Molecular approach to dissect adaptive traits in native European Populus nigra L.: construction of a genetic linkage map based on AFLP, SSR and SNP markers

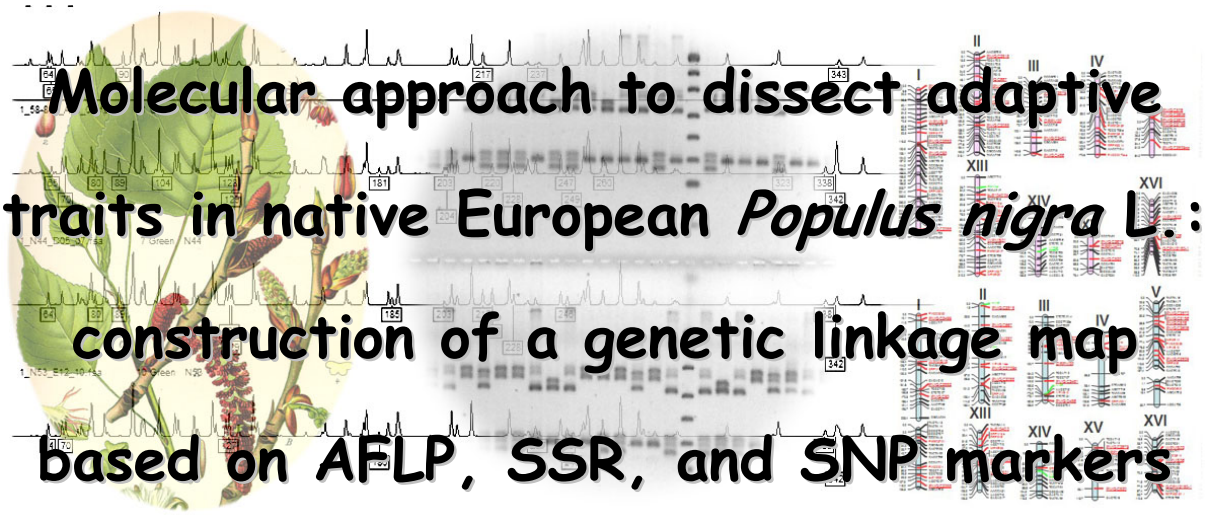
Muriel Gaudet

UNIVERSITÀ DEGLI STUDI DELLA TUSCIA

Dipartimento di Scienze dell'Ambiente Forestale e delle sue Risorse

CORSO DI DOTTORATO DI RICERCA

ECOLOGIA FORESTALE XVIII CICLO



Molecular approach to dissect adaptive traits in native European *Populus nigra* L.: construction of a genetic linkage map based on AFLP, SSR, and SNP markers

S.S.D. AGR 05

COORDINATORE: PROF. PAOLO DE ANGELIS

TUTORI:

PROF. GIUSEPPE SCARASCIA MUGNOZZA

DR. MAURIZIO SABATTI

DR. VÉRONIQUE JORGE

CANDIDATO: GAUDET MURIEL

A ***Jocelyne, Joël***

Lucienne, Erwin

Martine, Marlène, Léa

Acknowledgments - Ringraziamenti - Remerciements

❀ *At the end of this research work, which I have carried out at the Department of Forest Environment and Resources (**Di.S.A.F.Ri**) in University of Tuscia, I would like to acknowledge **Riccardo Valentini,** department director, and **Giuseppe Scarascia Mugnozza**, scientific responsible for Viterbo's operative unit within POPYOMICS program.*

❀ *I acknowledge all the research teams involved in E.U. program **POPYOMICS**.*

❀ *I thank **Gerald A. Tuskan** and **Stephan P. DiFazio** (Oak Ridge National Laboratory, U.S.A.), **Michele Morgante,** and **Giusi Zaina,** (University of Udine) for exchanging data and information useful for molecular analyses and discussion of results.*

❀ *Ringrazio **Maurizio Sabatti** per la sua disponibilità, la sua simpatia e per le pause tè e le discussioni alla macchinetta del caffè.*

❀ *Ringrazio **Paolo de Angelis** per i consigli e l'aiuto nella risoluzione di tanti piccoli problemi informatici e amministrativi.*

❀ *Ringrazio **i colleghi del dipartimento DISAFRI** per la loro piacevole compagnia, in particolare **Isabella Paolucci** per la sua collaborazione e il suo aiuto.*

❀ *Ringrazio i colleghi del CNR IBAF di Porano: **Daniela Taurchini, Marcello Cherubini, Claudia Mattioni e Fiorella Villani** per la loro accoglienza e simpatia che hanno reso il periodo di lavoro al sequenziatore molto piacevole.*

❀ *Ringrazio **Francesco Salani** per le sue discussioni in francese e per i sui "software génials" e **Luca Ricciotti** per la collaborazione scientifica e le avventurose missioni svolte insieme.*

❀ *Ringrazio* ***Giulia Fara*** *per il suo ottimo lavoro sugli SNPs e la sua amicizia.*

❀ *Ringrazio* ***Moica Piazzai*** *e* ***Paola Pollegioni*** *per la loro gentilezza e le piacevoli serate passate insieme.*

❀ *Ringrazio* ***tutti i nuovi amici italiani*** *per i gradevoli momenti passati insieme.*

❀ *Un très grand merci à* ***Véronique Jorge*** *(INRA Orléans, France) pour son aide précieuse lors de l'élaboration des données et pour sa sympathie qui m'ont été d'un grand soutien tout au long de cette thèse. Je remercie également* ***Catherine Bastien*** *(INRA Orléans, France) pour sa disponibilité et ses conseils.*

❀ *Merci à* ***Nathalie Boizot*** *pour sa tasse de thé toujours prête qui a été bien appréciable lors des mesures des plantes au mois de Décembre. Et surtout je profite de cette occasion pour la remercier, avec son mari, pour leur amitié.*

❀ *La fin de ce travail a été aussi l'occasion d'un bilan sur mon parcours et je voudrais remercier* ***Christian Breton, Christian Jay-Allemand, Philippe Label*** *et* ***Jean-Paul Charpentier*** *pour leur initiation au travail de recherche, leur formation aux techniques de biologie moléculaire et biochimie, et les entraînements aux présentations orales, qui se sont avérés très utile pour la réalisation de cette thèse.*

❀ *Une pensée particulière pour* ***Anne-Laure Ponsonaille****, toujours présente quelques soit la distance.*

❀ *Ringrazio profondamente* ***Bruna e Abbramo*** *per tutto.*

❀ *Merci :*

« Il a les yeux presque aussi clairs
Que les murs blancs du fond de l'Espagne
Un jour je bâtirai un empire
Avec tous nos instant de plaisirs
Pour que plus rien ne m'éloigne
Du garçon qui m'accompagne »

Table of content

List of Figures

List of Tables

List of abbreviations

AB-QTL	**A**dvanced **B**ackcross **QTL**
AFLP	**A**mplified **F**ragment **L**ength **P**olymorphism
AG	**AG**AMOUS
AS-PCR	**A**llele **S**pecific-**PCR**
BAC	**B**acterial **A**rtificial **C**hromosome
BC	**B**ack**c**ross
BLAST	**B**asic **L**ocal **A**lignment **S**earch **T**ool
bp	**b**ase **p**airs
CAPS	**C**leaved **A**mplified **P**olymorphic **S**equence
cDNA	**c**omplementary **DNA**
cM	**c**enti **M**organ
COS	**C**onserved **O**rthologous **S**et
dCAPS	**d**erived **C**leaved **A**mplified **P**olymorphic **S**equence
DNA	**D**eoxyribo**n**ucleic **A**cid
dNTP	**d**eoxy**n**ucleoside **T**ri**p**hosphate
DOE	**D**epartment **O**f **E**nergy
eQTL	**e**xpression **QTL**
EST	**E**xpressed **S**equence **T**ag
IAA	**I**ndol **A**cetic **A**cid
ISSR	**I**nter **S**imple **S**equence **R**epeat
LD	**L**inkage **D**esequilibrium
LOD	**L**og of the **Od**ds ratio
MALDITOF	**M**atrix-**A**ssisted **L**aser **D**esorption/**I**onization **T**ime-**O**f-**F**light
MAS	**M**arker-**A**ssisted **S**election
PCR	**P**olymerase **C**hain **R**eaction

PhyA	**Phy**tochrome **A**
PhyB	**Phy**tochrome **B**
PTAG	*Populus trichocarpa* **AG**AMOUS
QTL	**Q**uantitative **T**rait **L**oci
RAPD	**R**andom **A**mplified **P**olymorphic **D**NA
RFLP	**R**estriction **F**ragment **L**ength **P**olymorphism
RIL	**R**ecombinant **I**nbred **L**ines
SCAR	**S**equence-**C**haracterized **A**mplified **R**egion
SNP	**S**ingle **N**ucleotide **P**olymorphism
SSR	**S**imple **S**equence **R**epeat
STS	**S**equence **T**agged **S**ite
TILLING	**T**arget **I**nduced **L**ocal **L**esions **In** **G**enome
Tm	**m**elting **T**emperature

Introduction

Trees are used to produce a variety of wood-based products including timber, pulp, and paper. More recently, their use as source of renewable energy has also been highlighted, as has their value for carbon mitigation within the Kyoto Protocol. The domestication of trees has only just begun in comparison to food crops; the long generation time and complex nature of juvenile and mature phase are contributory factors. To accelerate the domestication process and to further understand some of the unique processes that occur in woody plants, such as dormancy and secondary wood formation, a "model" tree is needed (Taylor, 2002). The genus *Populus* was accepted as a model for trees because it has a relatively small genome, about 550 million base pairs, which is similar to rice, only 4× larger than *Arabidopsis*, and one fortieth the size of the *Pinus* genome. Other important attributes and resources include: worldwide distribution of multiple species; high growth rates; ease of clonal propagation; availability of genetic maps and structured pedigrees; publicly accessible molecular markers, gene sequences, bacterial artificial chromosome (BAC) libraries; high-throughput plant transformation and regeneration capabilities (Wullschleger *et al.*, 2002). Moreover, earlier in 2002, the United States Department of Energy (DOE) announced plans to sequence the genome of a *Populus trichocarpa* genotype. The project initiated in February 2002 (Wullschleger *et al.*, 2002; Taylor, 2002) and in September 2004 the sequence was available at the website http://genome.jgi-psf.org/Poptr1/Poptr1.home.html. Finally, *Populus* is unique in that it will not only act as a model for all woody species, but is in itself a forest tree of considerable commercial importance, especially for its fast growth (Taylor, 2002).

In this context the POPYOMICS project (contract QLRT-2001-00953) has been funded by the Fifth Framework Program of the European Union and has started on November 2002. The project is now reaching its end (May 2006). The main aim of the POPYOMICS project is to link physiology, molecular genetics and genomics to understand and improve yield of *Populus* for growth across Europe as a bioenergy and timber crop. To reach this purpose, the latest techniques in genetic mapping, genomics, and the physical sequence of poplar, as available, will be used to define genes which determine yield and disease resistance in *Populus*. The long-term ambition of POPYOMICS is to develop methods to select new genotypes of *Populus* with improved potential for above and below ground carbon sequestration optimizing wood production in short rotation forestry. The work of the project has been undertaken by nine partners, from

five European countries, within five workpackages. The range of techniques and experimental approaches is large, from field trials to candidate gene discovery and microarrays. The genetic resources consist of five pedigrees representing four *Populus* species: *P. trichocarpa*, *P. deltoides*, *P. nigra,* and *P. alba*. They have been selected in a variety of climatic conditions, ranging from Southern, Mediterranean to the extreme Northern Europe. The selected pedigrees were replicated and maintained in three European sites: United Kingdom, France, and Italy. The main objectives of the POPYOMICS project are listed below.

- **Mapping the five available pedigrees with microsatellite (SSR) markers**, which are particularly important since they are transferable among *Populus* species. These SSR markers may be used to link the maps and **form a consensus map of *Populus***.
- **Identifying quantitative trait loci (QTLs) and testing their robustness in contrasting genetic backgrounds**. QTLs for yield traits, for disease and pest resistance are being identified for the three sites (United Kingdom, France, and Italy). Then, the robustness of QTLs will be tested by examining genotype x environment interactions in detail.
- **Detecting candidate genes by a genomic approach and mapping them in *Populus*.** Transcript profiling approach as well as *Populus* microarray approach will be used to identify candidate genes which may co-locate to QTLs.
- **Studying linkage disequilibrium**. A collection of black poplar genotypes will be used to confirm the location of QTLs/candidate genes in mapping progeny.

A detailed description of the POPYOMICS project is available on the website http://www.soton.ac.uk/~popyomic/index.htm. The experimental work of this thesis was funded by the POPYOMICS project and developed within the frame of this research program.

P. nigra is a tree of social and economic interest and also of ecological importance as indicator species of riparian woodlands. It has a wide geographical distribution ranging from Central and Southern Europe to Central Asia and Northern Africa. *P. nigra* plays a central role in poplar breeding programs and has contributed to many successful inter-specific hybrids. During the last centuries, large areas of the natural habitat of *P. nigra* have been lost because of changed management of riverbanks, involving drainage, more intensive grazing, and more frequent tree felling. As a consequence, *P. nigra* is threatened with extinction (Arens *et al.*, 1998; Imbert and Lefèvre, 2003; Storme *et al.*, 2003). Currently, efforts are being made to restore the natural borders of rivers and their associated ecosystems. In softwood flooded plain forests, *P. nigra* is a keystone species because of its adaptation to water dynamics and sediment movement. The remaining *P. nigra* stands as well as the germplasm collected in gene banks could provide the

genotypes for establishing new populations of *P. nigra* (Storme *et al.*, 2003). Many studies on *P. nigra* are carried out with traditional methods combined with modern techniques. In fact, molecular genetics is a keystone to assess the remaining genetic diversity in order to preserve and to restore the *P. nigra* natural populations. For an effective protection and use of the remaining *P. nigra* genetic resources, a better knowledge of *P. nigra* genome is needed.

Mapping and sequencing of plant genomes would help to elucidate gene function, gene regulation and their expression (Mohan *et al.*, 1997). Linkage maps have been utilized for identifying chromosomal regions that contain genes controlling simple traits (controlled by a single gene) and quantitative traits using QTL analysis. The process of constructing linkage maps and conducting QTL analysis (to identify genomic regions associated with traits) is known as QTL mapping. DNA markers which are tightly linked to agronomically important genes may be used as molecular tools for marker-assisted selection (MAS) in plant breeding (Mohan *et al.*, 1997; Collard *et al.*, 2005).

In this context, the main objectives of this study are:

- **the construction of a genetic map of *P. nigra* from Italian natural populations** including bridge markers useful for comparing the other *Populus* maps and generating a *Populus* consensus map;
- **the construction of an accurate framework map,** which is efficient for the dissection of complex traits and for studying the genetic basis of QTLs;
- **the comparison of the *P. nigra* map obtained in this study to the *Populus* maps from other studies** to validate the accuracy of locus ordering and analyze the structure of *Populus* genome.

A *P. nigra* genetic map rich in SSRs leads to many perspectives. SSRs markers are ideal bridges for map comparison and direct links to genomic sequence. Moreover, genetic maps will be a powerful tool for exploring the function of candidate genes (Yin *et al.*, 2004b). Mapping genes and QTLs may open possibilities to develop strategies for MAS.

1 State of the art

1.1The Black poplar

1.1.1 Taxonomy and distribution of black poplar

1.1.1.1 Taxonomy

Black poplar, whose botanic name *is Populus nigra* L. (*P. nigra*), is an angiosperm of the genus *Populus* (2n = 38) in the *Aigeiros* section of the *Salicaceae* family. Compared to the willow (*Salix* spp.), which belongs also to the *Salicaceae* family, there are relatively few species (29) of poplars but they clearly fall into a number of groups which are morphologically and ecologically quite distinct. These groups are traditionally recognized as sections. With few exceptions, there is a large consensus in the literature on the characteristics and species composition of the sections. The major barriers to hybridization in the genus lie between sections. With the description of the section *Abaso* to accommodate *P. mexicana*, the number of sections has been brought to six (Table 1). Previously, *P. mexicana*, which only superficially resembles the *Aigeiros* cottonwoods, had been placed with them (Eckenwalder, 1996).

Table 1: Classification of *Populus* species in sections (Eckenwalder, 1996).
s.l. (*sensu lato*): indicates other species which are often recognized as distinct in the literature and which might be retained as subspecies. Some other species contain additional subspecies or varieties.

