAL, L_AF, AR, G_AFR, FFR, and AFOL (see **Table 5** for abbreviations).

The subsequent taxonomic classification following the key of Verga (2000) of individuals in each morphological group showed that: Group 1 contains four individuals *Prosopis flexuosa* and six putative hybrids; Group 2 is constituted of 79 individuals of *P. flexuosa* and one putative hybrid individual; and Group 3 is formed by two *P. chilensis* individuals and eight putative hybrids similar to *P. chilensis* (**Appendix 2**).

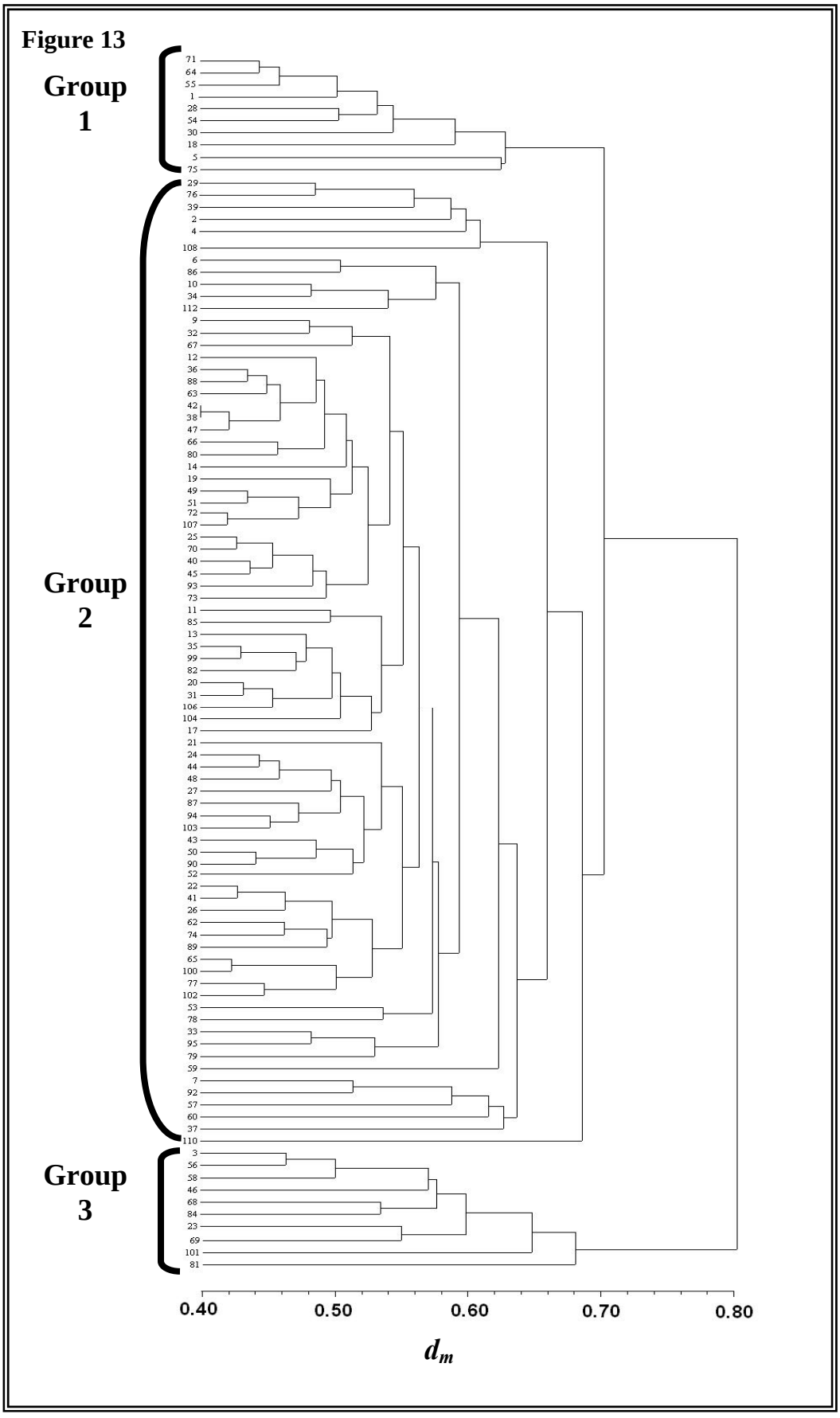

Figure 13: Dendrogram obtained from morphological distance (d_m) using UPGMA clustering method.

5.4. Characterization of the genetic structure and variation

5.4.1. Adult population

The allelic structures of each morphological group in the adult population at six microsatellite loci are presented in **Figure 14** and in **Appendix 5**. The different alleles found at each locus were named as the number of base pairs of the amplified fragment.

The results of the count for private alleles showed that Group 2 had 29 alleles which are not found neither in Group 1 nor in Group 3. Seven of the private alleles for Group 2 were found at locus Mo07, two at locus Mo08, nine at locus Mo09, four at locus Mo13, and seven at locus Mo16. On the other hand, Group 3 showed one private allele at locus Mo13, and Group 1 did not exhibit private alleles.

Rare or unique alleles were found in five individual trees. Tree 10 was homozygote for allele 173 at locus Mo16; tree 17 showed once the allele 222 at locus Mo13; tree 59 was homozygote for the allele 205 at locus Mo07; and trees 49 and 74 were heterozygote for the alleles 223 and 235, respectively, at locus Mo09. All these trees belong to Group 2.

The allelic multiplicity observed at six SSR loci, in three morphological groups in the adult population is presented in **Table 6**. The total number of observed alleles (N_a) over all groups and over all loci was 70, with an average number of alleles per locus (*A/L*) of 11.67. Group 2 showed the highest *A/L* (11.17), followed by Group 1 (5.33), and finally Group 3 (4.33). N_a observed at all loci, except at locus Mo05, was larger in Group 2 as compared to Groups 1 and 3.

The number of effective alleles of the pool was 4.75, ranging from 1.8 (locus Mo05) to 7.5 (locus Mo08). Allelic richness computed based on a minimum sample size of ten individuals, showed the highest mean value in Group 2 (6.64), followed by Group 1 (5.33), and the lowest in Group 3 (4.33).

Figure 14

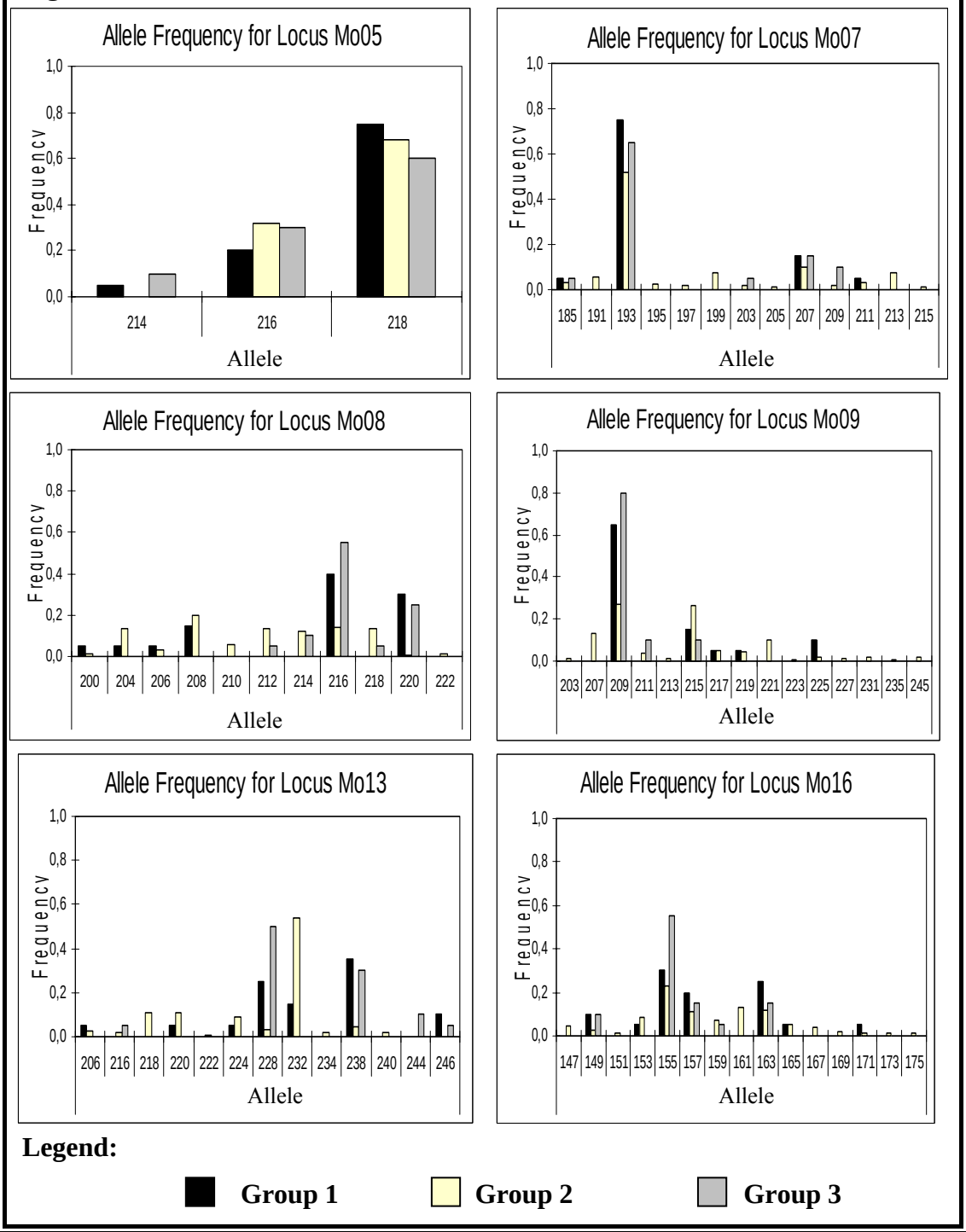

Figure 14: Allelic frequencies at microsatellite loci *Mo05*, *Mo07*, *Mo08, Mo09, Mo13* and *Mo16* in three morphological groups analyzed (***Group 1***, ***Group 2*** and ***Group 3***) in the adult population.

Table 6: Allelic multiplicity registered at six microsatellite loci in the morphological groups (**Groups**) and in the pool (**Pooled**), in the adult population (N_a: observed number of alleles; N_e: effective number of alleles; ***R***: allelic richness; **Mean**: average value).

		SSR Loci						
Groups		**Mo05**	**Mo07**	**Mo08**	**Mo09**	**Mo13**	**Mo16**	**Mean**
Group 1	N_a	3	4	6	5	7	7	5.33
	N_e	1.65	1.69	3.57	2.17	4.44	4.76	2.53
	R	3	4	6	5	7	7	5.33
Group 2	N_a	2	13	11	15	11	15	11.17
	N_e	1.77	3.35	7.27	5.69	3.09	8.28	3.70
	R	2	7.14	7.57	7.63	6.32	9.18	6.64
Group 3	N_a	3	5	5	3	5	5	4.33
	N_e	2.17	2.17	2.63	1.52	2.82	2.78	2.24
	R	3	5	5	3	5	5	4.33
Pooled	N_a	3	13	11	15	13	15	11.67
	N_e	1.80	2.99	7.50	4.79	4.15	7.28	3.73
	R	2.28	6.79	7.95	7.23	7.16	8.80	6.70

The estimates of the gene diversity H_e for the morphological groups in the adult population are summarized in **Table 7**. The average observed and expected heterozygosity (H_o and H_e), as well as the fixation index F were calculated for the morphological groups and for the pool (Pooled). The average observed heterozygosity found for Group 3 (H_o = 0.550) was higher than that in Group 2 (H_o = 0.515) or Group 1 (H_o = 0.483). H_e showed the highest mean value in Group 2 (H_e = 0.730), followed by Group 1 (H_e = 0.605), and the lowest in Group 3 (H_e = 0.554). Locus Mo16 was the most diverse locus in Groups 1 (H_e = 0.790) and 2 (H_e = 0.879), while in Group 3 Mo13 showed the highest diversity (H_e = 0.645).

Homozygote excess, exhibited by a positive F value, was observed over all loci and all groups, excepts at locus Mo09 in Group 1, and at loci Mo07, Mo09, Mo13 and Mo16 in Group 3. Mean F values were positive for all groups, with a value very close to zero for Group 3. Pooled F values (i.e., calculated by pooling data of all morphological groups together) ranged from 0.140 at locus Mo09 to 0.519 at locus Mo05. The results of the Pearson's χ^2 goodness-of-fit test for homozygotes excess over Hardy-Weinberg expectations revealed significant excess of homozygotes (heterozygotes) in Group 1 only at locus Mo05. Significant values of homozygotes excess were found at locus Mo05 in Group 3, and at all loci with the exception of locus Mo13 in Group 2.

Table 7: Gene diversity and fixation index in the morphological groups in the adult population. H_o: observed heterozygosity; H_e: expected heterozygosity; F: fixation index. *H-W*: results of the Pearson's χ^2 goodness-of-fit test for homozygotes excess over Hardy-Weinberg expectations (n.s.: not significant; *: significant at $p<0.05$; **: significant at $p<0.01$; ***: significant at $p<0.001$).

		SSR Loci						
Groups		**Mo05**	**Mo07**	**Mo08**	**Mo09**	**Mo13**	**Mo16**	**Mean**
Group1	H_o	0.100	0.400	0.600	0.600	0.600	0.600	0.483
	H_e	0.395	0.410	0.720	0.540	0.775	0.790	0.605
	F	0.747	0.024	0.167	-0.111	0.226	0.241	0.216
	H-W	*	n.s.	n.s.	n.s.	n.s.	n.s.	--
Group2	H_o	0.231	0.354	0.650	0.725	0.637	0.493	0.515
	H_e	0.436	0.702	0.863	0.824	0.676	0.879	0.730
	F	0.470	0.495	0.246	0.120	0.057	0.439	0.310
	H-W	***	***	***	*	n.s.	***	--
Group3	H_o	0.200	0.700	0.500	0.400	0.800	0.700	0.550
	H_e	0.540	0.540	0.620	0.340	0.645	0.640	0.554
	F	0.630	-0.296	0.194	-0.176	-0.240	-0.094	0.003
	H-W	*	n.s.	n.s.	n.s.	n.s.	n.s.	--
Pooled	H_o	0.214	0.394	0.630	0.680	0.650	0.526	0.516
	H_e	0.446	0.666	0.867	0.791	0.759	0.863	0.732
	F	0.519	0.408	0.273	0.140	0.144	0.390	0.312
	H-W	***	***	***	**	*	***	--

Number of observed alleles, observed and expected heterozygosity and their respective standard deviations, computed for Group 2 after 100 random resamplings performed with a sample size of ten individuals, are shown in **Table 8**.

Table 8: Allelic multiplicity, gene diversity, and their respective standard deviations computed for **Group 2** after 100 random resamplings performed with a sample size of ten (N_a: observed number of alleles; H_o: observed heterozygosity; H_e: expected heterozygosity; **SD**: standard deviation).

	SSR Loci						
	Mo05	**Mo07**	**Mo08**	**Mo09**	**Mo13**	**Mo16**	**Mean**
N_a	2.000	5.952	7.635	7.394	6.096	8.029	6.184
SD	0.000	1.280	1.071	1.477	1.153	1.226	-
H_o	0.213	0.337	0.672	0.727	0.633	0.484	0.511
SD	0.114	0.147	0.144	0.141	0.144	0.160	-
H_e	0.408	0.651	0.822	0.780	0.641	0.817	0.687
SD	0.084	0.119	0.031	0.055	0.099	0.046	-

Comparing these results to those obtained among 80 individuals of Group 2, the number of alleles showed a considerable reduction at all loci, except at locus Mo05. On the other hand, H_o and H_e showed values close to the ones estimated in the original sample in Group 2, among 80 individuals.

Total genetic differentiation between morphological groups is given in **Table 9**. In **Appendix 6** pairwise matrices of genetic differentiation between groups are presented by locus. Genetic distances, D (Nei, 1972) and d_0 (Gregorius, 1974), used to estimate genetic differentiation between morphological groups showed the highest differentiation between Group 3 and 2 (D = 0.388; and d_0 = 0.503). The lowest differentiation was observed for both distances between Group 3 and 1 (D = 0.075; and d_0 = 0.267), whereas between Group 1 and 2 the values were intermediate (D = 0.236; and d_0 = 0.426).

Table 9: Pairwise matrix of Nei's (1972) genetic distance ***D*** (below diagonal), and Gregorius' (1974) genetic distance d_0 (above diagonal), between morphological groups in the adult population based on six SSR loci.

Groups	Group1	Group2	Group3
Group1	--	0.426	0.267
Group2	0.236	--	0.503
Group3	0.075	0.388	--

Pairwise F_{st} is shown in **Table 10**. The highest differentiation occurs between Groups 2 and 3, whereas Group 1 is more similar to Group 3 than to Group 2.

Table 10: Pairwise F_{st} (below diagonal) among morphological groups in the adult population at six SSR loci, and ***p-values*** (above diagonal).

Groups	Group1	Group2	Group3
Group1	--	<0.001	0.064
Group2	0.072	--	<0.001
Group3	0.025	0.129	--

Genetic differentiation among groups was also estimated as D_j and δ (**Table 11**). Both parameters are graphically illustrated in **Figure 15** with "snail diagrams" for the morphological groups in the adult population. D_j and δ values for each locus and for the gene pool were computed with equal c_j. At all but two SSR loci (Mo05 and Mo07),

Group 1 showed the lowest differentiation values (D_j) as it can be seen in **Table 11** and in **Figure 15**.

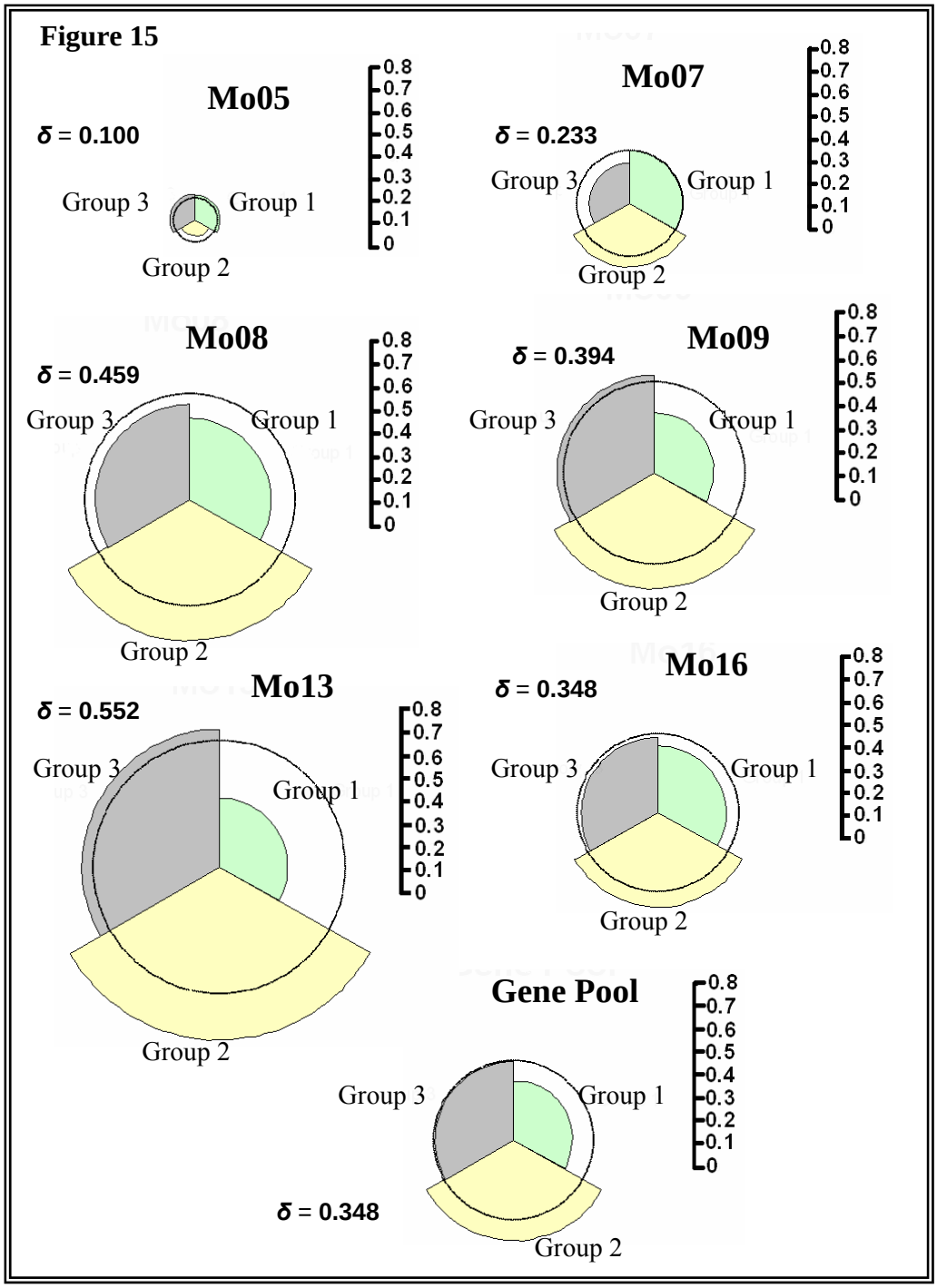

Figure 15: Graphic representation of the genetic differentiation (δ and D_j) among morphological groups in the adult population, at six SSR loci and for the gene pool. In each graphic, the circumference radius is equal to the average differentiation δ. The solid sectors (with different colors) represent the contribution of the single morphological groups to the total differentiation of the collection of groups. The radii of each sector are equal to the differentiation levels of the individual groups (D_j). The angles of the sectors represent the population weights (c_j), which are all identical in the present case (c_j=1/3).

D_j and δ estimates, as well as F_{st} are presented in **Table 11**. The estimate of δ computed with equal sample size for the gene pool (average), and over six microsatellite loci, indicates a mean genetic differentiation of 34.8%. Group 2 is the most strongly differentiated from the other groups (D_j = 0.441). The lowest D_j was observed for Group 1 (D_j = 0.258). Locus Mo13 reveals the highest genetic differentiation among groups based on δ, as well as F_{st}. The mean value of F_{st} shows an average genetic differentiation among groups of 9.7%.

Table 11: Allelic differentiation among morphological groups in the adult population: D_j, genetic differentiation; δ, average differentiation computed with equal sample size (c_j = 0.333); F_{st}, Wright's Fst; ***p-value***, level of significance for F_{st}.

Group		Group 1	Group 2	Group 3	δ	F_{st}	*p-value*
Locus Mo05	D_j	0.110	0.075	0.115	0.100	0.009	(0.622)
Locus Mo07	D_j	0.233	0.285	0.180	0.233	0.013	(0.371)
Locus Mo08	D_j	0.353	0.613	0.413	0.459	0.107	(<0.001)
Locus Mo09	D_j	0.259	0.500	0.422	0.394	0.123	(<0.001)
Locus Mo13	D_j	0.297	0.756	0.603	0.552	0.227	(<0.001)
Locus Mo16	D_j	0.293	0.418	0.333	0.348	0.029	(0.121)
Gene pool	D_j	0.258	0.441	0.344	0.348	0.097	(<0.001)

Table 12 shows the theoretical exclusion probabilities for paternity by locus, computed over all trees in the adult population. Locus Mo16 showed the highest exclusion probability for paternity, and locus Mo05 the lowest.

Table 12: Theoretical exclusion probabilities for paternity computed by locus over all trees in the adult population.

Locus	Mo05	Mo07	Mo08	Mo09	Mo13	Mo16
Paternity exclusion probability	0.186	0.485	0.733	0.615	0.589	0.736

5.4.2. Seed population

The relative allele frequencies for each morphological group in the seed population at three SSR loci are shown in **Figure 16** and in **Appendix 7**. As it was done in the adult population, the different alleles found at each locus were named as the number of base pairs of the amplified fragment.

Figure 16

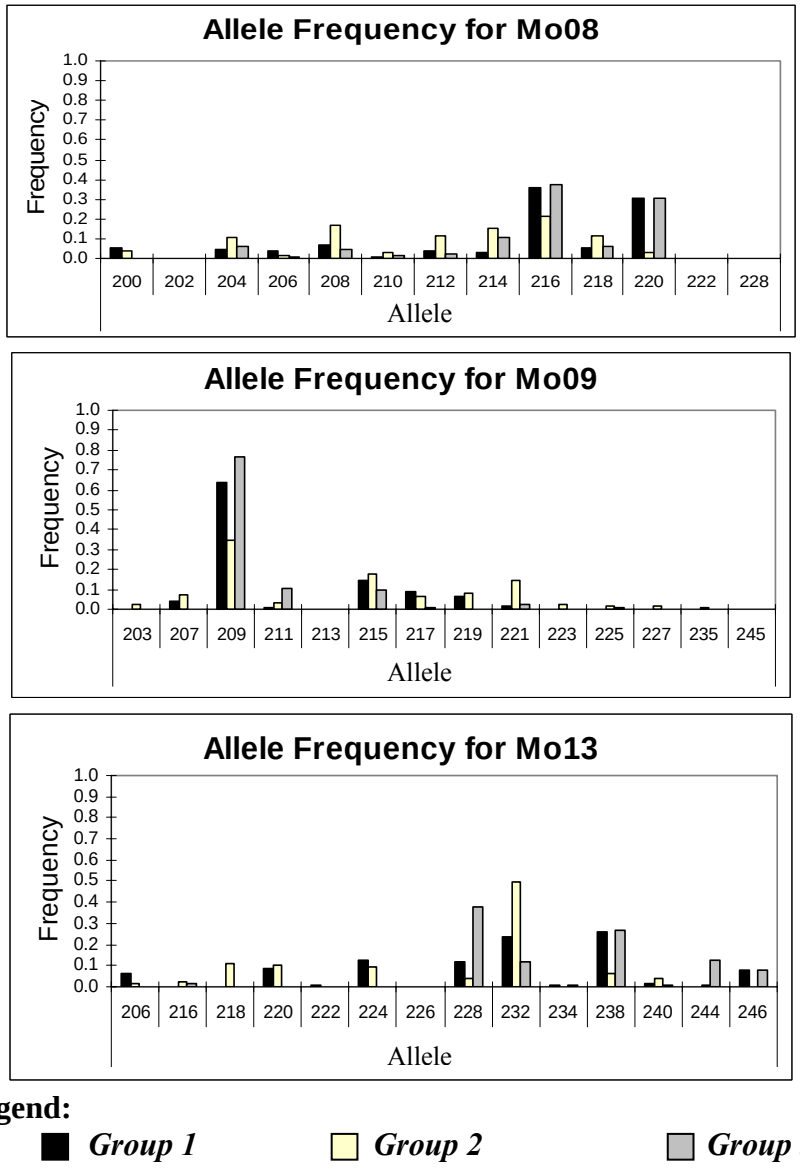

Figure 16: Allelic frequencies at microsatellite loci *Mo08, Mo09* and *Mo13* in three morphological groups analyzed (***Group 1***, ***Group 2*** and ***Group 3***) in the seed population.

Table 13 shows the allelic multiplicity within morphological groups, recorded at the three loci analyzed. A total of 41 different alleles were found in the seed population. The average number of alleles per locus (*A/L*) observed over all groups and at three loci was 13.67. Group 2 showed the highest *A/L* (13.0), followed by Group 1 (9.0), and finally Group 3 (8.33). The number of effective alleles of the pool was 4.55, ranging from 3.12 (locus Mo09) to 6.26 (locus Mo08). Allelic richness computed based on a minimum sample size of ten individuals, showed the highest mean value in Group 2 (7.15), followed by Group 1 (6.01), and the lowest in Group 3 (4.81).

The number of private alleles found in the seed population was 11. Ten were found in Group 2, two at locus Mo08, six at Mo09 and two at locus Mo13; and one was found in Group 3 at locus Mo08. Only one private allele (in Group 2 at locus Mo13) showed a relative frequency over 5%. Group 1 did not show private alleles.

Table 13: Allelic multiplicity registered at three microsatellite loci in the morphological groups (**Groups**) and in the pool (**Pooled**), in the seed population (N_a: observed number of alleles; N_e: effective number of alleles; ***R***: allelic richness; **Mean**: average value).

		SSR Loci			
Groups		**Mo08**	**Mo09**	**Mo13**	**Mean**
Group 1	N_a	10	7	10	9.00
	N_e	4.17	2.28	5.86	3.53
	R	6.51	4.64	6.89	6.01
Group 2	N_a	12	14	13	13.00
	N_e	7.16	5.28	3.49	4.87
	R	7.57	7.35	6.52	7.15
Group 3	N_a	10	6	9	8.33
	N_e	3.94	1.64	4	2.70
	R	5.78	3.39	5.27	4.81
Pooled	N_a	13	14	14	13.67
	N_e	6.26	3.12	5.57	4.55
	R	7.41	6.19	7.45	7.02

Comparing these results to those obtained in the adult population at the same loci (**Table 6**; loci ***Mo08***, ***Mo09*** and ***Mo13***), it can be seen that three new alleles absent in the adult population were scored in the seed population, two at locus Mo08 and one at locus Mo13. On the other hand, one allele found at locus Mo09 in the adult population

was not recorded in the seed population. A considerable increase in the number of alleles (N_a) for all loci was observed in the seed population for Groups 1 and 3 as compared to the adult population, although the total number of alleles per locus did not show a major rise. On the other hand, the allelic richness computed for all groups based on a minimum sample size of ten individuals did not show considerable variations across populations in any of the groups.

Table 14 summarizes the estimates of gene diversity within groups in the seed population. Calculations based on three SSR loci showed that the average expected heterozygosity (H_e) was higher than the average observed heterozygosity (H_o) for all groups and in the pool, too. Mean H_e ranged from 0.794 to 0.629, while average H_o ranged from 0.471 to 0.706. The mean fixation index (*F*) over three loci was positive, indicating an excess of homozygotes. *F* values were positive at all loci and for all groups, except at locus Mo09 in Group 1, where the fixation index was negative. Significant values of homozygotes excess were found at all loci and all groups, excluding loci Mo09 and Mo13 in Group 1, and Mo13 in Group 2.

Table 14: Gene diversity and fixation index in the morphological groups in the seed population. H_o: observed heterozygosity; H_e: expected heterozygosity; *F*: fixation index. *H-W*: results of the Pearson's χ^2 goodness-of-fit test for homozygotes excess over Hardy-Weinberg expectations (n.s.: not significant; *: significant at $p<0.05$; **: significant at $p<0.01$; ***: significant at $p<0.001$).

		SSR Loci			
Groups		**Mo08**	**Mo09**	**Mo13**	**Mean**
Group1	H_o	0.630	0.580	0.747	0.652
	H_e	0.760	0.561	0.829	0.716
	F	0.172	-0.034	0.099	0.079
	H-W	**	n.s.	n.s.	-
Group2	H_o	0.722	0.726	0.669	0.706
	H_e	0.860	0.810	0.713	0.794
	F	0.160	0.104	0.062	0.109
	H-W	***	***	n.s.	-
Group3	H_o	0.593	0.273	0.547	0.471
	H_e	0.746	0.391	0.750	0.629
	F	0.205	0.301	0.271	0.259
	H-W	***	**	***	-
Pooled	H_o	0.665	0.552	0.640	0.619
	H_e	0.840	0.680	0.821	0.780
	F	0.209	0.187	0.220	0.205
	H-W	***	***	***	-

When comparing to the adult population at loci Mo08, Mo09 and M013 (**Table 7**), a slight increment in the genetic diversity (H_e) can be observed in the seed population at loci Mo09 and Mo13 in Group 1, and at locus Mo13 in Group 2, while Group 3 showed more considerable increments of H_e at all loci. Within the pool, the diversity increased in the seed population only at locus Mo13. In the adult population, Group 3 revealed no significant excess of homozygotes over Hardy-Weinberg expectations, while in the seed population this group showed significant homozygotes excess at all loci, with high values of *F*.

Genetic differentiation among morphological groups in the seed population was estimated as D (Nei, 1972) and d_0 (Gregorius, 1974) at the three analyzed loci. The distances were compared with those found in the adult population at the same loci (**Table 15**). In the seed population, the highest levels of differentiation were observed between Group 3 and Group 2 (D = 0.479; d_0 = 0.555), and the lowest, between Group 1 and Group 3 (D = 0.084; d_0 = 0.258). When comparing across populations, a decline in D and d_0 values is observed over all groups. This decrease is higher between Group 2 and 3 as it can be seen in **Table 15**.

Table 15: Pairwise matrix of Nei's (1972) genetic distance ***D*** (below diagonal), and Gregorius' (1974) genetic distance $\boldsymbol{d_0}$ (above diagonal), between morphological groups of the adult and seed populations, at SSR loci Mo08, Mo09 and Mo13.

Adult Population			
Groups	**Group 1**	**Group 2**	**Group 3**
Group 1	-	0.575	0.333
Group 2	0.592	-	0.717
Group 3	0.086	0.904	-

Seed Population			
Groups	**Group 1**	**Group 2**	**Group 3**
Group 1	-	0.407	0.258
Group 2	0.287	-	0.555
Group 3	0.084	0.479	-

Genetic differentiation among groups the seed population was also estimated as $\boldsymbol{\delta}$, computed with equal $\boldsymbol{c_j}$, and $\boldsymbol{F_{st}}$. These parameters were contrasted with those observed in the adult population at the same loci (**Table 16**). In the seed population the highest genetic differentiation was observed at locus Mo13 ($\boldsymbol{\delta}$ = 0.481; $\boldsymbol{F_{st}}$ = 0.144). The estimate of $\boldsymbol{\delta}$ and $\boldsymbol{F_{st}}$ over three microsatellite loci in the seed population indicates a mean genetic differentiation of 36.6% and 10.5%, respectively. Reduction in the mean values of differentiation, as well as in single locus values, was observed in the seed population when comparing to the adult population.

Table 16: Genetic differentiation among groups of the adult and seed populations at three SSR loci: δ, average differentiation computed with equal sample size ($c_j = 0.333$); F_{st}, F-statistic; p, *p-value* for F_{st} after 10,000 replicates.

	Adult population				Seed population			
	Mo08	**Mo09**	**Mo13**	**Mean**	**Mo08**	**Mo09**	**Mo13**	**Mean**
δ	0.459	0.394	0.552	0.468	0.312	0.304	0.481	0.366
F_{st}	0.108	0.123	0.227	0.152	0.062	0.109	0.144	0.105
p	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

As a complement of the abovementioned parameters, genetic differentiation among groups in the seed population was also estimated as D_j (Gregorius and Roberds, 1986). **Table 17** shows D_j estimation calculated with equal sample size, together with δ. These parameters are graphically represented in **Figure 17**.

Table 17: Allelic differentiation among morphological groups in the seed population: D_j, genetic differentiation; δ, average differentiation estimated with equal sample size (Gregorius and Roberds, 1986).

Group		Group 1	Group 2	Group 3	δ
Locus Mo08	D_j	0.265	0.431	0.241	0.312
Locus Mo09	D_j	0.171	0.381	0.361	0.304
Locus Mo13	D_j	0.309	0.564	0.571	0.481
Gene pool	D_j	0.249	0.459	0.391	0.366

The estimates of genetic differentiation (D_j and δ) among groups in the seed population show a clear decline comparing to the same parameters estimated in the adult population at loci Mo08, Mo09 and Mo13 (**Table 11** and **Figure 15**). Group 2 showed the highest reduction in genetic differentiation when comparing across populations. In this group, D_j values decrease between 11.9% and 19.2% in the seed population compared to the adult population. Comparing the seed population to the adult population, average genetic differentiation declined 5.4% in Group 1, 16.4% in Group 2, and 8.8% in Group 3.

Figure 17

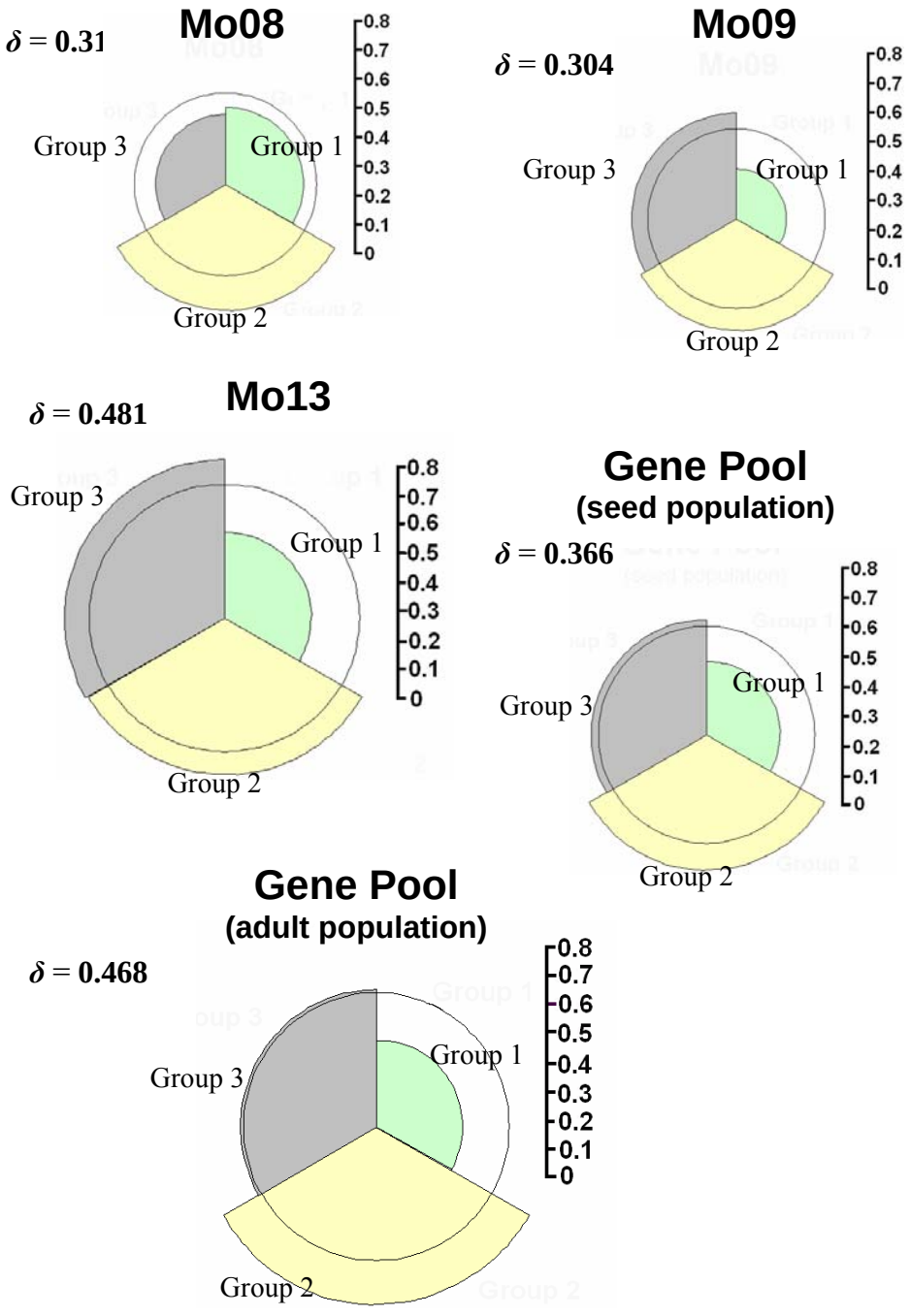

Figure 17: Graphic representation of the genetic differentiation D_j and δ among morphological groups in the seed population at SSR loci *Mo08, Mo09* and *Mo13*, and for the gene pool at the mentioned loci in the seed population and in the adult population. In each graphic, the circumference radius is equal to the average differentiation δ; solid sectors represent the contribution of the single morphological groups to the total differentiation and their radii are equal to the differentiation levels of the individual groups (D_j); angles of the sectors represent the population weights (c_j), which are all identical in the present case (c_j=1/3).

5.5. Mating system

Parameters of genetic diversity and differentiation were computed for the pooled pollen cloud of each morphological group and compared with those obtained in the adult population. Manuscript **III** presents in detail the results obtained in the mating system analysis of the hybrid swarm between *P. chilensis* and *P. flexuosa*.

Gene diversity (H_e) ranged from 0.34 to 0.86 in the adult population, and from 0.47 to 0.87 in the pooled pollen contribution (**Table 3** in manuscript **III**). The highest diversity over all loci (*mean* H_e) was detected for both populations in Group 2. Within each group, the mean genetic diversity (N_e, H_e) of the pooled pollen cloud was higher in comparison to the respective groups in the adult population.

Genetic distances D and d_0 showed the highest differentiation among Group 3 and 2 in the adult population and in the pooled pollen cloud (**Table 4** in manuscript **III**). Reductions in distance values were observed between all pairs of groups, except for D between Groups 1 and 3, when the adult population was compared to the pooled pollen contribution. Between Groups 1 and 3 D was similar when the adult population and the pooled pollen clouds were compared. The largest decrease in distance values was observed between Groups 2 and 3. In addition, total genetic differentiation among morphological groups (G_{ST} and δ), decreased for all loci when comparing the adult population to the pooled pollen cloud (**Table 5** in manuscript **III**). The genetic distance between a group and its complement ($\boldsymbol{D_j}$) was lower for the pooled pollen clouds in comparison to the adult populations for all groups and at all loci (**Table 6** in manuscript **III)**.

Genetic differentiation among pollen clouds of individual mother trees showed relative high distance values for most of the pairs analyzed (**Table 7** in manuscript **III**). However, when genetic differentiation between individual pollen clouds was compared within and among groups, no significance differences were registered in their distances.

The theoretical exclusion probability computed over all adult trees and over three loci reached 95.77 %. The highest exclusion probability for single locus was registered at locus Mo08 (73.21 %), followed by locus Mo09 (61.47 %) and finally locus Mo13

(58.86 %). For 96 offspring among the 516 analyzed, no compatible male parent was registered within the study stand. By contrast, for 78 seeds, all but one male parent could be excluded. The remaining offspring (342), showed more than one potential father. Out of these 342 seeds, 143 identified all potential fathers from the same morphological group. For the set of progenies that showed more than one potential father, in four cases the phenological analysis allowed the exclusion of one or more potential fathers. For those offspring with more than one potential father, a single father was assigned considering the one with the highest LOD Score. Among the 420 progenies that had their male parent within the stand, 170 resulted from hybridization events between groups (**Table 8** in manuscript **III**).

The results obtained from the estimation of the mating preferences showed that there is a strong preference of Group 3 to mate with individuals of the same group. Besides, this group shows preference for Group 1 but in a lower degree, and a strong repulsion towards Group 2. Group 1 shows preference to mate with individuals of the same group and also with Group 3, and repulsion towards Group 2. On the other hand, Group 2 shows preference values close to one for all groups, this indicates a trend towards random mating within this group (**Figure 3** in manuscript **III**)

The minimum estimate of the outcrossing rate (t) showed for individual trees values between 0.292 and 1.000. At group level, the minimum estimate of t for Group 1 was 0.938; for Group 2, 0.920; and for Group 1, 0.692. Outcrossing estimated by paternity analysis ranged from 0.389 and 1.000 for individual trees. Individual multilocus outcrossing rate (t_m), as estimated based on the mixed mating model ranged from 0.330 to 1.200. For Group 1 t_m was 0.979, for Group t_m = 2 0.954, and for Group 3 t_m = 0.777. For all methodologies used in the estimation, the lowest individual values of t were found among trees of Group 3. This group showed also the lowest value of outcrossing at group level estimations (**Table 9** in manuscript **III**).

Individual biparental inbreeding estimated assuming the mixed mating model, ranged from -0.088 to 0.342; while the values found by group were 0.045 for Group 1, 0.075 for Group 2, and 0.143 for Group 3. The lowest value of correlated mating found at individual trees was -0.046 in tree number 30, and the highest was 0.466 in seed tree 46.

6. GENERAL DISCUSSION

6.1. Development of microsatellite gene markers in *Prosopis*

In the present study, six polymorphic microsatellite markers were developed for the species *Prosopis chilensis*. Until today, no previous reports on SSRs isolation for this or any other *Prosopis* species were found. The strategy for microsatellite development adopted in this work, based on the protocol of Fischer and Bachmann (1998), has shown to be a relatively simple and fast alternative for SSRs isolation. This enrichment protocol requires only basic skills in molecular biology (cloning is the most difficult step), and limited laboratory equipment in addition to what is required for subsequent microsatellite screening. Besides, by starting from DNA that has already been extracted, cloned products ready to be sequenced can be obtained in about four days if the method has been already set up in the laboratory.

Three main procedures should be regarded as crucial steps for a successful SSRs isolation following the aforesaid protocol, i.e. (*i*) initial DNA quality; (*ii*) hybridization and capture; and (*iii*) cloning. DNA from which SSRs will be isolated is the starting point in microsatellite development. For this reason, high quality DNA is a basic prerequisite. It is essential to start with pure and non-degraded DNA, free of RNA, proteins, or any other contaminant including foreign DNA. Besides, for the development of microsatellite libraries it is important to have DNA from a single individual. The enrichment step, i.e. hybridization of the microsatellite probes to the restricted-ligated DNA fragments and posterior indirect capture, determines the efficiency of the protocol. Efficiency is expressed generally, as the number of clones containing SSR motifs. In this procedure, the annealing temperature used for the hybridization is a crucial factor to obtain a successful enrichment. The annealing temperature will depend on the microsatellite oligonucleotide probes used. In addition, the enrichment step can be repeated to increase efficiency. In the present work, a single enrichment was performed with high levels of efficiency. Finally, in order to be able to sequence the enriched fragments, they have to be cloned into a vector. Vector construction and posterior cloning represent the most complicated steps in the protocol applied for the development of SSR in *Prosopis*.

In the present study, the efficiency of microsatellite enrichment using the protocol of Fischer and Bachmann (1998) was 95.8%. This level of efficiency was registered considering as SSR segments those tandemly repeated units of one to six nucleotides, repeated at least three times, and of a total size of at least six base pairs. When only long-length SSR fragments (i.e., microsatellite fragments of at least five repeated units and a minimum length of ten base pairs) were considered, the level of enrichment reached 30.8%. Similar protocols for SSR isolation based on the establishment of microsatellite-enriched libraries reported levels of enrichment between 11% and 99% (Edwars et al., 1996; Paetkau, 1999; Rodriguez et al., 2002; Saneyoshi et al., 2003). However, no information on the selection criteria for the SSR segments was reported in the mentioned works. Other methods for microsatellite isolation, such as from AFLPs (Yamamoto et al., 2002), random amplified microsatellite PCR (Vander Stappen et al., 1999), RAPDs (Ueno et al., 1999) or expressed sequences tags (Woodhead et al., 2003) showed an average efficiency of enrichment of 65%.

Pandey et al. (2004), using the same protocol as in the present study to develop microsatellite markers in *Acer pseudoplatanus*, but performing two enrichments, reported a level of enrichment of 60% following the same criteria on the selection of SSR segments. Fischer (2000), using a similar protocol in *Allium cepa*, and adding a second enrichment, obtained 75% of sequences containing microsatellite motifs. Thus, the efficiency of microsatellite isolation in *Prosopis chilensis* can be considered within the normal levels of efficiency taking into account that only a single enrichment was performed, and assuming that the level of enrichment in a second enrichment procedure is similar to the one obtained in the first step. A second enrichment step after the first enrichment may help increasing the efficiency of enrichment considerably. Besides, a higher efficiency level can be achieved by using more oligoprobes for hybridization, which may increase the chance of getting more fragments with microsatellite repeats. In the present study only three oligonucleotide probes were used ($(CA)_{10}$, $(GAA)_8$, and $(AAC)_8$). In *Allium cepa* (Fischer, 2000) six different oligoprobes were used ($(CT)_{10}$, $(CA)_{10}$, $(TAA)_8$, $(GAA)_8$, $(AAC)_8$, and $(GGT)_8$), obtaining higher enrichment efficiency (15% higher) compared to the results obtained in *Acer pseudoplatanus* (Pandey et al., 2004) using only three probes ($(CA)_{10}$, $(GAA)_8$, and $(AAC)_8$).

In the current work, 16% of all detected SSR motifs showed specific complementary sequences with the oligoprobes used for the hybridization (10% with the $(CA)_{10}$ probe, 4% with $(GAA)_8$, and 1% with the probe $(AAC)_8$). In contrast to the results found in this study, Pandey et al. (2004) found that 82.6% of the microsatellite fragments detected contained complementary sequences to the oligoprobes used. Since Pandey et al. (2004) used the same probes as in the present study, the results are comparable. The considerable higher value obtained by Pandey et al. (2004) could be due to the performance of two successive hybridization steps in their procedure, while in the present work only a single hybridization was carried out. Nevertheless, the percentage of motifs with partial similarity to the oligoprobes, e.g. $(GAAA)_n$ or $(AACA)_n$, was 33%, and this level reached 68% when mononucleotide repeat motifs were included. This is a clear indication of unspecific hybridization of the oligoprobes to the template DNA, probably caused by lack of optimum hybridization conditions (e.g. annealing temperature and duration of annealing). Cordeiro et al. (1999) found that high levels of enrichment are achieved under optimal enrichment conditions. However, even small deviations from the optimal wash temperature and/or wash buffer salt concentrations can significantly affect the levels of enrichment.

The most abundant microsatellite repeats, representing 46.6% of the sequences with SSRs, were mononucleotide repeats, followed by dinucleotides with 40.5, trinucleotides with 11.1%, and finally motifs with more than four repeats with 1.8%. The percentage of dinucleotide repeats detected was higher as compared to trinucleotides, even though two trinucleotide probes and only one dinucleotide probe were used for the hybridization. Higher proportion of dinucleotide repeats than trinucleotide was also reported in *Araucaria cunninghanii* and *Pinus elite* (Scotti et al., 1999), and in *Acer pseudoplatanus* (Pandey et al., 2004). These results contrast to those found by Morgante et al. (2002). They found that trinucleotide microsatellite repeats are almost twice more frequent than dinucleotide repeats in plant genomes in studies performed in databases of *Arabidopsis*, rice, soybean and wheat. The lower level of trinucleotide SSR in the aforesaid forest tree species as compared to agricultural crops could be explained by different composition in sequences of genomic DNA in the studied plant species (Pandey, 2005), or by different sensitivity of the enrichment protocol to capture di- or trinucleotide repeapts.

In the present study, the search for homology and similarity performed comparing all sequences containing SSR motifs against the EMBL Nucleotide Plant Sequence Database (EMBL-Bank) showed that 88.7% of the sequences conform to nontranscribed regions of different plant species' genomes. Only 7.8% of the sequences had high similarity with chloroplast regions, the majority of these sequences contained mononucleotide repeated SSRs. Finally, a very low proportion of sequences containing SSRs, merely 3.5% (four sequences), showed high similarity with transcribed regions. One out of four sequences was located in an intron region, and three were in translated exon regions with suspected function in plant growth regulation. Recent studies (Morgante et al., 2002) claim that most SSR motifs in plants are located in transcribed regions, but results presented in this work do not support these findings. The identification of selected (nonneutral) SSR loci with similarity to known genes could be useful to test association of microsatellite alleles with adaptive trait variation in natural populations, for mapping studies and for the future development of QTL markers. On the other hand, SSRs in nontranscribed regions are particularly useful to reconstruct descent in studies of the reproduction system of plants. However, in the present work, no primers were developed for the four SSR regions that showed high similarity with transcriber regions.

By virtue of their high polymorphism, microsatellite loci are considered ideal markers for paternity analysis, gene mapping, plant breeding, conservation biology, and population genetics (Rafalski and Tingey, 1993; Gupta et al., 1996; Jarne and Lagoda, 1996). However, because DNA sequence knowledge is required to design appropriate primers for the PCR assay, the development and application of SSRs in plants has mostly been restricted to comparative few species (Peakall et al., 1998). It might be possible to facilitate more widespread use of SSRs in plants if we are able to transfer SSR loci across species. Cross-species transfer of microsatellite markers of genomic origin has been demonstrated in many genera. Peakall et al. (1998) demonstrated widespread cross-species amplification within *Glycine*. Transfer of SSR markers within genera of fruit trees (*Citrus*, *Malus*, *Pyrus*, *Prunus*) and timber species (*Quercus*, *Pinus*, *Acer*) have also been described (Kijas et al., 1995; Pierantoni et al., 2004; Wünsch and Hormaza, 2002; Isagi and Suhandono, 1997; González-Martínez et al., 2004; Pandey et al., 2004). Transfer rates are variable among species, for example, 17% to 100% in

Pinus (Echt et al., 1999), 22% to 78% in *Quercus* (Isagi and Suhandono, 1997), and 59% in *Prunus* (Cipriani et al., 1999) have been reported.

In this study, six SSR primer pairs developed in *Prosopis chilensis* were successfully transferred to four other *Prosopis* species (*P. alba*, *P. cladenia*, *P. flexuosa* and *P. nigra*). In addition, five out of six primers showed highly robust amplification in other two *Prosopis* species (*P. hassleri* and *P. ruscifolia*). All the aforesaid species, including *P. chilensis*, belong to the Algarobia Section. The results for cross-species amplification in species of other sections did not show such high levels of transferability. When the primers were tested for cross-species amplification in *P. torquata*, only four out of six primers showed amplification. Furthermore, in *P. ferox* only three primers amplified. These two species belong to the section Strombocarpa.

6.2. Morphological assignment

The taxonomy of the *Prosopis* genus is complicated because of the high morphological similarities found among certain species and the profuse interspecific hybridization that creates intermediate phenotypes (Bessega et al., 2000b). Although the Argentinean species can be taxonomically characterized using systematic keys (Burkart, 1940 and 1976), it is not possible to clearly identify hybrids. In a recent study Vega and Hernández (2005) assessed the occurrence of interspecific hybridization among *Prosopis* species in the Argentinean Chaco region. They used Burkart's systematic keys (Burkart, 1940 and 1976) to classify taxonomically individual trees. Adult trees, belonging to pure lines of *Prosopis alba*, *P. nigra* and *P hassleri*, were easily characterized using Burkart's systematic keys. On the other hand, the keys were tested on putative hybrids individuals, which were found not to conform to recognized species.

Within the *Algarobia* Section, a large number of two- and even three-way hybrids have been postulated and confirmed with a variety of morphometric, enzymatic and molecular studies in Argentina (Palacios and Bravo, 1981; Hunziker at al., 1986; Verga, 1995; Vega and Hernández, 2005) and in Mexico (Almanza et al., 1992). In this Section, conventional taxonomy based on classical systematics, and traditional methods of numerical taxonomy are able to unambiguously assign to one species or another *pure* individuals, i.e. individuals that exhibit clearly defined species morphological characteristics. However, the frequent interspecific hybridization generates a continuum

of intermediate individuals between the pure species, and clear boundaries between taxonomic species are lost. In this case, such hybrids can not be unequivocally characterized by traditional methods, and new classification procedures are needed.

In the present study, morphological characterization of individual trees was performed applying the "Morphological Distance" (d_m) method proposed by Verga (1995; Verga and Gregorius, in press.). With this procedure, in a sympatric area between *P. chilensis* and *P. flexuosa*, three distinct groups were identified, Groups 2 and 3 mainly representing the species *P. flexuosa* and *P. chilensis*, respectively, and Group 1 predominantly with intermediate forms. The "Morphological Distance" (d_m) is specified as the proportion of elements (morphological characters) which two collection (groups) do not share. The main difference with the traditional classification method is that the characterization is performed without any *a priori* grouping assumption. Clusters are generated from direct comparison between individuals, considering each of them as *"real"* biological unit (OUT = Operational Taxonomic Unit). In addition, in the "Morphological Distance" the differences are defined by means of the comparison of the frequency distributions of the character within each OUT.

Traditional taxonomic methods, e.g. taxonomic distance or the key for identification of hybrids developed by Verga (2000), define the differences by means of distances between the average expressions of the metric character. Thus, d_m method considers each individual data collected for the comparison between OTUs, and not even one peace of information is missed in the analysis. On the other hand, in the traditional methods that compare average values, all the information generated from the intra-OUT variability is missed (Verga and Gregorius, in press.).

In the present study, discrepancies were observed between the two methods used in the classifications of individuals. Among 80 individuals clustered together by the d_m method, 79 were classified taxonomically as *P. flexuosa*, and one as hybrid. The taxonomic classification of individuals within Group 1 showed that four individuals are *Prosopis flexuosa* and six putative hybrids. By the d_m method a clear separated cluster, Group 3, can be identified. According to the taxonomic classification this group is constituted by two *P. chilensis* individuals and eight putative hybrids similar to *P. chilensis*.

6.3. Characterization of genetic structure and variation

The genetic study showed that in the adult population, the highest average number of alleles per locus, estimated over six SSR loci, was registered in Group 2. The number of alleles found in this group was significantly higher than the number of alleles found in Groups 1 and 3 at all loci, except at locus Mo05. This locus presented only two different alleles among individuals of Group 2, and three alleles between individuals of Groups 1 and 3. However, since sample sizes between the groups are very heterogeneous (Groups 1 and 3 ten individuals each, and Group 2 80 individuals), these differences can be related to unequal sample size among groups. In this sense, allelic richness is a more appropriate parameter to compare allelic multiplicity between groups. Considering this parameter, the highest average allelic richness was registered in Group 2. This group showed also the higher values of allelic richness for loci Mo07, Mo08, Mo09 and Mo16. However, even when Group 2 showed the highest allelic richness at almost all loci, the differences with the other groups was not as high as the differences registered for the number of alleles.

The levels of genetic diversity in Group 2 were higher compared to the other two morphological groups. Group 2 showed also higher diversity values also when genetic diversity was computed after 100 random resamplings performed with a sample size of ten individuals. This shows that even when there is a different sampling effect on the estimations of genetic diversity parameters, Group 2 shows higher diversity in comparison with the other two groups.

Genetic differentiation among groups of the adult population reflects partially the results obtained in the morphological classification. In agreement with the morphological analyses, the highest genetic distance is found between Groups 2 and 3, which represent, respectively, the pure species *P. flexuosa* and *P. chilensis*. On the other hand, the lowest genetic differentiation is observed among Groups 1 and 3, while the morphological analysis showed that Groups 1 and 2 are more similar morphologically. Discrepancies between morphological and molecular data have also been observed in other studies with *Prosopis* species (Saidman and Vilardi, 1987; Ramírez et al., 1999).

In the adult population, overall Nei's genetic distance (*D*) between Groups 2 and 3 was relatively high (0.388) in comparison to the results found so far of differentiation among

species of the Section Algarobia based on isozyme and RAPD markers (Saidman et al., 2000; Bessega et al., 2000b; Saidman and Vilardi, 1987). In the present work, the relatively high levels of genetic distance found between Groups 2 and 3 are mainly due to the very high values reached at locus Mo13 (D = 2.435) and at locus Mo08 (D = 0.760). These two loci reflect a clear differentiation among groups, ranging and even exceeding the values proposed by Ayala et al. (1974) to distinguish taxonomic species. Furthermore, locus Mo13 exhibited between Groups 2 and 3 a higher distance than that reported by Verga (1995) between *Prosopis chilensis* and *P. flexuosa* at locus *ADH-A*. This isozyme locus is considered a diagnostic locus for the differentiation of *Prosopis chilensis* and *P. flexuosa*. In his study, Verga (1995) found among pure populations of *Prosopis chilensis* and *P. flexuosa* a genetic distance (d_0) of 0.895 at the isozyme locus *ADH-A*. In the present work, between Groups 2 and 3, d_0 reached 0.906 at locus Mo13 and 0.650 at locus Mo08. An analysis of the allele frequencies at the mentioned loci revealed a significant difference in the allelic frequency distribution among Groups 2 and 3. Consequently, these loci, especially locus Mo13, could be considered diagnostic loci for the differentiation among the species *Prosopis chilensis* and *P. flexuosa*. However, more investigations at these loci are necessary in a wider geographical scale and also in "pure" populations to prove that they are truly diagnostic loci. On the other hand, due to the fact that the differentiation is based on differences in the allelic frequencies and not in the presence of a single diagnostic allele, assignment of individual trees to one species may be difficult.

The morphological analysis characterized Group 1 as an intermediate morphological group between the pure species *Prosopis chilensis* and *P. flexuosa*. The genetic analysis revealed that Group 1 showed the lowest among-groups genetic differentiation ($\boldsymbol{D_j}$) at four of the six studied loci, and for the gene pool, too. In addition, the allelic frequencies at loci Mo08 and Mo13, mentioned before as putative diagnostic loci, were intermediate between the frequencies of Groups 2 and 3. These results support that the intermediate nature of Group 1 has a genetic basic.

In the adult population, as well as in the seed population, the within group genetic diversity (H_e) registered in this study showed much higher values than those reported in previous studies for *Prosopis*, and based on isozyme markers. According to Saidman et al. (2000), among species of Section Algarobia the mean heterozygosity found, based

on isozyme gene loci, was 0.21. This observation is not surprising, since SSRs show higher levels of polymorphism than isozyme loci for most species. Comparing within populations, Group 3 showed the lowest diversity and Group 2, the highest in adults and progenies. This might be considered an effect of the larger sample size in Group 2. However, the highest level of diversity in Group 2 are supported also by the allelic richness and in the adult population, by the computation of genetic diversity parameters after 100 random resamplings performed within Group 2 with a sample size of ten individuals.

Considering Group 3 representative of *Prosopis chilensis* and Group 2 of *P. flexuosa*, these results agree with early studies based on isozyme, revealing higher diversity in populations of *P. flexuosa* than *P. chilensis* in the Argentinean *Chaco Árido* (Verga, 1995). The lower diversity levels found in *P. chilensis* compared to *P. flexuosa* could be a result of a speciation process, such as a bottleneck or high selection pressure, suffered by *P. chilensis*. As a result, this species became more specialized, with relative higher ecological requirements and a restricted ecological niche, with populations limited to small isolated areas where the water supply is higher. Thus, its low variability can be considered a consequence of this high specialization. In contrast, *P. flexuosa* have a wider ecological niche and explore an ample variety of habitats. Therefore, a higher genetic variation is expected in *P. flexuosa* in order to adapt to a wider range of ecological requirements.

In the adult population, in all but one group (Group 3), the mean fixation index (*F*) was positive. Group 1 showed positive *F* values at all loci, except at locus Mo09. In Group 2 *F* was positive at all loci, and in Group 3 only at loci Mo05 and Mo08. However, significant homozygote excess over Hardy-Weinberg expectations was detected in Group 2 at all loci, except Mo13, and in Group 3 at locus Mo05. On the other hand, in the seed population the fixation index was positive at all analyzed loci, and over all groups, with the exception of locus Mo09 in Group 1. Homozygote excess was not significant in Group 1 at loci Mo09 and Mo13, and in Group 2 at locus Mo13. In all other cases, highly significant homozygotes excess was detected. These results indicate a general trend towards homozygote excess within groups, especially in the seed population.

As a general feature, a significant excess of homozygotes within populations were also found in all populations of species of the Algarobia Section so far studied (Saidman 1985, 1986, 1988a, 1990, 1993; Saidman and Vilardi, 1987, 1993; Saidman et al., 1997, 1998a; Verga, 1995; Bessega et al., 2000a, b). Such excess of homozygotes might be due to a certain degree of endogamy cause either by a family structure within the populations, or by some degree of selfing. In the present study, excess of homozygotes should be attributed mainly to selfing, especially in Group 3. This group showed highly significant excess of homozygotes in the seed population, while in the adult population no significant excess of homozygotes was detected. Moreover, in the study of the mating system this group showed high estimates for the selfing rate, with values as high as 67% in individual tree estimations. However, in Groups 1 and 2, the selfing rates explain only partially the homozygote excess. Therefore, the hypothesis supporting population substructure as a cause of homozygote excess can not be excluded. Other causes for positive fixation index (*F*) values can be the presence of undetected null alleles (Nascimento et al., 2005). This is particularly acceptable in Group 2, composed by *Prosopis flexuosa*, considering that the SSR primers were originally developed for *P. chilensis*. The appearances of null alleles have been reported in cross species transferability of SSR primers have shown, even between closely related species (Vornam et al., 2004). Another factor affecting the homozygotes/heterozygotes equilibrium can be the non-random sampling of the see-population and the limited number of seed trees investigated in the population.

The reduction in the excess of homozygotes in the adult population compared to the seed population suggests selection against homozygotes between the seedling stage and maturity. This selection would be responsible for maintaining relative high levels of variation within the groups.

Comparisons across populations (adult population and seed population) were performed considering only the three loci analyzed in the seed population (Mo08, Mo09 and Mo13). For this reason, the following discussions are based on the analyses of the mentioned loci. The total genetic differentiation among groups (measure as F_{st} and δ) and the genetic distances (D and d_0) between them, decrease in the seed population in comparison to the adult population. Hybridization and introgression might be considered plausible explanations for the decreased differentiation between groups in

seedling. These processes might be important in homogenizing allelic frequencies and reducing genetic differentiation between species. The results of the paternity analysis performed in the study of the mating system support this hypothesis, revealing 40.5% of *inter-group* matings. This reveals a virtual lack of clear reproductive barriers between the species *P. chilensis* and *P. flexuosa*. Post-zygotic selection mechanisms are expected in order to maintain certain morphological discontinuities between the groups. According to the reduction observed in the distances in one generation, the identity of each group would be lost in a few generations. Furthermore, hybridization was not shown to be able to erode morphological discontinuities even though it has been widely described between species of the Section Algarobia.

Several approaches to determine if hybridization does or does not imply effective gene flow among species of the Algarobia Section have been proposed (Saidman and Vilardi, 1987, 1993; Montoya et al., 1994; Bessega et al., 2000b). Montoya et al. (1994) studied sympatric and conspecific populations of *P. alba*, *P. flexuosa*, *P. nigra* and *P. ruscifolia* by means of Wright's F_{st} (Wright, 1951). They found insignificant gene flow among species, even for sympatric or neighboring ones. Other approaches arrived at the same conclusion. Even though indirect estimates of gene flow are no reliable measures of current migration rates (Bossart and Pashley-Prowell, 1998; Whitlock and McCauley, 1999), they indicate that interspecific gene flow is relatively low. Considering that premating isolation mechanisms are not well developed in some species of the Algarobia Section, the lack of effective gene flow among these species reported in other studies suggests the existence of more efficient postmating isolation mechanisms maintaining species boundaries within definite limits. Disruptive selection against intermediates can be postulated as a possible postmating mechanism for maintaining discontinuities between species. In this sense, in a previous study on hybrids swarms of *P. caldenia* and sympatric species, Saidman (1988b) found that despite the evidence indicating that hybridization does occur between *P. caldenia* and other two *Prosopis* species in sympatric zones, no adult hybrids were found in the study site. This suggests that plants from hybrid seeds do not reach the adult stage. The prevention of development of hybrid seeds or seedlings involving the species *P. caldenia* would represent a postzygotic isolation mechanism developed to maintain this species integrity.

6.4. Mating system analysis

Gene flow estimations based on parental analysis allow the estimation of effective gene movement because the method is sensitive to all biological factors involved in the mating success: viability of the pollen; compatibility between donor and receiving plants; competition among different pollen tubes. In the present study, paternity assignment was conducted by categorical allocation based on multilocus genotypes at three microsatellite loci. These loci showed a high theoretical multilocus exclusion probability (95.77 %). The highest exclusion probability for a single locus was registered at locus Mo08 (73.21 %), followed by locus Mo09 (61.47 %) and finally locus Mo13 (58.86 %).

From 100 adult trees in the adult population, only five pairs of individuals were observed with matching multilocus genotypes at the three studied loci. This reveals the hypervariable nature of these markers, and their high level of applicability in fingerprinting. However, unambiguous parental assignment, i.e. the exclusion of all but one potential father, was possible in only 15.12% of the cases. For 18.60% of the seeds no pollen donor was found within the studied stand. On the other hand, 66.28% of the seeds showed more than one potential father. The maximum number of potential fathers registered for a single offspring was 12. A possible explanation to the low values of unambiguous parental assignment can be the frequency distribution of alleles. Even when the number of alleles per locus is relatively high, there are a reduced number of alleles with high frequency, and many alleles show low frequencies. This causes a reduction in the effective number of alleles. In this sense, to yield a single nonexcluded parent more SSR loci would be needed.

Among the methods of parentage analysis, exclusion is the simplest techniques. This is a very interesting method because exclusion of all but one parent for each offspring in a population could be considered as the ideal of parental analysis. Nevertheless, this ideal is very difficult to achieve. As it was discussed above, in the present work the number of cases where all but one parent was excluded were low. When total exclusion is not possible, other methodology must be used to assign a progeny to a single parent among a group of nonexcluded parents. Categorical allocation is a method of parentage analysis that assigns progeny to nonexcluded parents based on likelihood scores derived

from their genotypes (Jones and Ardren, 2003). Categorical allocation is a reliable approach for parental assignment. In this study, for those offspring that showed more than one potential father, the most likely father was assigned based on the likelihood ratio computed by categorical allocation.

The results of the paternity analysis show that there is 40.48% of *inter-group* hybridization among 420 mating events studied. Considering the pure species *P. flexuosa* and *P. chilensis* represented, respectively, by the morphological groups 2 and 3, the results of the paternity analysis show that interspecific hybridization occurs. Interspecific pollen flow is asymmetric, and it can be assumed that Group 2 receives more pollen from Group 3 than the opposite. However, even when hybridization occurs, there is a preference to mate with individuals of the same type, especially in Group 3. Besides, Group 1, intermediate individuals, mates also with individuals of Groups 2 and 3, but with a higher mating rate with this last group.

It is possible to hypothesize that under natural conditions the two species, *P. chilensis* and *P. flexuosa*, coexist in the Chaco Árido region under equilibrium, even when gene flow occurs in sympatric areas. However, under certain environmental conditions interspecific hybridization raises and introgression occurs. In this sense, modifications of the environment as a result of human activities or climatic phenomena, such us intensive droughts, floods or climatic changes, produce open habitats favoring hybridization and introgression (Anderson, 1949). The continuous production of hybrids might favor the occurrence of "evolutionary experiments" through the production of novel genetic combinations (Grant, 1981). These novel genetic combinations frequently display special capabilities to exploit open or hybrid habitats. The possibility to hybridize gives a group of related species the ability to occupy new habitats and to persist in degraded and changing environments.

7. CONCLUSIONS

The present study is the first report on microsatellite markers development for the genus *Prosopis*. In addition, it is the first documentation on the use of nuclear microsatellites for the genetic analysis in *P. chilensis* and *P. flexuosa*, two important arboreal species in the Argentinean *Chaco Árido*. The microsatellite gene markers developed in the present work showed high levels of polymorphism. Moreover, strong across species transferability was found when they were tested on seven other *Prosopis* species. Owing to these characteristics, they may represent a very useful molecular tool with applications in many different fields of study, and in a wide variety of *Prosopis* species. In this sense, further studies on transferability of the developed markers to other *Prosopis* species, as well as their characterization, are required to extend their application to other species of the genus.

In a hybrid swarm located in a sympatric area between *Prosopis chilensis* and *P. flexuosa* in the Argentinean *Chaco Árido*, three morphological groups could be identified, two representative for the taxonomic species *P. chilensis* and *P. flexuosa*, and another group with intermediate phenotypes between the pure species. The genetic analysis based on the developed microsatellite markers shows that the groups varied also in their genetic structures. Whereas the morphological groups representing the "pure" species are clearly differentiated among them at the studied microsatellite loci, the intermediate group shows an intermediate genetic structure.

The analysis of the mating system confirms the occurrence of interspecific hybridization between *P. chilensis* and *P. flexuosa*. In addition, based on the results achieved in the mating system analysis, a hypothetical model of gene flow within the studied hybrid swarm can be postulated. **Figure 18** represents a graphical illustration of this model. The lack of clear prezygotic isolation mechanisms between these species makes possible a strong gene flow among them and the formation of interspecific hybrids. Gene flow is asymmetric, *P. flexuosa* seems to receive more pollen from *P. chilensis* than the opposite. While *P. chilensis* shows a more define isolation with regard to *P. flexuosa* and clear assortative mating, *P. flexuosa* shows a trend toward random mating within the stand. In addition, hybrids are fertile and by repeated backcrossing with the parental species they incorporate genes of one species into the gene pool of the other

species. This process of introgression is also asymmetric. A higher mating rate of intermediate individuals with *P. chilensis* generates greater levels of introgression towards this species. However, backcrossing occurs also between hybrids and *P. flexuosa*.

The intermediate group (hybrids) represents a very interesting aspect of this complex system. Even when matings occur in all directions, hybrids show also some degree of isolation, especially towards *P. flexuosa*. In this sense, a group of hybrids with genetic isolation mechanisms would represent very interesting material for breeding programs. Individuals included in such groups, could be used to restitute plant covered in degraded areas.

Figure 18

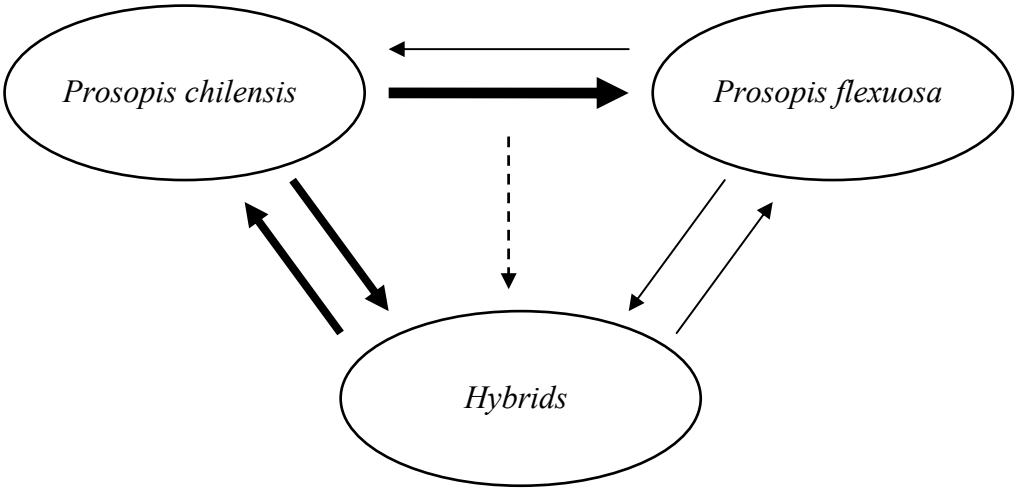

Figure 18: Hypothetical model of gene flow dynamics in the studied hybrid swarm. Solid arrows represent gene flow among groups, their width, the gene flow intensity; the dash arrow symbolizes the process of hybrids formation.

Deeper studies are required in order to explain better the prezygotic mechanisms that allow partial isolation shown by *Prosopis chilensis* and the hybrids. This involves more complete studies on the phenology, considering factors concerning the dynamics of flower phenology, such us the length of the overlapping period or flowering intensity. In addition, studies regarding the behavior of pollen vectors, or possible incompatibility mechanisms should be considered.