	Section					
	Abaso	***Turanga***	***Leucoides***	***Aigeiros***	***Tacamahaca***	***Populus***
Species	*P. mexicana*	*P. euphratica*	*P. lasiocarpa*	*P. nigra, s.l.*	*P. trichocarpa*	*P. alba*
		P. pruinosa	*P. glauca, s.l.*	*P. deltoides*	*P. laurifolia*	*P. tremula, s.l.*
		P. ilicifolia	*P. heterophylla*	*P. fremontii*	*P. ciliata*	*P. tremuloides*
					P. szechuanica	*P. adenopoda*
					P. yunnanensis	*P. gamblei*
					P. suaveolens, s.l.	*P. sieboldii*
					P. simonii, s.l.	*P. simoroa*
					P. balsamifera	*P. guzmanantlensis*
					P. angustifolia	*P. monticola*
						P. grandidentata

The status of *P. nigra* and the relationships between the sections *Aigeiros* (cottonwoods) and *Tacamahaca* (balsam poplars) are subjected to questions. These two sections are the only ones known to be freely intercrossable. Although vegetatively and ecologically readily distinguishable, there are no clear differences in flowers and inflorescences between cottonwoods and balsam poplars. They could be accommodated in a single section where they would have separate

subsections anyway (as the aspens and white poplars are in the section *Populus* i.e. subsections *Tremulae* and *Albidae*). The present evidence, including phylogenetic analyses, seems to favor keeping them apart, but resolution of this issue affects the placement of *P. nigra*. Actually, *P. nigra*, the type species of the section *Aigeiros*, is not clearly more similar to the Northern American cottonwoods placed with it in section *Aigeiros* than some species of the balsam poplars in the section *Tacamahaca*. It also has peculiar crossability relationships, successful only in one direction, with both Northern American cottonwoods and balsam poplars (Eckenwalder, 1996).
To solve these dilemmas for the classification of poplar at sections and species level new researches are needed and these could lead to some evolution of the *Populus* spp. taxonomy.

1.1.1.2 Distribution

Populus is tropical in origin and the greatest diversity of the genus still lies far South of the boreal region. *P. nigra* has a large distribution area throughout Europe and is also found in Northern Africa and Central and West Asia. The distribution area extends from the Mediterranean in the South to approximately 64° latitude in the North and from the British Isles in the West to Kazakhstan and China in the East. The distribution area also includes the Caucasus and large parts of the Middle East (Fig. 1) (Vanden Broeck, 2003).

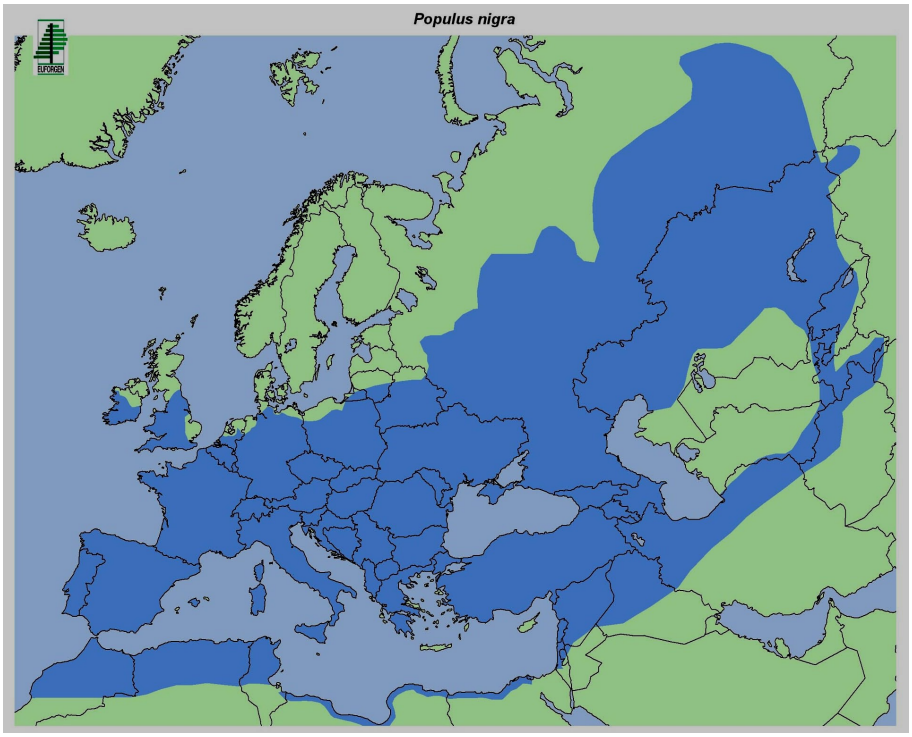


Fig. 1: Distribution range of European *P. nigra*
The blue area represents the distribution of European *P. nigra* from Vanden Broeck (2003)

In Italy *P. nigra* is present from the sea level up to 1000-1200 m in the Alps and up to 1500-1600 m in the Apennines. *P. nigra* is an azonal species not linked to particular climatic area but being a riparian species, it is linked to the soil moisture.

1.1.2 Biology and ecology

P. nigra is a typical tree species of the alluvial forests of many European and Siberian rivers. Being heliophilous, *P. nigra* usually forms local populations by colonizing open areas on alluvial soils through seeds, cuttings or root fragments (Fig. 2).

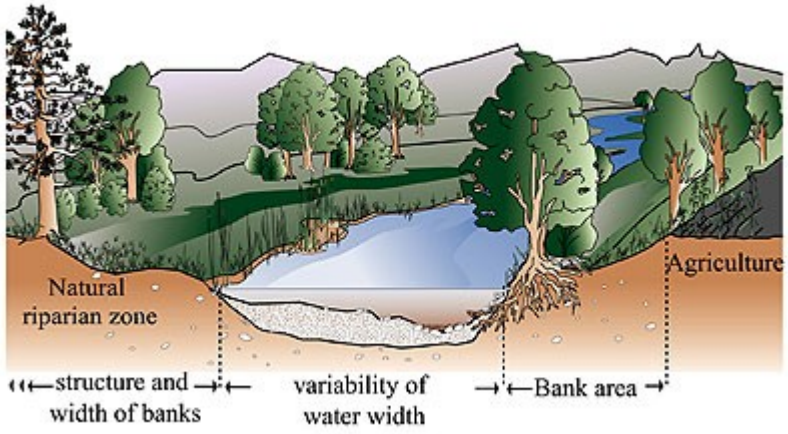


Fig. 2: Illustration of the alluvial area structure.[online web]
From URL: http://www.modul-stufen-konzept.ch/seiten-e/oekomor-e.htm
Example of a river bank area with the representation of the natural riparian zones which are the *P. nigra* habitat.

It is characterized by a great diversity of population type, from isolated trees to huge pure or mixed stands. *P. nigra* has a rapid growth. Individual trees may live over 400 years.
This tree has a linear *habitus*, its height can reach 20-25 m with a diameter up to 1 m. The trunk, very right, is easily ramified and the canopy is wide and dense with large and dispersed branches (Fig. 3).

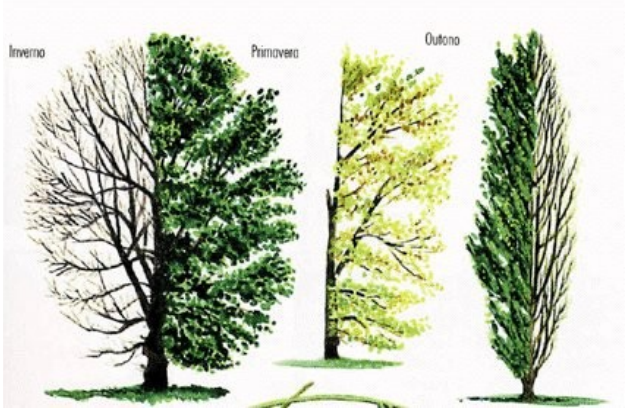


Fig. 3: Picture of *Populus nigra* L and *Populus nigra* var. *italica* at different seasons. [online: web].
From URL:
http://centros.edu.xunta.es/iesaslagoas/slorenf/arb7.htm
P. nigra is represented at winter, spring and autumn season and on the right *italica* variety at spring and winter season.

The bark on young branches and small trunks is nearly smooth and grayish-green in color. Bark on larger trunks begins to split forming rough, fissured, dark gray patterns (Fig. 4).

Fig. 4: *P. nigra* bark at old age. [online: web]
From URL: http://www.lhi.org.uk/images/thumb_NE00107_Black_Poplar.jpg/

The size, shape, and character of *P. nigra* leaves are rather variable within a single tree (Fig. 5). The blade of juvenile leaves is usually flat but also sinuate along margins. The blade margins are shallowly often irregularly crenate. Adult leaves show a blade which is more or less similar to a rhombic form terminated by a narrow point. The upper surface of the blade is dark green shiny, dully light green underneath, glabrous, leathery and flat. The blade margin is shallowly crenate to dentate. At the end of growing season leaves are often infested by rust fungi while healthy leaves get yellow before leaf fall (Mottl and Uradnieek, 2003).

Fig. 5: Photo of *P. nigra* leaves [online: web]
From URL: http://web.mit.edu/cfox/www/flowers/2003-05-31/2158_Lg.jpg.4.html

P. nigra is a dioecious species with trees that are either male or female. They reach the reproductive stage when they are 10-15 year old, but will not begin to produce remarkable quantities of seeds until they are more than 20 year old (Braatne *et al.*, 1996). Approximately 1-2 weeks prior to leaf initiation in the early spring (March-April), during the flood peak period along the rivers in temperate Europe, male and female trees produce flowers clustered in pendulous catkins. The catkins (<10 cm long) tend to be borne in the upper tree crown and are reddish-purple in appearance on males and slightly larger and green in females (Fig. 6). Wind-dispersed pollen landing on receptive stigma will fertilize ovule within 24 hours of arrival and the subsequent ripening and seed maturation process lasts 4-6 weeks. During this period the female catkins lengthen and swelling green fruit capsules appear along their length. Approximately 20-50 fruit capsules will ripen on each catkin producing up to 225 seeds per catkin (about 4-5 seeds per capsule). The period of seed release in *P. nigra* is strategically timed to coincide with the abatement of floodwaters in spring when, for a short period, ideal conditions for seed germination and seedling establishment are present. Vast quantities of seeds will also be dispersed by the river extending the period of seed dispersal by 2-3 weeks.

Fig. 6: Picture of *P. nigra* leaves, catkins, flowers, and seeds. [online: web]
From URL: http://caliban.mpiz-koeln.mpg.de/~stueber/thome/band2/tafel_020.html. (Otto, 1885)

In common with many colonizer species, *P. nigra* depends on the wind for pollination and is capable of vegetative regeneration. The generation of ramets is not spontaneous in this species. Asexual reproduction is promoted only by flood disturbances when through extended periods of submergence and/or mechanical damage to parent plants, dormant primordial in roots and shoots are stimulated to produce new shoots and roots (Barsoum, 2001).

P. nigra is affected at all stages of its life cycle by hydrological conditions and is reliant on them for regeneration. Seeds are disseminated through wind and water, have a short viability and need very specific soil-water conditions for germination. Successful regeneration occurs in years when soil moisture remains high enough for roots to grow down at the same rate as water recedes from the saturated waterfront, but not so high that anoxic conditions prevail. It follows that in many years, successful regeneration does not occur, and that in naturally occurring stands a strong age structure frequently exists, reflecting the history of flooding. Regeneration is generally poor within old established stands; the riparian forest naturally evolves towards hardwood formations (Vanden Broeck, 2003).

1.1.3 Importance and uses

1.1.3.1 Economic importance of Populus nigra

P. nigra is a tree of social and economic interest. It is important as a crossing parent in the production of healthy and fast-growing *Populus x canadensis* Moench (*P. deltoides x P. nigra*) hybrids which are grown widely in Europe and in other parts of the world (Vanden Broeck, 2003). It is predominantly used as a parent pool in breeding programs in many parts of the world: 63% of the poplar cultivars used in forest plantations descend from *P. nigra* either as a pure species or from inter-specific hybrids. In some Eastern European countries up to 50% of the production of poplar wood comes from *P. nigra.* In Europe a surface of about one million ha is estimated with an annual production of poplar wood in excess of ten million cubic metres which goes for a variety of uses.

- The wood of hybrid poplars is mainly used as raw material for the industries: furniture, packaging, particleboard, plywood, and matches. It is particularly favored for the manufacture of fruit boxes because the wood has no fragrance. On the other hand, *P. nigra* wood is used as round wood for rural construction and for the daily needs of rural people. In Turkey, there are approximately 130000 ha of poplar plantations, of which 70000 ha are hybrid poplar and 60000 ha consist of various clones of *P. nigra*. This species provides about 57% of Turkey's annual poplar wood production (3.5 million cubic

metres) and more than 80% of *P. nigra* wood (1.75 million cubic metres) is used for domestic needs (Vanden Broeck, 2003; Toplu, 2005).

- *P. nigra* is used as a pure species for soil protection and afforestation in polluted industrial zones (Vanden Broeck, 2003).

Since the Kyoto Conference (1997), there has been an increasing interest about renewable energy sources and possible alternatives to fossil fuels that could contribute to a significant reduction in greenhouse gas emission and enhance the overall sustainability of modern society. In order to reduce the carbon dioxide emission levels, many fast growing hardwoods such as poplar species are tested for the biomass production. *P. nigra* and its descend inter-specific hybrid are among the most promising (Benetka *et al.*, 2002; Laureysens *et al.*, 2005). *P. nigra* can be hybridized with *P. deltoides*, *P. trichocarpa* and other exotic *Populus* species (Fig. 7) providing adaptability to various soil and climate conditions, rooting ability, high resistance to bacterial canker caused by *Xanthomonas populi*, fair resistance to *Marssonina brunnea* and to poplar mosaic virus (Vanden Broeck, 2003).

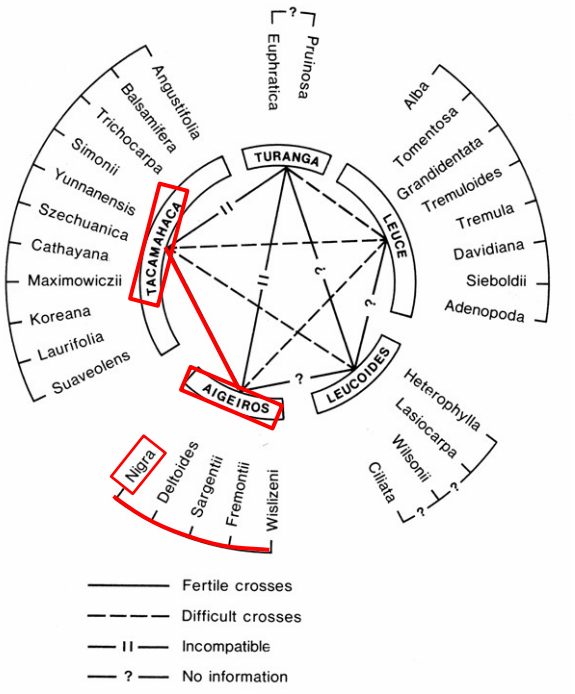


Fig. 7: Crossability of *Populus* species from Zsuffa (1975)
The red outline indicates the fertile crosses of *P. nigra*.

For the commercial poplar cultivation the *P. x canadensis* clones proved an extraordinary success. They were introduced into many European, Asian, North and South American countries. Until the 1960s, all poplar species used for commercial cultivation were either new hybrids or varieties of

combinations of *P. deltoides* and *P. nigra*. A number of pests and diseases threaten the present clones and the creation of healthy, fast growing clones is a constant process which depends on *P. nigra* as one of the crossing parents. It specifically contributes resistance to bacterial canker (*Xanthomonas populi*) in the *P. x canadensis* hybrids. However, only the cross with *P. nigra* as father provides an hybrid offspring. In cross-breeding trials with the reciprocal combination *P. nigra x P. deltoides*, the embryos died off in an early stage (Hofmann, 2001). Therefore, *P. nigra* is used in breeding programs in many parts of the world. In Italy, the breeding program is an example, at species level, of how a subdivided breeding population gives flexibility in maintaining genetic variation (Bisoffi and Gullberg, 1996) (Fig. 8).