8. ABSTRACT

The *Prosopis* genus (family Fabaceae) occurs worldwide in arid and semi-arid regions. It includes 44 species involving trees and shrubs. Argentina constitutes the main center of diversity for the genus *Prosopis*, with approximately 28 species. In Argentina, the phytogeographic provinces of the Chaco, Monte, and Espinal cover over one million square kilometers. Within these phytogeographic provinces, *Prosopis chilensis* (Molina) Stuntz and *Prosopis flexuosa* DC., two hardwood arboreal species, constitute important natural resources due to their multipurpose nature. These trees provide timber, fuel wood and forage. In addition, they stabilize the soil and prevent erosion. *Prosopis chilensis* and *P. flexuosa* play a vital role in the ecology and the economy of many arid and semi-arid zones in Argentina, being an important resource of the local populations.

In spite of their economic and ecological importance, the exploration of the genetic resources of both *Prosopis* species is in an early stage, and there is little information about the genetic structure and reproductive system of the species. Highly polymorphic molecular markers are indispensable for the study of the genetic resources of these species, especially for the study of the reproduction system. Microsatellite markers (SSRs = Simple Sequence Repeats) offer a great potential for the analysis of gene flow and mating system due to their usually high polymorphism and co-dominant mode of inheritance.

The aim of the present study was first to develop microsatellite markers for the genus *Prosopis*; and secondly to use the developed markers in a study of the reproduction system of a hybrid swarm between *Prosopis chilensis* and *Prosopis flexuosa*.

In order to develop microsatellite markers, SSR loci were isolated for the species *Prosopis chilensis* using an enrichment protocol developed by Fischer and Bachmann (1998). This method is a non-radioactive strategy for SSRs isolation based on the establishment of microsatellite-enriched libraries using microsatellite-oligonucleotide probes and magnetic capture of the fragments containing the SSR motifs. Two trinucleotide ($(GAA)_8$, $(ACC)_8$) and one dinucleotide ($(CA)_{10}$) probes were used for the hybridization to the genomic DNA of *Prosopis chilensis*. One enrichment step was

performed. The enriched fragments were directly ligated into a vector and transformed into competent *Escherichia coli* cells. From the SSR-enriched library, 120 positive colonies were sequenced. The average length of the sequenced fragments was 580 bp (± 96 bp). Out of 120 sequences, 18 were selected for primer development considering those sequences with SSRs with at least five repeats. All primer pairs were tested and nine out of 18 amplified products of the expected size. These nine primers were labeled with a fluorescent dye at its 5'-end, and variation was characterized in a natural population of *Prosopis chilensis* (20 trees) and *P. flexuosa* (20 trees) of the Argentinean Arid Chaco. Six out of nine primer pairs amplified reproducible and well-scorable bands which were polymorphic in both species.

The novel SSR markers developed showed a Polymorphism Information Content (PIC) between 0.14 and 0.70 in *P. chilensis* and, between 0.41 and 0.85 in *P. flexuosa*. The number of alleles (N_a) varied from two to six, and from two to 13 in *P. chilensis* and *P. flexuosa* respectively. Expected heterozygosity (H_e) ranged from 0.14 to 0.73 in *P. chilensis*, and from 0.46 to 0.86 in *P. flexuosa*. A test for cross-species amplification in seven other *Prosopis* species revealed a broad cross-species affinity, all primers amplified in at least five species. The number of alleles found varied from one to five.

The results of the enrichment strategy used for the development of SSR markers in *Prosopis,* was evaluated. An analysis of the 120 clones sequenced showed that 95.8% (115 out of 120) had at least one SSR motif matching the search criteria (i.e.: microsatellites fragments composed of tandemly repeated units of one to six nucleotides, repeated at least three times and of a total size of at least six bases). When regarding SSR segments of at least five repeated units and a minimum length of ten base pairs, the level of enrichment was 30.8%. Sixteen percent (84 fragments) of all detected motifs showed specific complementary sequences with the oligoprobes used in the enrichment. A comparison of the 115 sequences containing SSR motifs against the EMBL Nucleotide Plant Sequence Database revealed that 7.8% of the clones (9 clones) had high similarity with chloroplast regions. Most of the clones (102 clones) with similarity to nuclear DNA conform to non-transcribed regions of different species. Only 4 clones showed high similarity with transcribed regions.

The developed SSR markers were applied to study the genetic variation, genetic structure and aspects of the reproductive system of a hybrid swarm between *Prosopis chilensis* and *P. flexuosa* located in the Provincial Park and Forest Natural Reserve Chancaní, in the west of Córdoba province, Argentina. The study was carried out in a 4700 m^2 plot, included in a continuous forest located southeast of the aforesaid reserve. In the studied plot, all *Prosopis* trees which flowered during the flowering season 2003-2004 were mapped. A total of 100 flowering *Prosopis* trees (*adult population*) were identified during that period. Morphological assignment was performed by means of qualitative and quantitative morphological characters from leaves and fruits. Three morphological groups were obtained: *Group 2* with individuals belonging to the taxonomic species *Prosopis flexuosa*; *Group 3* with individuals with characteristics of *P. chilensis*; and *Group 1* with morphologically intermediate individuals between both species.

Total DNA was extracted from fresh leaves of each individual tree of the adult population. Genetic structure and variation was characterized in the adult population using the six novel SSR markers. Population genetic parameters were calculated. The total number of alleles over all groups and at six SSR loci was 70, with an average number of alleles per locus of 11.67. Allelic richness showed the highest mean value in Group 2 (6.64), followed by Group 1 (5.33), and the lowest in Group 3 (4.33). The estimate of the gene diversity (H_e) showed the highest mean value in Group 2 (H_e = 0.730), followed by Group 1 (H_e = 0.605), and the lowest in Group 3 (H_e = 0.554).

For the study of the reproduction system, twenty three open-pollinated mother trees were selected in the adult population. Seeds were harvested during 2003-2004. A total of 516 seeds (*seed population*) were collected. Seed population was composed of 23 sets of seeds divided into groups according to the morphological classification of the mother trees. DNA was isolated from cotyledons after germination. Population genetic parameters were calculated at three microsatellite loci in the seed population and in the effective pollen cloud, and compared with the results obtained in the adult population. A total of 41 different alleles were found in the seed population. Three new alleles absent in the adult population were scored. Genetic parameters used to estimate genetic differentiation between morphological groups showed the highest differentiation

between Group 3 and 2 in the seed population as well as in the effective pollen cloud. The lowest differentiation was observed between Group 3 and 1. Reduction in genetic differentiation was observed over all groups comparing the seed population and effective pollen cloud to the adult population. Group 2 showed the highest reduction in genetic differentiation parameters when comparing across populations.

Paternity analysis was used in order to asses more precisely the gene flow between morphological groups. Paternity assignment was conducted by categorical allocation based on multilocus genotypes. Potential fathers whose flowering period did not overlap with that from the mother tree were excluded. For 96 offspring among the 516 analyzed, no compatible male parent was registered within the studied plot. By contrast, for 78 seeds all but one male parent could be excluded. The remaining offspring (342), showed more than one potential father. For these offspring, a single father was assigned considering the one with the highest LOD Score. Among the 420 progenies that had their male parent within the stand, 170 resulted from hybridization events between groups According to this result the degree of *inter-group* hybridization reaches 40.5%.

A further study of the mating system showed that there is a strong preference of Group 3 to mate with individuals of the same group. In addition, this group shows preference for Group 1 but in a lower degree, and a strong repulsion towards Group 2. Group 1 shows preference to mate with individuals of the same group and also with Group 3, and repulsion towards Group 2. On the other hand, Group 2 shows preference values close to one for all groups, this indicates a trend towards random mating within this group.

Results point towards strong gene flow between morphological groups. A considerable reduction in the genetic differentiation among morphological groups was observed in just one generation (from the adult population to the seed population). The mating system analysis indicates that gene flow is asymmetric. For the formation of intermediate individuals (Group 1), Group 2 receives more pollen from Group 3 than the opposite. Hybridization is accompanied by introgression. A higher mating rate of intermediate individuals with Group 3 generates greater levels of introgression towards this group. However, backcrossing occurs also between Group 1 and 2.

The outcrossing rate (t) was estimated for the flowering period 2003-2004 using three highly polymorphic SSR loci. Different methodologies were applied in the estimation of outcrossing, i.e., minimum estimate of t based on foreign alleles, paternity analysis, rare alleles, and based on the mixed mating model. The minimum estimate of t showed individual values between 0.292 and 1.000. At a group level, the minimum estimated of t for Group 1 was 0.938; for Group 2, 0.920; and for Group 3, 0.692. Values of outcrossing estimated by paternity analysis ranged from 0.389 and 1.000. Individual multilocus outcrossing rate (t_m), as estimated based on the mixed mating model, ranged from 0.330 to 1.200. Multilocus outcrossing rate for Group 1 was 0.979, for Group 2 0.954, and for Group 3 0.777. For all methodologies used in the estimation of the outcrossing rate, the lowest individual values were found among trees of Group 3. Individual biparental inbreeding ranged from -0.088 to 0.342; while the values found by group were 0.045 for Group 1, 0.075 for Group 2, and 0.143 for Group 3. Correlated mating showed relatively low values, suggesting that individuals within progeny sets are mainly half rather than full sibs.

9. ZUSAMMENFASSUNG

Die Gattung *Prosopis* (Familie Fabaceae) kommt weltweit in Halbtrocken- und Trocken-Gebieten vor. Sie umfasst 44 Baum- und Straucharten. Mit ca. 28 Arten ist Argentinien das Diversitätszentrum der Gattung *Prosopis*. Die untersuchten Baumarten, *Prosopis chilensis* (Molina) und *Prosopis flexuosa* DC kommen dort in den phytogeographischen Provinzen Chaco, Monte und Espinal vor, die sich zusammen über eine Fläche von mehr als einer Million Quadratkilometer erstrecken. Aufgrund ihrer vielfältigen Nutzungsmöglichkeiten stellen diese beiden Baumarten wichtige natürliche Ressourcen dar. Neben der Produktion von Nutzholz, Brennholz und Futter tragen diese beiden Arten zur Stabilisierung von Böden bei und beugen damit der Erosion vor. Als wichtige Ressource der lokalen Bevölkerung spielen *Prosopis chilensis* und *P. flexuosa* eine wesentliche Rolle in der Ökologie und Ökonomie vieler Trocken- und Halbtrocken-Gebiete in Argentinien.

Trotz ihrer ökologischen und ökonomischen Bedeutung befindet sich die Erforschung der genetischen Ressourcen dieser beiden Arten in einem Anfangsstadium, und es gibt nur wenige Informationen über die genetischen Strukturen und das Reproduktionssystem dieser beiden Arten. Für die Untersuchung genetischer Ressourcen und insbesondere für Analysen des Reproduktionssystems sind hoch variable molekulare Marker unverzichtbar. Aufgrund ihrer allgemein hohen Variabilität und ihres kodominanten Vererbungsmodus zeigen Mikrosatelliten Marker (SSRs = Simple Sequence Repeats) ein großes Potential zur Untersuchung von Genfluss- und Paarungssystem.

Das Ziel der vorliegenden Untersuchung war es, zunächst Mikrosatelliten Marker für die Gattung *Prosopis* zu entwickeln, und sie dann für die Untersuchung des reproduktiven Systems in einem Hybridschwarm von *P. chilensis* und *P. flexuosa* zu nutzen.

Für die Identifizierung von Mikrosatelliten Markern bzw. SSR-Genorten wurde die isolierte DNS der Art *P. chilensis* nach einem Protokoll von Fischer und Bachmann (1998) angereichert. Diese nicht-radioaktive Methode zur Identifizierung von SSRs beruht auf der Etablierung einer angereicherten Mikrosatelliten-Bibliothek unter Verwendung von Oligonukleotiden und der magnetischen Isolierung von DNS-Fragmenten mit SSR-Motiven.

Für die Hybridisierung der genomischen DNS von *P. chilensis* wurden zwei Trinukleotid- Sonden ($(GAA)_8$, $(ACC)_8$) und eine Dinucleotid-Sonde ($(CA)_{10}$) eingesetzt. Anschließend wurde ein Anreicherungsschritt durchgeführt. Die angereicherten Fragmente wurden direkt in einen Vektor ligiert und in kompetente *Escherichia coli* Zellen transformiert. Von der angereicherten SSR-Bibliothek wurden 120 positive Klone sequenziert. Die durchschnittliche Länge der sequenzierten Fragmente betrug 580 bp (± 96 bp). Für die Primer-Entwicklung wurden aus den 120 Sequenzen 18 ausgewählt, die mindestens fünf Wiederholungen von SSR-Motiven aufwiesen. Von den getesteten Primerpaaren zeigten neun Amplifizierungsprodukte mit der erwarteten Größe. Von diesen neun Primerpaaren wurde jeweils ein Primer mit einem Fluoreszenzfarbstoff am 5'-Ende markiert, und die genetische Variation an den entsprechenden SSR-Genorten wurde in jeweils einer Population von *P. chilensis* (20 Bäume) und *P. flexuosa* (20 Bäume) aus der trockenen Provinz Chaco charakterisiert. Sechs der neun Primerpaare amplifizierten reproduzierbare und gut auswertbare Fragmente, mit Polymorphismen in beiden Arten.

Die neu entwickelten SSR-Marker zeigten einen PIC (Polymorphism Information Content) zwischen 0,14 und 0,70 in *P. chilensis* und zwischen 0,41 und 0,85 in *P. flexuosa*. Die Anzahl der Allele (N_a) variierte zwischen zwei und sechs in *P. chilensis* und zwischen zwei und 13 in *P. flexuosa*. Die erwartete Heterozygotie (H_e) lag zwischen 0,14 und 0,73 in *P. chilensis,* während sich die Werte bei *P. flexuosa* zwischen 0,46 und 0,86 bewegten. Ein Test bezüglich der Nutzung dieser SSR-Primer auch in anderen *Prosopis* Arten zeigte deren weitgehende Anwendbarkeit über Artengrenzen hinweg. Alle Primer lieferten in mindestens fünf Arten Amplifizierungsprodukte. Die Anzahl der gefundenen Allele variierte zwischen eins und fünf.

Weiterhin wurden die Ergebnisse dieser Anreicherungsstrategie zur Entwicklung von SSR Markern in *Prosopis* evaluiert. Die Analyse der 120 sequenzierten Klonen zeigte, dass 95,8% der Klone (115 von 120) mindestens ein SSR-Motiv enthielten, das mit den Suchkriterien übereinstimmte (d.h. aus Tandem-Wiederholungen von eins bis sechs Nukleotiden zusammengesetzte Mikrosatelliten Fragmente mit einer Gesamtgröße von mindestens sechs Basen und mindestens drei Wiederholungen). Bei den SSR-Fragmenten mit mindestens fünf wiederholten Einheiten und einer minimalen Länge von zehn Basenpaaren lag das Niveau der Anreicherung bei 30,8%.

Sechzehn Prozent (84 Fragmente) aller detektierten Motive zeigten spezifisch komplementäre Sequenzen zu den Oligonukleotid-Sonden, die für die Anreicherung benutzt wurden. Ein Vergleich von 115 Sequenzen, die SSR Motive enthielten, mit der „EMBL-Nucleotide Plant Sequence Database“ ergab, dass 7,8 % der Klone (9 Klone) eine hohe Übereinstimmung zu Regionen der Chloroplasten-DNS zeigten. Die meisten Klone (102) zeigten Übereinstimmungen mit nicht-transkribierten Regionen der Kern-DNS. Nur vier Klone zeigten hohe Ähnlichkeiten zu transkribierten Regionen.

Die entwickelten SSR Marker wurden angewendet, um die genetische Variation, die genetischen Strukturen und das Reproduktionssystem in einem Hybrid-Schwarm von *P. chilensis* und *P. flexuosa* zu untersuchen. Der Hybrid-Schwarm liegt im Wald-Naturschutz-Gebiet Chancaní, im Westen der Provinz Córdoba, Argentinien. Die Untersuchung wurde auf einer Fläche von 4700 m^2 im Süd-Westen des zuvor bezeichneten Waldgebiets durchgeführt. Auf der Untersuchungsfläche wurden alle, in der Saison 2003-2004 blühenden, *Prosopis* Bäume kartiert. Insgesamt wurden 100 blühende *Prosopis*-Bäume (Elternpopulation) in dieser Periode beobachtet. Die morphologische Zuordnung einzelner Bäume erfolgte mit Hilfe von qualitativen und quantitativen Merkmalen der Blätter und Früchte. Dabei bildeten sich drei Gruppen: Gruppe 2 mit Individuen der taxonomischen Art *P. flexuosa*, Gruppe 3 mit Individuen, die als *P. chilensis* charakterisiert wurden, und Gruppe 1 mit morphologisch intermediären Individuen beider Arten.

Aus frischen Blättern jedes einzelnen Baumes der Elternpopulation wurde die Gesamt-DNS extrahiert. Die genetischen Strukturen und die genetische Variation wurden in der Elternpopulation mit Hilfe von sechs SSR-Markern charakterisiert. Auf deren Grundlage wurden genetische Parameter berechnet. Über alle Gruppen hinweg wurden an den sechs SSR-Genorten 70 Allele beobachtet. Die durchschnittliche Anzahl von Allelen pro Genlocus betrug 11,67. Der Parameter „Allelic richness“ zeigte den höchsten Mittelwert in Gruppe 2 (6,64) gefolgt von Gruppe 1 (5,33), während Gruppe 3 den geringsten Wert (4,33) zeigte. Schätzungen der genetischen Diversität H_e (gene diversity) zeigten den höchsten Mittelwert in Gruppe 2 (H_e= 0,730) gefolgt von Gruppe 1 (H_e = 0,605) und Gruppe 3 mit dem geringsten Wert (H_e = 0,554).

Für die Untersuchung des Reproduktionssystems wurden in der Altbaumpopulation 23 Bäume ausgewählt, deren Samen aus freier Abblüte hervorgegangen sind. Die Samen

wurden im Zeitraum von 2003-2004 gesammelt. Insgesamt wurden 516 Samen (Samenpopulation) gesammelt. Die Samenpopulation setzte sich aus 23 Gruppen von Samen, entsprechend der morphologischen Klassifikation der 23 Samenbäume, zusammen. Nach der Keimung der Samen wurde die DNS aus den Keimblättern (Kotyledonen) isoliert. Für die Samenpopulationen und die effektiven Pollenwolken wurden populationsgenetische Parameter auf der Basis von drei Mikrosatelliten-Genorten berechnet und mit den Ergebnissen aus der Untersuchung der Altbäume verglichen. Insgesamt wurden in der Samenpopulation 41 Allele gefunden, wobei drei dieser Allele nicht in der Altbaumpopulation vorkamen. Die genetischen Parameter zur Schätzung der genetischen Differenzierung zwischen den morphologischen Gruppen zeigten, sowohl in der Samenpopulation als auch in der effektiven Pollenwolke, die höchste Differenzierung zwischen Gruppe 2 und Gruppe 3. Die geringste Differenzierung wurde zwischen Gruppe 1 und 3 beobachtet. Eine Abnahme der genetischen Differenzierung von den Altbäumen zu den betreffenden Samenpopulationen und den effektiven Pollenwolken konnte für alle Gruppen beobachtet werden. Über die Populationen hinweg zeigte Gruppe 2 die stärkste Abnahme bezüglich der genetischen Differenzierung.

Um den Genfluss zwischen den morphologischen Gruppen genauer zu bestimmen, wurden Vaterschaftsanalysen durchgeführt. Die Zuordnung der Vaterschaft erfolgte nach dem Ausschlussprinzip auf der Grundlage von Multi-Locus-Genotypen. Auch potentielle Väter, deren Blühperiode sich nicht mit der Samenbäume überschnitt, wurden ausgeschlossen. Für 96 von 516 Nachkommen konnte kein kompatibler Pollenspender innerhalb der Untersuchungsfläche gefunden werden. Im Gegensatz dazu konnten für 78 Nachkommen alle Väter bis auf einen ausgeschlossen werden. Für die übrigen Nachkommen (342) kamen mehrere potentielle Väter in Frage. Für diese Nachkommen wurde die Vaterschaft einzelner Bäume auf der Basis von Wahrscheinlichkeitswerten („LOD-scores“) bestimmt. Unter den 420 Nachkommen, deren Polleneltern potentiell innerhalb des Bestandes liegen können, stammen 170 aus einer Hybridisierung zwischen den Gruppen. Nach diesem Ergebnis erreicht der Anteil der Hybridisierung zwischen den Gruppen eine Größenordnung von 40,5 %.

Weitergehende Untersuchungen des Paarungssystems zeigten, dass innerhalb der Gruppe 3 eine starke Paarungspräferenz für Bäume der gleichen Gruppe besteht. Weiterhin zeigt diese Gruppe schwächer ausgeprägte Präferenzen für Paarungen mit

Individuen der Gruppe 1, während eine starke Repulsion im Hinblick auf Gruppe 2 besteht. Auch Gruppe 1 zeigt Präferenzen für Paarungen innerhalb der Gruppe und mit Gruppe 3 sowie eine Abweisung von Paarungen mit Individuen der Gruppe 2. Gruppe 2 hingegen zeigt Paarungspräferenzen von nahe eins 1 für alle Gruppen. Dies deutet auf einen Trend zur Zufallspaarung für Individuen dieser Gruppe hin.

Die Ergebnisse lassen einen starken Genfluss zwischen morphologischen Gruppen vermuten. Eine beträchtliche Reduktion der genetischen Differenzierung zwischen morphologischen Gruppen wurde in nur einer Generation (von der Altbaumpopulation zur Samenpopulation) beobachtet. Die Analysen des Paarungssystems deuten auf asymmetrischen Genfluss hin. Für das Zustandekommen von intermediären Individuen (Gruppe 1) bekommt Gruppe 2 mehr Pollen von Gruppe 3 als umgekehrt. Die Hybridisierung wird begleitet von Introgression. Eine höhere Rate von Paarungen intermediärer Individuen mit Gruppe 3 erzeugt einen höheren Anteil von Introgression in Richtung dieser Gruppe. Rückkreuzungen kommen aber auch zwischen Gruppe 1 und 2 vor.

Die Fremdbefruchtungsrate (t) wurde für die Blühperiode 2003-2004 auf der Basis von drei hochvariablen Mikrosatelliten-Genorte geschätzt. Dabei wurden verschiedene Methoden angewendet wie etwa die Schätzung von t auf Grundlage von: Fremdallelen, Vaterschaftsanalysen, seltenen Allele sowie des „mixed-mating-model“. Die Minimum-Schätzung von t zeigte Werte für einzelne Individuen zwischen 0,292 und 1,0. Auf dem Gruppen-Niveau betrug der t-Wert 0,938 für Gruppe 1, 0,920 für Gruppe 2 und 0,692 für Gruppe 3. Schätzungen der Fremdbefruchtungsrate basierend auf der Vaterschaftsanalyse schwankten zwischen 0,389 und 1,0. Die Multi-Locus-Fremdbefruchtungsraten für einzelne Individuen (t_m) geschätzt nach dem „mixed-mating-model“, bewegten sich zwischen 0,330 und 1,2. Die Multilocus-Fremdbefruchtungsrate war 0,979 für Gruppe 1, 0,954 für Gruppe 2 und 0,777 für Gruppe 3. Unabhängig von der Methode wurden die geringsten individuelle Werte für die Bäume der Gruppe 3 gefunden. Die „biparental inbreeding“-Werte für einzelne Individuen lagen zwischen -0,088 und 0,342, während die Werte in den Gruppen bei 0,045 für Gruppe 1, 0,075 für Gruppe 2 und 0,143 für Gruppe 3 lagen. Die „correlated-mating“-Werte zeigten relativ geringe Werte die nahe legen, dass die Individuen innerhalb der Nachkommenschaften eher Halb- als Vollgeschwister sind.

10. RESUMEN

El género *Prosopis* (familia Fabaceae) comprende unas 44 especies arbóreas y arbustivas de gran importancia en las zonas áridas y semiáridas de todo el planeta. Argentina constituye el principal centro de diversidad del género *Prosopis* con 28 especies aproximadamente, 13 de ellas endémicas. En Argentina, las provincias fitogeográficas del Chaco, Monte y Espinal cubren más de un millón de kilómetros cuadrados. Dentro de estas provincias fitogeográficas, dos especies arbóreas del género *Prosopis*, *P. chilensis* (Molina) Stuntz y *P. flexuosa* DC., constituyen un importante recurso natural dado que se consideran especies de uso múltiple. Estos árboles proveen madera de alta calidad, así como combustible y forraje. Además tienen la capacidad de estabilizar el suelo y prevenir la erosión. Ambas especies, *Prosopis chilensis* y *P. flexuosa*, desempeñan un rol vital en la ecología y economía de numerosas regiones áridas y semiáridas de la Argentina, siendo muchas veces el principal recurso para los pobladores locales.

A pesar de su gran importancia económica y ecológica, el estudio de los recursos genéticos de *Prosopis chilensis* y *P. flexuosa*, se halla aún en una etapa inicial y existe poca información concerniente a la estructura genética y al sistema reproductivo de ambas especies. Para el estudio de los recursos genéticos de estas especies, y especialmente para el estudio de su sistema reproductivo, se hace imprescindible contar con marcadores moleculares altamente polimórficos. Dentro de este marco, los marcadores de microsatélites o *SSRs* (por su siglas en inglés *Simple Sequence Repeats* = Secuencias Simples Repetidas) ofrecen un gran potencial para el análisis del flujo génico y el estudio del sistema de apareamiento dado a su modo de herencia co-dominante, y a su generalmente alto grado de polimorfismo.

El objetivo del presente trabajo fue, primero desarrollar marcadores de microsatélites para el género *Prosopis*; y a continuación aplicar los marcadores desarrollados en el estudio del sistema reproductivo de un enjambre híbrido entre *Prosopis chilensis* y *P. flexuosa*, localizado en el Chaco Árido Argentino.

Con la finalidad de desarrollar los marcadores de microsatélites, se identificaron y aislaron loci de SSR en la especie *Prosopis chilensis* mediante un protocolo de aislamiento de microsatélites desarrollado por Fischer y Bachmann (1998). Éste es un método no radiactivo para el aislamiento de segmentos de microsatélites basado en el establecimiento de bibliotecas enriquecidas de microsatélites (*microsatellite-enriched libraries*) utilizando sondas de oligonucleótidos homólogas a las secuencias de SSR que se quieren aislar. Luego de una o varias hibridizaciones, las sondas que contienen los fragmentos de ADN con microsatélites son aisladas mediante separación magnética. En el presente estudio se utilizaron dos sondas de trinucleótidos de secuencias $(GAA)_8$ y $(ACC)_8$, y una de dinucleótido con secuencia $(CA)_{10}$ para aislar microsatélites a partir de ADN genómico de *Prosopis chilensis*. Se realizó un paso de enriquecimiento y posteriormente los fragmentos enriquecidos en SSRs fueron clonados directamente en un vector y transformados en una bacteria de células competentes de *Escherichia coli*. A partir de la biblioteca enriquecida de microsatélites obtenida, se secuenciaron 120 clones positivos. El tamaño promedio de los fragmentos secuenciados fue 580 (± 96) pares de bases (bp). De 120 secuencias obtenidas, 18 fueron seleccionadas para el desarrollo de primers (cebadores) considerando aquellas secuencias que contenían fragmentos de microsatélites con al menos cinco motivos repetidos. Todos los primers diseñados fueron testeados. Nueve de los 18 primers desarrollados amplificaron fragmentos de tamaño esperado. Estos nueve primers fueron marcados con partículas fluorescentes en su extremo 5'-, y su variación fue caracterizada en una población natural de *Prosopis chilensis* (20 individuos) y otra de *P. flexuosa* (20 individuos) del Chaco Árido Argentino. Seis de los nueve pares de primers probados amplificaron bandas polimórficas, claras y reproducibles en ambas especies.

Los nuevos marcadores de microsatélites desarrollados mostraron un Contenido de Información Polimórfica (PIC) entre 0,14 y 0,70 en *P. chilensis* y entre 0,41 y 0,85 en *P. flexuosa*. El número de alelos encontrados (N_a) varió de dos a seis, y de dos a 13 en *P. chilensis* y *P. flexuosa* respectivamente. La heterocigosis esperada (H_e) mostró valores comprendidos entre 0,14 y 0,73 en *P. chilensis*, y entre 0,46 y 0,86 en *P. flexuosa*. Ensayos de transferibilidad de los marcadores desarrollados a otras siete especies de *Prosopis* revelaron una amplia afinidad entre especies, todos los primers

amplificaron en al menos cinco de las siete especies testeadas. El número de alelos encontrados en estos casos varió de uno a cinco.

El protocolo utilizado en el aislamiento y desarrollo de marcadores de microsatélites en *Prosopis* fue evaluado con el objetivo de determinar su eficiencia. Un análisis de los 120 clones secuenciados demostró que el 95,8% de ellos (115 de un total de 120) poseía al menos un motivo de microsatélite de acuerdo al criterio de búsqueda establecido (a saber: fragmentos de microsatélites compuestos de unidades repetidas en tándem de uno a seis nucleótidos, repetidas al menos tres veces y de un tamaño total mínimo de seis bases). Considerando sólo aquellos segmentos de SSRs con al menos cinco unidades repetidas y de un tamaño mínimo de diez pares de bases, el nivel de eficiencia alcanzaba el 30,8%. Un 16% (84 fragmentos) del total de fragmentos con SSRs detectados mostraron complementariedad con las secuencias de las sondas de oligonucleótidos utilizadas en el aislamiento. Una comparación de las 115 secuencias que contenían motivos de microsatélites con el Banco de Datos de Secuencias de Nucleótidos de Plantas del EMBL (*EMBL Nucleotide Plant Sequence Database*) reveló que el 7,8% de los clones (9 clones) poseía una alta similitud con secuencias de regiones de cloroplasto de otras especies vegetales. Por otra parte, la mayoría de los clones (102 clones) mostraron similitud con regiones no-transcriptivas del ADN nuclear de otras especies vegetales. Sólo cuatro clones mostraron una alta similitud con regiones del ADN que transcriben para proteínas en otras plantas.

Los marcadores de microsatélites desarrollados fueron aplicados posteriormente a estudios de variación genética, estructura genética y en el estudio de aspectos concernientes al sistema reproductivo de un enjambre híbrido entre *Prosopis chilensis* y *P. flexuosa* situado en el Parque Provincial y Reserva Forestal Natural Chancan, al Oeste de la provincia de Córdoba, Argentina. El estudio se llevó a cabo en una parcela 4700 m^2, incluida dentro de un monte continuo localizado el Sureste de la mencionada reserva. Dentro de la parcela de estudio se identificaron todos los individuos de *Prosopis* que florecieron durante la temporada de floración 2003-2004. Durante el mencionado período fueron identificados un total de 100 árboles de *Prosopis* (población adulta = *adult population*) que florecieron en la parcela. Mediante un relevamiento espacial de estos individuos se confeccionó un plano de la parcela. A partir de

caracteres cuali- y cuantitativos de hoja y fruto se llevó a cabo un análisis morfológico para clasificar a los individuos. Este análisis permitió establecer tres grupos morfológicos existentes dentro de la parcela: *Grupo 2* compuesto por árboles pertenecientes a la especie taxonómica *Prosopis flexuosa*; *Grupo 3* con individuos con características de la especie *P. chilensis*; y *Grupo 1* con individuos con características morfológicas intermedias entre ambas especies.

De cada individuo de la población adulta se extrajo ADN total a partir de muestras de hojas frescas. Mediante los seis marcadores de microsatélites desarrollados se caracterizó la estructura genética y la variación de la población de árboles adultos, y se calcularon diversos parámetros de genética poblacional. El número total de alelos encontrados considerando todos los grupos y en los seis loci de SSR analizados fue 70, con un número promedio de alelos por locus de 11,67. La riqueza alélica (*allelic richness*) mostró el valor medio más elevado en el Grupo 2 (6,64), seguido por el Grupo 1 (5,33), y el valor más bajo en el Grupo 3 (4,33). La diversidad génica estimada, (H_e) también mostró el valor medio más alto en el Grupo 2 (H_e = 0,730), seguido por el Grupo 1 (H_e = 0,605), con el valor más bajo en el Grupo 3 (H_e = 0,554).

Para el estudio del sistema de reproducción se seleccionaron dentro de la población de árboles adultos 23 árboles semilleros de los cuales, durante la temporada 2003-2004, se recolectó semilla originada por polinización abierta. En total se cosecharon 516 semillas (población de semillas = *seed population*). Esta población estaba compuesta por 23 sets de semillas agrupadas de acuerdo a la clasificación morfológica del árbol semillero que las originó. De cada individuo de la población de semillas se extrajo ADN total a partir de cotiledones recién germinados. En la población de semillas y en la nube de polen efectiva se calcularon diferentes parámetros de genética de poblaciones a partir del análisis de tres loci de microsatélites. Estos resultados fueron comparados con los obtenidos en la población de árboles adultos. Un total de 41 alelos diferentes fueron encontrados en la población de semillas. Tres nuevos alelos ausentes en la población de árboles adultos pudieron ser identificados entre las semillas. Los parámetros genéticos calculados para estimar la diferenciación genética entre grupos morfológicos arrojaron que la mayor diferenciación se registró entre los Grupos 3 y 2 en la población de semillas así como en la nube de polen. La diferenciación más baja se observó entre los

Grupos 3 y 1. Cuando se compararon la población de semillas y la nube de polen efectiva se observó una reducción en le diferenciación genética en todos los grupos. El Grupo 2 mostró el mayor descenso en los parámetros de diferenciación genética cuando se comparó entre poblaciones.

Con el objetivo de poder estimar de manera más precisa el flujo genético entre grupos morfológicos se llevó a cabo un estudio basado en análisis de paternidad. La asignación de paternidad se realizó mediante "*categorical allocation*" a partir de genotipos multilocus. Aquellos padres potenciales que no registraron solapamiento en su período de floración con el del árbol semillero fueron descartados. Para 96 descendencias (semillas) entre las 516 analizadas, no se registró padre compatible dentro de la parcela estudiada. Por el contrario, para 78 semillas todos excepto un padre pudieron ser excluidos. Las restantes progenies (342), registraron más de un padre potencial dentro de la parcela. Para estas semillas, la asignación de un solo padre fue realizada considerando como padre más probable aquel con el máximo "*LOD Score*". Entre las 420 progenies que registraron su padre potencial dentro de la parcela, 170 resultaron de hibridación entre grupos. De acuerdo a estos resultados, el porcentaje de hibridación entre grupos alcanza el 40,5%.

Subsiguientes estudios del sistema de apareamiento mostraron que existe una fuerte preferencia del Grupo 3 a aparearse con individuos de su mismo grupo. Al mismo tiempo, este grupo muestra preferencia por el Grupo 1, pero en menor grado, y evidencia una fuerte repulsión hacia el Grupo 2. El Grupo 1 reveló preferencias por individuos de su mismo grupo y también por el Grupo 3, y repulsión hacia el Grupo 2. Por otro lado, el Grupo 2 mostró valores de preferencia de apareamiento cercanos a uno para todos los grupos, esto indica una tendencia al apareamiento al azar dentro de este grupo.

Los resultados obtenidos apuntan a que existe un fuerte flujo genético entre los diferentes grupos morfológicos. Una considerable reducción en la diferenciación genética entre grupos morfológicos se pudo observar en sólo una generación (desde la población de árboles adultos a la población de semillas). El análisis del sistema de apareamiento indica que el flujo génico entre los grupos es asimétrico. En la formación

de individuos con características intermedias (Grupo 1), el Grupo 2 recibe mayor cantidad de polen del Grupo 3 que en el sentido opuesto. La hibridación está acompañada por un proceso de introgresión. A su vez, una mayor frecuencia de apareamientos de individuos intermedios con el Grupo 3 genera mayores niveles de introgresión hacia este grupo. No obstante, también ocurren retrocruzas entre el Grupo 1 y el 2.

La tasa de fecundación cruzada (*t*) se estimó para el período de floración 2003-2004 mediante el uso de tres loci de microsatélites altamente polimórficos. En la estimación de esta tasa se utilizaron diferentes metodologías, a saber: tasa de fecundación cruzada mínima estimada basada en alelos extraños, estimación de *t* mediante análisis de paternidad, mediante alelos raros, y basada en el modelo de "*mixed mating*" de Ritland y Jain (1981). La tasa de fecundación cruzada mínima estimada arrojó valores individuales entre 0,292 y 1,000. A nivel de grupo, la estimación mínima de *t* para el Grupo 1 fue 0,938, para el Grupo 2 de 0,920, y para el Grupo 3 de 0,692. Los valores de *t* estimados mediante análisis de paternidad variaron entre 0,389 y 1,000. La tasa de fecundación cruzada individual basada en genotipos multilocus (t_m) estimada mediante el modelo de "*mixed mating*", osciló entre 0,330 y 1,200. La tasa de fecundación cruzada multilocus (t_m) para el Grupo 1 fue de 0,979, para el Grupo 2 de 0,954, y para el Grupo 3 de 0,777. En todas las metodologías utilizadas para el cálculo de *t*, los valores más bajos de este parámetro a nivel individual fueron registrados entre individuos del Grupo 3. La endogamia biparental ("*biparental inbreeding*") individual varió entre -0,088 y 0,342; mientras que los valores hallados a nivel de grupo fueron 0,045 para el Grupo 1, 0,075 para el Grupo 2, y 0,143 para el Grupo 3. El apareamiento correlacionado ("*correlated mating*") mostró valores relativamente bajos, indicando que los individuos dentro de cada set de progenies son principalmente medio hermanos y en menor medida hermanos enteros.

11. REFERENCES

Alden H A. 1995. Hardwoods of North America. General Technical Report FPL–GTR–83. Madison, WI: United States Department of Agriculture, Forest Service, Forest Products Laboratory. 136 p.

Ali S, Muller S R, Epplen J T. 1986. DNA fingerprinting by oligonucleotide probes specific for simple sequence repeats. Human Genetics 74: 239–243.

Almanza S G, Moya E G, Wendt T L, Cossio F V G. 1992. Potencial de hibridación natural en el mesquite (*Prosopis laevigata* y *P. glandulosa* var. *torreyana*, Leguminosae) de la altiplanicie de San Luis Potosí. Acta Botánica Mexicana 20: 101-117.

Alves R M, Artero A S, Sebbenn A M, Figueira A. 2003. Mating system in a natural population of *Theobroma grandiflorum* (Wild. ex Spreng) Schum., by microsatellite markers. Genetics and Molecular Biology 26: 373-379.

Anderson E. 1949. Introgressive hybridization. Wiley, New Cork, USA.

Ayala F J, Tracey M L, Hedgecock D, Richmond R C. 1974. Genetic differentiation during the speciation process in *Drosophila*. Evolution 28: 576-592.

Balboa O, Parraguez J M. 1986. Phenology studies of *Prosopis* species growing in Chile. The current state of knowledge on *Prosopis juliflora*. II International Conference on *Prosopis*, Recife, Brazil. Pp. 259-267.

Barreneche T, Bodenes C, Lexer C, Trontin J-F, Fluch S, Streiff R, Plomion C, Roussel G, Steinkellner H, Burg K, Favre J-M, Glossl J, Kremer A. 1998. A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. Theoretical and Applied Genetics 97: 1090-1103.

Bessega C F, Ferreira L I, Julio N, Montoya S, Saidman B O, Vilardi J C. 2000a. Mating system parameters in species of genus *Prosopis* (Leguminosae). Hereditas 132: 19-28.

Bessega C F, Ferreira L I, Vilardi J C, Saidman B O. 2000b. Unexpected low genetic differentiation among allopatric species of section Algarobia of *Prosopis* (Leguminosae). Genetica 109: 255-266.

Bessega C F, Saidman B O, Vilardi J C. 2000c. Isozyme and RAPD studies in *Prosopis glandulosa* and *P. velutina* (Leguminisae, Mimosoideae). Genetics and Molecular Biology 23: 1-5.

Bossart J L, Pashley-Prowell D P. 1998. Genetic estimates of population structure and gene flow: limitations, lessons and new directions. Trends in Ecology & Evolution 13: 202-206

Bowcock A M, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd J R, Cavalli- Sforza L L. 1994. High resolution of human evolution with polymorphic microsatellites. Nature 368: 455-457.

Brondani R P V, Brondani C, Tarchini R, Grattapaglia D. 1998. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. Theoretical and Applied Genetics 97: 816-827.

Brown A H D, Weir B S. 1983. Measuring genetic variability in plant populations. In: Isozymes in Plant Genetics and Breeding. Tanksley S D, Orton P A (eds.). Elsevier Science Publ., Amsterdam. Pp. 219-239.

Brown L Y, Brown S A. 2004. Alanine tracts: the expanding story of human illness and trinucleotide repeats. Trends in Genetics 20: 51-58.

Bullock P, Miller J, Botchan M. 1986. Effects of poly [d(pGpT)-d(pApC)] and poIy [d(pCpG)'d(pCpG)J repeats on homologous recombination in somatic cells. Molecular and Cellular Biology 6: 3948-3953.

Burkart A, Simpson B B. 1977. The genus *Prosopis* and annoted key to the species of the world. In: Mesquite: Its biology in two desert ecosystems. B B Simpson (ed.). Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania, USA. Pp. 201-215.

Burkart A. 1940. Materials for a monograph of *Prosopis* genus (*Leguminosae*). Darwiniana 4: 57-128.

Burkart A. 1952. Las leguminosas argentinas. Editorial Acme, Buenos Aires.

Burkart A. 1976. A monograph of the genus *Prosopis* (Leguminosae subfam. Mimosoideae). (Part 1 and 2). Catalogue of the recognized species of *Prosopis*. Journal of the Arnold Arboretum 57: 219-249 and 450-525.

Butcher P A, Glaubitz J C, Moran G F. 1999. Applications for microsatellite markers in the domestication and conservation of forest trees. In: Forest Genetic Resources N° 27, FAO (ed.), Rome, 1999.

Byrne M, Marquez-Garcia M I, Uren T, Smith D S, Moran G F. 1996. Conservation and genetic diversity of microsatellite loci in the genus Eucalyptus. Australian Journal of Botany 44: 331-341.

Cabido M R, Pacha M J. 2002. Vegetación y flora de la Reserva Natural Chancaní. Agencia Córdoba Ambiente S.E., Córdoba, Argentina.

Cabrera A. 1976. Regiones Fitogeográficas Argentinas. In: Enciclopedia Argentina de Agricultura y Jardinería. Tomo II. ACME (ed.). Fascículo 1. Pp: 85.

Chase M R, Moller C, Kesseli R, Bawa K S. 1996. Distant gene flow in tropical trees. Nature 383: 398-399.

Cipriani G, Lot G, Huang W G, Marrazzo M T, Peterlunger E, Testolin R. 1999. AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L) Batsch]: isolation, characterization and cross-species amplification in Prunus. Theoretic and Applied Genetics 99: 65-72.

Collevatti R G, Grattapaglia D, Hay J D. 2001. High resolution microsatellite based analysis of the mating system allows the detection of significant biparental inbreeding in *Caryocar brasiliense*, an endangered tropical tree species. Heredity 86: 60-67.

Cony M A. 1996. Genetic variability in *Prosopis flexuosa* D.C., a native tree of the Monte phytogeographic province, Argentina. Forest Ecology and Management 87: 41-49.

Cony M A. 2000. Importancia de los estudios fisiológicos, ecofisiológicos y genéticos sobre especies del género *Prosopis* para la recuperación de ecosistemas áridos degradados de América Latina. Memorias del Seminario Internacional Bosques secos y desertificación. Pp: 225-276.

Cooke R I. 1984. The characterization and identification of crop cultivars by electrophoresis. Electrophoresis 5: 59-72.

Cooper S, Bull C M, Gardner M G. 1997. Characterization of microsatellite loci from the socially monogamous lizard *Tiliqua rugosa* using a PCR-based isolation technique. Molecular Ecology 6: 793-795.

Cordeiro G M, Maguire T L, Edwards K J, Henry R J. 1999. Optimization of a microsatellite enrichment technique in *Saccharum* spp. Plant Molecular Biology Reporter 17: 225-229.

Crow J F, Kimura M. 1970. An introduction to population genetics theory. Harper and Row, New York, Evanston and London.

Cruz G. 1999. Production, characterization and uses of *Prosopis* seed galactomannan. PhD thesis, ETH Nº 13153, Switzerland.

Curtu A-L, Finkeldey R, Gailing O. 2004. Comparative sequencing of a microsatellite locus reveals size homoplasy within and between European oak species (*Quercus* spp.). Plant Molecular Biology Reporter 22: 339-346.

D'Antoni H L, Solbrig O T. 1977. Algarrobos in South American cultures: past and present. In: Mesquite: Its biology in two desert ecosystems. B B Simpson (ed.). Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania, USA. Pp. 189-200.

Dawson I K, Waugh R, Simons A J, Powell W. 1997. Simple sequence repeats provide a direct estimate of pollen mediated gene dispersal in the tropical tree *Gliricidia sepium*. Molecular Ecology 6: 179-183.

Dayanandan S, Dole J, Bawa K, Kesseli R. 1999. Population structure delineated with microsatellite markers in fragmented populations of a tropical tree, *Carapa guianensis* (Meliaceae). Molecular Ecology 8: 1585-1592.

Devey M E, Bell J C, Smith D N, Neale D B, Moran G F. 1996. A genetic linkage map for *Pinus radiata* based on RFLP, RAPD and microsatellite markers. Theoretical and Applied Genetics 92: 673– 679.

Devey M E, Sewell M M, Uren TL, Neale DB. 1999. Comparative mapping in loblolly and radiata pine using RFLP and microsatellite markers. Theoretical and Applied Genetics 99: 656-662.

Di Rienzo A, Peterson A C, Garza J C, Valdes A M, Slatkin M, Freimer N B. 1994. Mutational processes of simple-sequence repeat loci in human populations. Proceedings of the National Academy of Sciences of the USA 91: 3166-3170.

Dick C W, Etchelecu G, Austerlitz F. 2003. Pollen dispersal of tropical trees (*Dinizia excelsa*, Fabaceae) by native insects and African honeybees in pristine and fragmented Amazonian rainforest. Molecular Ecology 12: 753-764.

Dobzhansky T. 1970. Genetics of the evolutionary process. Columbia University Press, New York, USA.

Dorocicz I R. 1999. A survey of the microsatellite content of barley and several applications of microsatellite polymorphism. National Library of Canada, Ottawa, ON, Canada. 146 pp.

Dow B D, Ashley M V. 1996. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak *Quercus macrocarpa*. Molecular Ecology 5: 615-627.

Dow B D, Ashley M V. 1998. High levels of gene flow in bur oak revealed by paternity analysis using microsatellites. Journal of Heredity 89: 62-70.

Echt C S, DeVerno L L, Anzidei M, Vendramin G G. 1998. Chloroplast microsatellites reveal population genetic diversity in red pine, *Pinus resinosa* Ait. Molecular Ecology 7: 307-316.

Echt C S, Nelson C D. 1997. Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). Theoretical and Applied Genetics 94: 1031-1037.

Echt C S, Vendramin G G, Nelson C D, Marquardt P. 1999. Microsatellite DNA as shared genetic markers among conifer species. Canadian Journal of Forest Research 29: 365-371.

Edwards A, Civitello A, Hammond H A, Caskey C T. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. The American Journal of Human Genetics 49: 746-756.

Edwards A, Hammond H A, Jin L, Caskey C T, Chakraborty R. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12: 241-253.

Edwards K J, Barker J H A, Daly A, Jones C, Karp A. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20: 758-760.

Eisen J A. 1999. Mechanistic basis for microsatellite instability. In: Goldstein and Schlötterer (eds.). Microsatellites. Evolution and applications. Oxford University Press: 35-48.

Elias T S. 1981. Mimosoideae. In: Advances in Legume Systematics. Polhill R M, Raven P H (eds.) Royal Botanic Gardens, Kew, UK. Pp. 143-151.

Ender A, Schwenk K, Städler T, Streit B, Schierwater B. 1996. RAPD identification of microsatellites in *Daphnia*. Molecular Ecology 5: 437-441.

Estoup A, Cornuet J-M. 1999. Microsatellite evolution: inferences from population data. In: Goldstein D B, Schlötterer C (eds.) Microsatellites. Evolution and Applications. Oxford University Press. Pp. 49-65.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.01: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47-50.

Excoffier L, Smouse P, Quattro J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131, 479-491.

Fahima T, Roder M S, Wendehake K, Kirzhner V M, Nevo E. 2002. Microsatellite polymorphism in natural populations of wild emmer wheat, *Triticum dicoccoides*, in Israel. Theoretic and Applied Genetics 104: 17-29.

Felger R S. 1977. Mesquite in Indian cultures of southwestern North America. In: Mesquite: Its biology in two desert ecosystems. B B Simpson (ed.). Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania, USA. Pp. 150-176.

Field D, Wills C. 1998. Abundant microsatellite polymorphism in *Saccharomyces cerevisiae*, and the different distribution of microsatellites in eight prokaryotes and *S. cerevisiae*, result from strong mutation pressures and a variety of selective forces. Procedures of the Natural Academy of Science, USA Vol. 95, pp. 1647-1652.

Finkeldey R, Hattermer H H. in press. Tropical forest genetics. Springer-Verlag Berlin, Heidelberg, Germany.

Fischer D, Bachman K. 1998. Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). *BioTechniques* 24, 796–802.

Fischer D. 2000. Entwicklung und Anwendung von STMS-Markern für Kultivaridentifizierung in *Allium cepa* L., erschlossen durch eine Magnetpartikel-basierte 121 Precloning-Anreicherungsmethode für große Genome. Ph.D. Dissertation, Johann- Wolfgang-Goethe Universität, Frankfurt, Germany.

Fisher P J, Gardner R C, Richardson T E. 1996. Single locus microsatellite isolated using 5'-anchored PCR. Nucleic Acids Research 24: 4369-4371.

Fondon III J W, Garner H R. 2004. Molecular origins of rapid and continuous morphological evolution. Proceedings of the Natural Academy of Sciences, USA 101: 18058-18063.

Forbes S H, Hogg J T, Buchanan F C, Crawford A M, Allendorf F W. 1995. Microsatellite evolution in congeneric mammals: Domestic and Bighorn sheep. Molecular Biology and Evolution 12: 1106-1113.

Gailing O, Leinemann L, Finkeldey R. 2005. Molecular tools for the conservation of forest genetic resources. Institute of Forest Genetics and Forest Tree Breeding, Georg-August University Goettingen, Germany. Online publication: http://aunp.forst.unigoettingen.de/outputs/papers/Molectools_Gailing.pdf

Galera F M. 2000. Las especies del género Prosopis (Algarrobos) de América Latina con especial énfasis en aquellas de interés económico. Impreso en los talleres gráficos de Graziani Gráfica, Córdoba, Argentina.

García C, Arroyo J M, Godoy J A, Jordano P. 2005. Mating patterns, pollen dispersal, and the ecological maternal neighbourhood in a *Prunus mahaleb* L. population. Molecular Ecology 14: 1821-1830.

Gentles A J, Karlin S. 2001. Genome-scale compositional comparisons in eukaryotes. Genome Research 11: 540-546.

Genise, J, Palacios R A, Hoc P S, Carrizo R, Moffat L, Mom M P, Agullo M A, Picca P, Torregrosa S. 1990. Observaciones sobre la biología floral de *Prosopis* (*Leguminosae*, *Mimosoideae*) II. Fases florales y visitantes en el distrito Chaqueño Serrano. Darwiniana 30: 71-85.

Gerber S, Chabrier P, Kremer A. 2003. FAMOZ: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. Molecular Ecology Notes 3, 479-481.

Gillet E M, Hattemer H H. 1989. Genetic analysis of isoenzyme phenotypes using single tree progenies. Heredity 63: 135-141.

Gillet E M. 1997. Maximum likelihood estimators of the gametic contributions to single plant progenies. Biometrics 53: 504-523.

Gillet E M. 1999. DNA markers – Concepts and characteristics. In: Which DNA Marker for Which Purpose? Final Compendium of the Research Project Development, optimization and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV Research Programme Molecular Tools for Biodiversity. Gillet EM (ed.). 1999. URL http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm

Gillet E M. 2004. GSED Version 1.1j: Genetic structure from electrophoresis data. Institüt für Forstgenetik und Forstpflanzenzüchtung, Universität Göttingen, Germany.

Gilmour D S, Thomas G H, Elgin S C. 1989. Drosophila nuclear proteins bind to regions of alternating C and T residues in gene promoters. Science 245: 1487-1490.

Godoy J A, Jordano P. 2001. Seed dispersal by animals: exact identification of source trees with endocarp DNA microsatellites. Molecular Ecology 10: 2275-2283.

González-Martínez S C, Gerber S, Cervera M T, Martínez-Zapater J M, Alía R, Gil L. 2003. Selfing and sibship structure in a two-cohort stand of maritime pine (*Pinus pinaster* Ait.) using nuclear SSR markers. Annals of Forest Science 60: 115-121.

González-Martínez S C, Robledo-Arnuncio J J, Collada C, Díaz A, Williams C G, Alía R, Cervera M T. 2004. Cross-amplification and sequence variation of microsatellite loci in Eurasian hard pines. Theoretic and Applied Genetics 109: 103-111.

Goudet J. 2001. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html.

Grant V. 1981. Plant speciation. Columbia University Press. New York, USA.

Gregorius H-R, Roberds J H. 1986. Measurement of genetical differentiation among subpopulations. Theoretical and Applied Genetics 71: 826-834.

Gregorius H-R. 1974. Genetischer Abstand zwischen Populationen. I. Zur Konzeption der genetischen Abstandsmessung. Silvae Genetica 23: 22-27.

Gregorius H-R. 1988. The meaning of genetic variation within and between subpopulations. Theoretical and Applied Genetics 76: 947-951.

Grodzicker T, Williams J, Sharp P, Sambrook J. 1974. Physical mapping of temperature-sensitive mutations of adenoviruses. Cold Spring Harbor Symposium on Quantitative Biology 39: 439-446.

Groover A T, Devey M E, Fiddler T A, Lee J M, Megraw R A, Mitchell-Olds T, Sherman B K, Vujcic S L, Williams C G, Neale D B. 1994. Identification of quantitative trait loci influencing wood specific gravity in loblolly pine. Genetics 138:1293-1300.

Gugerli F, Senn J, Anzidei M, Madaghiele A, Büchler U, Sperisen C, Vendramin G G. 2001. Chloroplast microsatellites and mitochondrial *nad*1 intron 2 sequences indicate congruent phylogenetic relationships among Swiss stone pine (*Pinus cembra*), Siberian stone pine (*Pinus sibirica*), and Siberian dwarf pine (*Pinus pumila*). Molecular Ecology 10: 1489-1497.

Gupta P K, Balyan H S, Sharma P C, Ramesh B. 1996. Microsatellites in plants: a new class of molecular markers. Current Science 70: 45-54.

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 113: 163-185.

Hamada H, Petrino M G, Kakunaga T. 1982. A novel repeated element with Z-DNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proceedings of the National Academy of Sciences, USA 79: 6465-6469.

Hammock E A D, Young L J. 2005. Microsatellite instability generates diversity in brain and sociobehavioral traits. Science 308: 1630-1634.

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, Sunderland, pp: 43-63.

Hancock J M. 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: Goldstein D B, Schlötterer C (eds.) Microsatellites. Evolution and Applications. Oxford University Press. Pp. 1-9.

Hartl D L, Clark A G. 1997. Principles of population genetics. Third Edition. Sinauer Associates INC., Sunderland, MA.

Hattemer H H. 1991. Genetic analysis and population genetics. In: Fineschi S, Malvolti M E, Cannata F, Hattemer H H (eds.). Biochemical markers in the population genetics of forest trees. SPB, The Hague, pp 5-22.

Hokanson S C, Szewc-McFadden A K, Lamboy W F, McFerson J R. 1998. Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a Malus x domestica borkth. core subset collection. Theoretical and Applied Genetics 97: 671-683.

Hunziker J H, Saidman B O, Naranjo C A, Palacios R A, Poggio L, Burghardt A D. 1986. Hybridization and genetic variation of Argentine species of *Prosopis*. Forest Ecology and Management 16: 301-315.

Indrioko S, Gailing O, Finkeldey R. 2006. Molecular phylogeny of *Dipterocarpaceae* in Indonesia based on chloroplast DNA. Plant Systematics and Evolution (in press.) DOI 10.1007/s00606-006-0435-8

Isagi Y, Suhandono S. 1997. PCR primers amplifying microsatellite loci of *Quercus myrsinifolia* Blume and their conservation between oak species. Molecular Ecology 6: 897-899.

Ishii N, Hijikata M, Osumi T, Hashimoto T. 1987. Structural organization of the gene for rat enoyl-CoA hydratase:3-hydroxyacyl-CoA dehydrogenase bifunctional enzyme. Journal of Biological Chemistry 262: 8144-8150.

Isoda K, Watanabe A, Kuramoto N, Kondo T. 2006. Linkage mapping of *Pinus densiflora* and QTL analysis for resistant to the pine wilt disease. In: International Plant & Animal Genome XIV Conference: Abstract P513, 14th–18th Jan., 2006, San Diego, CA.

Jakse J, Javornik B. 2001. High throughput isolation of microsatellites in hop (*Humulus lupulus* L.). Plant Molecular Biology Reporter 19: 217-226.

Jamieson A, Taylor S S. 1997. Comparison of three probability formulae for parentage exclusion. Animal Genetics 28, 397-400.

Jarne P, Lagoda P J L. 1996. Microsatellites, from molecules to populations and back. Trends in Ecology & Evolution 11: 424-429.

Jones A G, Ardren W R. 2003. Methods of parentage analysis in natural populations. Molecular Ecology 12:2551-2523.

Joseau M J. 2006. Caracterización morfológica y genética de poblaciones del género *Prosopis* del Chaco Semiárido del Norte de Córdoba y Sur de Santiago del Estero. Tesis para optar al grado académico de Doctor en Ciencias Agropecuarias. Escuela para Graduados, Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Argentina.

Karagyozov L, Kalcheva I D, Chapman V M. 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Research 21: 3911-3912.

Karhu A. 2001. Evolution and applications of pine microsatellites. Academic Dissertation presented in the Department of Biology, Faculty of Science, University of Oulu, Finland.

Karlin S. 2002. Amino acid runs in eukaryotic proteome and disease associations. Proceedings of the National Academy of Science, USA 99: 333-338.

Karlin U O, Díaz R. 1984. Potencialidad y manejo de algarrobos en el árido subtropical. SECyT, Ministerio de Educación y Justicia, Programa nacional de recursos naturales renovables.

Kashi Y, King D G. 2006. Simple sequence repeats as advantageous mutators in evolution. Trends in Genetics 22: 253-259.

Kashi Y, Soller M. 1999. Functional roles of microsatellites and minisatellites. In: Goldstein D B, Schlötterer C (eds.) Microsatellites. Evolution and Applications. Oxford University Press. Pp. 10-23.

Katti M V, Ranjekar P K, Gupta V S. 2001. Differential distribution of simple sequence repeats in eukaryotic genome sequences. Molecular Biology and Evolution 18: 1161-1167.

Keys R N, Smith S. 1994. Mating system parameters and population genetic structure in pioneer populations of *Prosopis velutina* (Leguminosae). American Journal of Botany 81: 1013-1020.

Kijas J M H, Fowler J C S, Garbett C A, Thomas M R. 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. BioTechniques 16: 657-662.

Kijas J M H, Fowler J C S, Thomas M R. 1995. An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species. Genome 38: 349-355-

Kimura M, Crow J F. 1964. The number of alleles that can be maintained in a finite population. Genetics 49: 725-738.

Kimura M, Ohta T. 1978. Stepwise mutation model and distribution of allelic frequencies in a finite population. Proceedings of the National Academy of Sciences of the USA 75: 2868-2872.

King D G, Trifonov E N, Kashi Y. 2006. Tuning knobs in the genome: Evolution of simple sequence repeats by indirect selection. In: Lynn H. Caporale (ed.) The Implicit Genome, Oxford University Press. Pp. 77-90.

Kirst M, Grattapaglia G. 1999. Development of genotyping systems for *Eucalyptus* species based on multiplex SSR. In: International Plant & Animal Genome VII Conference: Abstract P459, 17th–21st Jan., 1999, San Diego, CA.

Kirst M., Brondani R V, Ciampi A Y, Gaiotto F A, Squilassi M G, Rezende G D, D. Grattapaglia D. 1999. Fingerprinting *Eucalyptus* genotypes with highly informative microsatellite markers. In: International Plant & Animal Genome VII Conference: Abstract P458, 17th–21st Jan., 1999, San Diego, CA.

Koppen W. 1931. Grundriss der Klimakunde. Walter de Gruyter Co., Berlin. 388 pp.

Koreth J, O'Leary J J, O'D McGee J. 1996. Microsatellites and PCR genomic analysis. Journal of Pathology 178: 239-248.

Künzler P, Matsuo K, Schaffner W. 1995. Pathological, physiological, and evolutionary aspects of short instable DNA repeats in the human genome. Biological Chemistry Hoppe-Seyler 376: 201-211.

Lench N J, Norris A, Bailey A, Booth A, Markham A F. 1996. Vectored PCR isolation of microsatellite repeat sequences using anchored dinucleotide repeat primers. Nucleic Acids Research 24: 2190-2191.

Levinson G, Gutman G A. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. Molecular Biology and Evolution 4: 203-221.

Lewis G P, Elias T S. 1981. Mimosaea. In: Advances in Legume Systematics. Polhill R M, Raven P H (eds.) Royal Botanic Gardens, Kew, UK. Pp. 155-168.

Li Y-C, Korol A B, Fahima T, Nevo E. 2004. Microsatellite within genes: Structure, function, and evolution. Molecular Biology and Evolution 21: 991-1007.

Litt M, Luty J A. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. The American Journal of Human Genetics 44: 397-401.

Lue N F, Buchman A F, Kornberg R D. 1989. Activation of RNA polymerase II transcription by a thymidine rich upstream element in vitro. Proceedings of the National Academy of Science, USA 86: 486-490.

Lunt D H, Hutchinson W F, Carvalho G R. 1999. An efficient method for PCR-based identification of microsatellite arrays (PIMA). Molecular Ecology 8: 893-894.

Mantovan N G. 2004. Variabilidad intra-específica de *Prosopis flexuosa* D.C. var. flexuosa en el Monte. Su estudio morfo-fisiológico. Tesis Doctoral, Presentado al Programa de Postgrado en Biología, Universidad Nacional de Cuyo, Argentina.

Marcucci Poltri S. 2006. Marcadores moleculares aplicados a programas de mejoramiento genético en Eucalyptus. In: Norverto C A (ed.). 2006. Mejores Árboles para más Forestadores. El Programa de Producción de Material de Propagación Mejorado y el Mejoramiento Genético en el Proyecto Forestal de Desarrollo. Secretaría de Agricultura, Ganadería, Pesca y Alimentos, República Argentina.

Marquardt P E, Epperson B K. 2004. Spatial and population genetic structure of microsatellite in white pine. Molecular Ecology 13: 3305-3315.

Mayr E. 1963. Animal species and evolution. Harvard University Press, Cambridge, MA, USA.

Mendel G. 1866. Versuche über Pflanzenhybriden. Verhandlungen des naturforschenden Vereins in Brünn 4: 3-47.

Meyer W, Michell T G, Freedman E Z, Vilgalys R. 1993. Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of *Cryptococcus neoformans*. Journal of Clinical Biology 31: 2274–2280.

Montoya S, Saidman B O, Vilardi J C, Bessega C. 1994. Diferenciación y flujo genético entre especies de la sección Algarobia, Genero *Prosopis* (Leguminosae). Actas XXIV Congreso Argentino de Genética, La Plata.

Mooney H A, Simpson B B, Solbrig O T. 1977. Phenology, morphology, physiology. In: Mesquite: Its biology in two desert ecosystems. B B Simpson (ed.). Dowden, Hutchinson and Ross, Stroudsburg, Pennsylvania, USA. Pp. 26-43.

Moore S S, Sargeant L L, King T J, Mattick J S, Georges M, Hetzel D J S. 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. Genomics 10: 654-660.

Morgante M, Hanafey M, Powell W. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nature Genetics 30:194-200.

Morgante M, Olivieri A M. 1993. PCR- amplified microsatellites as markers in plant genetics. The Plant Journal 3: 175-182.

Mullis K B, Fallona F A. 1987. Specific synthesis of DNA in vitro via a polymerase catalyzed chain reaction. Methods in Enzymology. 155: 335- 350.

Nascimento de Sousa S, Finkeldey R, Gailing O. 2005. Experimental verification of microsatellite null alleles in Norway spruce (*Picea abies* [L.] Karst.): Implications for population genetic studies. Plant Molecular Biology Reporter 23. 113-119.

Naylor L H, Clark E M. 1990. $d(TG)_n$ $d(CA)_n$ sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. Nucleic Acids Research 18: 1595-1601.

Nei M. 1972. Genetic distance between populations. American Naturalist 106 (949): 283-292.

Nei M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the Natural Academy of Science, USA. Pp: 3321-3323.

Nevo E, Beharav A, Meyer R C, Hackett C A, Forster B P, Russell J R, Powell W. 2005. Genomic microsatellite adaptive divergence of wild barley by microclimatic stress in 'Evolution Canyon', Israel. Biological Journal of the Linnean Society 84: 205-224.

Olufowote J O, Xu Y, Chen X, Park W D, Beachell H M, Dilday R H, Goto M, McCouch S R. 1997. Comparative evaluation of within-cultivars variation of rice *(Oryza sativa* L) using microsatellites and RFLP markers. Genome 40: 370-378.

Orgel L E, Crick F H C. 1980. Selfish DNA: the ultimate parasite. Nature 284: 604-607.

Paetkau D. 1999. Microsatellites obtained using strand extension: an enrichment protocol. BioTechniques 26: 690-697.

Palacios R, Bravo L. 1981. Hibridización natural en *Prosopis*: Evidencias morfológicas y cromatográficas. Darwiniana 23: 3-35.

Palacios R. 1988. Taxonomía numérica (Descriptores). *Prosopis* en la Argentina. Facultad de Ciencias Agrarias, Universidad Nacional de Córdoba, Argentina. Pp. 91-96.

Pandey M, Gailing O, Fischer D, Hattemer H H, Finkeldey R. 2004. Characterization of microsatellite markers in sycamore (*Acer pseudoplatanus* L.). Molecular Ecology Notes 4: 253-255.

Pandey M. 2005. Development of microsatellites in sycamore maple (*Acer pseudoplatanus* L.) and their application in population genetics. Published online through the library of Lower Saxony State and Georg-August University of Göttingen, Germany. http://resolver.sub.uni-goettingen.de/purl/?webdoc-9

Pasiecznik N M, Felker P, Harris P J C, Harsh L N, Cruz G, Tewari J C, Cadoret K, Maldonado L J. 2001. The *Prosopis juliflora – Prosopis pallida* complex: a monograph. HDRA, Coventry, UK. Pp. 162.

Pasiecznik N M, Harris P J C, Smith S J. 2004. Identifying tropical Prosopis species: a field guide. HDRA, Coventry. Printed by Emmerson Press, Kenilworth, UK.

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A. 1998. Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Molecular Biology and Evolution 15: 1275-1287.

Peakall R, Smouse P E. 2005. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Australian National University, Canberra, Australia.

Pearson W R, Lipman D J. 1988. Improved tools for biological sequence comparison. Proceedings of the Academy of Natural Sciences, USA 85: 2444-2448.

Pearson W R. 1990. Rapid and sensitive sequence comparison with FASTP and FASTA. Methods in Enzymology 183: 63-98.

Petit R J, Mousadik A, Pons O. 1998. Identifying populations for conservation on the basis of genetic markers. Conservation Biology 12: 844-855.

Pierantoni L, Cho K H, Shin I S, Chiodini R, Tartarini S, Dondini L, Kang S J, Sansavini S. 2004. Characterization and transferability of apple SSRs to two European pear F-1 populations. Theoretical and Applied Genetics 109: 1519-1524.

Plaschke J, Ganal M W, Roder M S. 1995. Detection of genetic diversity in closely related bread wheat using microsatellite markers. Theoretic and Applied Genetics 91: 1001-1007.

Poulsen G B, Kahl G, Weising K.1993. Abundance and polymorphism of simple repetitive DNA sequences in *Brassica napus* L. Theoretical and Applied Genetics 85. 994-1000.

Prasad M, Varshney R K, Roy J K, Balyan H S, Gupta P K. 2000. The use of microsatellite for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theoretical and Applied Genetics 100: 584-592.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germoplasm analysis. Molecular Breeding 2: 225-235.

Rafalsky J A, Tingey S V. 1993. Genetic diagnostics in plant breeding: RAPDs, microsatellites and machines. Trends Genetics 9: 275-280.

Ramírez L, de la Vega A, Razkin N, Luna V, Harris P J C. 1999. Analysis of the relationships between species of the genus *Prosopis* revealed by the use of molecular markers. Agronomie 19: 31-43.

Rassmann K, Schlötterer C, Tautz D. 1991. Isolation of simple-sequence loci for use in polymerase chain reaction-based DNA fingerprinting. Electrophoresis 12: 113 118.

Ritland K, Jain S. 1981. A model for the estimation of outcrossing rate and gene frequencies using *n* independent gene loci. Heredity 47: 35-52.

Rockman M V, Hahn M W, Soranzo N, Loisel D A, Goldstein D B, Wray G A. 2004. Positive selection on MMP3 regulation has shaped heart disease risk. Current Biology 14: 1531-1539.

Rodrigues N B, LoVerde P T, Romanha A J, Oliveira G. 2002. Characterization of new *Schistosoma mansoni* microsatellite loci in sequences obtained from public DNA database and microsatellite enriched genomic libraries. Memórias do Instituto Oswaldo Cruz, Rio de Janerio, vol 97 (Suppl I): 71-75.

Saidman B O, Bessega C F, Ferreira L I, Julio N, Vilardi J C. 2000. The use of genetic markers to assess population structure and relationships among species of the genus *Prosopis* (Leguminosae). Boletín de la Sociedad Argentina de Botánica 35 (3-4): 315-324.

Saidman B O, Bessega C F, Ferreira L I, Vilardi J C. 1998a. Random amplified polymorphic DNA (RAPDs) variation in hybrid swarms and pure population of the genus *Prosopis* (Leguminosae). International Foundation for Science (IFS). Proceedings of the Workshop "Recent Advances in Biotechnology for Tree Conservation and Management", Florianópolis, 122-134.

Saidman B O, Vilardi J C, Montoya S, Dieguez M J, Hopp H E. 1998b. Molecular markers: a tool for the understanding of the relationship among species of *Prosopis* (Leguminosae, Mimosoideae). In. Puri S (ed.). Tree Improvement: Applied Research and Technology Transfer. Science Publisher Inc., USA. Pp. 311-324.

Saidman B O, Vilardi J C, Montoya S, Poggio L. 1997. Genetic variability and ploidy level in species of *Prosopis* (Leguminosae). Boletín de la Sociedad Argentina de Botánica 32: 217-225.

Saidman B O, Vilardi J C. 1987. Analysis of the genetic similarities among seven species of *Prosopis*. Theoretic and Applied Genetics, 75: 109-1116.

Saidman B O, Vilardi J C. 1993. Genetic variability and germplasm conservation in the genus *Prosopis*. In: Puri S (ed.). Nursery Technology of Forest Tree Species of Arid and Semiarid Regions. Winrock-Oxford & IBH Publishing Co., PVT, New Delhi, Bombay, Calcutta. Pp. 187-198.

Saidman B O. 1985. Estudio de la variación alozímica en el género *Prosopis*. Tesis Doctoral, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires.

Saidman B O. 1986. Isoenzymatic studies of alcohol dehydrogenase and glutamate oxalacetate transaminase in four South American species of *Prosopis* and their natural hybrids. Silvae Genetica 35: 3-10.

Saidman B O. 1988a. La electroforesis de isoenzymas para la medición de la variabilidad genética en especies de *Prosopis*. In: *Prosopis* en Argentina. Documento preliminar elaborado para el I Taller Internacional Sobre Recurso Genético y Conservación de Germoplasma en *Prosopis*. Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Argentina. FAO-PIRB. Pp. 107-118.

Saidman B O. 1988b. Isozyme studies on hybrids swarms of *Prosopis caldenia* and sympatric species. Silvae Genetica 39: 5-8.

Saidman B O. 1990. Isoenzyme studies on hybrid swarms of *Prosopis caldenia* and sympatric species. Genetics 39: 5-8.

Saidman B O. 1993. Las isoenzimas en el estudio de la variación genética y las afinidades entre especies de *Prosopis*. Boletín Genético del Instituto Fitotécnico de Castelar 16: 25-37.

SaikiI R K, Gelfand D H, Stoffel S, Scharf S J, Higuchi G, Horn G T, Mullis K B, Erlich H A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491.

Saneyoshi U, Yoshihiko T, Izumi W. 2003. Development of microsatellite markers in *Primula sieboldii* E. Morren, a threatened Japanese perennial herb. Conservation Genetics 4: 809-811.

Sawyer L A, Hennessy J M, Peixoto A A, Rosato E, Parkinson H, Costa R, Kyriacou C P. 1997. Natural variation in *Drosophila* clock gene and temperature compensation. Science. 278**:** 2117–2120.

Scalfi M, Troggio M, Piovani P, Leonardi S, Magnaschi G, Vendramin G G, Menozzi P. 2004. A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.). Theoretical and Applied Genetics 108: 433-441.

Schlotterer C, Tautz D. 1992. Slippage synthesis of simple sequence DNA. Nucleic Acids Research 20: 211-215.

Schug M D, Mackay T F C, Aquadro C F. 1997. Low mutation rates of microsatellite loci in *Drosophila melanogaster*. Nature Genetics 15: 99-102.

Schug M D, Wetterstrand K A, Gaudette M S, Lim R H, Hutter C M, Aquadro C F. 1998. The distribution and frequency of microsatellite loci in *Drosophila melanogaster*. Molecular Ecology 7: 57-70.

Scotti L J, Cross M, Shepherd M, Maguire T, Henry R J. 1999. Increasing the efficiency of microsatellite discovery from poorly-enriched Libraries in coniferous forest species. Plant Molecular Biology Reporter 17:351-354.

Shinde D, Lai Y, Sun F, Arnheim N. 2003. *Taq* DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: $(CA/GT)_n$ and $(A/T)_n$ microsatellites. Nucleic Acids Research 31: 974-980.

Simpson B B. 1977. Mesquite, its biology in two desert shrub ecosystems. Dowden, Hutchinson and Ross, Stroudsberg, Pennsylvania, U.S.A. 250 pages.

Smith D N, Devey M E. 1994. Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome 37: 977-983.

Smith G P. 1976. Evolution of repeated DNA sequences by unequal crossover. Science 191: 528–535.