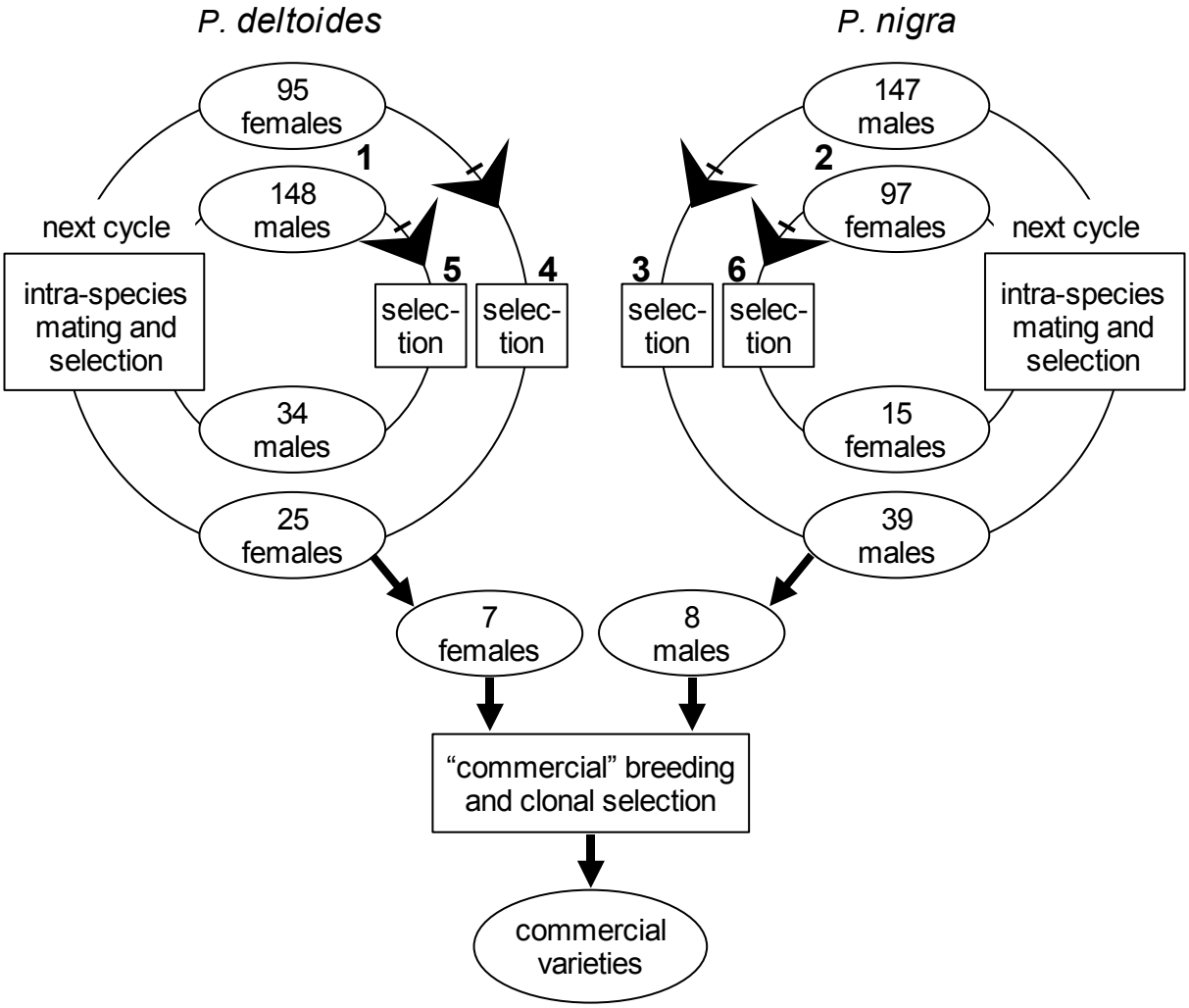


Fig. 8: The Italian breeding program for poplar
From Bisoffi and Gullberg (1996)
(1) 1958-1981: collection, provenance and progeny testing, and scoring for growth, phenology, and *Melampsora* resistance that result in 300 *P. deltoides* selected clones; (2) 1982-1984: collection of 300 *P. nigra* clones covering the whole of Italy; (3) 1987: common tester progeny trial of *P. nigra* males: *6 P. deltoides* x 147 *P. nigra* males; (4) 1988: polycross test of *P. deltoides* female: 95 *P. deltoides* x *P. nigra* pollen mix; (5) 1989-1991: common tester progeny trial of *P. deltoides* males: 6 *P. deltoides* females x 148 *P. deltoides* males. (6) 1990: polycross test *ofP. nigra* females: 97 *P. nigra* females x *P. nigra* pollen mix.

Recently, in Turkey, to improve the quantity and the quality of wood harvest from poplar plantations, classical breeding programs with *P. deltoides*, native *P. nigra,* and *P. x canadensis* are carried out for commercial release of the most productive hybrid clones. Useful traits of parents, such as fast growth, desired wood quality, and resistance to frost, can be combined in a hybrid, and some clones are selected from intra- or inter-specific crossings. Domestic *P. nigra* provenances are included in the breeding program because of their adaptability to continental conditions (Toplu, 2005).

1.1.3.2 Environmental interest of *Populus nigra*

P. nigra is also of ecological importance. It is a pioneer species of riparian ecosystems and contributes to the natural control of flooding and water quality. The riparian ecosystems are very dynamic environment and are characterized by a high level of diversity of the fauna and flora.

However, populations of *P. nigra* face severe threats. It is one of the most threatened tree species in Europe. Three main factors have been recognized (De Vries and Turok, 2001):

- **the alteration of riparian ecosystems throughout the species' distribution area**. Agriculture and urbanization of floodplain areas have displaced native poplar stands while other human activities, including regulation of floods through hydraulic engineering practices, have favored later successional hardwood forests over poplar stands in the remaining wild areas. Although the species may demonstrate locally highly successful regeneration, some regions of Europe have witnessed significant reductions in populations or the complete disappearance of *P. nigra*.
- **the autochthonous *P. nigra* resources have been overexploited** for the use of wood and faster growing hybrid poplars have been planted to replace them.
- **there is a potential threat of introgression** from cultivated clones, and other poplar species. Very few clones are extensively cultivated and these contribute to a large extent to the pollen and seed pools. The risk concerns not only introduced hybrids, but also pure *P. nigra* varieties with a wide distribution such as the male clone *P. nigra* cv *italica*, distributed all over continental Europe.

Currently, there is a great interest in Europe to restore the bottomlands. Not only for the natural control of flooding but also because the bottomlands will serve as ecological corridors through which larger natural areas are connected. Strategies for restoration and conservation of the riparian ecosystem should be based on firm scientific footing.

Two main projects, "EUFORGEN *P. nigra* network" and "EUROPOP", are working on the evaluation of the existing biodiversity of *P. nigra* natural populations. The objectives of these projects are the conservation and the restoration of the remaining natural riparian ecosystems. Traditional methods will be combined with modern techniques and the high level of standardization will ensure a synthesis of the genetic diversity of *P. nigra* in Europe. This information is essential for the evaluation of the existing biodiversity in river populations so guidelines and strategies for *in situ* and *ex situ* conservation can be provided. Static *ex situ* conservation is a widely applied strategy for short-term conservation to preserve genotypes in collections or gene banks. Many *ex situ* collections have been already carried out such as in Turkey since 1962 by the Poplar and Fast-Growing Forest Trees Research Institute in Izmit (Toplu, 2005). In Italy, in 1981 the Istituto di Sperimentazione per la Pioppicultura and the Centro di Sperimentazione Agricola e Forestale started a joint program on the identification and collection of spontaneous *P. nigra* individuals (Bisoffi *et al.*, 1987). Recently, nine European gene bank collections were analyzed with molecular markers in order to estimate the number of hybrids, the extent of clone duplications and the genetic diversity within and between the gene bank collections. This work allowed evaluating the quality of the gene bank and the existing genetic diversity of *P. nigra* in these nine European collections (Storme *et al.*, 2003).

However, when the objective is the long-term gene conservation and maximization of the adaptive potential of a species, a dynamic *in situ* conservation is preferable. This can be achieved through *in situ* conservation of native stands (including their restoration), long-term breeding programs or both. Successful *in situ* conservation of *P. nigra* in Europe depends on the location and protection of its natural habitats. A preliminary assessment of the genetic diversity among adult trees in the candidate populations is recommended to conserve a high amount of diversity and a low number of clonal duplicates (Vanden Broeck, 2003). In this purpose the genetic diversity of *P. nigra* population was evaluated, for example, along Rhine river (Arens *et al.*, 1998), in the Upper Severn area of the UK (Winfield *et al.*, 1998), along the Drome river in France (Imbert and Lefèvre, 2003). These studied showed a higher conservation of the genetic variability in France than in UK and in Netherlands where in the populations analyzed there are few different genotypes and many clonal duplicates.

Conclusion: *P. nigra* is a tree of social and economic interest. It is also of ecological importance as an indicator species of riparian woodlands. It dominates the early successional stage of floodplain woodlands in many temperate areas. Its current rarity in some cases is due to the loss of its natural habitat for the anthropic activities as the drainage of rivers and management of river banks. Another threat to the gene pool of *P. nigra* might come from the potential inter-specific hybridization. Fortunately, there is now a great interest to preserve and restore *P. nigra* natural populations and its natural habitat. In order to achieve these objectives many studies on *P. nigra* are carried out with traditional methods combined with modern techniques where the field of molecular genetic is a keystone. **A deep knowledge of *P. nigra* genome is therefore an essential starting point to develop molecular markers to study the genetic diversity. In this context, a genetic map provides important tools for both the assessment of *P. nigra* diversity and breeding programs.**

1.2 Genetic mapping

1.2.1 Introduction

Just like a state map allows finding a specific place, genetic maps allow scientists to search for a specific gene somewhere within a vast genome of plants or animals. To continue the comparison, state maps have cities and towns that serve as landmarks and genetic maps have landmarks known as genetic markers, or "markers" for short. Therefore the construction of the genetic map consists in placing beacons or markers on the genome (Fig. 9). The figure 9 is a good illustration of the markers which are beacons to represent the genome, here the human chromosome 11.

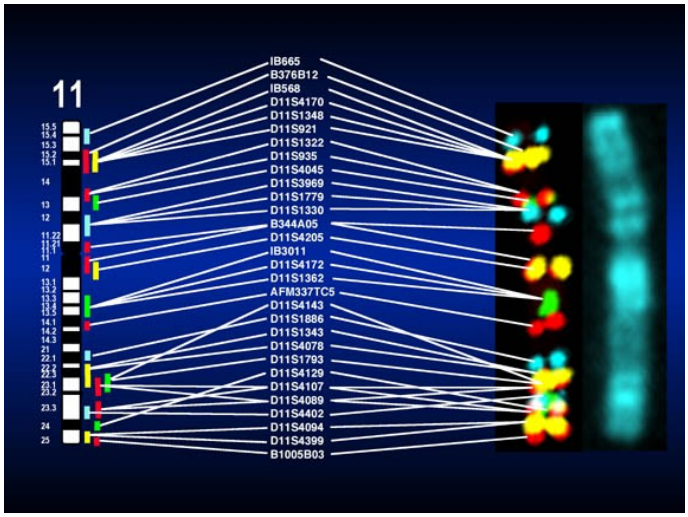


Fig. 9: The Visible Genetic Map of Human Chromosome 11. [online: web]
From URL: http://www.csmc.edu/csri/korenberg/chroma11.html. (Korenberg, J.R)
The markers were labeled with fluorescent dyes and hybridized simultaneously on the Human chromosome 11.

The markers on the genetic map allow then to access to the genes. The map is an important tool to study the genome structure and to detect, localize, and identify genes. In particular, the decomposition of a complex trait, such as yield for plants, in its discrete components (Quantitative Trait Loci or QTL) can be realized. The linkage among markers and QTLs of agronomic interest permits to use these markers in breeding programs. The main objective of QTL mapping is to find the genes responsible for the traits however this is a complex and long work. Therefore, finding markers linked to traits, and indirectly to genes, is faster in a first step.

In the last decade, genetic mapping has been particularly developed with the technological advancement of molecular biology. The PCR (Polymerase Chain Reaction), which can be considered the basic technique of molecular biology, has permitted to set up many molecular markers

(Fig. 10). The sequencers have also become powerful tools for genetic mapping because they allow the development of new markers but also gaining time, precision, and working at large scale (Fig. 11). Therefore, the availability of this large number and kind of molecular markers has allowed the construction of saturated genetic maps in many plant species. For example: the highly saturated map of tomato (Haanstra *et al.*, 1999) including 67 RFLP (Restriction Fragment Length Polymorphism) and 1078 AFLP (Amplified Fragment Length Polymorphism) markers, spanning 1482 cM (centi Morgan); the sorghum map of 1713 cM encompassing 2926 loci constituted of AFLP, RFLP and SSR (Simple Sequence Repeat) (Menz *et al.*, 2002); the high resolution rice map containing 1383 DNA markers distributed along 1575 cM (Kurata *et al.*, 1994). The genomes of more marginal species such as forest trees have been also mapped: Norway spruce (*Picea abies*) (Acheré *et al.*, 2004), loblolly pine (*Pinus taeda*) (Remington *et al.*, 1999), eucalyptus (Grattapaglia and Sederoff, 1994), European chestnut (Casasoli *et al.*, 2001), and different poplar species and hybrids (Cervera *et al.*, 2001).

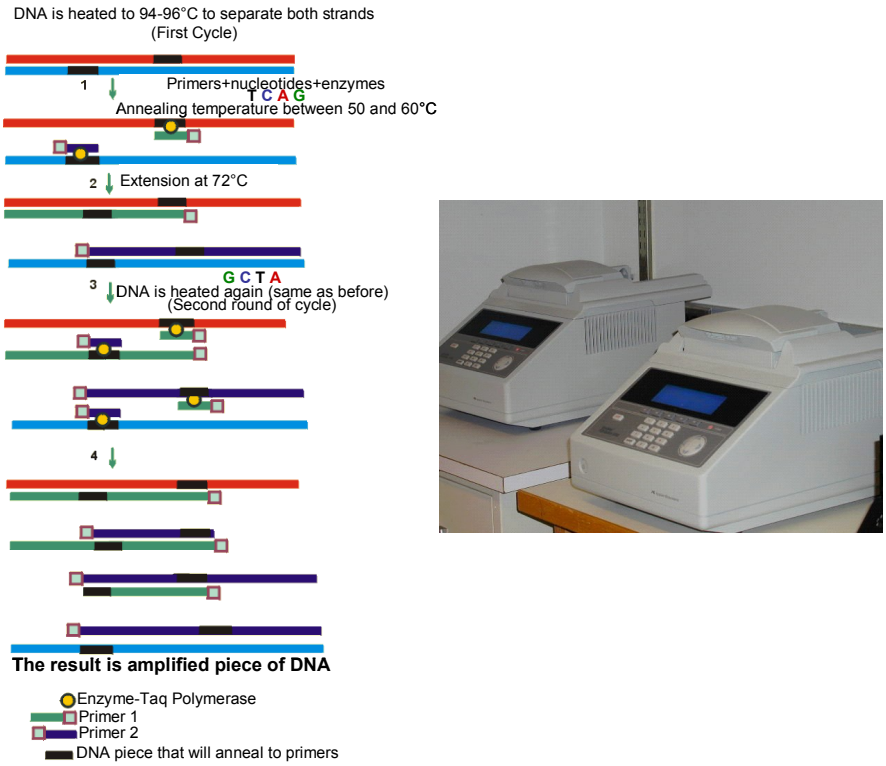


Fig. 10: Scheme of PCR reaction. [online: web]
From URL: http://www.science2discover.com/images/PCR.gif and http://134.174.23.167/zonrhmapper/images/PCR.JPG.
A specific region of DNA can be amplified thanks to primers (small pieces of DNA which bind to a complementary sequence of the DNA), nucleotides (small molecules which compose the DNA) and, the Taq Polymerase enzyme (which has the property to add the nucleotides at the 3' end of the primers). The PCR reaction consists in 3 consecutive steps. The double strand DNA has the property to separate its strands at high temperature (94°C), it is the denaturation step. After the DNA denaturation, the temperature is lowered (50 – 60°C) to permit the binding of primers, it is the annealing step. Then the temperature is increased to reach the optimal condition of work for the Taq (72°C), it is the elongation step. These 3 main steps are repeated and a large quantity of DNA copies is obtained (exponential amplification). The photo on the right represents an example of thermocyclers necessary to perform the PCR reactions.

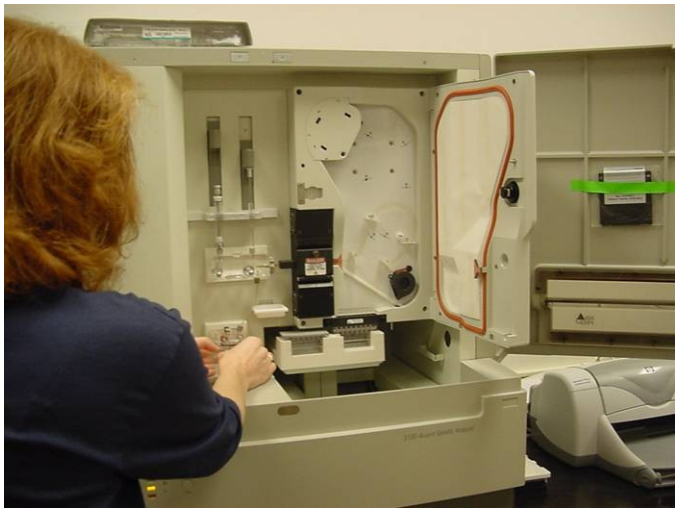
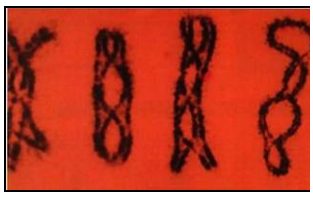


Fig. 11: Photo of a sequencer. [online: web]
From URL: http://newton.bhsu.edu/biology/images/DNAsquencersetup.jpg
This is an example of a capillary sequencer which allows to sequence the nucleic acids but also to separate the DNA fragments obtained by PCR. The sequencer generates new molecular markers and separates PCR products with high throughput, allowing a rapid analysis of molecular markers on a large number of individuals.