Smouse P E, Dyer R J, Westfall R D, Sork V L. 2001. Two-generation analysis of pollen flow across landscape. I. Male gamete heterogeneity among females. Evolution 55: 260-271.

Sneath P H A, Sokal R R. 1973. Numerical Taxonomy. Freeman, San Francisco.

Solbrig O T, Bawa K S. 1975. Isozyme variation in species of *Prosopis* (Leguminosae). Journal of the Arnold Arboretum, 56: 398-412.

Solbrig O T, Cantilo P D. 1975. Reproductive adaptations in *Prosopis* (*Leguminosae, Mimosoideae*). Journal of the Arnold Arboretum 56: 185-210.

Soranzo N, Provan J, Powell W. 1999. An example of microsatellite length variation in the mitochondrial genome of conifers. Genome 42: 158-161.

Sork V L, Davis F W, Dyer R J, Smouse P E. 2002. Mating patterns in a savanna population of Valley Oak (*Quercus lobata* Neé). USDA Forest Service General Technical Report PSW-GTR-184.2002.

Streiff R, Ducousso A, Lexer C, Streinkellner H, Gloessl J, Kremer A. 1999. Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Leibl. Molecular Ecology 8: 831-841.

Streisinger G, Okada Y, Emrich J, Newton J, Tsugita A, Terzaghi E, Inouye M. 1966. Frameshift mutations and the genetic code. Cold Spring Harbor Symposium on Quantitative Biology 31: 77-84.

Suen T C, Hung M C. 1990. Multiple *cis*- and *trans*-acting elements involved in regulation of the *neu* gene. Molecular and Cellular Biology 10: 6306-6315.

Tabbener H E, Cottrell J E. 2003. The use of PCR based DNA markers to study the paternity of poplar seedlings. Forest Ecology and Management 179: 363-376.

Tautz D, Renz M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Research 12: 4127-4137.

Tautz D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Research 17: 6463-6471.

Templeton A R. 1989. The meaning of species and speciation: A genetic perspective. In Otte and Endler (eds.), Speciation and its Consequences. Sunderland, MA: Sinauer. Pp. 3-27.

Tortorelli L. 1956. Maderas y bosques argentinos. Acme Agency Press, Buenos Aires, Argentina.

Tóth G, Gáspári Z, Jurka J. 2000. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Research 10: 967-981.

Trewartha G T. 1943. An introduction to weather and climate, 2nd edition. MacGraw-Hill, New York, NY. 545 pp.

Turner P C, McLennant A G, Bates A D, White M R H (eds). 1997. Instant notes in molecular biology. BIOS scientific publishers ltd, NY. pp 57-86.

Ueno S, Yoshimaru H, Tomaru N, Yamamoto S. 1999. Development and characterization of microsatellite markers in *Camellia japonica* L. Molecular Ecology 8**:**335 336.

Ujino T, Kawahara T, Tsumura Y, Nagamitsu T, Yoshimaru H, Wickneswari R. 1998. Development and polymorphism of simple sequence repeat DNA markers for *Shorea curtisii* and other Dipterocarpaceae species. Heredity 81: 422-428.

Valdes A M, Slatkin M, Freimer N B. 1993. Allele frequencies at microsatellite loci: the stepwise mutation mode1 revisited. Genetics 133: 737-749.

Valle G. 1993. TA-repeat microsatellites are closely associated with ARS consensus sequences in yeast chromosome III. Yeast 9: 753-759.

van Treuren R, Kuittinen H, Karkkainen K, Baena-Gonzalez E, Savolainen O. 1997. Evolution of microsatellites in *Arabis petraea* and *Arabis lyrata***,** outcrossing relatives of *Arabidopsis thaliana***.** Molecular Biology and Evolution 14: 220-229.

Vander Stappen J, Weltjens I, Volckaert G. 1999. Microsatellite markers in *Stylosanthes guianensis*. Molecular Ecology 8: 514 517.

Vega M V, Hernández P. 2005. Molecular evidence for natural interspecific hybridization in *Prosopis*. Agroforestry System 64: 197-202.

Vendramin G G, Lelli L, Rossi P, Morgante M. 1996. A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. Molecular Ecology 5: 585-598.

Verga A R, Carranza C, Ledesma M, Joseau M J, Díaz P. 2000. Identificación de fuentes de semilla superior de *Prosopis sp* para enriquecimiento del bosque nativo en el Chaco árido. Reunión Nacional del Algarrobo. Mendoza, 14 al 17 de noviembre.

Verga A R, Gregorius H-R. In press. Genetic consistency of morphological differentiation in the *Prosopis chilensis*-*P. flexuosa* complex. Silvae Genetica (in press.)

Verga A R. 1995. Genetische Untersuchungen an *Prosopis chilensis* und *P. flexuosa* (Mimosaceae) im trockenen Chaco Argentiniens. Göttingen Research Notes in Forest Genetics 19. German w. sum., 96+VIII.

Verga A R. 2000. Clave para la identificación de híbridos entre *Prosopis chilensis* y *P. flexuosa* sobre la base de caracteres cuantitativos. Multequina 9: 17-22, 2000.

Verstrepen K J, Jansen A, Lewitter F, Fink G R. 2005. Intragenic tandem repeats generate functional variability. Nature Genetics 37: 986-990.

Viard F, El-Kassaby Y A, Ritland K. 2001. Diversity and genetic structure in populations of *Pseudotsuga menziesii* (Pinaceae) at chloroplast microsatellite loci. Genome 44: 336-344.

Viguera E, Canceill D, Ehrlich S D. 2001. Replication slippage involves DNA polymerase pausing and dissociation. The EMBO Journal 20: 2587-2595.

Vornam B, Decarli N, Gailing O. 2004. Spatial distribution of genetic variation in a natural beech stand (*Fagus sylvatica* L.) based on microsatellite markers. Conservation Genetics 5: 561-570.

Weber J L, May P E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. The American Journal of Human Genetics 44: 388-396.

Weber J L, Wong C. 1993. Mutation of human short tandem repeats. Human Molecular Genetics 2:1123-1128.

Weber J L. 1990. Informativeness of human $(dC\text{-}dA)_n$ $(dG\text{-}dT)_n$ polymorphisms. Genomics 75: 24-30.

Weir B S. 1990. Genetic data analysis II. Sinauer Associates, Sunderland, Massachusetts 445pp.

Weising K, Gardner R C. 1999. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome 42: 9-19.

Weising K, Nybom H, Wolf K, Kahl G. 2005. DNA fingerprinting in plants. Principles, methods, and applications. Taylor and Francis, Boca Raton.

Weldon D. 1986. Exceptional physical properties of Texas mesquite wood. Forest Ecology and Management 16: 149:153.

White M J D. 1978. Modes of speciation. Freeman, San Francisco, USA.

Whitlock M C, McCauley D E. 1999. Indirect measures of gene flow and migration: Fst≠1/(4*Nm*+1). Heredity 82: 117-125.

Wierdl M, Dominska M, Petes T D. 1997. Microsatellite instability in yeast: dependence on the length of the microsatellite. Genetics 146: 767-779.

Wolff R K, Plaetke R, Jeffreys A J, White R. 1989. Unequal crossing over between homologous chromosomes is not the major mechanism involved in the generation of new alleles at VNTR loci. Genomics 5: 382-384.

Woodhead M, Russell J, Squirrell J, Hollingsworth P M, Cardle L, Ramsay L, Gibby M, Powell W. 2003. Development of EST-SSRs from the Alpine lady fern, *Athyrium distentifolium*. Molecular Ecology Notes 3**:** 287 290.

Wright S. 1951. The genetical structure of populations. Annals of Eugenics 15: 323-354.

Wu K, Jones R, Dannaeberger L, Scolnik P A. 1994. Detection of microsatellite polymorphisms without cloning. Nucleic Acids Research 22: 3257–3258.

Wünsch A, Hormaza J I. 2002. Molecular characterization of sweet cherry (*Prunus avium* L.) genotypes using peach *Prunus persica* (L.) Batsch SSR sequences. Heredity 89: 56-63.

Wyman J, Bruneau A, Tremblay M-F. 2003. Microsatellite analysis of genetic diversity in four populations of *Populus tremuloides* in Quebec. Canadian Journal of Botany 81: 360-367.

Xu Y, Ma R C, Xie H, Liu J T, Cao M Q. 2004. Development of SSR markers for the phylogenetic analysis of almond trees from China and the Mediterranean region. Genome 47: 1091-1104.

Yamamoto T, Kimura T, Shoda M, Ban Y, Hayashi T, Matsuta N. 2002. Development of microsatellite markers in the Japanese pear (*Pyrus pyrifolia* Nakai). Molecular Ecology Notes 2: 14 16.

Zane L, Bargelloni L, Patarnello T. 2002. Strategies for microsatellite isolation: a review. Molecular Ecology 11: 1-16.

Zietkiewicz E, Rafalski A, Labuda D. 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20: 176–183.

12. LIST OF PUBLICATIONS

Mottura Martín C, Gailing Oliver, Verga Aníbal R, Finkeldey Reiner. 2004. Efficiency of microsatellite enrichment in *Prosopis chilensis* using magnetic capture. Plant Molecular Biology Reporter 22: 251-258.

Mottura Martín C, Gailing Oliver, Finkeldey Reiner. 2004. Establishment of microsatellite markers in *Prosopis chilensis* and *P. flexuosa*. Botanikertagung 2004, Technische Universität Carolo-Wilhelmina zu Braunschweig und Deutsche Botanische Gesellschaft. Braunschweig, Germany. Pp. 443 (Poster).

Mottura Martín C, Gailing Oliver, Verga Aníbal R, Finkeldey Reiner. 2005. Development and characterization of microsatellite markers for *Prosopis chilensis* and *P. flexuosa* and cross-species amplification. Molecular Ecology Notes 5: 487-489.

Mottura Martín C, Gailing Oliver, Verga Aníbal R, Finkeldey Reiner. 2005. Microsatellite development in *Prosopis chilensis*. Plant and Animal Genome XIII, 2005. San Diego, California, USA. Pp. 112 (Poster).

Mottura Martín C, Gailing Oliver, Córdoba Ana A, Verga Aníbal R, Finkeldey Reiner (in prep.). Mating system analysis of a hybrid swarm between *Prosopis chilensis* and *P. flexuosa* (Fabaceae) based on microsatellite markers.

I. Efficiency of microsatellite enrichment in *Prosopis chilensis* using magnetic capture

Martín C. Mottura [1, 2], Oliver Gailing [1], Aníbal R. Verga [2], Reiner Finkeldey [1]

[1] Institute of Forest Genetics and Forest Tree Breeding, Georg-August University, Göttingen, Büsgenweg 2, Göttingen, 37077, Germany

[2] IFFIVE-INTA (Institute of Plant Physiology and Plant Pathology – National Institute of Agricultural Technology) Camino 60 Cuadras Km 5.5, Córdoba, 5119, Argentina

Abstract

Microsatellites (i.e., simple sequence repeat [SSRs]) are highly variable genetic markers that are widely used at an intraspecific level in population genetic studies. Here, we employed an enrichment strategy for microsatellite isolation by using microsatellite oligoprobes and magnetic capture of the fragments (Fischer and Bachmann, 1998) in *Prosopis chilensis* (Mol.) Stuntz (Fabaceae). We analyzed the obtained level of enrichment by sequencing 120 enriched genomic fragments. A total of 521 SSR motives were detected. According to specific search criteria (SSR motifs ≥ 3 repeat units and ≥ 6 bp length), 95.8 % of the clones contained SSR motifs. Of these, 7.8% showed homology to chloroplast sequences and 92.2% to nuclear sequences. When regarding only nuclear SSRs with five or more repeat units and a minimum length of 10 bp, the level of enrichment was 30.8%. A FASTA search against the European Molecular Biology Laboratory (EMBL) database univocally revealed four clones in transcribed regions, 102 clones in genomic regions with unknown function and 9 clones in chloroplast regions. Among the loci with longer repeat units (≥ 10 bp, ≥ 5 repeat units), three were in transcribed regions and 65 in other genomic regions. We discuss the applicability of these markers for population genetic studies.

Key words: magnetic capture, microsatellite enrichment, microsatellites, molecular marker, Prosopis chilensis, SSRs.

Abbreviations: SSR = simple sequence repeat.

Introduction

Nuclear microsatellites (Litt and Luty, 1989; Weber and May, 1989), also known as simple sequence repeats (SSRs) (Tautz, 1989), are some of the most useful genetic markers because of their abundance and highly polymorphic nature. They have proven to be extremely valuable tools for the analysis of gene flow, mating system and population genetics in many plant species, including forest trees, because of their usually high variation and codominant inheritance. A major limitation for the application of microsatellite markers is their usually low transferability between species. Generally, new markers must be developed for each species, although it might be possible to transfer SSRs between closely related species of the same genus.

For population genetic studies, such as for the study of the mating system, a limited number of highly variable microsatellite loci are sufficient. The establishment of microsatellite enriched libraries is a fast and low-cost method for microsatellite marker development. Several enrichment protocols have been reported based on: *(1)* capture by streptavidin-coated magnetic beads (e.g., Kijas et al., 1994; Fischer and Bachmann, 1998; Hamilton et al., 1999); *(2)* microsatellite probes attached to small nylon membranes (e.g., Karagyzov et al., 1993; Edwards et al, 1996); *(3)* other not frequently used procedures such as magnetic capture of phagemid DNA (Paetkau, 1999). The efficiency of these protocols is estimated as the number of clones that contain microsatellite motifs. The estimates range from 10 to 95% (Cordeiro et al., 1999; Jakse and Javornik, 2001). Few detailed reports discuss the length of these SSR motifs and their position in the genome. We used an enrichment strategy reported by Fischer and Bachmann (1998) for the isolation of microsatellites in *Prosopis chilensis* (Mol.) Stuntz (Fabaceae). This method is based on affinity capture of single stranded restriction fragments annealed to biotynilated microsatellite oligonucleotides and subsequent magnetic separation. We evaluate the results of the enrichment by applying specific criteria: the length of the SSR motifs, specificity of the microsatellite oligoprobes, and position of the SSR loci in the genome. Sequences with SSR motifs were compared with genomic database sequences of other plant species European Molecular Biology Laboratory (EMBL) database. The number of sequences containing chloroplast and nuclear SSR loci and the number of SSR motifs in transcribed and nontranscribed regions was determined.

Results and Discussion

A total of 120 clones with an average length of 580 bp (± 96 bp) were sequenced. A search for SSR motifs was performed using the program BUSCADOR (available upon request). We looked for microsatellites composed of tandemly repeated units of 1 to 6 nucleotides, repeated at least three times and of a total size of at least six bases. According to Jurka and Pethiyagoda (1995), most simple repeats and their complementary counterparts can be represented by several different basic unit patterns. For example, the pattern $(GCC)_n$ also represents $(CCG)_n$, $(CGC)_n$, $(GGC)_n$, $(GCG)_n$ and $(CGG)_n$. Furthermore, whenever tandemly repeated patterns with different unit sizes are identical, they are listed under the smallest unit size. For example, $(CACACA)_n$ or $(CACA)_n$ patterns are considered as $(CA)_n$.

Of the clones sequenced, 95.8% (115 out of 120) showed at least one SSR motif matching the search criteria, 30.8% when only clones with long-length SSR segments (at least five repeat units and a minimum length of ten bp) were regarded. Similar protocols have shown levels of enrichment between 11% and 99% (Edwards et al., 1996; Paetkau, 1999; Rodrigues et al., 2002; Saneyoshi et al., 2003; Pandey et al., 2004), but there is no information on the selection criteria for the SSR segments. Most of the SSR segments found (444 out of 521) had only three or four repeat units or a length six to nine bp for mononucleotide repeats.

Sixteen percent (84) of all the detected motifs showed specific complementary sequences with the oligoprobes [11% (58) with the (CA) probe, 4% (21) with (GAA) and 1% (5) with (AAC)]. The percentage of motifs with partial similarity to the oligoprobes, e.g., $(GAAA)_n$ and $(AACA)_n$ was 33% (68% when mononucleotide repeat motifs are included). In total, ten SSR loci with longer repeat units (≥ 5 repeat units) showed the motif (CA), four loci showed the repeat motif (GAA) and one locus the motif (AAC) five or more times repeated.

All sequences (115 clones) that contained SSR motifs according to the search criteria were compared against the EMBL Nucleotide Plant Sequence Database (EMBL-Bank) using the program FASTA 3 (Pearson and Lipman, 1988; Pearson, 1990) for homology and similarity searching (http://www.ebi.ac.uk). Of the 115 clones, 7.8 % (9 clones) had high similarity with chloroplast regions, with an average of 4.3 SSRs per clone, and contained mainly mononucleotide repeats. Most of the clones (102 clones) with similarity to nuclear DNA conform to nontranscribed regions of different species. Only

four clones showed high similarity (> 60% identity over > 180 bp length with an expectation value < 1*e $^{-10}$) with transcribed regions. One of them was located in an intron region, three were in translated exon regions (**Table 2**). In three cases the clones contain only a part of the transcribed region (partial overlap); in clone R5 17, the total exon region is included.

In 106 clones with similarity to nuclear sequences, 466 SSR motifs were detected, including 216 mononucleotide repeats. A frequency distribution of SSR segments according to the length of microsatellite segments (number of repeats) is given in **Figure 1a - b**. Sixty-eight (14,6%) of these segments revealed microsatellite motifs with at least five repeat units and a minimum length of ten bp. Most of them (44 segments) were mononucleotide repeats, followed by dinucleotide repeats (17) and trinucleotide repeats (7) (**Figure 1a - b**, **Table 1**).

Figure 1

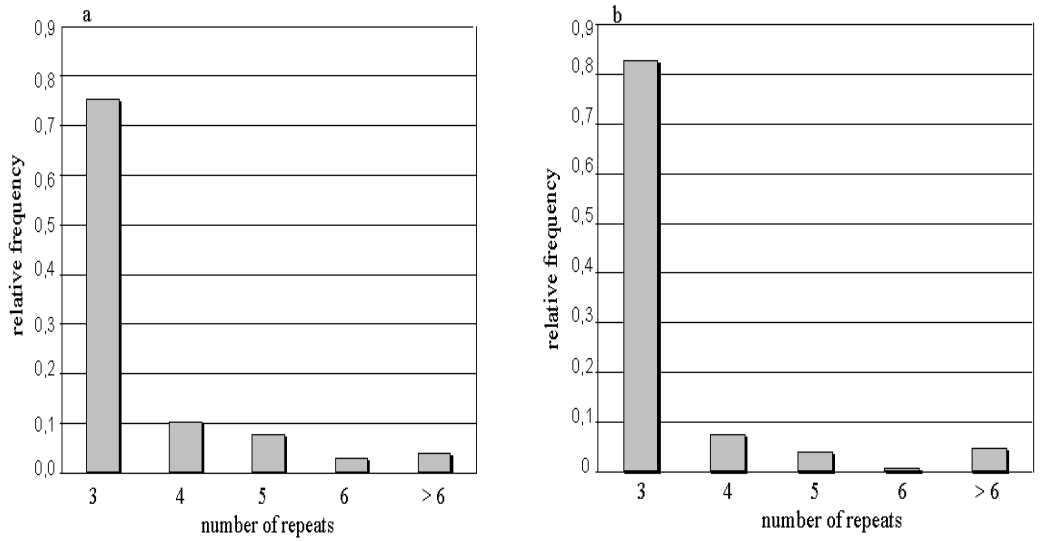

Figure 1: The relative frequency of simple sequence repeat (SSR) motifs with different number of repeats is shown. *(a)* Mononucleotide motifs are included. *(b)* Mononucleotides are excluded. For the definition of the categories (3 to >6) see Table 1, column 1.

Table 1: Analysis of microsatellites found in 106* *Prosopis chilensis* sequences from one *P. chilensis* SSR-enriched library.

Characteristics of SSRs	Number of SSRs of each motif																							
	Total n° of SSRs	Monont. repeat		Dinucleotide repeat				Trinucleotide repeat										Tetranucleotide repeat					Pentant. Repeat	Hexant repeat
		(A/T)	(C/G)	CA	AG	AT	CG	GAA	CTA	TAA	GGA	ACG	TCA	GCA	GCC	AAC	CCA	TGAT	GCAT	ATAC	AAAG	CCAA	AAAGA	AGAAAA
N° of repeats																								
3 (for monont. motifs 6 to 7 bp long)	351	96	48	42	66	46	3	16	3	8	2	2	3	3	1	3	1	1	1	1	1	1	2	1
4 (for monont. motifs 8 to 9 bp long)	47	22	6	4	6	4	1	1	1				1								1			
5 (for monont. motifs 10 to 11 bp long)	36	16	10	1	2	1		3		1						1	1							
6 (for monont. motifs 12 to 13 bp long)	14	10	2		2																			
> 6 (for monont. motifs > 13 bp long)	18	6		9	1		1	1																
Total n° of SSRs	466	150	66	56	77	51	5	21	4	9	2	2	4	3	1	4	2	1	1	1	2	1	2	1
Percentage	100	32,2	14,2	12,0	16,5	10,9	1,1	4,5	0,9	1,9	0,4	0,4	0,9	0,6	0,2	0,9	0,4	0,2	0,2	0,2	0,4	0,2	0,4	0,2
Longest SSR found (in bp)		17	13	42	16	10	16	21	12	15	9	9	12	9	9	15	15	12	12	12	16	12	15	18
N° of SSRs for which PCR primers were designed		3		9	2	1	1	2																

*Sequences with similarity to chloroplast regions were not included in this table.

Table 2: Clones with similarity* to transcribed regions.

Clone	SSR motif	EMBL matching mRNA for	mRNA region	Gene	Suspected functions	Species	Expectation value	Overlap (nt)	Similarity (%)
R4 16	$(GAA)_7$	CDPK-related protein Ca/calmodulin-dependent protein	exon	CRK (or PK421)	Plant growth and development	*Daucus carota*	1,1 *e^{-12}	185	77,8
						Nicotiana tabacum	1,2 *e^{-11}	200	74,5
R5 17	$(A)_7$	60s ribosomal protein L44	exon	*Gene encoding 60s ribosomal protein L44*	*Synthesis of 60s ribosomal protein L44*	*Oryza sativa*	3,4 *e^{-38}	486	66,9
R5 24	$(AC)_{17}$	growth regulator like protein	exon	Gene encoding protein At4g16650/dl4350w	*Growth regulation*	*Arabidopsis thaliana*	4,1 *e^{-19}	227	79,3
R7 25	$(CCAA)_3$	RING finger protein	intron	Gene encoding RING finger protein	Plant growth and development	*Lotus corniculatus*	1,8 *e^{-20}	542	62

*Only clones with more than 60% identity over more than 180 bp length with an expectation value < 1*e^{-10} were considered.

Three microsatellite regions (di- , tri- and tetranucleotide repeats) are in transcribed regions with suspected function in plant growth regulation (**Table 2**). Recent studies (Morgante et al., 2002) claim that most of SSR motifs in plants are located in transcribed regions, but results presented here do not support these findings. The identification of selected (nonneutral) SSR loci with similarity to known genes could be useful to test association of microsatellite alleles with adaptive trait variation in natural populations (test of neutrality; Watterson, 1978) and for mapping studies. SSRs in nontranscribed regions are particularly useful to reconstruct descent in studies of the reproduction system of plants.

Acknowledgements

This research was supported by the Ministry for Science and Culture of Lower Saxony, Germany, within the program Nachwuchswissenschagtler/innen aus außereuropäischen Ländern nach Niedersachsen. The authors thank Madhav Pandey (Forest Genetics, Georg-August University, Göttingen) for his technical support and Dr. Patrik Hoegger (Forest Botany, Georg-August University, Göttingen) for his help during the cloning and transformation stages.

References

Cordeiro G M, Maguire T L, Edwards K J, Henry R J. 1999. Optimization of a microsatellite enrichment technique in *Saccharum* spp. Plant Molecular Biology Reporter 17: 225-229.

Edwards K J, Barker J H A, Daly A, Jones C and Karp A. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20: 758-760.

Fischer D, Bachman K. 1998. Microsatellite enrichment in organisms with large genomes (*Allium cepa L.*). BioTechniques 24: 796-802.

Hamilton MB, Pincus EL, Di Fiore A and Fleischer RC. 1999. Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. BioTechniques 27: 500-507.

Jakse J, Javornik B. 2001. High throughput isolation of microsatellites in hop (*Humulus lupulus* L.). Plant Molecular Biology Reporter 19: 217-226.

Jurka J, Pethiyagoda C. 1995. Simple repetitive DNA sequences from primates: compilation and analysis. Journal of Molecular Evolution 40: 120-126.

Karagyozov L, Kalcheva I D, Chapman V M. 1993. Construction of random small-insert genomic libraries highly enriched for simple sequence repeats. Nucleic Acids Research 21: 3911-3912.

Kijas JMH, Fowler JCS, Garbett CA and Thomas MR. 1994. Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. BioTechniques 16: 657-662.

Litt M, Luty J A. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics 44: 397-401.

Marchuk D, Drumm M, Saulino A, Collins F S. 1990. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Research 19: 1154.

Morgante M, Hanafey M, Powell W. 2002. Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nature Genetics 30: 194-200.

Paetkau D. 1999. Microsatellites obtained using strand extension: an enrichment protocol. BioTechniques 26: 690-697.

Pandey M, Gailing O, Fischer D, Hattemer H H, Finkeldey R. 2004. Characterization of microsatellite markers in sycamore (*Acer pseudoplatanus* L.). Molecular Ecology Notes 4: 253-255.

Pearson W R, Lipman D J. 1988. Improved tools for biological sequence comparison. Proceedings of the Natural Academy Sciences of the USA 85:2444- 2448.

Pearson W R. 1990. Rapid and Sensitive Sequence Comparison with FASTP and FASTA Methods in Enzymology 183:63 - 98.

Rodrigues N B, LoVerde P T, Romanha A J, Oliveira G. 2002. Characterization of new *Schistosoma mansoni* microsatellite loci in sequences obtained from public DNA

database and microsatellite enriched genomic libraries. *Mem Inst Oswaldo Cruz*, Rio de Janerio, vol 97 (Suppl I): 71-75.

Saneyoshi U, Yoshihiko T, Izumi W. 2003. Development of microsatellite markers in *Primula sieboldii* E. Morren, a threatened Japanese perennial herb. Conservation Genetics 4: 809-811.

Tautz D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Research 17: 6463-6471.

Watterson G A. 1978. The homozygosity test of neutrality. Genetics 88: 405-417.

Weber J L, May P E. 1989. Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. American Journal of Human Genetics 44: 388-396.

II. Development and characterization of microsatellite markers for *Prosopis chilen*sis and *P. flexuosa* and cross-species amplification

Martín C. Mottura [1, 2], Oliver Gailing [1], Aníbal R. Verga [2], Reiner Finkeldey [1]

[1] Institute of Forest Genetics and Forest Tree Breeding, Georg-August University, Göttingen, Büsgenweg 2, Göttingen, 37077, Germany

[2] IFFIVE-INTA (Institute of Plant Physiology and Plant Pathology – National Institute of Agricultural Technology) Camino 60 Cuadras Km 5.5, Córdoba, 5119, Argentina

Abstract

Prosopis chilensis and *P. flexuosa* (Fabaceae) are closely related hardwood arboreal species that are widely distributed in the arid regions of Argentina. The development of highly polymorphic markers, such as microsatellites, is desirable for genetic studies of these species. Here, we present the development and characterization of six polymorphic microsatellite markers in *Prosopis chilensis* and *P. flexuosa*. These markers showed a polymorphism information content between 0.14 and 0.85 and the number of alleles varied from two to 13 considering both species. All markers revealed a broad cross-species affinity when tested in seven other *Prosopis* species. All primers amplified in at least five species.

Keywords: cross-species amplification, microsatellites, *Prosopis chilensis, Prosopis flexuosa*, SSRs.

Manuscript

The genus *Prosopis* L. (Fabaceae) constitutes a very important resource in many arid and semiarid regions of the world. In the Argentinean Arid Chaco, *Prosopis chilensis* (Mol.) Stuntz and *P. flexuosa* D.C., two hardwood arboreal species, are of great value as multipurpose species. Some previous studies on genetic variation, gene flow and on the reproduction system were conducted with morphological characters, isozymes or random amplified polymorphic DNA (RAPDs) markers (Verga, 1995; Bessega *et al.*, 2000; Saidman *et al.*, 2000). These studies show a low level of genetic differentiation among species and strongly suggest the existence of interspecific hybrids. The development of highly polymorphic microsatellite (SSR = simple sequence repeat) markers will provide valuable information on the mating system and on microevolutionary processes that occur within hybrids. The aim of this study was to develop SSR markers for *P. chilensis* and *P. flexuosa* and to test them for cross-amplification in seven other *Prosopis* species found in Argentina.

Total genomic DNA of *Prosopis chilensis* was extracted from five-day-old green cotyledons of a single tree using DNeasy® PlantMiniKit (QIAGEN) following the manufacturer's instructions. Microsatellites were isolated from total genomic DNA according to an enrichment protocol developed by Fischer and Bachmann (1998). DNA was digested with *RsaI* restriction enzyme (Roche) and 21mer (5´-CTCTTGCTTACGCGTGGACTA-3´) and 25mer (5´phosphorylated TAGTCCACGCGTAAGCAAGAGCACA-3´) adapters were ligated to the fragments. The products of the restriction-ligation reaction were then heat-denaturated and allowed to hybridize to three single-stranded 5´-biotinylated microsatellite oligonucleotides [$(CA)_{10}$ $(GAA)_8$ $(AAC)_8$]. These hybrids were captured by streptavidin-coated magnetic beads (Dynabeads®; Dynal GmbH). The magnetic beads were eluted, and the obtained DNA solution served as template for polymerase chain reaction (PCR) using the 21mer oligonucleotide as primer. The amplification product was directly ligated into pBluescript SK+ (Stratagene) by TA cloning according to Marchuk *et al.* (1990). The ligated vector fragments were transformed into competent *Escherichia coli* SURE cells (Stratagene) and plated onto LB agar medium. Recombinant plasmids were identified by means of blue-white screening and colony PCR was performed. Fragments were cleaned with MinElutePCR Purification Kit (QIAGEN). Sequencing was performed

using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and run on ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems/HITACHI). The efficiency of the enrichment was evaluated by the application of different criteria (Mottura *et al.*, 2004). When regarding only nuclear SSRs with five or more repeat units and a minimum length of 10 bp, the level of enrichment was 30.8%.

Out of 120 simple and non-interrupted sequences, 18 were selected for primer development considering those sequences with microsatellites with at least five repeats. Primer pairs were designed using the software PRIMER3 (Rozen and Skaletsky, 2000) in microsatellite flanking regions of the following SSR motifs: three $(A)_n$, nine $(AC)_n$; two $(CT)_n$; one $(AT)_n$; one $(GC)_n$; and two $(GAA)_n$. All primer pairs were tested and nine out of 18 primers amplified products of the expected size. Each forward primer was labeled with a fluorescent dye (6-FAM or HEX) at its 5`end. Variation was characterized in a natural population of *Prosopis chilensis* (20 trees) and *P. flexuosa* (20 trees) of the Argentinean Arid Chaco. DNA was extracted using DNeasy® PlantMiniKit (QIAGEN). PCR amplifications were carried out in a 15-μL reaction volume containing 10 ng template DNA, 1.5 mM of $MgCl_2$, 10 mM Tris-HCl pH 9.0, 50 mM KCl, 0.15 mM of each dNTPs, 0.5 units of Taq DNA polymerase and 0.6 μM of each primer on a PTC-200 (Mj Research) thermal cycler with the following cycling profile: initial denaturation at 94°C for 5min followed by 30 cycles at 94°C for 45 s denaturation, primer-specific annealing temperature (**Table 1**) for 45 s and at 72°C for 45 s extension, followed by a final extension step at 72°C for 10 min. Amplified products were analyzed on the ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems/HITACHI) and precisely sized using GENESCAN 3.7 and GENOTYPER 3.7 computer software (Applied Biosystems). Six primer pairs amplified reproducible and well-scorable bands which were polymorphic in both species (**Table 1**), two primer pairs amplified fragments that showed no variation among the tested populations and one generated non-reproducible bands.

Details about the six newly developed microsatellite loci and variability measures across 20 individuals of each species are summarized in **Table 1**. The number of alleles ranged from three to seven in *P. chilensis* and from two to 13 in *P. flexuosa*. Expected Heterozygosity (H_E) and polymorphism information content (PIC) in *P. chilensis* ranged from 0.14 to 0.73 and from 0.14 to 0.70, respectively, and in *P. flexuosa* from 0.46 to

0.86 and from 0.41 to 0.85. These results are in agreement with earlier studies supporting a higher genetic variation in *P. flexuosa*. Hardy-Weinberg equilibrium (HWE) probability tests (GENEPOP 3.3, http:// wbiomed.curtin.edu.au/genepop; Raymond and Rousset, 1995) showed significant ($P < 0.05$) deviation from HWE for loci Mo05 and Mo13 in *P. chilensis* and for loci Mo07, Mo08, Mo09 and Mo16 in *P. flexuosa*. Strong deviations from HWE ($P < 0.001$) were observed in *P. chilensis* for Mo05 and in *P. flexuosa* for loci Mo07 and Mo16, possibly due to the presence of null alleles. No highly significant ($P < 0.01$) linkage disequilibrium (GENEPOP 3.3, option 2) was found among loci. Only two pairs of markers showed linkage disequilibrium ($P < 0.05$) (Mo09/Mo13, $P = 0.04$ in *P. chilensis*, and Mo08/Mo13, $P = 0.03$ in *P. flexuosa*). Regular segregation was tested for all loci in single tree progenies following the method developed by Gillet and Hattemer (1989). All markers showed regular meiotic segregation and random fertilization of the eggs by pollen.

The six primer pairs were tested for cross-species amplification in seven other *Prosopis* species. DNA was extracted from eight individuals per species and PCR was conducted as described previously. Results for cross-species amplification are summarized in Table 2. All primers showed a highly robust amplification in most of the species tested. The number of alleles varied from 1 to 5 (**Table 2**).

Table 1 Characterization of microsatellite loci in *Prosopis chilensis* and *P. flexuosa* (20 individuals of each species were used in the characterization)

Locus name	Primer sequences (5´-3´)	T_a (°C)	Repeat motif of the cloned allele	Size of the cloned allele (bp)	*Prosopis chilensis*						*Prosopis flexuosa*						Gene Bank Accession Number
					Observed alleles size range (bp)	N_a	H_O	H_E	PIC	HWE test	Observed alleles size range (bp)	N_a	H_O	H_E	PIC	HWE test	
Mo05	**F:** AATTCTGCAGTCTCTTCGCC **R:** GATCCCTCGTGACTCCTCAG	64	$(CT)_3T(CT)_2$	218	214-218	3	0.20	0.54	0.49	<0.001	216-218	2	0.28	0.46	0.41	NS	AJ879505
Mo07	**F:** GAAGCTCCCTCACATTTTGC **R:** CTATTTGCGCAACACACAGC	59	$(GC)_8$	197	189-213	5	0.58	0.46	0.45	NS	185-213	8	0.41	0.78	0.77	<0.001	AJ879506
Mo08	**F:** TATCCTAAACGCCGGGCTAC **R:** TCCCATTCATGCATACTTAAACC	59	$(AC)_9$	218	208-222	6	0.50	0.50	0.48	NS	204-222	10	0.60	0.85	0.84	0.003	AJ879507
Mo09	**F:** ATTCCTCCCTCACATTTTGC **R:** CATTATGCCAGCCTTTGTTG	59	$(TG)_{17}$	233	209-211	2	0.15	0.14	0.14	NS	203-245	13	0.75	0.84	0.84	0.018	AJ879508
Mo13	**F:** TTGATTAGAGTTGCATGTGGATG **R:** TGCAGTCCCAAGTGTCAGAG	58	$(GT)_{10}CT(GT)_2$	228	218-246	6	0.90	0.68	0.66	0.046	216-240	8	0.45	0.56	0.55	NS	AJ879509
Mo16	**F:** CATTGCCCCAATATCACTCC **R:** GGGTCCATCCAGAGTAGTGG	60	$(CA)_{12}$	157	149-163	5	0.82	0.73	0.70	NS	147-175	11	0.43	0.86	0.85	<0.001	AJ879510

T_a, annealing temperature; bp, base pairs; N_a, total number of different alleles observed; H_O, observed heterozygosity; H_E, expected heterozygosity; PIC, Polymorphism Information Content; HWE test, *P-Values* for Hardy-Weinberg Equilibrium test; NS, not significant ($P > 0.05$).

Table 2 Cross-species amplification status of 6 microsatellites markers developed for *Prosopis chilensis* tested in 7 *Prosopis* species that occur naturally in Argentina (in each case 8 samples per species were used in the analysis)

Prosopis	**Microsatellite marker**					
species	**Mo05**	**Mo07**	**Mo08**	**Mo09**	**Mo13**	**Mo16**
P. alba	216-218 (2)	193 (1)	216 (1)	209 (1)	228-240 (4)	151-159 (4)
P. caldenia	218-220 (2)	185-193 (2)	208-218 (5)	213-217 (3)	216 (1)	141-157 (4)
P. ferox	218 (1)	—	—	209-215 (2)	216-228 (2)	—
P. hassleri	—	185-209 (5)	216-218 (2)	207-227 (4)	228-240 (4)	147-157 (4)
P. nigra	218 (1)	185-193 (2)	204-216 (4)	207-209 (2)	220-236 (5)	155-163 (4)
P. ruscifolia	218 (1)	185-195 (4)	212-218 (3)	209 (1)	230-236 (3)	—
P. torquata	218 (1)	—	—	207-209 (2)	228-238 (2)	155-163 (2)

The given numbers in the table indicate the amplified allele size range and (number of alleles); " — " indicates no amplification or unclear banding pattern.