1.2.2 Principle of genetic mapping

The construction of a genetic linkage map is based on the segregation study of simple genetic traits (morphological, biochemical, and molecular), the markers, in a progeny. The genetic markers must have a Mendelian segregation, be polymorphic, and easy to follow in each individual. When chromosomes pair in the first division of meiosis, crossovers occur between two non-sister chromatids generating an exchange of genetic material between the maternal and paternal chromosomes (crossing-over). If there are genetic markers (alleles) on the chromosomes, it is possible to observe new combinations of alleles at different loci as a result of these crossing-overs (Fig. 12).



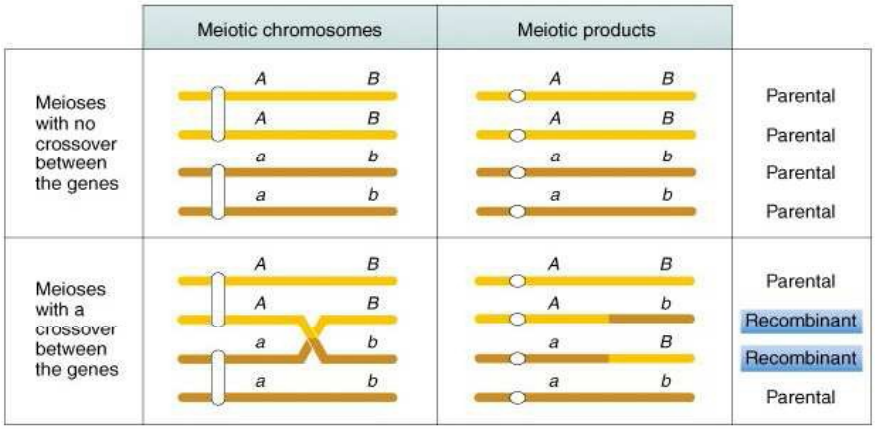


Fig. 12: Scheme of gamete formation with and without crossing-over. [online: web]
From URL: http://www.saintemarie-caen.asso.fr/svt/Term/TPS9_fichiers/image008.jpg (photo) and http://fig.cox.miami.edu/Faculty/Dana/F05_08.JPG (table)
The table represents the possible gametes obtained after the meiosis. The gametes descending from a crossover between loci are recombinant, the other ones are the parental type (non recombinant). On the left the photo illustrates chromosome pairs with crossing-overs.

The recombination rate between 2 loci is proportional to their distance: the greater is the distance between loci, the higher is the probability of recombination. In the same way two loci closer have less probability to recombine. Therefore the distance separating loci can be deduced from the recombination rate. This parameter is estimated for each pair of loci by analyzing allele distribution in the gametes or in the progenies. Then loci are ordered one each other to construct the genetic map of parents. The precision of the map depends on the number of meiosis analyzed (Fig. 13).

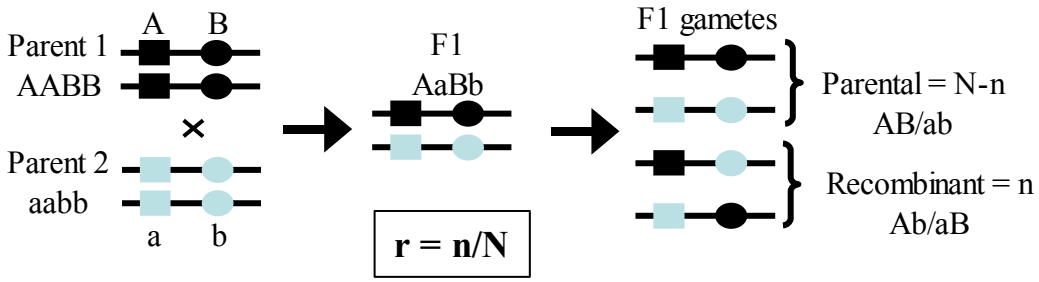


Fig. 13: Example of cross used for mapping population and genotype of gametes formed during F_1 individual meiosis.
A and B represent 2 loci with 2 alleles, A/a and B/b. The distance between these 2 loci can be calculated by the recombination rate r = n/N, where N is the total number of gametes and n is the number of recombinant gametes. If the frequencies of the 4 genotypes (AB, ab, Ab, aB) is the same, r = 0.5, that means the 2 loci segregate independently and therefore they are unlinked.

The map construction needs three main elements:

- a progeny within which it is possible to follow the segregation of genetic markers;
- genetic markers to characterize individuals of the progeny;
- statistical analysis of data segregation.

1.2.2.1 Pedigrees used for genetic mapping in plants

The first step of the genetic map construction is the choice of parents of the cross and the type of progenies to analyze. The parents are chosen in order to have a maximum of detectable segregation in the progeny. The main pedigrees used for genetic mapping in plants are the following:

- **F_2 progeny** descend from a self fertilization of F_1 hybrids. In this case, there are two efficient meiosis, the female and male gametes can be recombinant. Two very divergent lines were often chosen to obtain the F_1: maize (Sibov *et al.*, 2003), cotton, (Rong *et al.*, 2005).
- **Recombinant Inbred Lines (RIL)** stem from a F_2 family after five or six self fertilization cycles. At each generation an individual of each line is self fertilized to give

rise to the next generation. Individuals of the same line are genetically identical. The last generation lines are highly homozygous and each line presents particular allelic combinations. The segregation are observed among these lines in these species: wheat (Ellis *et al.*, 2005), sunflower (Al Chaarani *et al.*, 2005).

- **Backcross (BC)** results from the crossing between a F_1 individual and one of its parents. If the parent is genetically fixed (homozygote for all these loci), the meiosis do not carry segregation. On the other hand, the F_1 individuals form four types of gametes (Fig. 13) responsible of the segregation observed like for example in cacao (Crouzillat *et al.*, 1996) and pepper (Rao *et al.*, 2003).
- **Doubled haploids** come from the regeneration of plants from microspores or macrospores. This is possible for some species such as *Poaceae*, *Solanaceae*, etc. The regenerated plants are diploids because they have undergone a chromosomal doubling induced or spontaneous. At the genetic level each plant corresponds to a meiosis product therefore, a genetic map can be constructed with a progeny of these individuals. This is equivalent to the recombinant inbred lines but the plants are produced in one generation, for example in sweet pepper (Sugita *et al.*, 2005) and in cotton (Song *et al.*, 2005).
- **Endosperm (or megagametophyte)** of conifer, the nutritive tissue of the seed, is haploid with the same genetic constitution of the female's gametes. The megagametophytes of an individual form a mapping population as reported for pine, (Remington *et al.*, 1999), Norway spruce (Paglia *et al.*, 1998).
- **F_1 family hybrids** (outbred crosses) are used when the other mapping population, previously viewed, can not be obtained. Two cases are concerned: i) auto-incompatible species or for which it is impossible to have pure lines, such as diploid clones of potatoes, ii) perennial species with long life cycle such as trees, where the constitution of mapping population could take long time. For the F_1 family hybrids, the segregation is observed at heterozygous loci of each parent taken separately. As the trees have particularly high level of heterozygosity, the F_1 progeny can be used for map construction like in poplar (Zhang *et al.*, 2003) and in European beech (Scalfi *et al.*, 2004).

Therefore the choice of a pedigree for genetic mapping depends on the species characteristics (reproduction biology, time of generation, cost) but also on the kinds of markers used.

1.2.2.2 Markers used for genetic mapping

Genetic markers, in general, can be classified as morphological markers, biochemical markers (isozymes, proteins) and molecular markers (at DNA level). The ideal genetic marker should be:

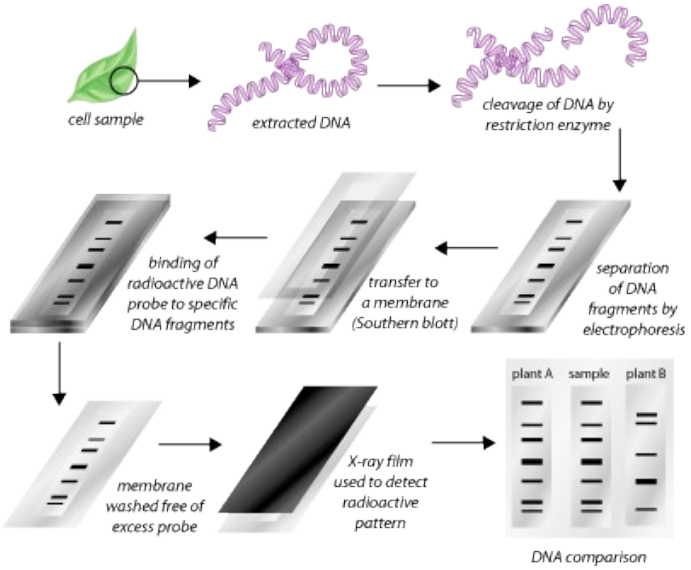
- **polymorphic**;
- **multi-allelic**;
- **co-dominant**: the heterozygote presents simultaneously the characters of the two homozygote parents; so it can be differentiated from each parental homozygote;
- **no epistatic**: there is no interaction between loci. The genotype can be read from the phenotype independently of the genotype of the other loci;
- **neutral**: the allelic substitution has not phenotypic effects, so there is no selective effect;
- **insensitive to environment**: the genotype can be inferred from the phenotype whatever the environment.

Morphological markers badly respond to these criteria. They are poorly polymorphic, generally dominants, often interfere with other characters and can be influenced by the environment. On the other hand, biochemical and molecular markers have, mostly, all the required qualities.

Two main techniques use the protein markers: the isozymes and the total proteins. Isozymes are based on the staining of proteins with identical function, but different electrophoretic mobility. The change of the polypeptide sequence induces physical properties alteration and, consequently, different migration, corresponding to the different alleles of the enzyme. Their expression are co-dominant and multi-allelic. These markers are used in many genetic linkage maps (Lespinasse *et al.*, 2000; Casasoli *et al.*, 2001) but the number available and the limited allelic variation of isozymes do not allow a sufficient coverage of the genome. Total proteins allow analyzing more loci. The allelic variation of genes coding for protein result in variation of mass and/or iso-electric point which can be revealed by electrophoresis. Because of technical and interpretation (alleles definition) difficulties, these protein markers are poorly used. In bibliography there are some examples on the use of protein markers such as in *Pinus* (Gerber *et al.*, 1993). Morphological and protein markers have been the first genetic markers used for the linkage analysis but linkage maps were limited in size until the advent of molecular markers.

Indeed the development of molecular markers, revealing the polymorphism at DNA level, permits to obtain many markers more easily. Also for these markers there was a technical evolution. First the RFLP (Restriction Fragment Length Polymorphism) were commonly used in genetic mapping (Helentjaris *et al.*, 1986; Helentjaris, 1987; Gebhardt *et al.*, 1989). These markers are

based on the polymorphism of the fragment length produced by the DNA digestion with restriction enzymes. After separation by gel electrophoresis they are detected by hybridization with a labeled probe (Fig. 14). These markers are co-dominant and multi-allelic but the method is time consuming and difficult.



After extraction, the DNA digested by restriction enzymes is separated on a gel. Then the fragments are transferred on a membrane (Southern blot) and finally hybridized with a labeled probe. DNA fragments are revealed and analyzed.

After RFLPs, molecular markers based on PCR were developed and revolutionized the world of mapping. The RAPD (Random Amplified Polymorphic DNA) are one of the first PCR-based molecular markers. This technique substantially reduce time, labor, and cost required for molecular mapping. RAPDs involve the use of a single DNA primer for simultaneous amplification of multiple random sequences (Fig. 15).

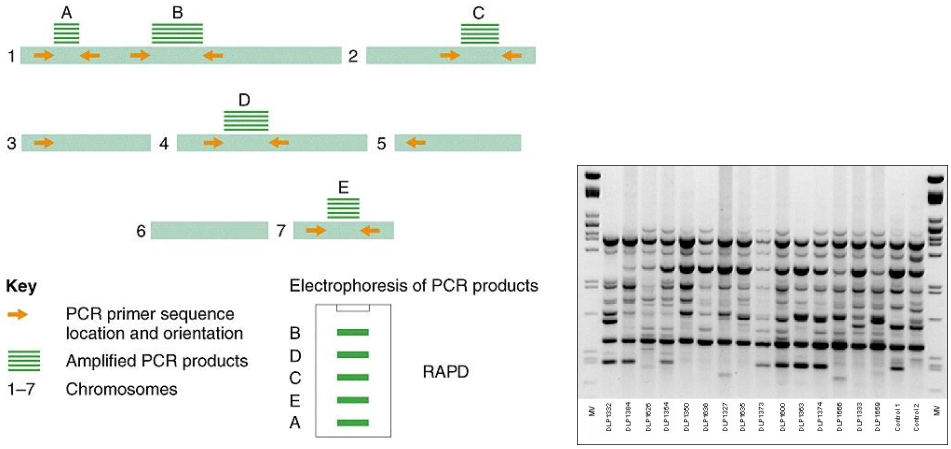


Fig. 15: Description of RAPD technique. [online web]
From URL: http://www.usask.ca/agriculture/plantsci/classes/plsc416/projects_2002/pawlin/resources/rapds.html (Scheme) and http://www.cipotato.org/market/PgmRprts/pr95-96/program2/prog25.htm (photo)
The RAPD analysis consists in amplifying unknown target sequences. Short primers (10 base pairs) with an arbitrary sequence are designed and used for PCR amplification of genomic DNA. The products of amplification are then separated by electrophoresis. The photo represents an example of results obtained after gel electrophoresis and staining.

The RAPD markers are dominants (presence or absence of DNA fragment) which is not ideal for genetic mapping, but they are so simple and quick that they were widely used in diverse species: rice (Kurata *et al.*, 1994), eucalyptus (Grattapaglia and Sederoff, 1994), white spruce (Gosselin *et al.*, 2002), douglas fir (Jermstad *et al.*, 1998), chestnut (Casasoli *et al.*, 2001), poplar (Bradshaw *et al.*, 1994; Yin *et al.*, 2001). A major disadvantage of RAPD technology is the inconsistent reproducibility of the results (Jones *et al.*, 1997). Another technique, AFLP (Amplified Fragment Length Polymorphism) was developed by Vos *et al.* (1995) and provides greatly enhanced performance in terms of reproducibility and efficiency. Now the AFLPs are more used than RAPDs even if they also provide dominant markers (Fig. 16). AFLPs allow the construction of dense maps even for large genomes such as trees.

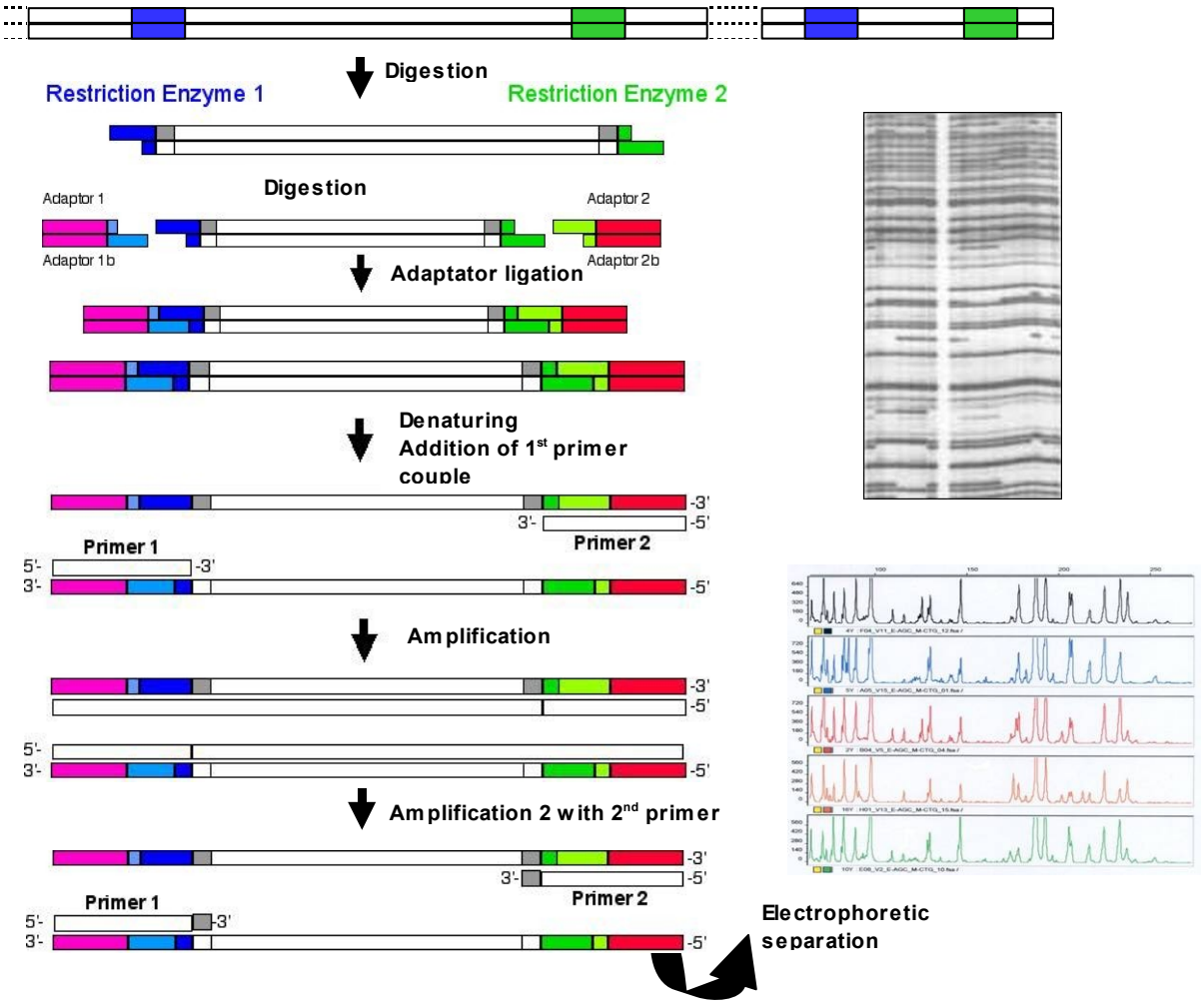
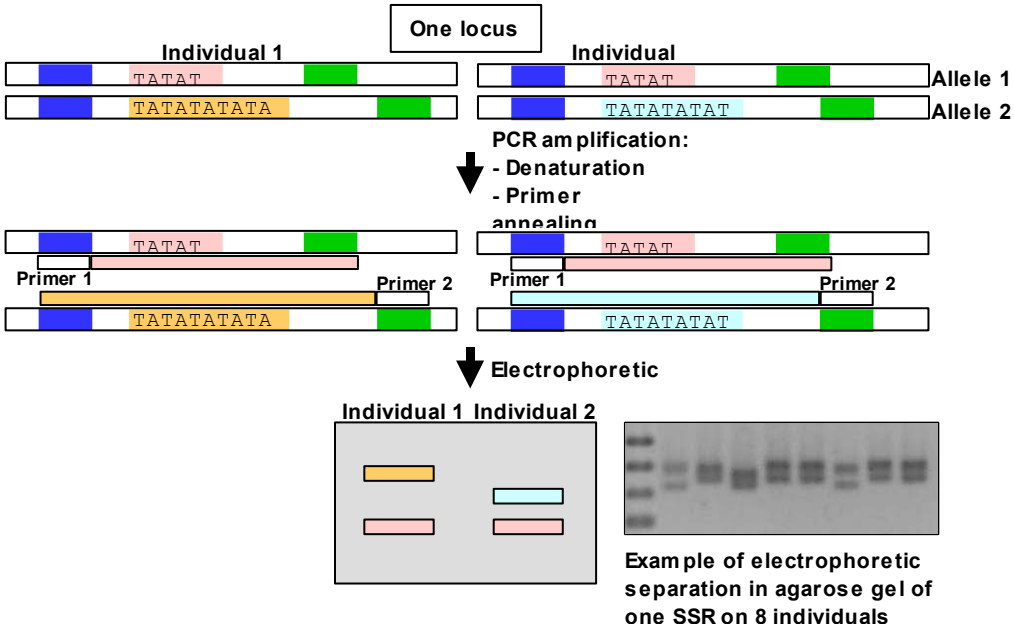


Fig. 16: Description of AFLP technique
From URL: http://www.scri.sari.ac.uk/SCRI/Web/MultimediaFiles/AFLP.JPEG (photo gel) and http://www.chelab.it/Images/News/upl/OLIO-AFLP-Taggiasca_small.jpg (photo electopherogram)
The procedure of AFLP is divided into three main steps: i) digestion of total genomic DNA with restriction enzymes and ligation of restriction half-site specific adaptors to all restriction fragments, ii) selective amplification of some PCR fragments with two primers that have corresponding adaptor and restriction site specific sequences, iii) electrophoretic separation followed by visualization of the band pattern.

The more recent published maps were constructed with AFLP and SSR (Simple Sequence Repeat) (Yin *et al.*, 2004b; Kenis and Keulemans, 2005; Tsuro *et al.*, 2005). The SSR, or microsatellite markers, are co-dominant markers and are defined by a variable number of repetitions of a very small number of nucleotides within a sequence. As these regions are very variable, the number of repeats for a given microsatellite may differ between individuals. SSRs can show a large number of different alleles for one locus. They are abundant and there is an even distribution across the genome. Another important advantage of these markers is the potential transferability among the species. Actually, they are obtained thanks to primers corresponding to the flanking regions of the microsatellite which are unique for each locus (Fig. 17).



A couple of primers are used to amplify a specific SSR with the PCR reaction. The high variability of the microsatellite regions allows the detection of many allele at the locus.

The development of SSR markers requires a high initial investment because the knowledge of the DNA sequence is necessary. However, the important progress of the sequencing techniques and the cost reduction have permitted the availability of SSR databases for many species. Over the past few years, the sequencing of many entire genomes allowed a new kind of marker, the SNPs (Single Nucleotide Polymorphism). The SNPs consist in single base changes or small insertions and deletions (indels) between homologous DNA fragments and they are present in all parts of the genome, coding or non coding regions (Fig. 18). In principle, SNPs could be bi-, tri- or tetra-allelic polymorphisms. However, tri-allelic and tetra-allelic SNPs are rare almost to the point of non-existence and for this reason SNPs are sometimes simply referred to bi-allelic co-

dominant markers. This is somewhat misleading because SNPs are only a subset of all possible bi-allelic polymorphisms (e.g., multiple base variations) (Brookes, 1999).

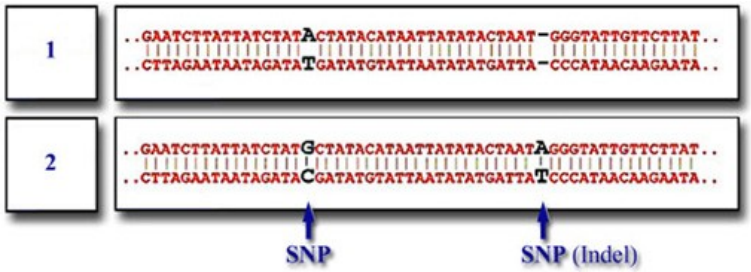


Fig. 18: Description of SNP [online web]
From URL: http://bldg6.arsusda.gov/~pooley/soy/cregan/snp.html
SNPs are polymorphisms due to single nucleotide substitutions (transitions > transversions) or single nucleotide insertions/deletions.

As SNPs are predominantly bi-allelic, they are considered less informative than SSRs. It is estimated that 2.25-2.5 SNPs are required to provide the same genotyping information as one SSR marker (Paris *et al.*, 2003). This single base polymorphism can be revealed by several approaches of different technical complexity. The most obvious result is obtained by direct sequencing but with a large number of individuals and/or SNPs it is relatively expensive. The most simple approach is by AS-PCR (Allele Specific-PCR) (Liu *et al.*, 1997; Bundock *et al.*, 2005) and the most advanced and expensive are the MALDITOF (Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight) mass spectrometry (Sun *et al.*, 2000; Paracchini *et al.*, 2002) and the analysis of the extension products on DNA chips (microarray) (Rostoks *et al.*, 2005; Huentelman *et al.*, 2005). The choice of the method depends on the number of individuals and SNPs to analyze and on the available technical resources.
There are other types of marker often derived from the previously techniques described such as the PCR-RFLP. In this case, a DNA region is amplified by PCR and then digested by restriction enzymes. One locus is targeted and a simple gel can be sufficient to reveal the genotype.

Conclusion: The choice of a genotyping method for the genetic map construction depends on the available resources but also on the information level of the markers. Therefore the maps are generally constructed with several kinds of markers such as AFLP and SSR. AFLPs allow to obtain a large quantity of markers in short time at a relative low cost and are ideal markers to saturate the map while SSRs are less numerous and more time consuming but more informative for segregation analysis. Another advantage of using different type of markers is to increase the probability to have a good coverage of the genome.

1.2.2.3 Steps of mapping construction and software

After genotyping the mapping population, the map construction consists in four steps:

1. **Verification of the Mendelian segregation**: a χ^2 test is performed in order to verify the Mendelian segregation. Markers having a significant test are called distorted. The segregation distortion can have a biological cause (for example, linkage with lethal loci) or a statistical cause (low number of individuals used for calculation). The distorted markers can be integrated in the genetic map if their position does not influence the statistical order reliability of the other ordered markers.
2. **Detection of the genetic linkage among markers:** this step consists in the formation of linkage groups with the 2 point analysis. Every possible marker couples are separately tested for linkage and joined on the linkage groups. The co-segregation of 2 markers is assessed by comparing the frequencies of parental and recombinant gametes to the theoretical frequencies expected without linkage (independent markers). Two loci, A and B, for which the proportion of parental gametes (1-θ) significantly exceeds the proportion of recombinant gametes (θ) are said to be genetically linked. There are 2 possible statistical tests:
    - **the χ^2 test** which evaluates the individual segregations of markers and the linkage between the 2 markers;
    - **the method of LOD** (logarithm of the odds ratio or likelihood ratio) is the most used. The LOD score measures the decimal logarithm of the likelihood ratio between the linkage and the independence hypothesis among markers. **r** is the recombination ratio among markers:

$$\mathbf{LOD = \log_{10} (e^{L}(r)/ e^{L}(r_0))}$$

where $\mathbf{e^{L}(r)}$ is the maximum likelihood evaluated at r and $\mathbf{e^{L}(r_0)}$ is the maximum likelihood evaluated at r = 0.5 (independence)

For example, a LOD score of 4 means that the linkage between the 2 markers is 10^4 = 10000 folds more probable than the independence of the 2 markers. The estimated r is used to calculate the genetic distance between the 2 markers.

The principle of the maximum likelihood is to find the value of a variable of function allowing maximizing this function. This can be done by iteration or by determining the value of the variable for which the derivative of the function is equal to 0.

Once the linkage between all pairs of markers has been tested by two point linkage analysis, markers can be grouped into a series of linkage groups. When the map is saturated, the number of linkage groups corresponds to the haploid number of chromosomes. The markers order and the recombination rates between adjacent markers can then be determined by multipoint linkage analysis.

3. **Determination of markers order within each linkage group**: for m markers there are m!/2 possible orders. To obtain the more probable order, once again, the method of the maximum likelihood is used. For each possible map the probability that the data lead to this map is calculated and it represents the likelihood. The map with the highest likelihood will be retained. The algorithms of likelihood calculation depend on the software used but they generally apply a three point or multipoint analysis, which takes into account the double recombination events.
4. **Estimation of distance among markers**: distances among loci on genetic maps are measured in units called Morgans (M), or centi-Morgans (cM). One cM distance between 2 loci is equivalent to a 1% probability of recombination between them. However, if 2 loci are not very closely linked, not all recombination events will be detected because when a double recombination event occurs the original phase (parental gametes) is observed. Various functions have been proposed to convert recombination frequencies into genetic distance. Morgan (1910, 1928) proposed the first "mapping function". He assumed equivalence between recombination frequency and map distance. That is **r = M**, where r is the probability of recombination between two loci. This relationship is approximately correct for closely linked loci. Over greater chromosomal distances recombination frequencies are not strictly additive. Other mapping functions have been proposed. The Haldane (1919) and Kosambi (1944) functions are the two most used mapping functions:
    - **the Haldane function** supposes that the crossing-over probability in a region of the chromosome is independent of the crossing-over occurrence in a neighboring segment of the chromosome. This distance (in Morgan) is defined by:

$$\mathbf{d = -\tfrac{1}{2} \ln (1-2r)}$$

    - **the Kosambi function** considers the genetic interference which means that the probability to obtain a crossing-over in a region of a chromosome depends on the existence of a crossing-over in a neighboring region. It is defined by (in Morgan):

$$\mathbf{d = \tfrac{1}{4} \ln [(1+2r)/(1-2r)]}$$

The Kosambi function is generally more used because it is closer to the biological reality. Nevertheless, as the interference is not constant along the chromosome, any mapping function give an accurate estimation of the genetic distance (Crow and Dove, 1990).

The mapping software perform the linkage analysis, the markers ordering, and the calculation of the distance. They use various algorithms and are adapted for specific pedigrees. A list of mapping software is available at URL: http://linkage.rockefeller.edu/soft/list1.html. The software generally used in plant genetics is MAPMAKER (Lander *et al.*, 1987) which work with the F_2 progeny, backcross, doubled haploid and recombinant lines. It allows also controling the markers order using the EM (Expectation-Maximization) algorithm. JOINMAP (Stam, 1993) and CARTHAGENE (Schiex *et al.*, 1995) software allow merging genetic maps of different pedigrees (consensus map). CARTHAGENE is more recent than JOINMAP and it is less used in plant genetic mapping.

1.2.2.4 Genome length estimation

A genetic map is saturated when the number of the linkage groups is equal to the haploid number of chromosome and when all new markers added to the map are linked to one of the groups. The length of the genome coverage by the markers can be calculated by summing up the dimensions of the linkage groups. For a non saturated map, the genome length is estimated from the segregation data under the assumption of random markers distribution according to the method developed by Hulbert *et al* (1988) and modified by Chakravarti *et al* (1991):

$$\mathbf{G_{(Z)} = 2MX_{(Z)}/K_{(Z)}}$$

where $\mathbf{G_{(Z)}}$ is the genome length in cM estimated at a $\mathbf{LOD \geq Z}$, $\mathbf{M}$ the number of informative meioses studied defined by $\mathbf{M = n(n-1)/2}$ where $\mathbf{n}$ is the number of markers linked, $\mathbf{X_{(Z)}}$ the maximum distance observed between 2 markers linked and $\mathbf{K_{(Z)}}$ the number of marker couples linked at $\mathbf{LOD \geq Z}$. This method is widely applied in plants (Paglia *et al.*, 1998; Arcade *et al.*, 2000; Myburg *et al.*, 2003; Scalfi *et al.*, 2004).

The genome length is a relative quantity and it varies according to the genome regions. Diverse factors are susceptible to affect the recombination rate (De Vienne, 1997):

- **The sex**: parents recombination rate differences were observed on many plant species, comparing maps from reciprocal backcrosses. For example, in tomato and barley the recombination rate is greater in the female than in the male parent whereas in *Arabidopsis* and maize it is the contrary.

- **The genetic remoteness between the parents**: several studies indicate that the progenies from inter-specific crosses lead to smaller genetic distances than those obtained with intra-specific crosses. For example, in potato, the map realized from an intra-specific cross has a length increased of 65% compared to the map from an inter-specific cross. The explanation could be the reduction of homology among DNA strands in the inter-specific crosses, resulting in a reduction of the crossing-over frequencies.
- **Individual genetic variation of the recombination rate genetically controlled**: this is probably the reason why different progenies of the same species do not necessary lead to identical map length, the variation could be over 20%.

The relationship between the genetic distance and the physical distance (calculated in base pairs) differs among plant species (Table 2).

Table 2: Physical and genetic genome length of 15 plants species from Chagné *et al* (2002).

Species	Physical length of haploid genome (Mb)	Genetic length (cM) (MAPMAKER)	Chromosome number (n)	Average length of chromosome (cM)	Ratio physical length/genetic (Mb/cM)
Arabidopsis thaliana	150	675	5	135	0.22
Prunus persica	300	712	8	90	0.42
Oryza sativa	150	1490	12	125	0.3
Populus deltoides	550	2300	19	121	0.23
Eucalyptus grandis	600	1370	11	125	0.43
Brassica rapa	650	1850	10	185	0.35
Quercus robur	900	1200	12	100	0.75
Lycopersicon	980	1280	12	107	0.76
Solanum tuberosum	1540	1120	12	93	1.37
Zea mays	2500	1860	10	186	1.34
Lactuca sativa	2730	1950	9	217	1.4
Triticum tauschii	4200	1330	7	190	3.15
Hordeum vulgare	5500	1250	7	178	4.4
Pinus taeda	21000	1700	12	141	12.35
Pinus pinaster	25500	1850	12	154	13.78

The weak variation of the map genetic length among species contrasts with the large variation of the DNA amount per cell. In angiosperms, the DNA amount can vary up to a factor of 600: *Arabidopsis* has only 0,3 pg of DNA per diploid cell (2n) whereas *Fritillus* (*Liliaceae*) reaches 255 pg per diploid cell. As the number of the expressed genes could be in the same magnitude order within the superior plants, these differences of DNA amount would be essentially due to wide variation of repeated and non coding DNA. These regions are not included in the calculation of the genome length. The crossing-over frequencies per unit of physical length decrease when the genome length increases. It has been clearly observed that regions of heterochromatin, composed of highly repeated sequences, correspond to zones with a low recombination rate. In the same way, coding regions have a higher recombination rate. The average number of base pairs per cM depends on the considered species (De Vienne, 1997).