Acknowledgements

This research was supported by the Ministry of Science and Culture of Lower Saxony, Germany. The technical assistance of Mr. T. Seliger is gratefully acknowledged.

References

Bessega C, Saidman B O, Vilardi J C. 2000. Isozymes and RAPD studies in *Prosopis glandulosa* and *P. velutina* (Leguminosae, Mimosoideae). Genetics and Molecular Biology 23: 1-5.

Fischer D, Bachmann K. 1998. Microsatellite enrichment in organisms with large genomes (*Allium cepa L.*). Biotechniques 24: 796-802.

Gillet E, Hattemer H H. 1989. Genetic analysis of isoenzyme phenotypes using single tree progenies. Heredity 63: 135-141.

Marchuk D, Drumm M, Saulino A, Collins F S. 1990. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Research 19: 1154.

Mottura M C, Gailing O, Verga A R, Finkeldey R. 2004. Efficiency of microsatellite enrichment in *Prosopis chilensis* using magnetic capture. Plant Molecular Biology Reporter 22: 251-258.

Raymond M, Rousset F. 1995. GENEPOP Version 1.2: population genetics software for exact test and ecumenicism. Journal of Heredity 886: 248-249.

Rozen S, Skaletsky H J. 2000. PRIMER 3 on the www for general users and for biologist programmers. In: Bioinformatics Methods and Protocols: Methods in Molecular Biology (eds. Krawertz S, Misener S), pp. 365-386. Humana Press, Totowa, NJ, USA

Saidman B O, Bessega C, Ferreyra L I, Julio N, Vilardi J C. 2000. The use of genetic markers to assess population structure and relationships among species of the genus *Prosopis* (Leguminosae). Boletín de la Sociedad Argentina de Botánica 35 (3-4): 315-324.

Verga A R. 1995. Genetische Untersuchungen an *Prosopis chilensis* und *P. flexuosa* (Mimosaceae) im trockenen Chaco Argentiniens. Göttingen Research Notes in Forest Genetics. Abteilung für Forstgenetik und Forstpflanzenzüchtung der Universität Göttingen ISSN 0940-7103, No. 19.

III. Mating system analysis of a hybrid swarm between *Prosopis chilensis* and *P. flexuosa* (Fabaceae) based on microsatellite markers.

Martín C. Mottura[12], Oliver Gailing[1], Ana A. Córdoba[2], Aníbal R. Verga[2], Reiner Finkeldey[1]

[1] Institute of Forest Genetics and Forest Tree Breeding, Georg-August University, Göttingen, Büsgenweg 2, Göttingen, 37077, Germany

[2] IFFIVE-INTA (Institute of Plant Physiology and Plant Pathology – National Institute of Agricultural Technology) Camino 60 Cuadras Km 5.5, Córdoba, 5119, Argentina

Abstract

Three highly polymorphic microsatellite markers (SSRs = Single Sequence Repeats) were applied in the study of the mating system of *Prosopis chilensis* and *P. flexuosa* in a sympatric area in the Argentinean Chaco Árido. We collected 516 seeds during the flowering season 2003-2004 from 23 open-pollinated mother trees among 100 adult trees of a hybrid swarm between *P. chilensis* and *P. flexuosa*. Three morphological groups were previously identified in the hybrid swarm, i.e. the two pure taxonomic species and a morphologically intermediate group. Genetic variation and differentiation parameters were estimated for each group and compared across the adult trees and the pollen clouds. Significant reductions in the genetic distance were observed in the pollen clouds compared to the adult trees. In addition, paternity analysis was used to assess gene flow between morphological groups. The results show high degrees of inter-group hybridization (40.48%), and asymmetric gene flow between groups. Estimates of outcrossing rate indicate that the species are predominantly outcrossing, but that selfing rates are high for a few trees.

Introduction

The study of hybrid complexes can provide excellent models to address evolutionary questions. Such models allow inference on speciation mechanisms and population dynamics (Bessega et al., 200b). In addition, they permit to study interspecific gene flow and the consequences of hybridization.

The genus *Prosopis* (Family Fabaceae) occurs worldwide in arid and semiarid regions. It includes 44 species grouped in five Sections (Burkart, 1976). *Prosopis chilensis (Mol.) Stuntz* and *P. flexuosa D.C.* are closely related hardwood arboreal species belonging to the section Algarobia. In the Argentinean *Chaco Árido*, these species constitute important natural resources. *Prosopis* trees are of great value because they are considered multipurpose species (Karlin and Díaz, 1984; Cony, 2000). They provide timber, fuel wood and forage. In addition, they stabilize the soil and prevent erosion.

Natural *Prosopis* forests have been over-exploited without much attention being paid to their conservation (Saidman et al., 1998). The irrational exploitation of this natural resource has caused the loss of many elite trees, degradation of the local populations of *Prosopis*, and sometimes even the extinction of extensive *Prosopis* forests (Cony, 2000; Karlin and Díaz, 1984; Verga et al., 2000).

Reforestation plans and forest enrichment strategies, as well as conservation programs are necessary in order to recover degraded environments and preserve the resource (Verga et al., 2000). This requires deep knowledge of the existing variability and structure of local populations and descendants, reproductive patterns and adaptive strategies of the species (Cony, 2000; Verga et al., 2000).

Previous studies of genetic variation, interspecific geneflow and the reproductive system conducted with morphological characters, isozymes or RAPDs markers (Solbrig and Bawa, 1975; Saidman, 1990; Saidman and Vilardi, 1987; Verga, 1995; Saidman et al., 1998; Saidman et al., 2000; Bessega et al., 2000b) indicated low levels of genetic differentiation among the species of section *Algarobia* and the existence of interspecific hybrids between at least seven species of this section (*P. alba*, *P. alpataco, P. caldenia, P. chilensis*, *P. flexuosa*, *P. nigra* and *P. ruscifolia*) (**Fig. 1**). In addition, hybrid swarms are described in areas where the natural habitat has been disturbed by human impact.

Because hybrids are fertile and isolation mechanisms seem to be weak, these species are considered to integrate a syngameon (Grant, 1981) (**Fig. 1**). According to Grant (1981), a syngameon is "the most inclusive unit of interbreeding in a hybridizing species group". Even though *Prosopis* hybrids are widespread, fertile and ancient, *Prosopis* species have maintained and are maintaining phenotypic and ecological cohesion within and distinction between, and have possibly maintained themselves as distinct evolutionary lineages for million of years. However, the continuous production of hybrids might favor the occurrence of evolutionary experiments through the production of novel genetic combinations, and hybridization might have been important in the evolution of this group (Saidman et al., 1998).

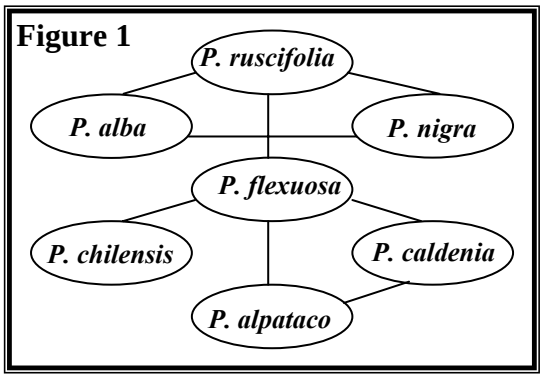

Figure 1: Representation of the *P. alba* - *P. alpataco* - *P. caldenia* - *P. chilensis* - *P. flexuosa* - *P. nigra* - *P. ruscifolia* syngameon in Argentina. Lines connecting the various species reflect the observed natural hybridization.

The occurrence of morphologically intermediate individuals between *Prosopis chilensis* and *P. flexuosa* in wide areas of sympatry indicates that these species cannot be considered as biological species. It is unknown for how long these two species have been in contact, and if a differentiation process between the groups of hybrids and the parental species had occurred. Hybrids may display characteristics that allow them to occupy niches different from the parental ones. Such specifically adapted populations could be used in forest genetic improvement programs to obtain basic material to restitute the plant covered in degraded areas.

In a companion paper (Mottura et al., in preparation), the morphological assignment of individuals trees in a hybrid swarm between *Prosopis chilensis* and *P. flexuosa* in the Argentinean Arid Chaco, was carried out, and the genetic structure and differentiation among the morphological groups obtained is described. The present study describes evidence for natural hybridization between *P. chilensis* and *P. flexuosa*, and aspects of the mating system in a hybrid swarm based on variation patterns in adults and their progenies at three SSR loci. Paternity analysis in combination with phenological data was used to estimate descriptive parameters on the mating system prevailing in the hybrid swarm. Selfing and outcrossing rates were estimated together with other mating parameters such us biparental inbreeding and correlated mating.

Materials and methods

Study site

The present study took place in the Provincial Park and Forest Natural Reserve Chancaní (*Parque Provincial y Reserva Forestal Natural Chancaní*). This natural reserve is located in the *Departamento Pocho*, on the western slope of the "*Sierras de Pocho-Guasapampa*", in the west of Córdoba province, Argentina. The Chancaní Reserve is placed in the *Chaco* phytogeographic region (Cabrera, 1976). The area is classified as Dry Subtropical Arid (Koppen, 1931; Trewartha, 1943). The annual average rainfall fluctuates between 400 and 600 millimeters, with a water deficit ranging from 400 to 500 millimeters per year (Cabido and Pacha, 2002).

The mating system study was carried out during the flowering season 2003-2004, in a 4700 m^2 plot included in a continuous forest located Southeast of the abovementioned reserve (Lat. 31º 23' S, Long. 65º 27' O, Alt. 360 m). The area belongs to an abandoned farm. Nowadays, it presents around 30 years of natural regeneration of the native forest (Cabido and Pacha, 2002). In this environment, which has been highly modified by man, both species of *Prosopis* (*P. chilensis* and *P. flexuosa*) occur together and many individuals with intermediate morphological characteristics between the two species can be observed.

Field measurements and observations

In the study plot, all *Prosopis* trees which flowered during 2003-2004 were mapped (**Fig. 2**). A total of 100 flowering *Prosopis* trees (*adult population*) were identified during that period. The position of each tree in the plot was measured in the field using polar coordinates (distances and angles). Subsequently, to build the map (**Fig. 2**) polar coordinates were converted to Cartesian coordinates.

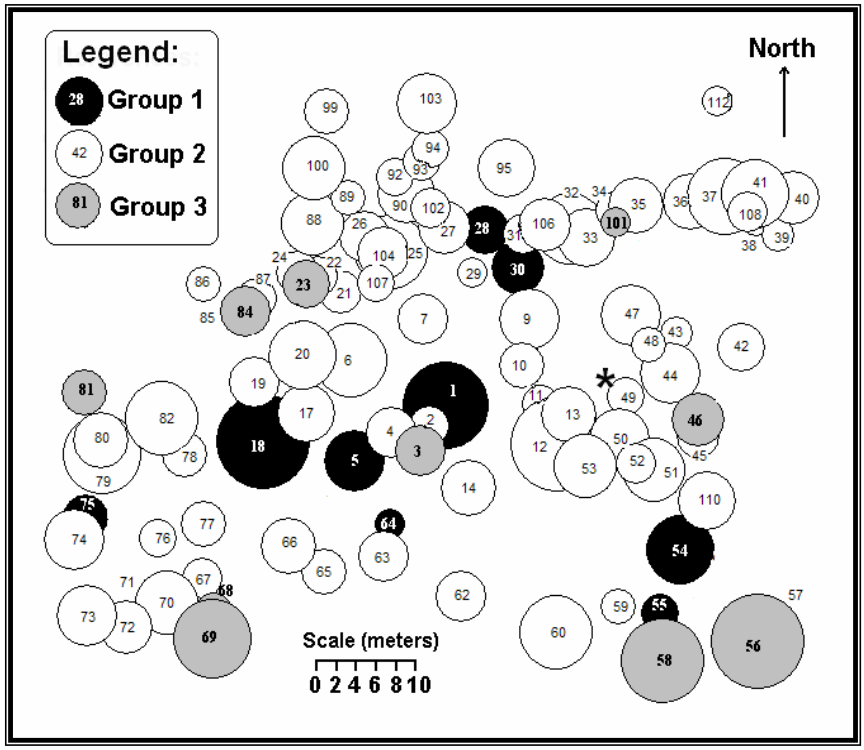

Figure 2: Map of the study plot with 100 mapped *Prosopis* trees that are represented by circles. The different colors refer to the morphological groups obtained after morphological analysis. The diameter of the circle is proportional to the tree crown size. Group 1: intermediate individuals; Group 2: *P. flexuosa*; Group 3: *P. chilensis*.

The morphological assignment of individuals trees at this site is described in detail in a companion paper (Mottura et al., in preparation), as well as the analysis of the genetic structure and differentiation at six SSR loci, among the morphological groups obtained. According to the morphological analysis performed, trees were classified in three morphological groups: Group 1, ten individuals with intermediate morphological characteristics between the species *P. chilensis* and *P. flexuosa*; Group 2, 80 individuals of *P. flexuosa*; and Group 3, ten individuals with characteristics of *P. chilensis*.

For the estimations of the crown size of each tree, the projection of the crown on the ground was measured in two directions, North-South and East-West. Then, the diameter of the crown was estimated as the arithmetic mean between the two values, and crown size was computed considering the crown as a perfect circumference.

Observations on flowering phenology were performed during the flowering period 2003-2004. Phenological observations were carried out for the phenological substage "*c*", or completely open and receptive flower (Genise et al., 1990). Phenological events were registered for each individual tree every ten days. Based on the phenological observations, overlapping of flowering periods between trees was estimated.

Sampled material

Fresh leaves from each individual of the adult population were collected for DNA isolation. In addition, twenty three open-pollinated mother trees were selected among the adult population. Seeds were harvested during the flowering season 2003-2004. The most representative individuals, from a morphological standpoint, of each morphological group were selected as mother trees. A total of 516 seeds (*seed population*) were collected. The seed population was composed of 23 single tree progenies with a variable size of the progeny arrays (between 8 and 48) (**Table 1**). The seed population was divided into groups according to the morphological classification of the mother trees.

Genetic analysis

Three microsatellite (SSR = Simple Sequence Repeat) loci originally developed for *Prosopis chilensis* (Mottura et al., 2005) were used in the genetic analysis (**Table 2**). Total DNA was extracted from leaves of each individual tree of the adult population using the DNeasy PlantMiniKit (Qiagen, Hilden, Germany). PCR amplification was performed as described in Mottura et al. (2005). SSR fragments were analyzed on an automatic sequencer (ABI 3100 Genetic Analyzer from Applied Biosystems). The length of the fragments was determined using the computer programs Gene scan 3.7 and Genotyper 3.7 (Applied Biosystems). On the seed population DNA was isolated from cotyledons after germination using the DNeasy 96 Plant Kit (Qiagen, Hilden, Germany).

Amplification and detection procedures for three SSR loci (**Table 2**) were performed as described for the adult population.

Table 1: Composition of the seed population (***Group***: morphological group assigned to the mother tree).

Mother tree	*Group*	N° of seeds harvested	Seed population
1	1	32	**Group 1**
18	1	9	
30	1	20	
54	1	20	
	Total	81	
4	2	32	**Group 2**
6	2	10	
10	2	10	
11	2	20	
13	2	24	
14	2	24	
31	2	19	
47	2	24	
49	2	20	
51	2	24	
59	2	32	
65	2	24	
	Total	263	
3	3	32	**Group 3**
46	3	20	
58	3	24	
69	3	48	
81	3	20	
84	3	18	
101	3	10	
	Total	172	

Table 2: SSR loci used in the genetic analysis of the hybrid swarm.

SSR Locus	Primer sequence (5´-3´)	T_a	SSR motif
Mo08	**F:** TATCCTAAACGCCGGGCTAC	59	$(AC)_9$
	R: TCCCATTCATGCATACTTAAACC		
Mo09	**F:** ATTCCTCCCTCACATTTTGC	59	$(TG)_{17}$
	R: CATTATGCCAGCCTTTGTTG		
Mo13	**F:** TTGATTAGAGTTGCATGTGGATG	58	$(GT)_{10}CT(GT)_2$
	R: TGCAGTCCCAAGTGTCAGAG		

F, forward primer; **R**, reverse primer; T_a, annealing temperature (in ° C).

The male gametic frequency of alleles in the effective pollen clouds ("pollen allele frequencies") of individual mother trees was estimated by the Maximum-Likelihood method of Gillet (1997). The effective pollen cloud for a particular morphological group (*pooled pollen cloud*) was computed pooling together individual pollen clouds using weights in proportion to the number of seeds investigated.

Population genetic parameters were calculated in the adult population and pooled pollen cloud using the software GSED Version 1.1j (Gillet, 2004), Arlequin ver. 3.01 (Excoffier et al., 2005), and GenAlEx6 (Peakall and Smouse, 2005). For each locus were calculated effective number of alleles (N_e) according to Crow and Kimura (1970); and expected (H_e) heterozygosity following Hartl and Clark (1997). Genetic differentiation among morphological groups and among populations was characterized using the following parameters: genetic distance D (Nei, 1972), G-statistics G_{ST} (Nei, 1973), genetic distance d_0 (Gregorius, 1974), and genetic differentiation D_j and average differentiation δ (Gregoruis and Roberts, 1986). In addition, pairwise genetic distance (D and d_0) were computed for individual pollen clouds.

Homogeneity test (G_h-*test* or log likelihood ratio test) was used to the heterogeneity among allele frequencies for statistical significance. Rare alleles (with expected frequency less than 5 %) were lumped together before the test was performed (Sokal and Rohlf, 1995).

Paternity analysis

Paternity analysis was used in order to assess more precisely the gene flow between morphological groups. Paternity assignment was conducted by categorical allocation based on multilocus genotypes using the software FAMOZ (Gerber et al., 2003). Genotypes of mother trees were known *a priori*. Selfing and outcrossing were assumed to have the same probability of occurrence. For a particular seed, the most likely father was assigned based on the Likelihood ratio (LOD score) computed. The potential father with the maximum LOD score was regarded as the most likely pollen donor for the respective seed. Potential fathers whose flowering period did not overlap with the flowering period of the mother tree were excluded. Seed to which no potential father could be assigned were discarded from the following analysis.

Considering the mating events detected by paternity assignment, mating preferences were calculated for each morphological group. First, mating frequencies within each group ($P_i(G_i x G_j)$) were calculated as:

$$P_i(G_i x G_j) = \frac{n_{ij}}{n_i}$$

where n_{ij} are the number of matings of type $Group_i$ x $Group_j$ within $Group_i$, and n_i are the number of within population mating with $Group_i$ seed trees. The first term (G_i) of the mating frequencies denotes the maternal contribution in the mating performed.

Then mating preferences for each group ($Z_{i(G_i x G_j)}$) were calculated as:

$$Z_{i(G_i x G_j)} = \frac{P_i(G_i x G_j)}{N_j / N}$$

where N_j is the number of adult trees in $Group_j$, and N is the total number of adults in the adult population.

In addition, a sensitivity analysis was conducted in order to test the confidence of the results obtained using the abovementioned method for paternity assignment. For those progenies with one or more potential father within the stand, and considering all potential fathers found by categorical allocation based on multilocus genotypes, paternity assignment was performed using six different methodologies:

I. *Fatherhood probability proportional to LOD score*: for each potential pollen parent a fatherhood probability was computed proportional to the LOD score found by categorical allocation.

II. *Fatherhood probability even*: fatherhood probabilities were computed considering all potential fathers with equal weight.

III. *Nearest tree*: the most likely father was assigned among all potential fathers considering that closer to the seed tree.

IV. *Fatherhood probability proportional to distance*: fatherhood probabilities were computed as the inverse of the distance between the seed tree and the potential pollen parent.

V. *Fatherhood probability proportional to crown size*: fatherhood probabilities were calculated proportional to the crown size of each potential pollen parent.

VI. *Fatherhood probability proportional to overlapping period*: fatherhood probabilities were computed proportional to the overlapping of flowering periods between the seed tree and the potential pollen donor.

Considering the mating events detected by the different methods, mating preferences were computed for each morphological group as it was mentioned above.

Mating system parameters

The estimation of the mating system parameters for the flowering period 2003-2004, was performed using three highly polymorphic SSR loci (**Table 2**). For each individual mother tree, the proportion of progeny with a "foreign" allele (i.e., an allele not shown by the mother tree) at any gene locus, was calculated as a *minimum estimate of the outcrossing rate*. Besides, outcrossing rates for individual trees were estimated based on the results of the paternity assignment. In addition, selfing rate (s) was estimated based on the presence of rare alleles in the adult population (Finkeldey and Hattemer, in press).

Furthermore, outcrossing rate (t) from single- and multilocus genotypes, correlated mating (r_p), and biparental inbreeding ($t_m - t_s$) were estimated assuming the mixed mating model (Ritland and Jain, 1981) using the computer program MLTR Version 3.0 (Ritland, 2004). Estimations were computed for each morphological group and for individual trees within each group. Computations were performed using the Newton-Raphson numerical method. Standard errors were estimated based on 100 bootstraps. Resampling for each bootstrapping was conducted within families.

Results

Genetic analysis

Parameters of genetic diversity and differentiation were computed for the adult population and for the pooled pollen cloud. The adult population was composed of 100 trees (**Fig. 2**) divided in three morphological groups: Group 1, ten individuals with intermediate characters between the species *P. chilensis* and *P. flexuosa*; Group 2, 80 individuals of *P. flexuosa*; and Group 3, ten individuals with characteristics of *P. chilensis*. The male gametic frequency of alleles in the effective pollen clouds of the individual mother trees were pooled together in proportion to the number of seeds investigated, and considering the morphological classification of the respective mother tree. As a result, three pooled pollen clouds were obtained, one from each morphological group.

Table 3 summarizes the genetic diversity within populations. The maximum number of effective alleles was found for both populations in Group 2 at locus Mo08, and the lowest, at locus Mo09 in Group 3 for both populations, too. Gene diversity (H_e) ranged from 0.34 to 0.86 in the adult population, and from 0.47 to 0.87 in the pooled pollen contribution. The highest diversity over all loci (*mean* H_e) was detected for both populations in Group 2. Within each group, the mean genetic diversity (N_e, H_e) of the pooled pollen cloud was higher in comparison to the respective groups in the adult population.

Table 3: Genetic diversity within morphological groups and for the pool in the adult population and in the estimated pooled pollen clouds (***Ne***, number of effective alleles; and $\boldsymbol{H_e}$, expected heterozygosity).

		Adult Population				**Pooled Pollen Cloud**			
Groups		**Mo08**	**Mo09**	**Mo13**	**Mean**	**Mo08**	**Mo09**	**Mo13**	**Mean**
Group 1	***Ne***	3.57	2.17	4.44	3.11	4.57	2.85	4.03	3.67
	$\boldsymbol{H_e}$	0.72	0.54	0.77	0.68	0.78	0.65	0.75	0.73
Group 2	***Ne***	7.27	5.69	3.09	4.71	7.64	5.11	3.39	4.82
	$\boldsymbol{H_e}$	0.86	0.82	0.68	0.79	0.87	0.80	0.70	0.79
Group 3	***Ne***	2.63	1.51	2.82	2.15	4.75	1.87	4.25	3.06
	$\boldsymbol{H_e}$	0.62	0.34	0.64	0.53	0.79	0.47	0.76	0.67
Pooled	***Ne***	7.50	4.79	4.15	5.15	6.42	3.42	4.37	4.43
	$\boldsymbol{H_e}$	0.87	0.79	0.76	0.81	0.84	0.71	0.77	0.77

Genetic distances, D (Nei, 1972) and d_0 (Gregorius, 1974), used to estimate genetic differentiation between morphological groups showed the highest differentiation among Group 3 and 2 in the adult population and in the pooled pollen cloud (**Table 4**). The lowest differentiation was observed between Group 3 and 1, whereas between Groups 1 and 2 the values were intermediate. Reductions in distance values were observed between all pairs of groups, except for D between Groups 1 and 3, when the adult population was compared to the pooled pollen contribution. Between Groups 1 and 3 D was similar when the adult population and the pooled pollen clouds were compared.

Table 4: Pairwise matrix of Nei's (1972) genetic distance ***D*** (below diagonal), and Gregorius' (1974) distance $\boldsymbol{d_0}$ (above diagonal) among groups in the adult population and in the pooled pollen clouds.

	Adult Population				**Pooled Pollen Cloud**		
Groups	**Group 1**	**Group 2**	**Group 3**	**Groups**	**Group 1**	**Group 2**	**Group 3**
Group 1	--	0.575	0.333	**Group 1**	--	0.291	0.247
Group 2	0.592	--	0.717	**Group 2**	0.121	--	0.418
Group 3	0.086	0.904	--	**Group 3**	0.087	0.288	--

Table 5 shows the total genetic differentiation among morphological groups computed by Nei's G_{ST} (Nei, 1973), and average differentiation δ (Gregoruis and Roberts, 1986). Locus Mo13 revealed the highest G_{ST} and δ values in the adult population. In the pooled pollen cloud, locus Mo09 showed the maximum G_{ST}, while the highest δ was found at locus Mo13. Mean values recorded for G_{ST} of were 0.110 and 0.037 in the adult population and in the pooled pollen contribution, respectively. Since mean δ was 0.468 in the seed population and 0.278 in the pooled pollen cloud. As it was observed in the genetic distance values, G_{ST} and δ decreased for all loci when comparing the adult population to the pooled pollen cloud. The genetic distance between a group and its complement ($\boldsymbol{D_j}$) was lower for the pooled pollen clouds in comparison to the adult populations for all groups and at all loci (**Table 6)**.

Table 5: By locus and average total genetic differentiation among groups of the adult population and pooled pollen clouds: G_{ST}, G-statistics (Nei, 1973); δ, average differentiation (Gregoruis, 1984).

	Adult Population				Pooled Pollen Cloud			
Locus	**Mo08**	**Mo09**	**Mo13**	**Mean**	**Mo08**	**Mo09**	**Mo13**	**Mean**
G_{st}	0.083	0.104	0.143	0.110	0.018	0.053	0.044	0.037
δ	0.459	0.394	0.552	0.468	0.218	0.283	0.333	0.278

Table 6: Amounts of genetic differentiation (D_j) among groups, by locus and for the gene pool, in the adult population and in the pooled pollen cloud.

	Adult Population			Pooled Pollen Cloud		
Locus	**Group 1**	**Group 2**	**Group 3**	**Group 1**	**Group 2**	**Group 3**
Mo08	0.353	0.613	0.413	0.215	0.290	0.149
Mo09	0.259	0.500	0.422	0.178	0.316	0.356
Mo13	0.297	0.756	0.603	0.171	0.404	0.424
Gene Pool	0.303	0.623	0.479	0.188	0.337	0.310

The homogeneity test (G_h-*test*) showed at all loci highly significant (p<0.001) deviation from the null hypothesis of homogeneous pollen allele frequencies across mother trees in the hybrid swarm.

Genetic distances among pollen clouds of individual mother trees are shown in **Table 7 (Annex)**. Within groups, the lowest differentiation was registered among pollen clouds of Group 1 (mean D = 0.35, mean d_0 = 0.45), followed by Group 3 (mean D = 0.45, mean d_0 = 0.49), and Group 2 (mean D = 0.52, mean d_0 = 0.52). Between morphological groups, the lowest values were recorded among Group 1 and 3 (mean D = 0.43, mean d_0 = 0.50). Differentiation among Groups 1 and 2 showed similar values (mean D = 0.52, mean d_0 = 0.50), while between Groups 2 and 3 differentiation was higher (mean D = 0.85, mean d_0 = 0.61).

Paternity analysis

The theoretical exclusion probability for paternity (Jamieson and Taylor, 1997) computed over all trees of three morphological groups and over three loci reached 95.77 %. The highest exclusion probability for single locus was observed at locus Mo08 (73.21 %), followed by locus Mo09 (61.47 %) and finally locus Mo13 (58.86 %). For

96 offspring (18.60 %) among the 516 analyzed, no compatible male parent was observed within the study stand. Their pollen parents were considered to be located outside the studied plot (**Table 8**). All but one male parent could be excluded for 78 seeds (15.12 %). The remaining offspring, 342 (66.28 %), showed more than one potential father. Out of these 342 seeds, 143 had all potential fathers from the same morphological group. The maximum number of potential fathers registered for a single offspring was 12. For the set of progenies that showed more than one potential father, in four cases (1.17%) the phenological analysis allowed the exclusion of one or more potential fathers because no overlapping in the flowering period with the seed tree was observed. Finally, for those offspring with more than one potential father, a single father was assigned considering the one with the highest LOD Score. **Table 8** shows the mating events detected by paternity assignment for each individual mother tree. Among the 420 progenies that had their male parent within the stand, 170 resulted from hybridization events between groups According to this result the degree of *inter-group* hybridization reaches 40.48%.

Figure 3 illustrates the mating preferences for each morphological group estimated based on the mating events detected by paternity analysis.

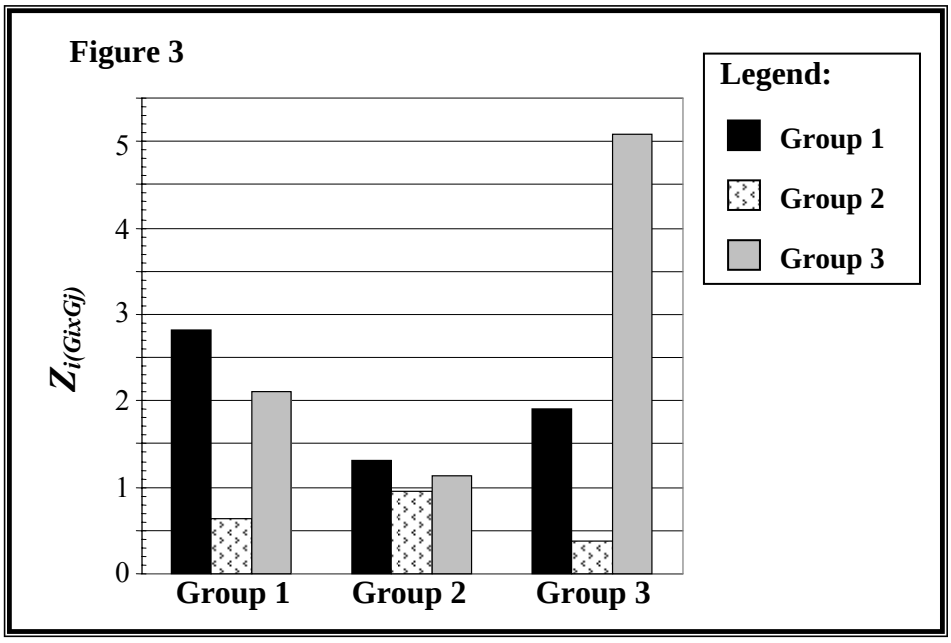

Figure 3: Mating preferences computed for each morphological group considering the mating events detected by paternity assignment conducted by categorical allocation, and consider the potential father with the maximum LOD score as the most likely pollen donor.

The results obtained showed that there is a strong preference of Group 3 to mate with individuals of the same group. In addition, this group shows preference for Group 1 but in a lower degree, and a strong repulsion towards Group 2. Group 1 shows preference to mate with individuals of the same group and also with Group 3, and repulsion towards Group 2. On the other hand, Group 2 shows preference values close to one for all groups, this indicates a trend towards random mating within this group.

Table 8: Number of offspring resulting from all possible mating events detected by paternity analysis for each individual mother tree analyzed.

Mother Tree		N° of offspring resulting from matings				
N°	**Morphological assignment**	**with Group 1**	**with Group 2**	**with Group 3**	**outside the plot**	**Total**
1	Group 1	8	14	7	3	32
18	Group 1	1	6	0	2	9
30	Group 1	8	7	2	3	20
54	Group 1	3	9	6	2	20
Total	Group 1	20	36	15	10	81
4	Group 2	10	3	10	9	32
6	Group 2	0	8	2		10
10	Group 2	1	8	0	1	10
11	Group 2	0	18	0	2	20
13	Group 2	4	15	1	4	24
14	Group 2	2	17	1	4	24
31	Group 2	0	15	0	4	19
47	Group 2	2	18	0	4	24
49	Group 2	0	17	1	2	20
51	Group 2	0	21	0	3	24
59	Group 2	1	15	5	11	32
65	Group 2	8	6	4	6	24
Total	Group 2	28	161	24	50	263
3	Group 3	6	5	7	14	32
46	Group 3	0	15	1	4	20
58	Group 3	4	2	17	1	24
69	Group 3	13	7	13	15	48
81	Group 3	1	5	14	0	20
84	Group 3	2	2	13	1	18
101	Group 3	0	5	4	1	10
Total	Group 3	26	41	69	36	172

Figure 4 (**I to VI**) shows mating preferences for each morphological group estimated by six different methodologies for paternity assignment. This sensitivity analysis revealed similar behaviors on mating preferences for all groups in all methods used. The results obtained in the sensitivity analysis support the mating preferences computed in

Figure 3, where the potential father with the maximum LOD score was assigned as the most likely pollen donor for a given progeny.

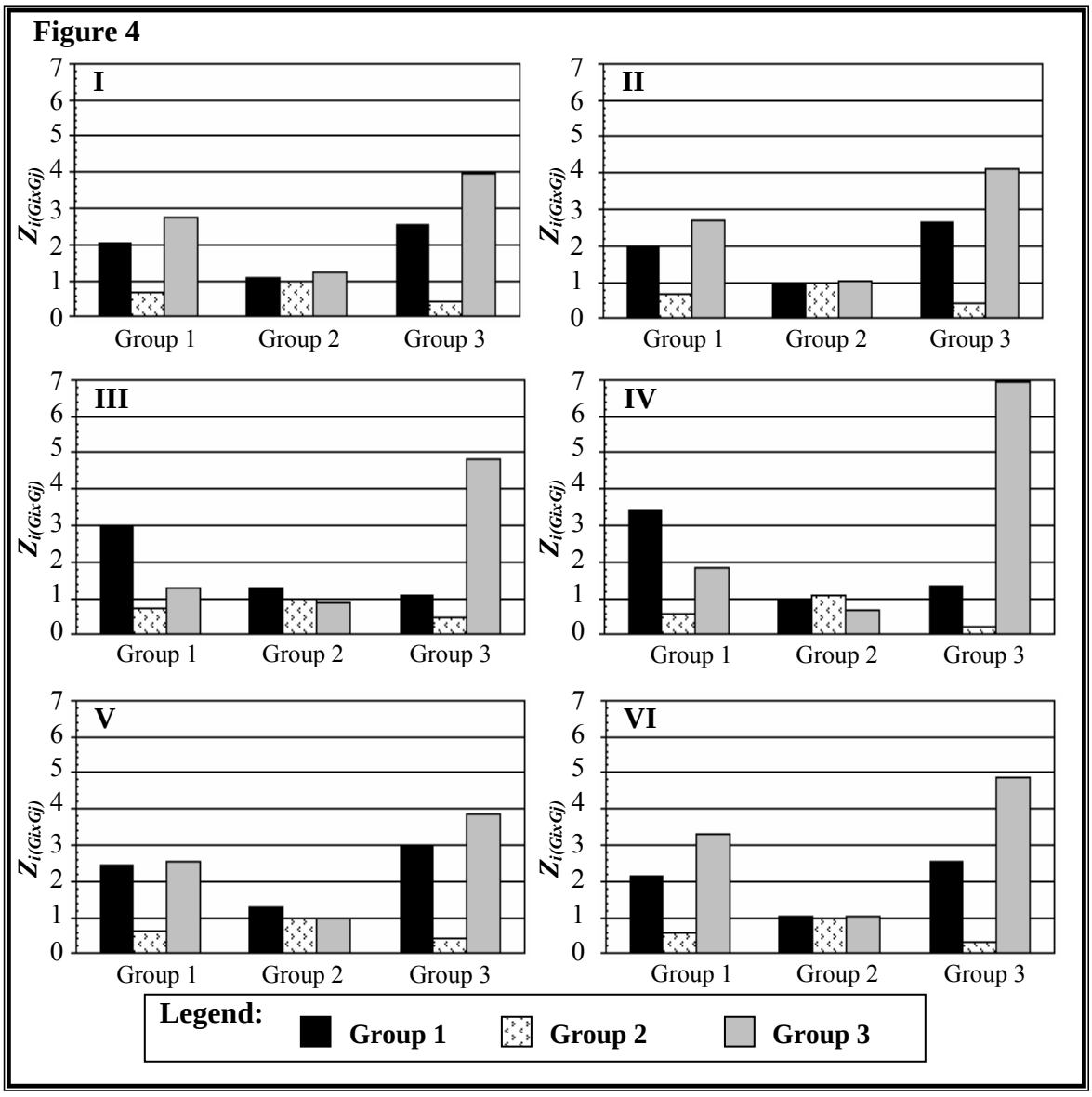

Figure 4: Mating preferences computed for each morphological group considering the mating events detected by paternity assignment using six different methodologies (**I** to **VI** refer to the methods used for paternity assignment listed in *Materials and methods – Paternity analysis*).

Mating system parameters

Mating system parameters estimated by different methods are shown for each single tree, for each group and for the pooled population in **Table 9**. The minimum estimate of the outcrossing rate (*t*), computed as the proportion of progeny with an allele not shown

by the mother tree, showed individual values between 0.292 and 1.000. At a group level, the minimum estimate of *t* for Group 1 was 0.938; for Group 2, 0.920; and for Group 3, 0.692. Individual values of outcrossing estimated by paternity analysis ranged from 0.389 and 1.000.

Table 9: Estimates of mating system parameters for each seed tree, for each group and for the pooled population by different methods (***t***: outcrossing rate; t_m: multilocus outcrossing rate; t_s: single-locus outcrossing rate; t_m-t_s: rate of biparental inbreeding; r_p: correlated mating). The number of seeds analyzed in each mother tree is specified in **Table 1**.

	Mother tree N°	Minimum estimate of *t*	Estimated *t* by paternity analysis	Estimates based on the mixed mating model			
				t_m	t_s	t_m-t_s	r_p
Group 1	1	0.906	0.969	1.032	1.019	0.013	0.032
	18	1.000	1.000	1.200	0.941	0.259	0.333
	30	0.950	1.000	1.011	1.051	-0.040	-0.046
	54	0.950	0.950	0.950	0.840	0.110	0.061
	Group	0.938	0.975	0.979	0.934	0.045	0.059
Group 2	4	1.000	1.000	1.200	0.871	0.329	0.263
	6	1.000	1.000	1.200	0.857	0.342	0.076
	10	0.900	1.000	0.941	0.740	0.201	0.306
	11	0.950	0.950	0.960	1.048	-0.088	0.261
	13	1.000	1.000	1.200	0.935	0.265	0.090
	14	0.708	0.875	0.765	0.589	0.176	0.154
	31	0.895	1.000	0.924	0.932	-0.008	0.015
	47	1.000	1.000	1.200	0.977	0.223	0.183
	49	0.850	0.850	0.851	0.665	0.186	0.418
	51	0.708	0.750	0.746	0.541	0.205	0.109
	59	1.000	1.000	1.200	0.941	0.259	-0.015
	65	1.000	1.000	1.200	1.200	0.000	0.207
	Group	0.920	0.951	0.954	0.879	0.075	0.154
Group 3	3	0.719	0.875	0.822	0.676	0.146	0.050
	46	0.950	0.950	0.976	0.713	0.263	0.466
	58	0.292	0.667	0.330	0.224	0.105	0.357
	69	0.958	0.958	1.006	0.790	0.216	0.134
	81	0.500	1.000	0.575	0.529	0.046	0.214
	84	0.389	0.389	0.404	0.300	0.103	0.322
	101	0.700	1.000	0.736	0.750	-0.015	0.288
	Group	0.692	0.849	0.777	0.635	0.143	0.136
Pooled		0.847	0.921	0.914	0.699	0.214	0.178

Among groups, t estimated by paternity assignment was 0.975, 0.951 and 0.849 for Groups 1, 2 and 3, respectively. Individual multilocus outcrossing rate (t_m), as estimated by the program MTLR (Ritland, 2004) ranged from 0.330 to 1.200. t_m for Group 1 was 0.979, for Group 2 0.954, and for Group 3 0.777. For all methodologies used in the estimation of the outcrossing rate, the lowest individual values of t were found among trees of Group 3. This group showed also the lowest value of outcrossing at group level in estimations.

Individual biparental inbreeding ($t_m - t_s$) estimated assuming the mixed mating model, ranged from -0.088 to 0.342; while the values found by group were 0.045 for Group 1, 0.075 for Group 2, and 0.143 for Group 3. The lowest value of correlated mating found at individual trees was -0.046 in tree number 30, and the highest was 0.466 in the seed tree 46.

A single mother tree (tree 49, *** Fig. 2**) was identified in the scored adult population carrying a unique allele at the locus Mo09. This tree was heterozygous at this locus. A total of 20 seeds harvested in the flowering period 2003-2004 were investigated for the estimation of selfing based on rare alleles. Selfing rate (s) estimated by this method was 20 %, for this tree.

Discussion

Genetic analysis

Measurements of differentiation at the seed stage reflect the effects of pollen and seed dispersal, i.e. the effects of the mating and migration system. By considering gene frequencies among the male gametic frequency of alleles in the effective pollen clouds only, the level of differentiation which is exclusively due to the mating system can be computed (Gregorius, 1984). In the present study, total genetic differentiation among groups (measure as G_{st} and δ) as well as the genetic distance (D and d_0) between them, decreased in the pooled pollen contribution in comparison to the adult population. In addition, the levels of genetic differentiation of individual groups, measured as $\boldsymbol{D_j}$, revealed a clear decrease when the pooled pollen contribution was compared to the

adult population. Strong pollen flow among groups is expected to homogenize the allelic frequencies of the respective pollen clouds and to reduce genetic differentiation among groups. The subsequent gene flow analysis based on paternity approach confirms this hypothesis, revealing 40.5% of *inter-group* hybridization among 420 mating events studied.

More evidence for hybridization arises from the interpretation of genetic differentiation between pollen clouds of individual mother trees within and among groups. Relatively high values of differentiation among individual pollen clouds were observed (average D = 0.61; average d_0 = 0.54). Distances of individual pollen clouds within groups were comparable to distances among groups except for the distances between Group 2 and Group 3. Stronger differentiation of individual pollen clouds among groups than within groups would be expected if reproductive isolation would occur between groups.

The allelic differentiation among individual effective pollen clouds showed relatively high distances for most of the pairs analyzed. The main cause of differentiation among pollen clouds is non-random mating, that can be due to limited pollen dispersal, incompatibilities or assortative mating among other factors. Besides, homogeneity test revealed significant deviation from the null hypothesis of the absence of mating preferences, i.e. homogeneous pollen allele frequencies across seed trees within the hybrid swarm. The effective pollen clouds of single seed trees are expected to be homogeneous if random mating is assumed. Thus, both analyses showed that even if there is a high level of hybridization within the hybrid swarm, mating does not occur at random among individuals. Heterogeneous pollen dispersal is expected in populations that critically depend on pollinating insects since their behavior is determined by various factors that are likely to vary on a microenvironmental scale within stands. In addition, positive assortative mating would be expected among individuals of the same morphological group.

Paternity analysis

The parental analysis approach provides a direct estimate of gene movement. Paternity analysis allows the detection of male parents and therefore the recognition of the mating event. Besides, gene movement estimations based on parental analysis permit the estimation of effective gene flow because the method accounts for all biological factors involved in the

mating success: viability of the pollen; compatibility between donor and receiving plants; competition among different pollen tubes. We estimated 40.5% of *inter-group* hybridization occurring among individuals in the studied site.

Considering the pure species *P. flexuosa* and *P. chilensis* represented, respectively, by the morphological groups 2 and 3, the results of the paternity analysis show that interspecific hybridization occurs. Analyzing within the aforesaid groups the mating events detected (see **Table 8**), discarding those in which no potential father could be assigned, the results show that 24 out of 213 (11.27%) of the seeds from mother trees of Group 2 were pollinated by trees of Group 3. On the other hand, 41 of 136 (30.15%) seeds belonging to mother trees from Group 3 were pollinated by trees from Group 2. The mating preferences obtained from the paternity analysis indicate that the morphological groups differ from each other markedly in their preference patterns. According to our results, trees of Group 2 show a trend towards random mating with trees of their own type and with individuals Groups 3. Trees of Group 3 show strong preference to mate with trees of their own type, and a significant repulsion towards Group 2. These results would suggest that Group 3 (*P. chilensis*) has a tendency of reproductive isolation from Group 2 (*P. flexuosa*). Therefore, even when there is a lack of clear reproductive barriers between the species *P. chilensis* and *P. flexuosa*, some isolation mechanisms remain active. In this sense, in the present work in only 1.2% of the cases observed, the phenological study allowed the exclusion of potential fathers due to non-overlapping in the flowering times with the seed tree. This would point to the existence of similar phenological patterns between individuals of different groups. However, prezygotic isolation mechanisms regarding flower phenology can not be fully excluded due to the fact that in this study, solely the occurrence or absence of overlapping was considered. Other factors concerning the dynamics of flower phenology, such us the length of the overlapping period or the flowering intensity, can play very important roles in prezygotic isolation.

Group 1 was morphologically characterized as a group of individuals with intermediate phenotypes between the pure taxonomic species *P. chilensis* and *P. flexuosa*. The study of the mating system showed that trees in this group had preference to mate with individuals of their own type, and with individuals of Group 3. However, the self-preference shown by Group 1 is only half of the self-preference shown by Group 3. In addition, the values of preference of Group 1 for Group 3 are similar to the value that this last group present for Group 1. On the other hand, Group 1 shows repulsion towards Group 2, although Group 2 shows a trend towards random mating with Group 1. This

suggests that backcrossing of intermediate individuals are more frequent with Group 3 (*P. chilensis*) than with Group 2(*P. flexuosa*).

A wide variation in the levels of interspecific crossing can be observed among some individual seed trees of Groups 2 and 3. Trees isolated between individuals of a different group (e.g., trees 46 or 101 from Group 3; see **Fig. 2**) or surrounded by a high number of trees of a different group (e.g., mother trees 4 or 59 from Group 2; see **Fig. 2**) show higher levels of interspecific crossing. This would indicate that the levels of hybridization are conditioned by the chances to mate with individuals of the same type (group). Assuming that in these species pollen dispersal is limited (Genisse et al., 1990; Bessega et al., 2000a), those individuals more or less isolated from trees of their own type, seemed to have reduced their chances to mate within their group, and the levels of hybridization raise. However, when there are neighbor trees belonging to the same group, matings are more frequent within the group. Thus, even when hybridization occurs, there is a preference to mate with individuals of the same type; this preference is particularly strong in Group 3.

Mating system parameters

Early studies on species of *Prosopis* belonging to Section Algarobia, pointed that these species are protogynous and they were largely considered obligate outcrosser (Burkart, 1937, 1952 and 1976; Simpson, 1977; Balboa and Parraguez, 1986). This view contrasts with more recent research on floral biology carried out by Geniesse et al. (1990) in three South American species of this section (*P. chilensis*, *P. flexuosa* and *P. pugionata*). These authors concluded that flowers are not protogynous even though they accept that the species are obligate outcrosser and that this condition is maintained by an autoincompatibility system.

Results of isozyme analysis showed significant excess of homozygotes in all populations of species of Section Algarobia so far studied (Bessega et al., 2000b). Such excess of homozygotes might be caused by a certain degree selfing. Different pollination studies on *P. glandulosa*, *P. laevigata* (Galindo-Almanza et al., 1992) and *P. velutina* (Keys, 1993; Keys and Smith, 1994) supported this possibility. Latest studies using isoenzymatic data, and based on the mixed mating model (Ritland and Jain 1981), estimate values of single and multilocus outcrossing rate from about 0.72 to 1.00

(Bessega et al., 2000a), indicating that the species of the Section Algarobia are mostly outcrossers but up to 28% of selfing can occur, with an average of 15%.

In the current study, outcrossing rate was estimated for one flowering period (2003-2004) using three highly polymorphic SSR loci and applying different methodologies for the estimation (minimum estimate of *t*, paternity analysis, rare alleles, and based on the mixed mating model). All estimators showed similar results. The multilocus outcrossing rate (t_m) estimated based on the mixed mating model, showed a mean value for the pool of 0.914, with values of 0.979, 0.954 and 0.777 for Groups 1, 2 and 3, respectively. This means that partial selfing (between 2.1 % and 22.3 %) can occur. These values are similar to those proposed by Bessega et al. (2000a) based on isozymes markers (t_m values of 0.809 and 0.882 for *Prosopis chilensis* and *P. flexuosa* respectively). The minimum estimate of outcrossing for an individual tree was 0.292, and was registered in Group 3. This group showed also the lowest outcrossing rate at all estimators. At the individual level, t_m estimates showed a wide variation, especially in Group 3, where the value ranged from 0.330 to 1.006. An assessment of the overall outcrossing rate of the studied species is preferably based on population estimates of t_m rather than on single tree estimates, which vary considerably and are known to be rather unreliable if based on mixed mating model (Luu, 2005). However, the wide variation of outcrossing values at individual level was consistence across the different methods, suggesting that the estimations based on the mixed mating model are reliable in this case. Variation in outcrossing rate among mother plants can be cause by micro-spatial variation in population density and/or ecological factors. More isolated plants might have higher selfing rates than those situated in more dense patches. Hence, considering Group 3 with a relatively high reproductive isolation from the other groups, the low density of trees of this group within the stand (only 10 individuals in 4700 m^2) could be a cause of higher selfing rates registered within the group.

Differences between t_m and t_s in Groups 2 and 3 suggest a possible biparental inbreeding due to either mating among relatives or positive assortative matings as the result of genetic structuring of the population. The levels of correlated mating (r_p) found in all groups (between 0.059 and 0.154) were significantly lower than those reported by Bessega et al. (2000a) for *Prosopis chilensis* (0.964) and *P. flexuosa* (0.990). This suggests that individuals within progeny sets are half rather than full sibs.

The outcrossing estimate based on the presence of a rare alleles showed similar results to those found based on alternative methods for tree 49.

Conclusions

Considering a hybrid swarm where three groups can be distinguish based on morphological characteristics, i.e. (*i*) Group 1, individuals with intermediate phenotypes between the pure taxonomic species *P. chilensis* and *P. flexuosa*; (*ii*) Group 2, individuals corresponding to the taxonomic species *P. flexuosa*; and (*iii*) Group 3, individuals with characteristics of *P. chilensis*, we can conclude that the species *P. chilensis* and *P. flexuosa* are highly related. The lack of clear reproductive barriers between them makes possible a strong gene flow among species. Gene flow is asymmetric. In the formation of hybrids (intermediate individuals), *P. flexuosa* receive more pollen from *P. chilensis* than the opposite. Hybridization is accompanied by introgression. In the process, hybrids, by repeated backcrossing with the parental species incorporate genes of one species into the gene pool of the other species. A higher mating rate of intermediate individuals with *P. chilensis* generates greater levels of introgression towards this species. However, backcrossing occurs also between hybrids and *P. flexuosa*. While *P. chilensis* shows a more defined isolation with regard to *P. flexuosa*, and clear assortative mating, *P. flexuosa* shows a trend toward random mating within the stand. In addition, hybrids show also some degree of isolation, especially with *P. flexuosa*. In this sense, a group of hybrids with genetic isolation mechanisms that allow the preservation of specific adaptational abilities to niches different from those where are their parental species, would represent very interesting material for breeding programs. Individuals included in such groups, could be use to restitute plant covered in degraded areas.

References

Balboa O, Parraguez J M. 1986. Phenology studies of *Prosopis* species growing in Chile. The current state of knowledge on *Prosopis juliflora*. II International Conference on *Prosopis*, Recife, Brazil. Pp. 259-267.

Bessega C F, Ferreira L I, Julio N, Montoya S, Saidman B O, Vilardi J C. 2000a. Mating system parameters in species of genus *Prosopis* (Leguminosae). Hereditas 132: 19-28.

Bessega C F, Ferreira L I, Vilardi J C, Saidman B O. 2000b. Unexpected low genetic differentiation among allopatric species of section Algarobia of *Prosopis* (Leguminosae). Genetica 109: 255-266.

Burkart A. 1937. Estudios morfológicos y etológicos en el Género *Prosopis*. Darwiniana 3: 27-59.

Burkart A. 1952. Las leguminosas argentines. Editorial Acme, Buenos Aires.

Burkart A. 1976. A monograph of the genus Prosopis (Leguminosae subf. Mimosoidae). Journal of Arnold Arboretum. 219-249, 450-525.

Cabido M R, Pacha M J. 2002. Vegetación y flora de la Reserva Natural Chancaní. Agencia Córdoba Ambiente S.E., Córdoba, Argentina.

Cabrera A. 1976. Regiones Fitogeográficas Argentinas. In: Enciclopedia Argentina de Agricultura y Jardinería. Tomo II. ACME (ED.). Fascículo 1. Pp: 85.

Cony MA. 2000. Importancia de los estudios fisiológicos, ecofisiológicos y genéticos sobre especies del género *Prosopis* para la recuperación de ecosistemas áridos degradados de América Latina. Memorias del Seminario Internacional Bosques secos y desertificación. Pg: 225-276.

Crow J F, Kimura M. 1970. An introduction to population genetics theory. Harper and Row, New York, Evanston, London.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.01: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47-50.

Finkeldey R, Hattermer H H. in press. Tropical forest genetics. Springer-Verlag Berlin, Heidelberg, Germany.

Galindo-Almanza S, García Moya E, Wendd T L, González Cossio F V. 1992. Potencial de hibridación natural en el mesquite (*Prosopis laevigata* y *P. glandulosa* var. Torreyana, Leguminosae) de la altiplanicie de San Luis Potosí. Acta Botánica mexicana 20: 101-117.

Genise, J, Palacios R A, Hoc P S, Carrizo R, Moffat L, Mom M P, Agullo M A, Picca P, Torregrosa S. 1990. Observaciones sobre la biología floral de *Prosopis* (*Leguminosae*, *Mimosoideae*) II. Fases florales y visitantes en el distrito Chaqueño Serrano. Darwiniana 30: 71-85.

Gerber S, Chabrier P, Kremer A. 2003. FAMOZ: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. Molecular Ecology Notes 3, 479-481.

Gillet E. 1997. Maximum likelihood estimators of the gametic contributions to single plant progenies. Biometrics 53: 504-523.

Gillet E M. 2004. GSED Version 1.1j: Genetic structure from electrophoresis data. Institüt für Forstgenetik und Forstpflanzenzüchtung, Universität Göttingen, Germany.

Grant V. 1981. Plant speciation. Columbia University Press. New York, USA.

Gregorius H-R, Roberds J H. 1986. Measurement of genetical differentiation among subpopulations. Theoretical and Applied Genetics 71: 826-834.

Gregorius H-R. 1974. Genetischer Abstand zwischen Populationen. I. Zur Konzeption der genetischen Abstandsmessung. Silvae Genetica 23: 22-27.

Gregorius H-R. 1984. Measurement of genetic differentiation in plant populations. Pp. 276-285 in: Gregorius H –R (ed.) Population genetics in forestry. Springer-Verlag, Berlin, Heidelberg, New York, Tokyo.

Gregorius H-R. 1989. Characterization and analysis of mating systems. Abteilung für Forstgenetik und Forstpflanzenzühtuhg. GeorgAugust-Universität Göttingen. Ekopan Verlag, Witzenhausen. Pp: 158.

Hartl D L, Clark A G. 1997. Principles of Population Genetics. 3rd Edition. Sinauer Associate Inc. Sunderland, MS.

Jamieson A, Taylor S S. 1997. Comparison of three probability formulae for parentage exclusion. Animal Genetics 28, 397-400.

Karlin U O, Díaz R. 1984. Potencialidad y manejo de algarrobos en el árido subtropical. SECyT. Ministerio de Educación y Justicia. Programa nacional de recursos naturales renovables.

Keys R N. 1993. Mating system and pollination biology of velvet mesquite (*Prosopis velutina* Wooton). PhD dissertation, University of Arizona, Tucson, AZ.

Keys R N, Smith S. 1994. Mating system parameters and population genetic structure in pioneer populations of *Prosopis vetulina* (Leguminosae). American Journal of Botany 81: 1013-1020.

Koppen W. 1931. Grundriss der Klimakunde. Walter de Gruyter Co., Berlin. 388 pp.

Luu H T. 2005. Genetic variation and the reproductive system of *Dipterocarpus* cf. *condorensis* Pierre in Vietnam. Cuviller Verlag, Göttingen, Germany. 155 pp.

Mottura M C, Finkeldey R, Verga A R, Gailing O. 2005. Development and characterization of microsatellite markers for *Prosopis chilensis* and *Prosopis flexuosa* and cross-species amplification. Molecular Ecology Notes 5, 487-489.

Nei M. 1972. Genetic distance between populations. American Naturalist 106 (949): 283-292.

Nei M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the Natural Academy of Science, USA. Pp. 3321-3323.

Peakall R, Smouse P E. 2005. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Australian National University, Canberra, Australia.

Ritland K. 2004. Multilocus Mating System Program (MLTR) Version 3.0.

Ritland K, Jain S. 1981. A model for the estimation of outcrossing rate and gene frequencies using *n* independent gene loci. Heredity 47: 35-52.

Saidman B O. 1990. Isoenzyme studies on hybrid swarms of *Prosopis caldenia* and sympatric species. Genetics, 39: 5-8.

Saidman B O, Bessega C F, Ferreira L I, Vilardi J C. 1998. Random amplified polymorphic DNA (RAPDs) variation in hybrid swarms and pure population of the genus *Prosopis* (Leguminosae). International Foundation for Science (IFS). Proceedings of the Workshop "Recent Advances in Biotechnology for Tree Conservation and Management", Florianópolis, 122-134.

Saidman B O, Bessega C F, Ferreira L I, Julio N, Vilardi J C. 2000. The use of genetic markers to assess population structure and relationships among species of the genus *Prosopis* (Leguminosae). Boletín de la Sociedad Argentina de Botánica 35 (3-4): 315-324.

Saidman B O, Vilardi J C. 1987. Analysis of the genetic similarities among seven species of *Prosopis*. Theoretic and Applied Genetics, 75: 109-1116.

Simpson B B. 1977. Mesquite, its biology in two desert shrub ecosystems. Dowden, Hutchinson and Ross, Stroudsberg, Pennsylvania, U.S.A. 250 pages.

Sokal R R, Rohlf F J. 1995. Biometry. Third Edition. W.H. Freeman and Company. New York.

Solbrig O T, Bawa K S. 1975. Isozyme variation in species of *Prosopis* (Leguminosae). Journal of the Arnold Arboretum, 56: 398-412.

Trewartha G T. 1943. An introduction to weather and climate, 2nd edition. MacGraw-Hill, New York, NY. 545 pp.

Verga A R. 1995. Genetische Untersuchungen an *Prosopis chilensis* und *P. flexuosa* (Mimosaceae) im trockenen Chaco Argentiniens. Göttingen Research Notes in Forest Genetics 19. German w. sum., 96+VIII.

Verga A R, Carranza C, Ledesma M, Joseau M J, Díaz P. 2000. Identificación de fuentes de semilla superior de *Prosopis sp* para enriquecimiento del bosque nativo en el Chaco árido. Reunión Nacional del Algarrobo. Mendoza, 14 al 17 de noviembre.

Annex

Tabe 7: Pairwise matrix of Nei's genetic distance ***D*** (below diagonal) and Gregorius's genetic distance d_0 (above diagonal) for each individual pollen cloud.

		Group 1				Group 2												Group 3						
		1	**18**	**30**	**54**	**4**	**6**	**10**	**11**	**13**	**14**	**31**	**47**	**49**	**51**	**59**	**65**	**3**	**46**	**58**	**69**	**81**	**84**	**101**
Group 1	**1**		0.51	0.21	0.42	0.35	0.42	0.63	0.74	0.40	0.43	0.44	0.45	0.60	0.63	0.40	0.29	0.41	0.54	0.45	0.34	0.32	0.52	0.44
	18	0.51		0.52	0.56	0.58	0.53	0.38	0.79	0.57	0.47	0.57	0.43	0.58	0.59	0.48	0.68	0.69	0.56	0.69	0.67	0.66	0.63	0.66
	30	0.08	0.43		0.45	0.38	0.46	0.55	0.71	0.42	0.41	0.40	0.45	0.67	0.59	0.35	0.26	0.35	0.57	0.51	0.33	0.38	0.54	0.50
	54	0.34	0.42	0.35		0.45	0.47	0.52	0.63	0.44	0.44	0.36	0.50	0.63	0.65	0.34	0.49	0.43	0.55	0.59	0.52	0.45	0.44	0.37
Group 2	**4**	0.15	0.51	0.23	0.28		0.56	0.62	0.74	0.61	0.48	0.59	0.64	0.75	0.64	0.48	0.40	0.43	0.66	0.48	0.28	0.28	0.44	0.40
	6	0.42	0.48	0.50	0.42	0.60		0.60	0.60	0.40	0.51	0.34	0.42	0.53	0.46	0.40	0.49	0.58	0.34	0.73	0.56	0.57	0.67	0.60
	10	0.77	0.19	0.54	0.55	0.85	0.56		0.60	0.61	0.37	0.51	0.43	0.47	0.53	0.42	0.73	0.62	0.60	0.82	0.66	0.68	0.72	0.65
	11	1.49	1.40	1.53	1.08	1.68	0.98	0.80		0.64	0.65	0.55	0.61	0.67	0.58	0.54	0.71	0.73	0.77	0.97	0.76	0.82	0.88	0.68
	13	0.30	0.54	0.29	0.49	0.61	0.28	0.51	0.78		0.45	0.38	0.32	0.47	0.54	0.39	0.50	0.60	0.42	0.75	0.60	0.65	0.67	0.53
	14	0.35	0.29	0.26	0.40	0.30	0.36	0.19	0.88	0.35		0.47	0.41	0.50	0.50	0.39	0.58	0.56	0.55	0.71	0.43	0.58	0.64	0.56
	31	0.40	0.45	0.29	0.28	0.50	0.35	0.36	0.79	0.24	0.28		0.44	0.58	0.60	0.37	0.50	0.50	0.55	0.71	0.54	0.61	0.63	0.59
	47	0.41	0.31	0.38	0.52	0.72	0.26	0.21	0.72	0.15	0.26	0.32		0.35	0.49	0.41	0.53	0.61	0.34	0.77	0.62	0.70	0.71	0.59
	49	0.68	0.54	0.69	0.85	1.18	0.37	0.33	0.92	0.24	0.37	0.49	0.08		0.53	0.51	0.78	0.76	0.38	0.94	0.79	0.82	0.84	0.70
	51	0.90	0.62	0.82	0.98	1.03	0.55	0.37	0.78	0.50	0.39	0.64	0.36	0.33		0.52	0.69	0.74	0.45	0.92	0.70	0.73	0.82	0.78
	59	0.26	0.24	0.24	0.23	0.35	0.34	0.22	0.67	0.25	0.18	0.18	0.24	0.41	0.39		0.50	0.46	0.55	0.69	0.57	0.53	0.59	0.47
	65	0.12	0.88	0.14	0.42	0.26	0.64	0.12	1.54	0.43	0.66	0.48	0.66	1.17	1.20	0.43		0.34	0.61	0.47	0.36	0.39	0.53	0.48
Gruop 3	**3**	0.28	0.77	0.20	0.31	0.28	0.79	0.92	1.80	0.56	0.50	0.38	0.77	1.29	1.20	0.42	0.18		0.65	0.50	0.36	0.39	0.46	0.46
	46	0.55	0.63	0.57	0.62	0.90	0.34	0.48	1.12	0.22	0.46	0.48	0.18	0.17	0.28	0.36	0.72	0.80		0.82	0.64	0.62	0.69	0.60
	58	0.22	0.84	0.28	0.42	0.18	1.28	1.48	4.81	1.00	0.85	0.81	1.43	2.63	2.51	0.56	0.22	0.30	1.70		0.54	0.30	0.31	0.53
	69	0.20	0.80	0.18	0.43	0.14	0.69	0.93	1.70	0.53	0.36	0.46	0.76	1.20	1.07	0.47	0.24	0.22	0.82	0.32		0.37	0.52	0.50
	81	0.16	0.80	0.20	0.33	0.11	0.85	1.16	2.45	0.65	0.59	0.54	0.98	1.55	1.25	0.39	0.15	0.19	0.83	0.09	0.16		0.34	0.39
	84	0.37	0.63	0.37	0.16	0.19	0.80	1.00	2.62	0.77	0.63	0.54	1.00	1.71	1.67	0.43	0.34	0.25	0.98	0.20	0.35	0.15		0.28
	101	0.33	0.78	0.42	0.17	0.27	0.64	0.97	1.09	0.49	0.63	0.51	0.64	1.03	1.37	0.38	0.35	0.32	0.65	0.38	0.39	0.23	0.11	

Appendix 1: The systematics of the genus *Prosopis* Linnaeus emend. Burkart according to Burkart (1976).

I. Section PROSOPIS

1. *P. cineraria* (L.) Druce
2. *P. farcta* (Solander ex Russell) MacBride
   var. *farcta*
   var. *glabra* Burkart
3. *P. koelziana* Burkart

II. Section ANONYCHIUM

4. *P. africana* (Guill., Perr. & Rich.) Taubert

III. Section STROMBOCARPA

Series Strombocarpae

5. *P. strombulifera* (Lam.) Bentham
   var. *strombulifera*
   var. *ruiziana* Burkart
6. *P. reptans* Bentham
   var. *reptans*
   var. *cinerascens* (A. Gray) Bentham
7. *P. abbreviata* Bentham
8. *P. torquata* (Cavanilles ex Lagasca) DC.
9. *P. pubescens* Bentham
10. P. palmeri Watson
11. *P. burkartii* Muñoz

Series Cavenicarpae

12. *P. ferox* Grisebach
13. *P. tamarugo* F. Philippi

IV. Section MONILICARPA

14. *P. argentina* Burkart

V. Section ALGAROBIA

Series Sericanthae

15. *P. sericantha* Gillies ex Hooker & Arnott
16. *P. kuntzei* Harms

Series Ruscifoliae

17. *P. ruscifolia* Grisebach
18. *P. fiebrigii* Harms
19. *P. vinalillo* Stuckert
20. *P. hassleri* Harms
    var. *hassleri*
    var. *nigroides* Burkart

Series Denudantes

21. *P. denudans* Bentham
    var. *denudans*
    var. *patagonica* (Spegazzini) Burkart
    var. *stenocarpa* Burkart
22. *P. ruizleali* Burkart
23. *P. castellanosii* Burkart
24. *P. calingastana* Burkart

Series Humiles

25. *P. humilis* Gillies ex Hooker & Arnott
26. *P. rojasiana* Burkart

Series Pallidae

27. *P. rubriflora* E. Hassler
28. *P. campestris* Grisebach
29. *P. pallida* (Humboldt & Bonpland ex Willd.) H.B.K.
30. *P. affinis* Sprengel
31. *P. articulata* S. Watson
32. *P. elata* (Burkart) Burkart
33. *P. tamaulipana* Burkart

Series Chilenses

34. *P. chilensis* (Molina) Stuntz emend. Burkart
    var. *chilensis*
    var. *riojana* Burkart
    var. *catamarcana* Burkart

35. *P. juliflora* Swartz DC.
    var. *juliflora*
    var. *inermis* (H.B.K.) Burkart
    var. *horrida* (Kunth) Burkart
36. *P. nigra* (Grisebach) Hieronymus
    var. *nigra*
    var. r*agonesei* Burkart
    var. *longispina* Burkart
37. *P. caldenia* Burkart
38. *P. laevigata* (Humboldt & Bonpland ex Willd.) M.C. Johnston
    var. *laevigata*
    var. *andicola* Burkart
39. *P. flexuosa* DC.
    forma *subinermi*s Burkart
40. *P. glandulosa* Torrey
    var. *glandulosa*
    var. *torreyana* (Benson) Johnston
    var. *prostrata* Burkart
41. *P. alpataco* R.A. Philippi
42. *P. alba* Grisebach
    var. *alba*
    var. *panta* Grisebach
43. *P. velutina* Wooton
44. *P. pugionata* Burkart

Appendix 2: List of 100 *Prosopis* trees identified in the studied plot, their morphological assignment and taxonomic classification. Morphological assignment was performed following the "Morphological Distance" (d_m) method. Taxonomic classification was carried out following the key for identification of interspecific hybrids between *Prosopis chilensis* and *P. flexuosa*, based on quantitative characters (Verga, 2000).

Individual	Morphological group assigned	Taxonomic classification
1	1	*Hybrid*
2	2	*Prosopis flexuosa*
3	3	*Hybrid*
4	2	*Prosopis flexuosa*
5	1	*Prosopis flexuosa*
6	2	*Prosopis flexuosa*
7	2	*Prosopis flexuosa*
9	2	*Prosopis flexuosa*
10	2	*Prosopis flexuosa*
11	2	*Prosopis flexuosa*
12	2	*Prosopis flexuosa*
13	2	*Prosopis flexuosa*
14	2	*Prosopis flexuosa*
17	2	*Prosopis flexuosa*
18	1	*Hybrid*
19	2	*Prosopis flexuosa*
20	2	*Prosopis flexuosa*
21	2	*Prosopis flexuosa*
22	2	*Prosopis flexuosa*
23	3	*Hybrid*
24	2	*Prosopis flexuosa*
25	2	*Prosopis flexuosa*
26	2	*Prosopis flexuosa*
27	2	*Prosopis flexuosa*
28	1	*Prosopis flexuosa*
29	2	*Prosopis flexuosa*
30	1	*Prosopis flexuosa*
31	2	*Prosopis flexuosa*
32	2	*Prosopis flexuosa*
33	2	*Prosopis flexuosa*
34	2	*Prosopis flexuosa*
35	2	*Prosopis flexuosa*

36	2	*Prosopis flexuosa*
37	2	*Prosopis flexuosa*
38	2	*Prosopis flexuosa*
39	2	*Prosopis flexuosa*
40	2	*Prosopis flexuosa*
41	2	*Prosopis flexuosa*
42	2	*Prosopis flexuosa*
43	2	*Prosopis flexuosa*
44	2	*Prosopis flexuosa*
45	2	*Prosopis flexuosa*
46	3	*Hybrid*
47	2	*Prosopis flexuosa*
48	2	*Prosopis flexuosa*
49	2	*Prosopis flexuosa*
50	2	*Prosopis flexuosa*
51	2	*Prosopis flexuosa*
52	2	*Prosopis flexuosa*
53	2	*Prosopis flexuosa*
54	1	*Hybrid*
55	1	*Hybrid*
56	3	*Hybrid*
57	2	*Prosopis flexuosa*
58	3	*Hybrid*
59	2	*Prosopis flexuosa*
60	2	*Hybrid*
62	2	*Prosopis flexuosa*
63	2	*Prosopis flexuosa*
64	1	*Hybrid*
65	2	*Prosopis flexuosa*
66	2	*Prosopis flexuosa*
67	2	*Prosopis flexuosa*
68	3	*Hybrid*
69	3	*Prosopis chilensis*
70	2	*Prosopis flexuosa*
71	1	*Hybrid*
72	2	*Prosopis flexuosa*
73	2	*Prosopis flexuosa*
74	2	*Prosopis flexuosa*
75	1	*Prosopis flexuosa*
76	2	*Prosopis flexuosa*
77	2	*Prosopis flexuosa*

78	2	*Prosopis flexuosa*
79	2	*Prosopis flexuosa*
80	2	*Prosopis flexuosa*
81	3	*Hybrid*
82	2	*Prosopis flexuosa*
84	3	*Prosopis chilensis*
85	2	*Prosopis flexuosa*
86	2	*Prosopis flexuosa*
87	2	*Prosopis flexuosa*
88	2	*Prosopis flexuosa*
89	2	*Prosopis flexuosa*
90	2	*Prosopis flexuosa*
92	2	*Prosopis flexuosa*
93	2	*Prosopis flexuosa*
94	2	*Prosopis flexuosa*
95	2	*Prosopis flexuosa*
99	2	*Prosopis flexuosa*
100	2	*Prosopis flexuosa*
101	3	*Hybrid*
102	2	*Prosopis flexuosa*
103	2	*Prosopis flexuosa*
104	2	*Prosopis flexuosa*
106	2	*Prosopis flexuosa*
107	2	*Prosopis flexuosa*
108	2	*Prosopis flexuosa*
110	2	*Prosopis flexuosa*
112	2	*Prosopis flexuosa*

Appendix 3: Protocol for microsatellite isolation in *Prosopis chilensis* using magnetic capture and direct cloning.

The following protocol was used to develop microsatellite markers in *Prosopis chilensis* at the Institute of Forest Genetics and Forest Tree Breeding, Georg-August University of Göttingen, Germany. The protocol is based on the method developed by Fischer and Bachmann (1998) for the isolation of microsatellite in organisms with large genome.

1. DNA isolation

Total genomic DNA was isolated from single individual, from five day-old green cotyledons using DNeasy® Plant Mini Kit (QIAGEN), following the manufacturer's instructions. Final elution was performed in 100 µl Buffer AE (QIAGEN).

After isolation, run an aliquot of the DNA sample on a 0.8 % agarose gel (TBE). Run together a molecular weight marker as a standard for quantification. Extraction should yield a high molecular weight band of DNA, with little to no smearing. A concentration of around 6 µg DNA in a 12µl solution is needed for subsequently procedures.

2. Restriction of nuclear genomic DNA

Total genomic DNA was digested by blunt-end-generating restriction endonuclease *RsaI*. This restriction enzyme recognizes the restriction site (GT/AC) generating blunt-end double-stranded DNA fragments.

Restriction mix:

12 µl Total genomic DNA (~ 6 µg DNA)
2.5 µl 10x restriction buffer
3 µl *RsaI* restriction enzyme (10 units/µl)
7.5 µl distilled water (H_2O)
25 µl total volume

- After prepare the restriction sample mix gently with the pipette and incubate at 37°C for 1.5 hours.

3. Adapter-to-primer ligation

Adapters with overlapping complementary sequences were ligated to the blunt-end termini of the restricted DNA fragments. The sequences of the two adapter-to-primers used in this procedure are given below:

Adapter 21-mer: 5'- CTCTTGCTTACGCGTGGACTA- 3'
Adapter 25-mer: 5'phosphorylated-TAGTCCACGCGTAAGCAAGAGCACA-3'

Prepare an *Adapter mix* (100 pmol/µl) as follows:

Adapter 21-mer (100 pmol/µl) + *Adapter 25-mer* (100 pmol/µl)

Add to the digested DNA (***Restriction mix*** = 25 µl total volume):

2.3 µl T4-Ligase (4 units/µl)
3.5 µl ATP (10 mM)
5.1 µl *Adapter mix* (10 pmol/ µl)

- Incubate de reaction at 37°C for two hours.
- Then incubate at 95°C for 10 minutes to inactivate the enzymes.
- Then leave briefly at room temperature. If necessary, the reaction can be stored at 4°C for two to three days or at -20°C for a longer period.
- After ligation purified the reaction with QIAEX II® kit (QIAGEN) for DNA extraction from agarose gels, or using any other DNA purification kit. The detailed procedure of cleaning using the aforesaid kit is given in the manufacturer's protocol. Elute the purified DNA in 10 µl distilled water (H_2O).