1.2.3 Interest of genetic maps for genome comparison

A genetic map with a high density of markers could be useful for the comparative mapping among phylogenetically related species, providing that there are orthologous loci among these species. Two loci are said orthologous when they derive from a speciation event and not from a duplication event (paralogous) (Fig. 19).

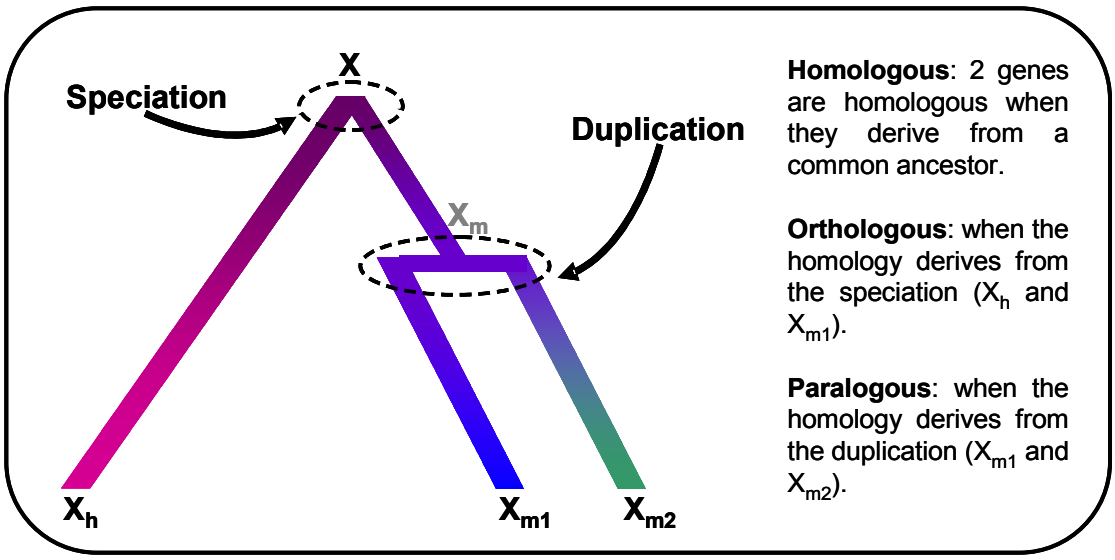


Fig. 19: Definition of homology relationship among markers after speciation and duplication
Example of the evolution of the gene X which is the ancestor of the gene X_h and X_m in the species h and m. The gene X_m is duplicated in X_{m1} and X_{m2}.

Then, it is possible to compare the genetic linkage among these loci and their order along a chromosome to study the synteny and the co-linearity. Synteny is the condition when the same loci are found in the same region among different species and the co-linearity is when these loci are in the same order. The comparative genetic mapping presents two major interests:

- the possibility to study the structure and evolution of the genome;
- the perspective to transfer molecular and genetic information from one species to another one such as the analysis and decomposition of complex traits.

1.2.3.1 Markers useful for comparative mapping

In order to compare genomes the maps would have orthologous and rather well conserved common markers. The markers stemming from expressed sequence could be more conserved among species. Actually, coding sequences evolve less because they control the important functions of the organism. RFLP using cDNA (coding DNA) probes are widely used in comparative mapping like markers coming from EST (Expressed Sequence Tag).

The SSRs are another kind of marker interesting for comparative mapping. The main advantage of these markers is their transferability among species. However, it depends on the phylogenetic distance, the genome complexity and the type of the composition of the repeated motif. The more the phylogenetic distance is large, the less the transferability is efficient. We have to note that the amplification of a SSR marker in another species does not mean necessarily that it is at the same locus but it must be verified by sequencing. Nevertheless, the SSRs are good candidate for map comparison (Casasoli, 2004).

1.2.3.2 Example of comparative genetic mapping

Comparative mapping presents a great interest for the comprehension of genome structure and evolution and to confirm and transfer information such as the decomposition of complex traits. A set of markers was defined and proposed to be used for comparative mapping between highly divergent genomes such as tomato and *Arabidopsis*. It is composed of 1025 genes referred to as a COS (Conserved Orthologous Set) markers (Fulton *et al.*, 2002).

A study among important crop species, shows that gene content and orders are highly conserved at the map level. In this study, seven species of the grass family were compared: rice, foxtail millet, sugar cane, sorghum, maize, the Triticeae cereals and oats (Devos and Gale, 1997). Another study shows a strong conservation of overall genic content across three Brassica genomes (*B. nigra*, *B. oleracea*, *B. rapa*) mirroring the conservation of genetic content observed over a much longer evolutionary span in cereals. The knowledge of the homologous regions of the genomes will increase the possibilities of transferring traits via homologous recombination in resynthesized hybrids and of selecting desirable hybrid chromosomes with the aid of genetic marker technology (Lagercrantz and Lydiate, 1996). Nevertheless, this may be not sufficient for the isolation of agronomic valuable genes. It was tested whether comparative mapping between *Arabidopsis* and maize of a small region surrounding the *DREB1A* gene in *Arabidopsis* could lead to the identification of an orthologous region in maize containing the *DREB1A* homologue. The results show a significant degree of orthology with the *Arabidopsis* region, but the extensive duplications and rearrangements in the *Arabidopsis* and maize genome as well as the evolutionary distance between *Arabidopsis* and maize, make orthology and co-linearity between these two species not sufficient to aid gene prediction and cloning in maize (Van Buuren *et al.*, 2002).

Analysis of comparative mapping was also carried out in tree species. In *Eucalyptus*, such an approach will provide valuable information on genome evolution and a powerful framework for comparative analysis of postzygotic reproductive barriers and other quantitative traits of

commercial importance in this genus. In a first analysis between *Eucalyptus grandis* and *Eucalyptus globulus* it has been found that all common markers were colinear and little evidence was found for gross chromosomal rearrangements (Myburg *et al.*, 2003). The map of maritime pine (*Pinus pinaster* Ait) was compared with the map of the loblolly pine (*Pinus taeda* L.). The synteny was maintained between the two species. The alignment of homologous linkage groups allowed the comparison of QTL location. The position of 2 QTLs controlling wood density and cell wall components were found to be conserved between the two species (Chagné *et al.*, 2003). This extensive synteny and co-linearity has also been found between loblolly pine and douglas fir (*Pseudotsuga menziesii* [Mirb.] Franco) (Krutovsky *et al.*, 2004). However, the comparison of the composite linkage map from *Picea mariana* and *Picea rubens* allowed the identification of one breakdown in synteny where one linkage group homologous to both *Picea* and *Pinus* corresponded to 2 linkage groups in douglas fir (Pelgas *et al.*, 2005). The last example of comparative genetic mapping in tree concerns *Quercus* and *Castanea*. The analysis performed between these 2 species would provide the means to investigate the correspondence of QTLs across the 2 genera and the opportunity to identify homologous chromosomal regions affecting important adaptive traits within the Fagaceae family. A first work suggests a conservation of macro-synteny between *Q. robur* and *C. sativa* (Barreneche *et al.*, 2004). In a second paper it was found a significant co-location of the QTLs controlling the timing of bud burst (Casasoli *et al.*, 2005).
These publications demonstrate the interesting information provided by the comparative genetic mapping. It permits a better understanding of the genome evolution and identification of the homologous chromosomal regions corresponding to important traits. The co-location of a QTL in different species confirms and consolidates the QTL analysis.

1.2.4 Interest of genetic maps for biodiversity studies

The studies on biodiversity often used enzymatic or anonymous molecular markers. Their position on the genome, coding or non coding regions, is unknown. The interest of genetic maps is to provide a list of molecular markers allowing representing the whole genome. The utilization of dominant markers such as AFLP (Arens *et al.*, 1998; Winfield *et al.*, 1998) is less efficient than co-dominant markers such as SSR (Fossati *et al.*, 2003; Imbert and Lefèvre, 2003; Storme *et al.*, 2003). The last ones present the advantage of being multi-allelic and highly variable. Thanks to these characteristics a genetic map rich in SSR could be a good source of markers for biodiversity studies.

1.2.5 Interest of genetic maps for breeding programs

One of the main objectives of plant breeders is to improve existing cultivars which are deficient in one or more traits by crossing such cultivars with lines which possess the desired trait. The conventional breeding programs are laborious, time consuming, involving several crosses, several generations, and careful phenotypic selection (Kumar, 1999). Linkage maps are a basis to find DNA markers which are tightly linked to agronomically important genes (Collard *et al.,* 2005). Once molecular markers closely link to desirables traits are identified MAS can be performed in early stages of plant development (Mohan *et al.*, 1997). MAS involves using the presence/absence of a marker to assist in phenotypic selection or as a substitute for it. This may make the selection more efficient, reliable, and cost-effective compared to the conventional plant breeding methodology (Collard *et al.*, 2005).

1.2.6 Genetic mapping of forest trees

Forest trees are wild species, preferentially or specifically allogamous, with very long generation time. Consequently, classical strategies used in genetic mapping are not adapted for them. Nevertheless, natural tree populations present a high level of heterozygosity and the possibility to obtain many full-sib progenies where markers can segregate. These properties lead to the development of a strategy mainly used in the mapping of forest trees: the pseudo-testcross strategy. This approach was described and applied the first time in *Eucalyptus* (Grattapaglia and Sederoff, 1994). It is based on the linkage analysis of dominant markers that are heterozygous in one parent and null in the other and, thus, segregate 1:1 in their F_1 progeny as in a testcross configuration (corresponding to the configuration observed when a heterozygote individual is crossed with a homozygote individual) (Fig. 20). As a consequence, two linkage maps are generated, one for each parent (i.e., two-way pseudo-testcross). The co-dominant markers must be considered as dominant markers and this strategy consists in following the segregation of the alleles from each parent separately by analyzing F_1 individuals (Fig. 21).

Two other strategies have been followed to construct linkage maps of trees: the "F_2 inbred model" and the "three-generation outbred model". The F_2 inbred model is based on a three-generation pedigree for which the grandparents are treated as inbred lines (represented by A_1/A_1 and A_2/A_2). In the F_2 generation, three genotypes occur at any loci: A_1/A_1, A_1/A_2, and A_2/A_2, segregating 1:2:1. Software programs, such as Mapmaker can assemble a combined parental map from the F_2 progeny data using the intercross mating type. It means that the data and the codification correspond to a classical F_2 progeny (see also. paragraph 1.1.2.1).

The three-generation outbred model is an extension of the pseudo-testcross strategy. Within a single outbred pedigree, any given co-dominant marker will segregate in one of three different ways. When one parent is heterozygous and the other is homozygous, segregation will be 1:1 (i.e., testcross mating type, configuration [1] and [3] Fig. 20). When both parents are heterozygous, segregation will be either 1:2:1 if both parents have the same genotype (configuration [4] and [5] Fig. 20, i.e., intercross mating type), or 1:1:1:1 if they have different genotypes (configuration [6] [7] [8] and [9], i.e., fully informative mating type). These segregation data are then subdivided into two independent data sets that separately contain the meiotic segregation data from each parent, and independent maps are constructed for each parent. A sex-average map is then constructed using an outbred mapping program, such as JoinMap, which uses fully informative and intercross markers to serve as common anchor-points between the 2 parental data sets (Cervera *et al.*, 2004).

These three mapping methods have been applied for many forest trees. For example, the pseudo-testcross strategy has been used for larch (Arcade *et al.*, 2000), Norway spruce (Acheré *et al.*, 2004), eucalypt (Brondani *et al.*, 2002), chestnut (Casasoli *et al.*, 2001), the European beech (Scalfi *et al.*, 2004) while the F_2 inbred model has been applied for poplar (Bradshaw *et al.*, 1994) and the three-generation outbred model for douglas-fir (Jermstad *et al.*, 1998).

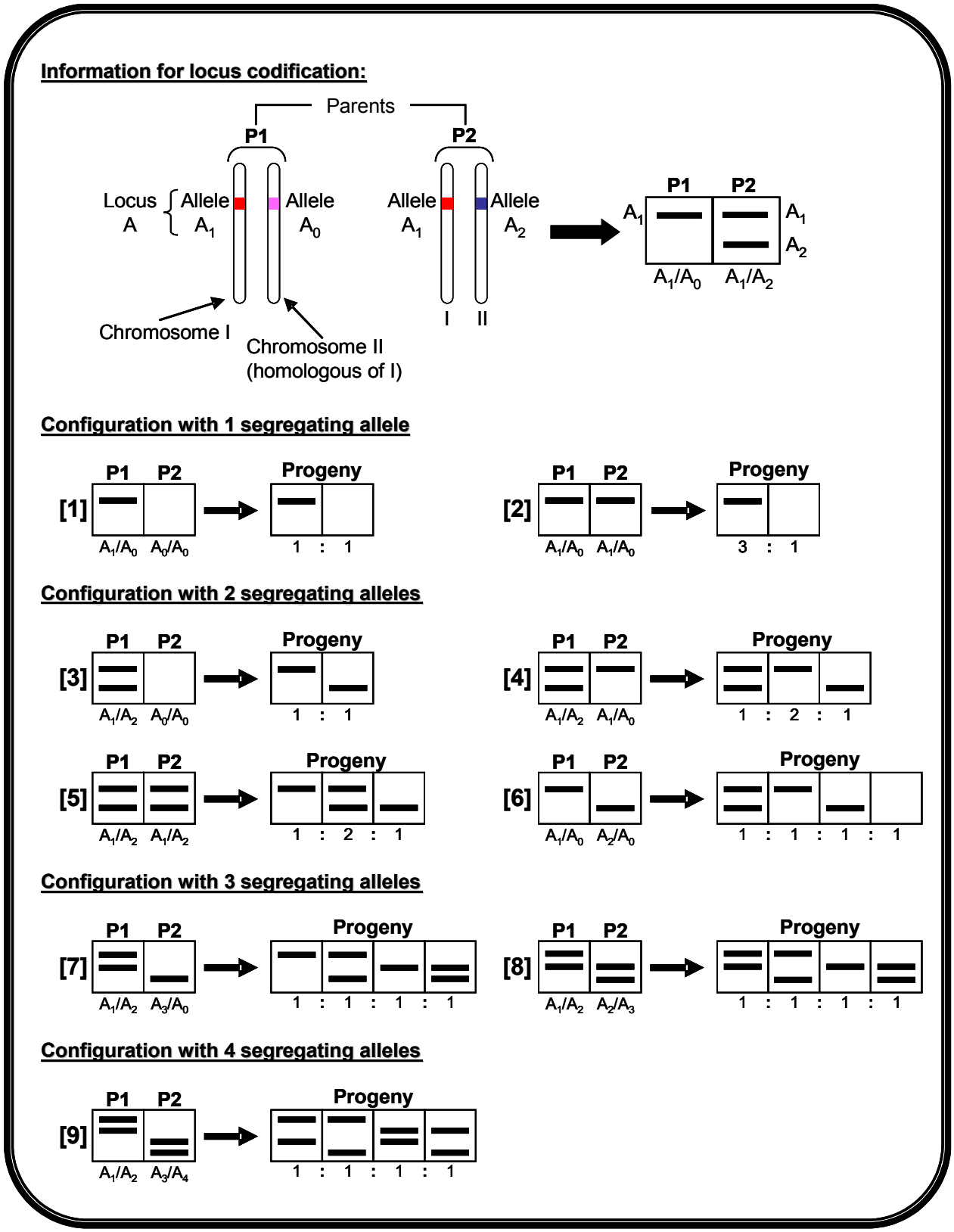


Fig. 20: Genotypic configuration observed in heterozygote crosses, modified from Lespinasse (1999)
Considering 2 heterozygote parents P1 and P2 at the locus A, 9 informative genotypic configurations for mapping can be distinguished. The configuration [1] is observed with the dominant markers as the configuration [2]. This last configuration cannot be used with Mapmaker software in pseudo-testcross strategy because it is impossible to determine the parental origin of the alleles in the progeny. The configurations [6], [7], [8] and [9] are fully informative.

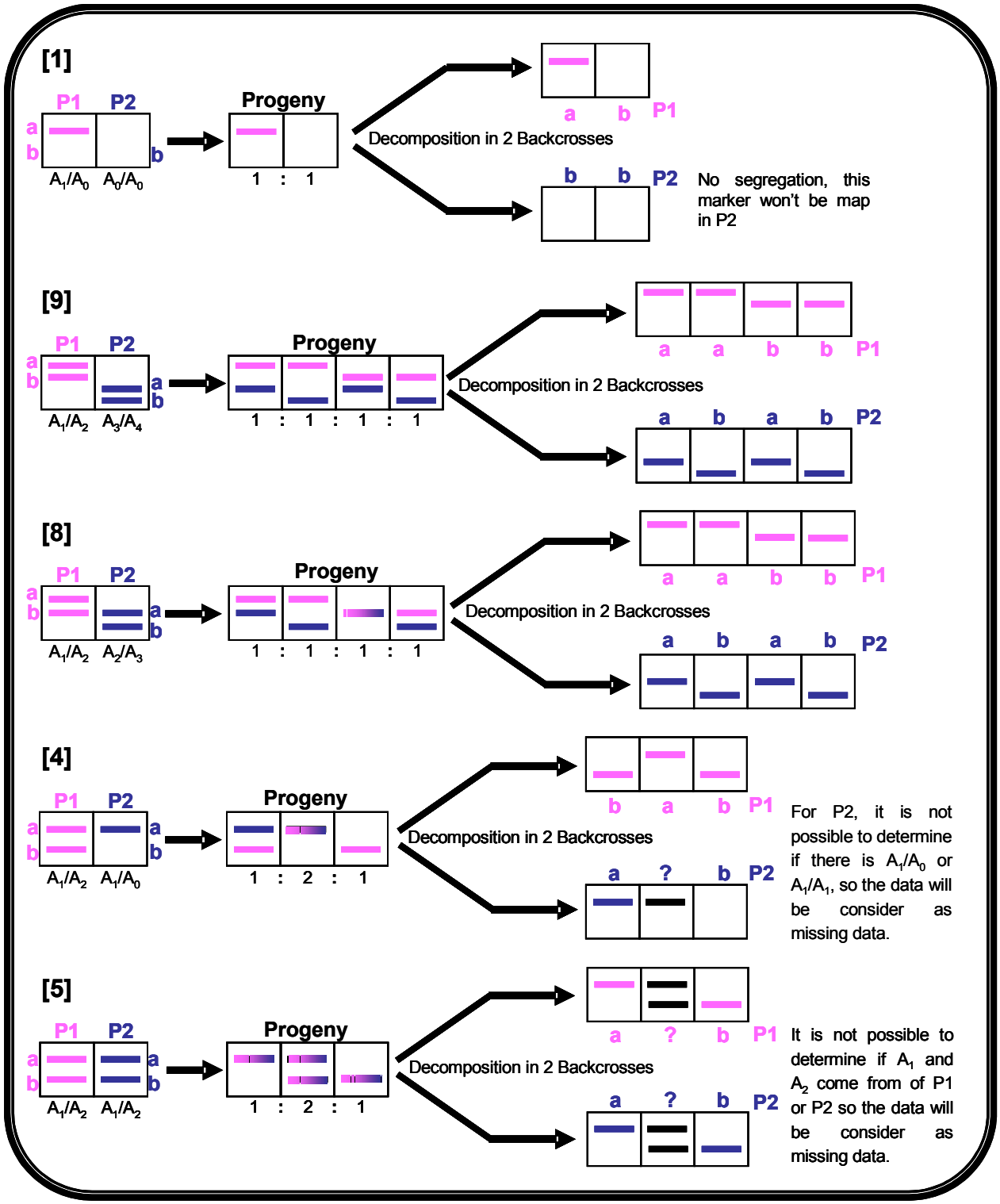


Fig. 21:Example of segregation data codification for the pseudo-testcross strategy, modified from Lespinasse (1999). Considering 2 heterozygote parents P1 and P2 at the locus A, the pseudo-testcross strategy consists in the decomposition of the segregation data in 2 backcrosses allowing the construction of 2 maps, one for P1 and one for P2. a and b correspond to different alleles of the locus A. The dominant markers are mapped only in the parent which has the visible alleles (heterozygote), corresponding to the configuration [1]. The configuration [8] and [9] can easily be decomposed in 2 backcrosses and the locus will be mapped in the 2 parents. In the configuration [4] and [5], this decomposition is more difficult. The allele origin from one parent is deduced from the information of the second parent alleles. Some phenotypes will be coded in missing data because, due to the dominance, the genotype is unknown.
The pink bands correspond to the P1 alleles whereas the blue bands correspond to the P2 alleles. When the progeny inherits the same allele from both P1 and P2, the band is in pink/blue.

1.2.7 Genetic mapping of poplar

Genetic linkage maps have been constructed for species belonging to three main *Populus* sections: *Populus* (*P. adenopoda* Maxim., *P. alba* L., *P. tremuloides* Michaux), *Tacamahaca* (*P. trichocarpa* T. & G., *P. cathayana* Rehder) and *Aigeiros* (*P. deltoides* Marshall, *P. nigra* L.). Maps were generated by analyzing marker segregation in intra-specific and inter-specific crosses between species of the same section, and hybrid progenies between *P. trichocarpa* and *P. deltoides* have also been used for genetic mapping. Even though the last two species have been classified in two different sections, they show close genetic relationships based on nuclear molecular markers (Cervera *et al.*, 2004). Up to now, 13 mapping works have been carried out leading to the construction of 24 maps of poplar: 8 for *P. deltoides*, 3 for the hybrid *P. trichocarpa* x *P. deltoides*, 3 for *P. trichocarpa*, 3 for *P. alba*, one for *P. tremuloides*, one for the hybrid *P. deltoides* x *P. cathayana*, one for *P. adenopoda*, one for the hybrid *P. euramericana*, one for *P. tomentosa* x *P. bolleana*, one for *P. bolleana* and one for *P. nigra*. Information regarding the design and construction of these maps is presented in Table 3. We can note that there are only 3 intra-specific crosses (*P. tremuloides*, *P. deltoides* and *P. alba*) out the 13 realized for poplar mapping and there is only one map of *P. nigra* coming from an inter-specific cross (*P. deltoides* x *P. nigra*).

The availability of these maps allows comparing the genomes of different poplar species. A first alignment of some maps (Bradshaw *et al.*, 1994; Frewen *et al.*, 2000; Cervera *et al.*, 2001) showed a complete agreement in linkage grouping and marker order among the *P. deltoides*, *P. nigra*, *P. trichocarpa*. SSR markers are useful for comparative mapping between different poplar species and validate the different map constructions (Cervera *et al.*, 2004).

Since September 2004, the entire genome sequence of *P. trichocarpa* is available. It was sequenced by the Joint Genome Institute (http://www.jgi.doe.gov/). This will allow intensive comparison of genome structure, gene order and, alignment of the genetic maps with the poplar genome sequence. This will help the location of large numbers of candidate genes on the genetic maps, and to compare their map positions with QTLs. The availability of the genome sequence also allows the development of SNPs. These markers are extremely valuable to saturate genetic linkage maps for more accurate localization of QTLs controlling traits of interest. SNP markers will also aid in the development of a "consensus" map for each *Populus* species or for "comparative mapping" within the *Populus* genus to identify common QTLs in different genetic backgrounds (Cervera *et al.*, 2004).

Table 3: Overview of the *Populus* genetic linkage maps. Modified from Cervera *et al.* (2004)
For each map, the cross, mapping strategy, kind and number of markers mapped, linkage groups found, and the genome length are given.

Cross (Individuals analyzed)	**Mapping strategy**	**Mapped species**	**Mapped markers**	**Total mapped markers**	**N of linkage groups**	**Genome Length (cM)**	**References**
P. tremuloides x *P. tremuloides* **93** F_2 hybrid Intra-specific cross	F_2 inbred model	*P. tremuloides*	Allozymes: 3 RFLP: 54	57	14	664 (K)	(Liu and Furnier, 1993)
(*P. trichocarpa* x *P. deltoides*) x (*P. trichocarpa* x *P.deltoides*) **90** F2 hybrid, family 331	F_2 inbred model	*P. trichocarpa* x *P. deltoides* hybrid	RFLP: 203 STS: 17 RAPD: 92	312	35	1261 (H)[a]	(Bradshaw *et al.*, 1994)[b]
(*P. deltoides* x *P. cathayana*) x (*P. deltoides* x *P. cathayana*) **ni** F_2 hybrid Inter-specific cross	F_2 inbred model	*P. deltoides* x *P. cathayana* hybrid	RAPD: 110	110	20	1899 (ni)	(Su *et al.*, 1998)
(*P. trichocarpa* x *P. deltoides*) x (*P. trichocarpa* x *P.deltoides*) **346** F_2 hybrid Inter-specific cross	Pseudo-testcross strategy	*P. trichocarpa*	AFLP: ni SSR[c]: 8 Genes: 5	ni	26	2002 (ni)	(Frewen *et al.*, 2000)
		P. deltoides	AFLP: ni SSR[c]: 8 Genes: 5	ni	24	1778 (ni)	
(*P. deltoides* x *P.deltoides* (F_1)) x *P.deltoides* (Male parent of F_1) **93** BC1 Intra-specific cross	Pseudo-testcross strategy	*P. deltoides*	AFLP: 137	137	19 5 triplets 19 doublets	2927 (K)[d]	(Wu *et al.*, 2000)
		P. deltoides (F_1)	ni	ni	ni	ni	

Table 3 (continued)

Cross (Individuals analyzed)	**Mapping strategy**	**Mapped species**	**Mapped markers**	**Total mapped markers**	**Nb of linkage groups**	**Genome Length (cM)**	**References**
P. deltoides x *P. nigra* **121** F_1 hybrid Inter-specific cross	Pseudo-testcross strategy	*P. deltoides*	AFLP: 394 SSR: 53 R. marker: 1	448	21 1 doublet	2304 (K)	(Cervera *et al.*, 2001)
		P. nigra	AFLP: 329 SSR: 40 Genes: 2	371	34 4 triplets 2 doublets	2791 (K)	
P. deltoides x *P. trichocarpa* **101** F_1 hybrid Inter-specific cross	Pseudo-testcross strategy	*P. deltoides*	AFLP: 305 SSR: 51 R. marker: 1 Genes: 2 (ur)	359	23 3 doublet	2304 (K)	
		P. trichocarpa	AFLP: 278 SSR: 60 STS: 1 Genes: 4 (ur)	343	23 3 triplets 2 doublets	2791 (K)	
P. adenopoda x *P. alba* **80** F_1 hybrid Inter-specific cross	Pseudo-testcross strategy	*P. adenopoda*	RAPD: 62	62	7 2 triplets 10 doublets	553 (K)[e]	(Yin *et al.*, 2001)
		P. alba	RAPD: 197	197	19 1 triplets 4 doublets	2300 (K)[e]	

Table 3 (continued)

Cross (Individuals analyzed)	**Mapping strategy**	**Mapped species**	**Mapped markers**	**Total mapped markers**	**Nb of linkage groups**	**Genome Length (cM)**	**References**
P. deltoides x *P. euramericana* **93** F_1 hybrid Inter-specific cross	Pseudo-testcross strategy	*P. deltoides*	RAPD: 85 AFLP: 219 ISSR: 6	310	31	3801 (K)	(Yin *et al.*, 2002)
		P. euramericana	RAPD: 68 AFLP: 162 ISSR: 11	241	34	3452 (K)	
(*P. trichocarpa* x *P. deltoides*) x (*P. trichocarpa* x *P.deltoides*) **44** F_2 hybrid, family 331 see Bradshaw, H. D. Jr. *et al.* 1994	Three-generation outbred model	*P. trichocarpa* x *P. deltoides* hybrid	SSR: ni	ni	23	850 (ni)	Sewell, M. M., ur
(*P. trichocarpa* x *P. deltoides*) x (*P. trichocarpa* x *P.deltoides*) **325** F_2 hybrid family 331 see Bradshaw, H. D. Jr. *et al.* 1994	Three-generation outbred model	*P. trichocarpa* x *P. deltoides* hybrid	SSR: ni	ni	23	850 (ni)	Sewell, M. M., ur
P. deltoides x *P. trichocarpa* **91** F_1 hybrid Inter-specific cross	Pseudo-testcross strategy	*P. deltoides*	AFLP: 139 RAPD: 107 SCAR: 1 RFLP:29 SSR: 16 Genes: 5	297	26 5 triplets 4 doublets	2845 (K)	Faivre-Rampant P., ur
		P. trichocarpa	AFLP: 92 RAPD: 98 STS: 1 RFLP:24 SSR: 19 Genes: 5	239	27 10 triplets 11 doublets	2095 (K)	

Table 3 (continued)

Cross (Individuals analyzed)	**Mapping strategy**	**Mapped species**	**Mapped markers**	**Total mapped markers**	**Nb of linkage groups**	**Genome Length (cM)**	**References**
(*P. tomentosa* x *P. bolleana*) x *P. tomentosa* **696** BC1 Inter-specific cross	Pseudo-testcross strategy	*P. tomentosa* x *P. bolleana*	AFLP: 236	236	19 1 triplet 20 doublets	1956 (K)	(Zhang *et al*., 2004)
		P. tomentosa	AFLP: 396	396	25 12 triplets 23 doublets	2683 (K)	
(*P. trichocarpa* x *P. deltoides*) x *P. deltoides* **180** F_1 hybrid (BC1) Inter-specific cross	Pseudo-testcross strategy	*P. trichocarpa* x *P. deltoides*	AFLP: 432 SSR: 105	544	19	2564 (H)	(Yin *et al*., 2004b)
		P. deltoides	AFLP: ni SSR: ni	158	33	1046 (H)	

Table 3 (continued)

Cross (Individuals analyzed)	Mapping strategy	Mapped species	Mapped markers	Total mapped markers	Nb of linkage groups	Genome Length (cM)	References
P. alba x *P. alba* **141** F_1 hybrid Intra-specific cross	Pseudo-testcross strategy	*P. alba*	AFLP: 203 SSR: 73 PCR marker: 12 Morphological marker: 1:	289	27	3012 (K)	Paolucci, I., ur
		P. alba	AFLP: 213 SSR: 67 PCR marker: 15	295	34	3427 (K)	

[a]: coverage based on 19 major linkage groups
[b]: the map has been extended to 512 markers (H.D. Bradshaw, unpublished data)
[c]: not indicated on which map
[d]: based on 19 major linkage groups obtained by alignment of linkage groups with intercross heteroduplex markers
[e]: based on framework
ur: unpublished results
ni: data not indicated
K: Kosambi units
H: Haldane units
R. marker: Resistance marker

2 Materials and Methods

As we have seen in the introduction, the construction of a genetic map consists in 3 main steps:

1. Constitution of a segregating progeny. An intra-specific controlled cross between 2 *P. nigra* genotypes was performed.
2. Selection of markers to characterize the progeny. Three types of markers were used in this works: AFLP, SSR, and SNP.
3. Statistical analysis of the marker segregation. The double pseudo-testcross strategy was used.

All the details of these steps will be described in this section.

2.1 Plant material

2.1.1 Parental provenance and mapping pedigree

Two *P. nigra* genotypes with divergent phenotypes have been chosen in natural Italian populations as parents to produce an intra-specific controlled cross (Fig. 22).

The female parent, "58-861", comes from Northern Italy (45°09'N, 7°01'E) near the Dora Riparia river and close to the Alps at 597m of altitude. The male parent, "Poli", comes from the Southern Italy (40°09'N, 16°41'E) near the Sinni river in the plain in front of the Ionio Sea at 7m of altitude. The full-sib family obtained in 2001 by crossing the two *P. nigra* parents is composed of 165 individuals and is maintained in the experimental farm of the Viterbo's university. Ninety-two F_1 individuals are being used as mapping pedigree.

Fig. 22: Provenance of the 2 *P. nigra* parents used to obtain the intra-specific controlled cross for mapping activity.

2.1.2 Plant characteristics

Due to their geographical origin, the parents of the cross differ for many traits. These differences are visible at morphological and phenological level (Fig 23).