4. Restriction-ligation test

In order to check the success of the restriction-ligation (steps **2** and **3**), run on a 0.8 % agarose gel (TBE), an aliquot (0.5 µl should be enough) of the purified product obtained from step **3**. Run together a molecular weight marker as a standard. A tenuous smear of DNA fragments should be observed between 300bp and 1000bp approximately.

Perform a PCR test with the restriction-ligation product using as a primer the *Adapter 21-mer*.

PCR Mix

9.8 µl H_2O
1.5 µl 10x PCR Buffer
1.0 µl $MgCl_2$
0.7 µl dNTPs (10mM)
2.0 µl primer *Adapter 21-mer* (10 pmol/µl)
0.1 µl TAQ DNA Polymerase (5 units/ µl)
1.0 µl DNA (purified restriction-ligation product obtained after step **3**)
16 µl total volume

PCR Program

94°C	5 minutes	
94°C	30 seconds	30 cycles
56°C	30 seconds	
72°C	1 minute	
72°C	7 minutes	

- Test the PCR product in 1% agarose gel (TBE). If the expected band is present, then start with the next step.

5. Hybridization of biotinylated microsatellite oligonucleotides to genomic DNA fragments

In this step the restricted-ligated and subsequently purified genomic DNA is hybridized to biotinylated microsatellite oligonucleotide probes. These oligonucleotide probes are artificial oligonucleotides constructed with and oligomer sequence complementary to the microsatellite sequence that want to be isolated. In the present protocol the oligonucleotide probes $(CA)_{10}$, $(GAA)_8$ and $(AAC)_8$ were used. Each oligonucleotide probes was in a 10 µM solution.

Preparation of oligonucleotide pool

The oligonucleotide pool is a mix of the selected biotinylated microsatellite oligonucleotide probes.

Oligonucleotides	Melting Temperature (°C)	Annealing Temperature (°C)	Mean Annealing Temperature (°C)
$(CA)_{10}$	80.4	75.4	75.0
$(GAA)_8$	78.0	73.0	75.0
$(AAC)_8$	82.2	77.0	75.0

- Mix 1.0 µl of 10 µM solutions from each biotinylated microsatellite oligonucleotide.
- Add to the above mixture 6.5 µl H_2O to obtain a final volume of 0.5µl.

Hybridization mixture

27.5 µl H_2O
19.5 µl 20x SSC Buffer
8.5 µl purified restriction-ligation product (DNA fragments with the adapters)
9.5 µl oligonucleotide pool (from the oligonucleotide pool prepared above)
65.0 µl final volume

Preparation of streptavidin coated Dynabeads (Dynabeads®, Dynal GmbH)

This preparation step can be performed during the incubation of the hybridization mix that is described below, after this step.

- Suspend the beads by gently shaking.
- Take 30 µl (0.3 mg Dynabeads particles, Dynabeads®, Dynal GmbH) of the suspended beads, and pipette into a tube.
- Wash twice the beads in 1 ml 1x TE buffer.
- After each wash place the tube in the MPC (Magnetic Particle Collecting Unit) to capture the beads. Be sure to use clear, tight closing tubes that fit well into the MPC. Wait 2 to 3 minutes until the beads are concentrated and separated from the buffer, and remove the supernatant buffer carefully with a pipette while the tube remains in the MPC. Keep the tube with the beads in it.
- Wash twice with 1 ml 6x SSC Buffer in the same way as in the previous step.
- Collect the beads in a tube.

Indirect capture hybridization

- Incubate the hybridization mixture (total volume 65.0 µl) at 98º C for 5 minutes.
- Allow slow (0.1º C/second) cooling to mean annealing temperature (75º C) in a thermocycler.
- Incubate at mean annealing temperature (75º C) for 20 minutes (in the meanwhile *streptavidin-coated Dynabeads* can be prepare, see step above)
- Quickly pipette the hybridization mix (65.0 µl) to dynabeads in 35 µl 6x SSC (pre-washed twice in 6x SSC as described above in ***Preparation of streptavidin coated Dynabeads***). Thus obtaining 100 µl capture mix in 6x SSC (= 1M salt, which is crucial for an efficient streptavidin-to-biotin bonding).
- Re-suspend the dynabeads and allow 20 minutes moderately agitated incubation.
- Wash twice at room temperature in 1 ml 2x SSC + 0.1 % SDS (5 minutes each wash) applying moderate agitation (2 Hz). After each wash, place the tube in the MPC to capture the beads, wait 2 to 3 minutes until the beads separate from the buffer and remove the supernatant with a pipette while the tube remains in the MPC. Keep the tube with the beads in it.
- Wash twice in 1x SSC at room temperature (5 minutes each wash) applying moderate agitation (2 Hz). After each wash, place the tube in the MPC to capture the beads, wait 2 to 3 minutes until the beads separate from the buffer and remove the supernatant with a pipette while the tube remains in the MPC. Keep the tube with the beads in it.
- Wash once (2 minutes) at mean annealing temperature (75º C) in 1 ml 1x SSC. After washing place the tube in the MPC to capture the beads, wait 2 to 3 minutes until the beads separate from the buffer and remove the supernatant with a pipette while the tube remains in the MPC. Remove the washing buffer immediately and do not leave open the tube to air-dry. Keep the tube with the beads in it.
- Wash once (5 minutes) at mean annealing temperature (75º C) in 1 ml 1x SSC. After washing place the tube in the MPC to capture the beads, wait 2 to 3 minutes until the beads separate from the buffer and remove the supernatant with a pipette while the tube remains in the MPC. Remove the washing buffer

immediately and do not leave open the tube to air-dry. Keep the tube with the beads in it.

6. Elution

- Add the following solution to the beads obtained in the previous step:
    - 20 µl 0.1M NaOH + 0.1M NaCl and mix well (prepare the mix with: 17.5 µl H_2O + 2 µl 1N NaOH + 0.5 µl 4M NaCl)*
    - 10 µl 2M HCl (8 µl H_2O + 2 µl 1N HCl)[1]*
    - 2.2 µl 0.1M Tris-HCl pH 7.5 (9 µl H_2O + 1 µl Tris-HCl pH 7.5)*

[1]***NOTE:** All solutions must be prepared before use with stock solutions as is mention in brackets.

- Mix
- Centrifuge at 12400 rpm (~12000 g) for 5 minutes until the beads deposit in the bottom of the tube
- Take 2.2 µl of the eluate after centrifugation. Take the eluate from the top of the solution that is in the tube without taking any bead particle.
- Pipette the 2.2 µl eluate in a Microcon spin filter (Microcon YM-100 color code "Blue", Membrane NMWL 100000; Cat. Nº 42412). This is to concentrate the DNA (2.2 µl eluate) and remove the salts and low molecular weight components by passage trough the Microcon spin filter.
- Dilute the sample (2.2 µl eluate) with 500 µl H_2O in the filter unit and position the filter on top of a 1.5 ml tube.
- Centrifuge at 12400 rpm (~12000 g) until about 10 µl remain in the filter (that is about 7 minutes in the Microcon filters).
- Repeat the dilution and centrifugation with the same filter once more.
- Invert the filter to a new 1.5 µl tube and centrifuge briefly at 3700 rpm to collect the sample, which yield about 10 µl of enriched DNA.

7. PCR amplification

With the enriched DNA obtained after elution, a PCR was performed using the *Adapter 21-mer* as a primer.

PCR Mix

30.5 µl H_2O
7.0 µl 10x PCR Buffer (with 15nM $MgCl_{2)}$
1.0 µl dNTPs (10mM)
10.0 µl primer *Adapter 21-mer* (10 pmol/µl)
0.5 µl TAQ DNA polymerase (5 units/ µl)
1.0 µl enriched DNA
50 µl total volume

PCR Program

94°C 5 minutes
94°C 30 seconds
54°C 30 seconds } 30 cycles
72°C 1 minute
72°C 7 minutes

- Test the PCR product in 1% agarose gel (TBE). A small aliquot, 1 to 2 µl should, be enough for this test. A continuous smear from PCR is expected as product. If the expected band is present, then start with the next step.

- To isolate DNA from any other product, after PCR amplification of the enriched DNA the PCR product was cleaned with QIAEX II® kit (QIAGEN) for DNA extraction from agarose gels. Any other DNA purification kit can be used as well. The detailed procedure of cleaning using the aforesaid kit is given in the manufacturer's protocol. Elute the purified DNA in 10 µl distilled water (H_2O). The purified product will be use for later cloning.

SECOND ENRICHMENT: At this point a ***second enrichment*** procedure can be performed. Use the purified eluate DNA obtained at the end of step **7** as restriction-ligation product (DNA fragments) in the ***Hybridization mixture*** at step 5, and repeated the procedure.

8. Cloning and transformation

Cloning was performed in a pBluescript II® SK (+) Phagemid Vector (STRATAGENE). The vector was prepared for cloning following the protocol for construction of *T-vectors* proposed by Marchuck et al. (1991).

DNA ligation reaction

Compounds	Ligation reaction	Control
H_2O	5 µl	7 µl
10x T4 DNA ligase reaction buffer	1 µl	1
Vector pBluescript	1 µl	1 µl
Insert DNA (product obtained after step **7**)	2 µl [2]*	0 µl
T4 ligase (1 unit/µl)	1 µl	1 µl
Final volume	10 µl	10 µl

[2]***NOTE**: The amount of insert DNA depends on the DNA concentration. There should be about twice as much insert DNA than Vector DNA.

- Vortex briefly and incubate the DNA ligation reaction at room temperature for 2 hours.

Bacterial transformation

Transformation was performed using bacterial SURE® Competent Cells (STRATAGENE). Competent bacterial cells must be stored at -80°C.

- Thaw the cells slowly on ice and keep them always on ice during the transformation.

- Prepare the following mixture:

  5 µl ligation reaction (from the step above)
  50 µl SURE® Competent Cells (STRATAGENE)
  55 µl final volume

- Mix gently (do not use vortex).

- Incubate on ice for 30 minutes.
- Heat-shock the cells by putting them in a heating block at 42°C for 45 seconds. Time and temperature are very important and must be very precise.
- Quickly incubate back the mixture on ice for 2 minutes.
- Add 500 µl of SOB media to the cells.
- Shake the mixture at 37°C for 30 minutes.
- Concentrate the cells: centrifuge the tubes and discard around 400 µl of the supernatant media. Resuspend the cells by shaking.
- Plate 100 µl of the cells per plate in a convenient medium (LB-ampicillin 100 µg/ml).
- Incubate at 37º C over night (the plate upside down).

9. Colony PCR

Detection of the positives colonies was performed by blue/white selection. White colonies are positive (with insert) and blue colonies are negative (without insert).

- Transfer the positive colonies (white) to grid plates (10 x 10) with LB-ampicillin medium using a sterilized pipette tip.
- Use the tip employed to transfer the colony to perform the PCR. After pick up and transfer the positive colony, some cells remain in the tip, place the tip in a tube with the following PCR mix:

PCR Mix

18.9 µl H_2O
3.0 µl 10x PCR Buffer (with 15nM $MgCl_{2)}$
1.0 µl dNTPs (5mM)
1.0 µl BF primer (10 µM) [3]*
1.0 µl BR primer (10 µM) [3]*
0.1 µl TAQ DNA polymerase (5 units/ µl)
DNA by introducing the pipette tip into the mix
25 µl total volume

[3]***NOTE**: Sequences of the primers

BF: 5`-TGTAAAACGACGGCCAGTGA- 3'

BR: 5`-AGCGGATAACAATTTCACACAGGA- 3'

PCR Program

94°C	5 minutes	
94°C	30 seconds	35 cycles
60°C	30 seconds	
72°C	1 minute 30 seconds	
72°C	10 minutes	

- Test the PCR product in 2% agarose gel (TBE). Run together a molecular weight marker as a standard. If the expected band is present (bands bigger than 500bp), then start with the next step.

- Incubate the grid plates (10 x 10) at 37º C over night (the plate upside down). These plates represent the microsatellite-enriched library. They can be store for 3 to 5 weeks at 4°C. For long time conservation of the microsatellite-enriched library, colonies should be transferred to tubes with LB-ampicillin medium with glycerol 15%. After incubation (37º C over night) stored at -80°C.

10. Sequencing of colony DNA

The PCR product received from colony PCR was purified using and sequencing was performed using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and carried out with the ABI PRISM® 3100 Genetic Analyser (Applied Biosystems / HITACHI).

- To the PCR product obtained after colony PCR, add distilled water (H_2O) to complete a final volume of 50 µl.

- Purified the 50 µl PCR + water product using MiniElute® PCR Purification Kit (QIAGEN) following the manufacturer's instruction. Perform the final elution in 10 µl EB Buffer (QIAGEN).

- Perform the sequenciation reaction as follows:

Sequenciation reaction mix

4.8 µl H_2O

1.5 µl 5x BigDye Terminator v1.1/3.1 Sequencing Buffer

0.7 µl Primer (5 pmol/µl) [4]*

1.0 µl BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems)

2.0 µl DNA (eluate obtained after purification with MiniElute® PCR Purification Kit)

10.0 µl final volume

[4]***NOTE**: The primer used for the sequenciation reaction was T-3 or T-7.

T-3: 5`-GCTCGAAATTAACCCTCACTAAAG- 3'

T-7: 5`-GAATTGTAATACGACTCACTATAG- 3'

PCR Program

96°C	1 minute	
96°C	10 seconds	35 cycles
45°C	10 seconds	
60°C	4 minutes	

Subsequent sequenciation was performed with an ABI PRISM® 3100 Genetic Analyser from Applied Biosystems / HITACHI.

References

Fischer D, Bachmann K (1998) Microsatellite enrichment in organisms with large genomes (*Allium cepa L.*). *Biotechniques*, **24**, 796-802.

Marchuk D, Drumm M, Saulino A, Collins F S (1990) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Research*, **19**, 1154.

Appendix 4: Complete sequences of the six microsatellite loci developed in *Prosopis chilensis*, as they appear publicly available in the GenBank® sequence database.

Microsatellite locus Mo05

Locus	AJ879505; 599 bp; DNA linear; PLN 11-AUG-2005
Definition	*Prosopis chilensis* microsatellite DNA, locus Mo05, allele 218.
Accession	AJ879505
Version	AJ879505.1; GI:68673100
Keywords	Microsatellite; repetitive DNA.
Source	*Prosopis chilensis* (Chilean mesquite)
Organism	*Prosopis chilensis* Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Mimosoideae; Mimoseae; *Prosopis*.
Reference	1
Authors	Mottura, M.C.; Finkeldey, R.; Verga, A.R.; and Gailing, O.
Title	Development and characterization of microsatellite markers for *Prosopis chilensis* and *P. flexuosa* and cross-species amplification
Journal	Molecular Ecology Notes 5 (3), 487-489 (2005)
Reference	2 (bases 1 to 599)
Authors	Mottura, M.C.
Title	Direct Submission
Journal	Submitted (02-FEB-2005) Mottura M.C., Institute of Forest Genetics, University of Goettingen, Buesgenweg 2, 37077 Goettingen, GERMANY
Comments	Genomic DNA, microsatellite region.
Features	Location/Qualifiers
source	1..599 organism ="Prosopis chilensis" molecular type ="genomic DNA" clone ="R4 - 5" tissue type ="seed" country = Argentina, Arid Chaco
satellite	464….474 note ="microsatellite locus Mo05" allele ="218" repeated type =TANDEM repeated unit sequence="ct(3)t(1)ct(2)"

Sequence

```
  1 tctcttgctt acgcgtggac taacttgctc catgagccac ctctgttttt tgttcaaaaa
 61 acagagatta acatagccat gaacctgaaa gttgcaagcc tcaagatctt tacaatggca
121 gagcaattgg acatttggac cttgtggtcc tcttattctt caattctatt tatttttcc
181 caataatatt actcttataa attaatgatg acttagataa gatatctata tgcaacgtga
241 ttttggcttt cctaattgat tgatttttt tccagctaca tgacatatat ttcatataaa
301 ataataaatt tcaaacaaat tctgcagtct cttcgcctcc aacccacaaa aaatgcctca
361 tatgaacata taatttcaaa acaatggtga tttcttcaaa gcatatacat tcttcgcccc
421 tttgcacttc cctccccaag taaccgattc ttttccacgt atgctctctt ctctattttg
481 gcttcccttt ttataagtga aacaatgaag gagggctgag gagtcacgag ggatcatctt
541 cttcatcaag aggagggaga aaagctggga agagctagac aaaacacngt tgcctaaat
//
```

Microsatellite locus Mo07

Locus	AJ879506; 662 bp; DNA linear; PLN 11-AUG-2005
Definition	*Prosopis chilensis* microsatellite DNA, locus Mo07, allele 197.
Accession	AJ879506
Version	AJ879506.1; GI:68673106
Keywords	Microsatellite; repetitive DNA.
Source	*Prosopis chilensis* (Chilean mesquite)
Organism	*Prosopis chilensis* Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Mimosoideae; Mimoseae; *Prosopis*.
Reference	1
Authors	Mottura, M.C.; Finkeldey, R.; Verga, A.R.; and Gailing, O.
Title	Development and characterization of microsatellite markers for *Prosopis chilensis* and *P. flexuosa* and cross-species amplification
Journal	Molecular Ecology Notes 5 (3), 487-489 (2005)
Reference	2 (bases 1 to 662)
Authors	Mottura, M.C.
Title	Direct Submission
Journal	Submitted (02-FEB-2005) Mottura M.C., Institute of Forest Genetics, University of Goettingen, Buesgenweg 2, 37077 Goettingen, GERMANY
Comments	Genomic DNA, microsatellite region.
Features	Location/Qualifiers
source	1..662 organism ="Prosopis chilensis" molecular type ="genomic DNA" clone ="R5 - 4" tissue type ="seed" country = Argentina, Arid Chaco
satellite	275….290 note ="microsatellite locus Mo07" allele ="197" repeated type =TANDEM repeated unit sequence="gc(8)"

Sequence

```
  1 tctcttgctt acgcgtggac taacaacagc aatttcactt tctcaagcct agaaaggaaa
 61 tgggccaatg tatagtctaa aaagtattgt gcgtggtgca ctaagtccaa aaagaagagc
121 atcagatttg ctgggttggg actcatgaaa tgccaaaata agatgcttat gcagtaagct
181 tgcaaggata agttaaaaca cacacacaca cacacgttgc ctccatttaa acctctcaga
241 agctccctca cattttgccc ttgcacaagc aagagcgcgc gcgcgcgcgc acagagagct
301 aagggaaggg gaagatatga tcgaagaaga aggaaagttc tcccctcaca tgcaaccgaa
361 cgtttagacc aaggtaagga ccttaaacct tattccatca acatttttcc ctgtggctgt
421 gtgttgcgca aatagataag aggttttcat ctggttttta aagagagaaa gggagggacg
481 catgcattaa gttaaaaatc atatttagat gatagatggg ggttgatttg agtgagaaga
541 agtgattttt agatctggag tttctgtttc cggggagtat ttatttaact tctcttacgg
601 gagctgaaat ctggaattca ggttataatg tttaattatg ttagtccacg cgtaagcaag
661 ag
//
```

Microsatellite locus Mo08

Locus	AJ879507; 549 bp; DNA; linear; PLN 11-AUG-2005
Definition	*Prosopis chilensis* microsatellite DNA, locus Mo08, allele 218.
Accession	AJ879507
Version	AJ879507.1; GI:68673109
Keywords	Microsatellite; repetitive DNA.
Source	*Prosopis chilensis* (Chilean mesquite)
Organism	*Prosopis chilensis* Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Mimosoideae; Mimoseae; *Prosopis*.
Reference	1
Authors	Mottura, M.C.; Finkeldey, R.; Verga, A.R.; and Gailing, O.
Title	Development and characterization of microsatellite markers for *Prosopis chilensis* and *P. flexuosa* and cross-species amplification
Journal	Molecular Ecology Notes 5 (3), 487-489 (2005)
Reference	2 (bases 1 to 549)
Authors	Mottura, M.C.
Title	Direct Submission
Journal	Submitted (02-FEB-2005) Mottura M.C., Institute of Forest Genetics, University of Goettingen, Buesgenweg 2, 37077 Goettingen, GERMANY
Comments	Genomic DNA, microsatellite region.
Features	Location/Qualifiers
source	1…549 organism ="Prosopis chilensis" molecular type ="genomic DNA" clone ="R5 - 16" tissue type ="seed" country = Argentina, Arid Chaco
satellite	363….380 note ="microsatellite locus Mo08" allele ="218" repeated type =TANDEM repeated unit sequence="ac(9)"

Sequence

```
  1 taacgtcttt ggtggaggaa aggaaaaaaa gagggattca ttagttgatg ttcttcccag
 61 taatgatgaa gctacttgcg ctttcaaaca aaccatcaaa cgcagaaaag aatttaattt
121 cattgctcag agactgtttc aaatttattc acctggactt tggcattacc aaaacaccca
181 ccctacacaa atcatattca attccattcc atctccccta cgttttggaa ttccatgcca
241 ctcttcttcc aggtcatttc cttttctgcc attattctat cctaaacgcc gggctacgta
301 ttatatacac accccaactt cttactttta catatacata catacatata tgatgtatga
361 atacacacac acacacacac cccacctggt tttggaaaga tatattggat cccaattaat
421 tcaaattaag ttatacttga gttataatca tctaaacaca gatcgagttt ctggtttaag
481 tatgcatgaa tgggaaaata tctctttaga agagttgtat ttatataaac cgaacagtgt
541 tcgaatagt
//
```

Microsatellite locus Mo09

Locus	AJ879508; 843 bp; DNA; linear; PLN 11-AUG-2005
Definition	*Prosopis chilensis* microsatellite DNA, locus Mo09, allele 233.
Accession	AJ879508
Version	AJ879508.1; GI:68673112
Keywords	Microsatellite; repetitive DNA.
Source	*Prosopis chilensis* (Chilean mesquite)
Organism	*Prosopis chilensis* Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Mimosoideae; Mimoseae; *Prosopis*.
Reference	1
Authors	Mottura, M.C.; Finkeldey, R.; Verga, A.R.; and Gailing, O.
Title	Development and characterization of microsatellite markers for *Prosopis chilensis* and *P. flexuosa* and cross-species amplification
Journal	Molecular Ecology Notes 5 (3), 487-489 (2005)
Reference	2 (bases 1 to 843)
Authors	Mottura, M.C.
Title	Direct Submission
Journal	Submitted (02-FEB-2005) Mottura M.C., Institute of Forest Genetics, University of Goettingen, Buesgenweg 2, 37077 Goettingen, GERMANY
Comments	Genomic DNA, microsatellite region.
Features	Location/Qualifiers
source	1…843 organism ="Prosopis chilensis" molecular type ="genomic DNA" clone ="R5 - 24" tissue type ="seed" country = Argentina, Arid Chaco
satellite	447….480 note ="microsatellite locus Mo09" allele ="233" repeated type =TANDEM repeated unit sequence="tg(17)"

Sequence

```
  1 ccagccaaga gagagagcta acatcttaaa cagttgtaaa acaatcaata gaggacccct
 61 gctaattgaa gagtgatctg agcccagatc actatttatc ccaatttata tcctaaaatc
121 cttatgacca aacgacaata acttgtctga gcactttcta aaaaattaaa tcaagttaca
181 agaaatttca gaattgattt cttggtcatg tttgacgtat tgccgtcttt ctcctttcca
241 ctagaccgaa ggcaaataag gacatgtaaa tatcaatcct aaatctaggc aaaattacat
301 agtccattcc tccctcacat tttgcgcaca agtataactc agtcaataat gtaacacagt
361 agccatattt cataaaagaa caatcacaat atgactatag ccattctata ataagaaggg
421 acttaaccat tatgattggc atgctctgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg
481 aaacgggaag caaatttctc accccaagta ggaaaataca acaaaggctg gcataatgtt
541 cgaaaaggta acaccccaca acagcatgag attacagtaa atttacttct taccttcgac
601 cagcaagtat cttggccatg ttaccattat tgtttgtgac aaacacgtca ctatcatcac
661 aaacaatgta gtcaatggca gcaagacggg aagagaatgg aaggaaaggc ttcaggtcct
721 cattagcaag catctcctta gtatacatgt tagggaaaag ttctctaaga gggcgcatag
781 tctcatcccc accatatatt tctcctgacg caacatagag gttagtccac gcgtaagcaa
841 gag
//
```

Microsatellite locus Mo13

Locus	AJ879509; 503 bp; DNA; linear; PLN 11-AUG-2005
Definition	*Prosopis chilensis* microsatellite DNA, locus Mo13, allele 228.
Accession	AJ879509
Version	AJ879509.1; GI:68673115
Keywords	Microsatellite; repetitive DNA.
Source	*Prosopis chilensis* (Chilean mesquite)
Organism	*Prosopis chilensis* Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Mimosoideae; Mimoseae; *Prosopis*.
Reference	1
Authors	Mottura, M.C.; Finkeldey, R.; Verga, A.R.; and Gailing, O.
Title	Development and characterization of microsatellite markers for *Prosopis chilensis* and *P. flexuosa* and cross-species amplification
Journal	Molecular Ecology Notes 5 (3), 487-489 (2005)
Reference	2 (bases 1 to 503)
Authors	Mottura, M.C.
Title	Direct Submission
Journal	Submitted (02-FEB-2005) Mottura M.C., Institute of Forest Genetics, University of Goettingen, Buesgenweg 2, 37077 Goettingen, GERMANY
Comments	Genomic DNA, microsatellite region.
Features	Location/Qualifiers
source	1…503 organism ="Prosopis chilensis" molecular type ="genomic DNA" clone ="R7 - 5" tissue type ="seed" country = Argentina, Arid Chaco
satellite	122….147 note ="microsatellite locus Mo013" allele ="228" repeated type =TANDEM repeated unit sequence="gt(10)ct(1)gt(2)"

Sequence

```
  1 actaacaata atcacttaca agagtggcaa acttatcatg agaagagttt tagtagaatg
 61 gttattcctc tctttttag tctgtgtttt gttttgatta gagttgcatg tggatgagag
121 agtgtgtgtg tgtgtgtgtg tctgtgtgac gtcaatttat ctagatctta tttgctggat
181 gcagttcctt cggaagatgg cacgagcata tgataaaagt gggggtagtg gcaagaaaac
241 actattatcc caacaagact tggagaagat gatggttgaa ggtaaaataa acgacatcat
301 gctctgacac ttgggactgc atcacctgtt cattacttat ttcttgtaat tggcgattga
361 ggggaagatt tatttggcag aggaattggg cagtatctga aatattttgt gctcaattaa
421 ttattaagca gctgctgaga acattgtatc aattttgttg agtgggtgat gattaatata
481 cagttagtcc acgcgtaagc aag
//
```

Microsatellite locus Mo16

Locus	AJ879510; 791 bp; DNA; linear; PLN 11-AUG-2005
Definition	*Prosopis chilensis* microsatellite DNA, locus Mo16, allele 157.
Accession	AJ879510
Version	AJ879510.1; GI:68673117
Keywords	Microsatellite; repetitive DNA.
Source	*Prosopis chilensis* (Chilean mesquite)
Organism	*Prosopis chilensis* Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Mimosoideae; Mimoseae; *Prosopis*.
Reference	1
Authors	Mottura, M.C.; Finkeldey, R.; Verga, A.R.; and Gailing, O.
Title	Development and characterization of microsatellite markers for *Prosopis chilensis* and *P. flexuosa* and cross-species amplification
Journal	Molecular Ecology Notes 5 (3), 487-489 (2005)
Reference	2 (bases 1 to 791)
Authors	Mottura, M.C.
Title	Direct Submission
Journal	Submitted (02-FEB-2005) Mottura M.C., Institute of Forest Genetics, University of Goettingen, Buesgenweg 2, 37077 Goettingen, GERMANY
Comments	Genomic DNA, microsatellite region.
Features	Location/Qualifiers
source	1…791 organism ="Prosopis chilensis" molecular type ="genomic DNA" clone ="R7 - 27" tissue type ="seed" country = Argentina, Arid Chaco
satellite	683….706 note ="microsatellite locus Mo016" allele ="157" repeated type =TANDEM repeated unit sequence="ca(12)"

Sequence

```
  1 ctaacgcttg gatggtttat gaagtgtggg atgtaagttc aatgaagtat tgaaaataag
 61 gataaaaatt agaattgaga ttgaaaaaaa aaaattaaaa taatgtttat catgaatgat
121 tacttattat ttgggactaa acaaatttat atatgtgaga ttggaaaata ctacgggggа
181 agtgggatat tcttcaaccg gcttatattt gattgcttaa ttcccagtaa taagaaattt
241 aatcattatt aaatattctt taatgattgt ttttcaatcc ggaatcccat tttttacttt
301 ttattttcca aaactttata cgcaaatata ttctaagaaa aacaaacaaa aacgcaggaa
361 aaatgaatga gaaatgattg ctgatttggg ggaagattac aaccaggaga tgggtgcatt
421 gcaacgaata aaaggtgtgg gctcataatc agatgaggag aaataaggga gtaacttgtg
541 gtgagtatgg tcgaccatca ccctaatttg agacttgaat gggaggctgt agctttcaag
601 ccgatgcgcc attaatcaaa gcattgcccc aatatcactc cgaaaacacc ctctcattcg
661 acttaatatt tttattaagt cccacacaca cacacacaca cacacattcc tattttcctt
721 ttaatttccc cctcatttag gcctctccat agctaccacc actactctgg atggaccccc
781 ctagtccacg c
//
```

Appendix 5: Allelic frequencies by morphological group (***Group 1***, ***Group 2*** and ***Group 3***) at microsatellite loci *Mo05*, *Mo07*, *Mo08, Mo09, Mo13* and *Mo16*, in the adult population

Locus	Allele	Group1	Group2	Group3
Mo05	**214**	0,050	0,000	0,100
	216	0,200	0,321	0,300
	218	0,750	0,679	0,600

Locus	Allele	Group1	Group2	Group3
Mo07	**185**	0,050	0,032	0,050
	191	0,000	0,057	0,000
	193	0,750	0,519	0,650
	195	0,000	0,025	0,000
	197	0,000	0,019	0,000
	199	0,000	0,076	0,000
	203	0,000	0,019	0,050
	205	0,000	0,013	0,000
	207	0,150	0,101	0,150
	209	0,000	0,019	0,100
	211	0,050	0,032	0,000
	213	0,000	0,076	0,000
	215	0,000	0,013	0,000

Locus	Allele	Group1	Group2	Group3
Mo08	**200**	0,050	0,013	0,000
	204	0,050	0,138	0,000
	206	0,050	0,031	0,000
	208	0,150	0,200	0,000
	210	0,000	0,056	0,000
	212	0,000	0,138	0,050
	214	0,000	0,125	0,100
	216	0,400	0,144	0,550
	218	0,000	0,138	0,050
	220	0,300	0,006	0,250
	222	0,000	0,013	0,000

Locus	Allele	Group1	Group2	Group3
Mo09	**203**	0,000	0,013	0,000
	207	0,000	0,131	0,000
	209	0,650	0,269	0,800
	211	0,000	0,038	0,100
	213	0,000	0,013	0,000
	215	0,150	0,262	0,100
	217	0,050	0,050	0,000
	219	0,050	0,044	0,000
	221	0,000	0,100	0,000
	223	0,000	0,006	0,000
	225	0,100	0,019	0,000
	227	0,000	0,013	0,000
	231	0,000	0,019	0,000
	235	0,000	0,006	0,000
	245	0,000	0,019	0,000

Locus	Allele	Group1	Group2	Group3
Mo13	**206**	0,050	0,025	0,000
	216	0,000	0,019	0,050
	218	0,000	0,106	0,000
	220	0,050	0,106	0,000
	222	0,000	0,006	0,000
	224	0,050	0,087	0,000
	228	0,250	0,031	0,500
	232	0,150	0,538	0,000
	234	0,000	0,019	0,000
	238	0,350	0,044	0,300
	240	0,000	0,019	0,000
	244	0,000	0,000	0,100
	246	0,100	0,000	0,050

Locus	Allele	Group1	Group2	Group3
Mo16	**147**	0,000	0,047	0,000
	149	0,100	0,027	0,100
	151	0,000	0,013	0,000
	153	0,050	0,087	0,000
	155	0,300	0,233	0,550
	157	0,200	0,113	0,150
	159	0,000	0,073	0,050
	161	0,000	0,133	0,000
	163	0,250	0,120	0,150
	165	0,050	0,053	0,000
	167	0,000	0,040	0,000
	169	0,000	0,020	0,000
	171	0,050	0,013	0,000
	173	0,000	0,013	0,000
	175	0,000	0,013	0,000

Appendix 6: Pairwise matrix of genetic distance between morphological groups of the adult population by locus. For all tables: Nei's (1972) genetic distance ***D*** below diagonal; Gregorius' (1974) genetic distance d_0 above diagonal.

Locus Mo05

Groups	Group1	Group2	Group3
Group1	--	0.121	0.150
Group2	0.018	--	0.100
Group3	0.024	0.011	--

Locus Mo07

Groups	Group1	Group2	Group3
Group1	--	0.316	0.150
Group2	0.029	--	0.310
Group3	0.016	0.037	--

Locus Mo08

Groups	Group1	Group2	Group3
Group1	--	0.606	0.350
Group2	0.690	--	0.650
Group3	0.101	0.760	--

Locus Mo09

Groups	Group1	Group2	Group3
Group1	--	0.469	0.250
Group2	0.254	--	0.594
Group3	0.029	0.329	--

Locus Mo13

Groups	Group1	Group2	Group3
Group1	--	0.650	0.400
Group2	0.856	--	0.906
Group3	0.185	2.435	--

Locus Mo16

Groups	Group1	Group2	Group3
Group1	--	0.393	0.300
Group2	0.180	--	0.457
Group3	0.126	0.206	--

Appendix 7: Allelic frequencies by morphological group (***Group 1***, ***Group 2*** and ***Group 3***) at microsatellite loci *Mo08, Mo09* and *Mo13* in the seed population.

Locus	Allele	Grupo1	Gurpo2	Grupo3
Mo08	**200**	0.056	0.042	0.000
	202	0.000	0.000	0.003
	204	0.043	0.108	0.058
	206	0.037	0.015	0.006
	208	0.068	0.165	0.044
	210	0.006	0.032	0.015
	212	0.037	0.114	0.023
	214	0.031	0.152	0.110
	216	0.358	0.217	0.372
	218	0.056	0.116	0.064
	220	0.309	0.032	0.305
	222	0.000	0.004	0.000
	228	0.000	0.002	0.000

Locus	Allele	Grupo1	Gurpo2	Grupo3
Mo09	**203**	0.000	0.025	0.000
	207	0.043	0.072	0.000
	209	0.636	0.346	0.767
	211	0.006	0.029	0.102
	213	0.000	0.004	0.000
	215	0.142	0.177	0.096
	217	0.093	0.061	0.009
	219	0.062	0.080	0.000
	221	0.019	0.144	0.020
	223	0.000	0.021	0.000
	225	0.000	0.013	0.006
	227	0.000	0.015	0.000
	235	0.000	0.010	0.000
	245	0.000	0.004	0.000

Locus	Allele	Grupo1	Gurpo2	Grupo3
Mo13	**206**	0.063	0.013	0.000
	216	0.000	0.027	0.017
	218	0.000	0.110	0.000
	220	0.089	0.101	0.003
	222	0.006	0.002	0.000
	224	0.127	0.093	0.000
	226	0.000	0.002	0.000
	228	0.120	0.038	0.375
	232	0.234	0.498	0.119
	234	0.006	0.000	0.006
	238	0.259	0.061	0.270
	240	0.019	0.040	0.006
	244	0.000	0.011	0.122
	246	0.076	0.004	0.081

CURRICULUM VITAE

Personal data

Name	Martín Carlos Mottura
Nationality	Argentinean
Date of birth	12th August 1975
Place of birth	San Guillermo, Santa Fe, Argentina
Email	martinmottura@gmail.com or martinmottura@hotmail.com

Education

1982-1988	Elementary School Mariano Moreno, San Guillermo, Santa Fe, Argentina
1989-1992	Agricultural High School Our Lady of the Rosary, Colonia Vignaud, Córdoba, Argentina. Degree: Agricultural Expert
1993-1994	Senior High School Goodland, Goodland, Kansas, USA
1995-2002	Faculty of Agricultural Science, National University of Córdoba, Argentina. Degree: Agricultural Engineer
2003-2006	Doctoral study at the Institute of Forest Genetics and Forest Tree Breeding, Faculty of Forest Science, Georg-August University of Göttingen, Germany

Training Practical at the Institute of Silviculture, Faculty of Agricultural Science, National University of Córdoba (2001-2002); Introduction to Plant Biotechnology (2002); Methods and Tools in Sequence Analysis (2004); Analysis of Population Genetic Data (2004-2005); Course on Scientific Writing (2005); Applied Statistics (2005).

Occupation

2002-present	Assistant researcher for the project "Conservation, genetic improvement and silviculture of *algarrobo (Prosopis spp*.) in the Argentinean Arid Chaco", National Institute of Agricultural Technology (INTA-IFFIVE), Córdoba, Argentina.