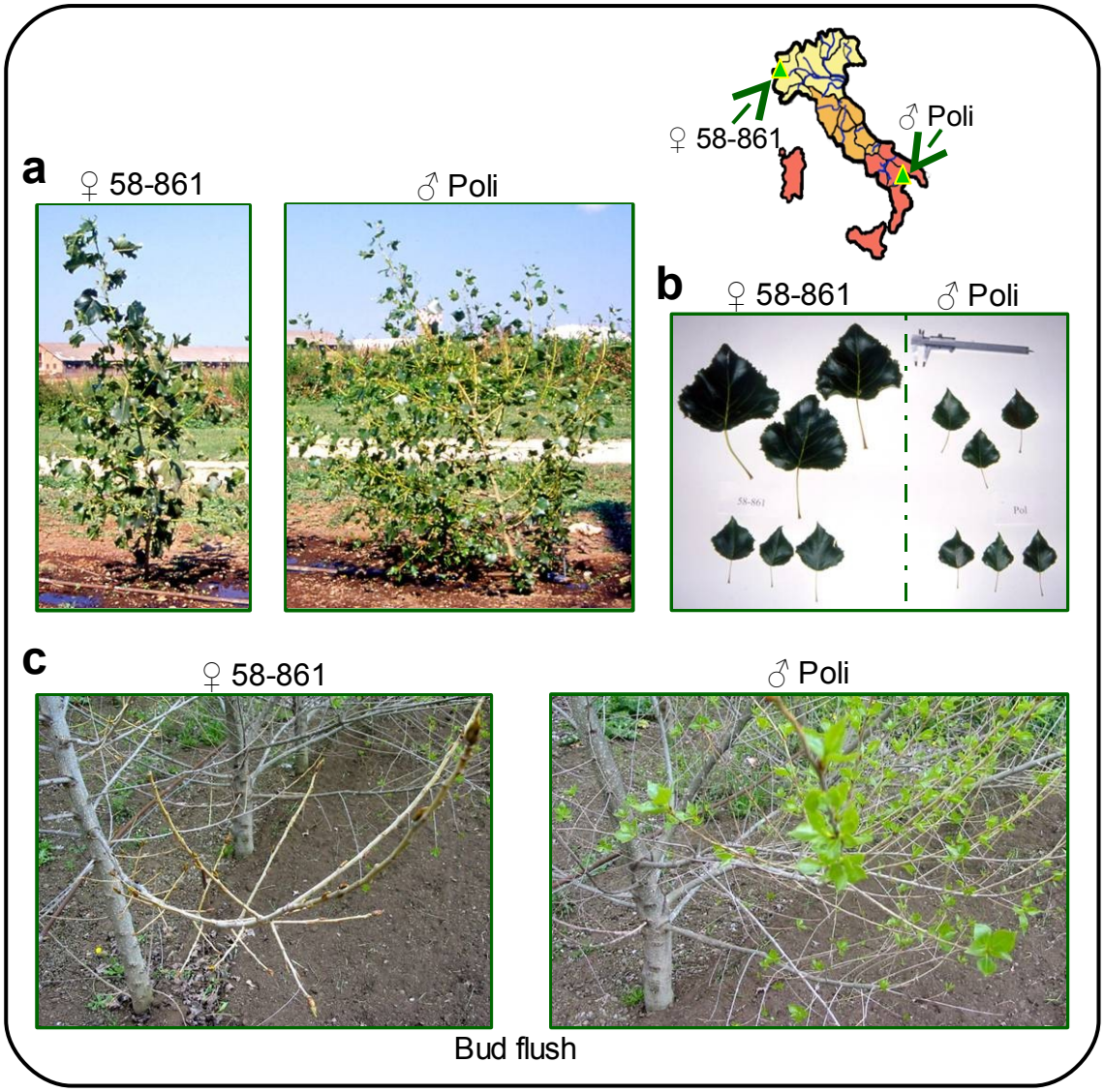


Fig. 23: Characteristics of the parents of the *P. nigra* intra-specific cross.
The female and the male parent are highly different at the morphological and phenological level. **a**, **b**: differences at branches and leaves level. **c**: differences at the bud flush level. The photo of bud flush was taken on 07/04/2004 in the experimental field of the Tuscia's university, Viterbo.

The differences in the main traits investigated in two locations are reported in Table 4 and Fig. 23. The male parent characteristics are typical of plants growing in hot and dry conditions as in its environment of provenance in Southern Italy. It has more branches (about 3 folds more sylleptic branches) and smaller leaves (leaf area is on average 3 to 4 folds smaller) than the female parent (Table 4, Fig. 23 a and b). Poli parent tends to invest more resources on branches

during the first growing year than the female 58-861 (fig 23 a). 58-861 parent, on average, opens its buds later (Table 4, Fig. 23 c) and sets the apical bud earlier than the Poli parent (Table 4) probably for an adaptation to the climatic conditions of its geographic origin.

Table 4: Main traits measured in two locations on the *P. nigra* parents used to obtain the mapping pedigree.

Trait	Site							
	Montelibretti (Roma)				Cavallermaggiore (CN)			
	58-861	Std. dev.*	Poli	Std. dev.*	58-861	Std. dev.*	Poli	Std. dev.*
Survival	83%		67%		83%		83%	
Crown characteristics								
N° sylleptic branches	29.25	16.46	82.5	6.14	17.2	9.15	41.6	12.36
N° proleptics branches	40.75	5.56	28	8.76	49.2	6.26	35.8	10.89
Single mature LA (cm^2)	119.43	19.38	25.15	1.1	70.57	12.31	21.6	2.28
Petiole length (cm)	8.11	0.74	3.57	0.19	7.32	0.71	3.49	0.13
Specific leaf area (cm^2/g)	129.9	18.94	128.58	19.08	125.98	17.96	109.63	13.3
Growth parameters								
Circum 1 year (mm)	52.6	10.69	47.89	5.94	42.7	13.61	21.98	6.66
Height 1 year (cm)	252.5	31.33	247	23.17	224.4	48.04	163.6	30.2
Circum 2 year (mm)	126.25	7.23	113	14.47	136.8	24.58	92	11.11
Height 2 year (cm)	488	22.32	460	102.45	635.8	73.73	470.75	34.88
Phenology								
Budset (day of the year)	-	-	-	-	257.8	5.4	281.6	2.61
Budburst (day of the year)	89.6	7.09	83.75	5.06	106.2	1.1	89.6	6.19

* Std. dev.: Standard deviation

The choice of highly divergent parents collected from opposite environments, for the intra-specific cross, maximizes segregating variation in the progeny at phenotypic and genotypic level.

2.2 Molecular analyses

All the PCR reactions were performed with the Taq DNA polymerase, reaction buffers and dNTP set from Amersham Biosciences (Italy), H_2O milliQ sterile (Millipore, Italy). The PCR Express Thermal cycler (Hybaid, Italy) and the Delphi 1000 Thermal cycler were indifferently used.

2.2.1 DNA extraction and quantification

Young leaves were collected from the parents and 92 F_1 progeny and stored at -80°C until analysis. DNA was extracted from all F_1 individuals and both parents using the GenElute™ Plant Genomic DNA Miniprep kit (Sigma-Aldrich, Italy).

Relative estimations of DNA concentration in each extract were performed by gel electrophoresis. Two µl of DNA extract were separated in a 0.8% agarose gel and visualized by

ethidium bromide staining. A series of dilutions of poplar genomic DNA, quantified spectrophotometrically, was used as a standard (Fig. 24).

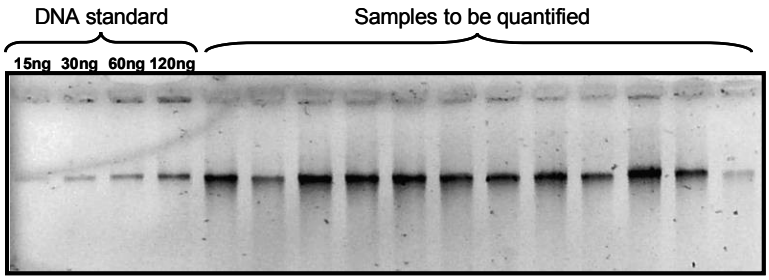


Fig. 24: Direct estimation of DNA concentration on agarose gel.
Two µl of DNA were separated in a 0.8% agarose gel and visualized by ethidium bromide staining to be quantified. The DNA quantity was estimated by comparing the band intensity of DNA extract to the standard.

The quality of the DNA for PCR amplification was tested with the Win3 marker which amplifies specific alleles of poplar species (Heinze, 1997). This marker allows also detecting introgression of a *P. deltoides* allele. The PCR conditions for the Win3 marker are described in Table 5. In addition to Win3, 4 SSR markers included in our analysis (WPMS9, WPMS18, WPMS20, and PMGC14) were also useful to detect introgression. WPMS9, WPMS18, and PMGC14 show *P. deltoides*-specific alleles of 234 bp, 220 bp, and 193/199 bp, respectively whereas WPMS20 presents one allele of 224 bp, which occurs at high frequency in *P. deltoides* but is rare in *P. nigra* (Fossati *et al.*, 2003).

Table 5: PCR mix (A) and thermal program (B) for Win3.

A

PCR reaction volume: 12.5µl		
PCR mix	**Vol. (µl)**	**Concentration**
DNA 10ng/µl	1.5	1.2ng/µl
Taq Polymerase 5U/µl	0.05	0.02U/µl
Reaction buffer 10X	1.25	1X
$MgCl_2$ 25mM	0.38	2.26mM
Each nucleotides 2mM	1.25	0.2mM
Each primer 2.5µM	1.25	0.25µM
H_2O	Up to 12.5µl	

B

Thermal program			
T°C	**Time**		**Cycles**
	min	**s**	
94	1	0	1
94	0	50	42
55	0	50	
72	1	20	
72	-	-	-
4	∞	∞	1

2.2.2 AFLP analysis and marker nomenclature

AFLP analyses were carried out using the "AFLP® Core Reagent Kit" (Invitrogen life technologies, Italy) following the manufacturer's instructions with modifications consisting in scaling down the reaction volumes to conserve reagents. A part of these modifications come from the protocol of the "Maize Mapping Project", www.maizemap.org/aflp.htm. As many of the reagents for the AFLP analysis are expensive, this protocol allows making the assay as cheap as possible, maintaining a high data quality. It was possible to apply this method because AFLPs were analyzed on an automated capillary sequencer (ABI3100, Applied Biosystem, Italy), which requires less material than a gel analysis. The reaction consists in 4 steps: i) the restriction digestion, ii) the ligation, iii) the preselective amplification and iv) the selective amplification.

1. The restriction digestion was performed in a volume reaction of 5µl containing 100ng DNA, 0.25U of each restriction endonucleases *Eco*RI and *Mse*I, restriction endonuclease buffer 1X and H_2O. This mixture was incubated for 3 h at 37°C. Then, the restriction endonucleases were inactivated by heating the mixture for 10 min at 70°C.
2. The ligation reaction consisted in adding 4.8µl of the "Adapter/Ligation solution from the "AFLP® Core Reagent Kit" and 0.2U of T4 DNA Ligase to the 5µl of digested DNA (step 1). This mixture was incubated 2 h at 20°C.
3. The preselective amplification reaction was performed with a 1:4 dilution of the 10µl ligation reaction product (step 2). The PCR conditions are described in Table 6. Preselective AFLP primers contained no selective nucleotides for *Eco*RI and one nucleotide (*C*) for *Mse*I (*Mse*I_*C*).

Table 6: PCR mix (A) and thermal program (B) for the preselective AFLP amplification reactions.

A

PCR reaction volume: 20µl		
PCR mix	**Vol. (µl)**	**Concentration**
DNA D/L 1:4*	2	-
Taq Polymerase 5U/µl	0.1	0.025U/µl
Reaction buffer 10X	2	1X
$MgCl_2$ 25mM	0.6	2.25mM
Each nucleotides 2mM	2	0.2mM
Each primer 10µM	0.6	0.3µM
H_2O	Up to 20µl	

B

Thermal program			
T°C	**Time**		**Cycles**
	min	**s**	
94	-	-	-
94	0	30	30
56	1	0	
72	1	0	
72	10	0	1
4	∞	∞	1

*: Dilution 1:4 of the 10µl Digestion/Ligation reaction product

4. The selective amplification reaction was performed with a 1:50 dilution of the 20µl preselective PCR product (step 3). The PCR reaction is described in Table 7 and 8. AFLP primers contained 2 selective nucleotides for *Eco*RI and 3 for *Mse*I (Table 9).

Table 7: PCR mix for the selective AFLP amplification reaction.

PCR reaction volume: 10µl		
PCR mix	**Vol. (µl)**	**Concentration**
DNA SA 1:50*	1	-
Taq Polymerase 5U/µl	0.05	0.025U/µl
Reaction buffer 10X	1	1X
$MgCl_2$ 25mM	0.3	2.25mM
Each nucleotides 2mM	1	0.2mM
Each primer 10µM	0.3	0.3µM
H_2O	Up to 10µl	

*: Dilution 1:50 of the 20µl preselective amplification reaction product

Table 8: Thermal program for the selective AFLP amplification reaction with touchdown.

Thermal program			
T°C	**Time**		**Cycles**
	min	**s**	
94	2	0	1
94	0	30	10
66-57*	1	0	
72	1	30	
94	0	30	20
56	1	0	
72	1	0	
72	5	0	1
4	∞	∞	1

*: - 1°C each cycle, start at 66°C and end at 57°C

The touchdown PCR consists in decreasing the annealing temperature of the reaction every cycle. Here the Tm was decreased 1°C every cycle, from 66°C to a "touchdown" at 57°C, during 10 PCR cycles. The Touchdown PCR simplifies the complicated process of determining optimal annealing temperatures. Any difference in Tm between the correct and incorrect annealing will give an advantage of 2-fold per cycle (or 4-fold per °C) to the correct product, all else being equal (Don *et al.*, 1991).

The *Eco*RI primer was named "E" followed by the 2 selective nucleotides and the *Mse*I primer was named "M" followed by the 3 selective nucleotides. The primer combination code is given in Table 9. Forty primer combinations were analyzed during this work.

*Eco*RI primers were labeled with the fluorescent dye Joe or 6-Fam (Table 9). PCR products were separated on a ABI3100 sequencer using the standard genotyping module modified for the following settings: 2KV injection voltage, 22 s injection time, 15KV run voltage, 2000 s run time. At the beginning, a dilution 1:20 of the PCR amplification product was performed. This dilution was again diluted 1:10 with loading buffer containing 98% of deionized formamide and 2% of internal standard, Genescan 500Rox (Applied Biosystems, Italy) or Mapmarker1000Rox (Biosense, Italy) (Table 8) and then denatured at 95°C for 5 min followed by a rapid cooling on ice.

The AFLP marker name refers to the primers used: the 5 letters refer to the selective nucleotides of the primer combination used, the first 2 correspond to the 2 selective nucleotides of the *Eco*RI primer and the last 3 refer to the 3 selective nucleotides of the *Mse*I primer. The obtained polymorphic peaks were numbered serially in ascending order of molecular weight; thus the numbers of the AFLP marker code refer to the relative position of the polymorphic peak on the electropherogram. In this way, the AFLP marker AGCAT01 has the lowest molecular weight and was obtained with the primer combination *Eco*RI_AG, *Mse*I_CAT (EagMcat).

Table 9: sequence and nomenclature of the 40 primer combinations used for the AFLP analyses.
In this table, the complete sequence of *Eco*RI and *Mse*I primers without selective nucleotide (A) as well as the complete sequence of the primers used for the preselective amplification (B) are presented. The selective nucleotide in the complete sequence is in bold and underlined. The codes of the primer combinations are listed in the first column (C). The *Eco*RI primer was named "E" followed by the 2 selective nucleotides and the *Mse*I primer was named "M" followed by the 3 selective nucleotides.

A

Primers	Sequence
Eco*RI	5' GACTGCGTACCAATTC 3'
Mse*I	5' GATGAGTCCTGAGTAA 3'

B

Primers used for preselective amplification	Sequence
***Eco*RI**	5' GACTGCGTACCAATTC 3'
***Mse*I_C**	5' GATGAGTCCTGAGTAA**C** 3'

C

Primer combination code	Fluorescent dye	Internal standard
EaaMcgt	Joe	Mapmarker1000
EagMcga	6-Fam	Mapmarker1000
EagMcat	6-Fam	Mapmarker1000
EagMctg	6-Fam	Mapmarker1000
EcaMcac	6-Fam	Mapmarker1000
EcaMctg	6-Fam	Mapmarker1000
EccMcct	Joe	Mapmarker1000
EccMcta	Joe	Mapmarker1000
EccMcat	Joe	Mapmarker1000
EccMctt	Joe	Mapmarker1000
EccMctc	Joe	Mapmarker1000
EccMcac	Joe	Mapmarker1000
EccMcga	Joe	Mapmarker1000
EccMcgt	Joe	Mapmarker1000
EctMcag	Joe	Mapmarker1000
EctMctc	Joe	Mapmarker1000
EctMcgt	Joe	Mapmarker1000
EtaMcga	Joe	Mapmarker1000
EagMctt	6-Fam	Mapmarker1000
EcgMcca	Joe	GS 500
EcgMcag	Joe	Mapmarker1000
EcgMcac	Joe	GS 500
EcgMcga	Joe	GS 500
EtcMcgt	Joe	Mapmarker1000
EaaMcat	Joe	Mapmarker1000
EaaMctt	Joe	Mapmarker1000
EaaMcct	Joe	Mapmarker1000
EaaMcca	Joe	Mapmarker1000
EtaMcat	Joe	Mapmarker1000
EtaMctt	Joe	Mapmarker1000
EtaMcta	Joe	Mapmarker1000
EtaMcca	Joe	Mapmarker1000
EtcMcat	Joe	Mapmarker1000
EtcMctt	Joe	Mapmarker1000
EtcMcct	Joe	Mapmarker1000
EtcMcta	Joe	Mapmarker1000
EcaMcag	6-Fam	Mapmarker1000
EcaMctc	6-Fam	Mapmarker1000
EcaMcct	6-Fam	Mapmarker1000
EcaMcta	6-Fam	Mapmarker1000

***** Sequence of primer without selective nucleotide.

2.2.3 SSR analysis and nomenclature

The sequences of the SSR primers used in this study come from 6 sources:

- The *Populus* Molecular Genetics Cooperative. The sequences are available on the following web page: http://www.ornl.gov/sci/ipgc/ssr_resource.htm. The name of the SSR markers were taken from the website and they have the **PMGC** or **GCPM** prefix; **152 of these SSR were tested** in our study.
- Oak Ridge National Laboratory (Tuskan *et al.*, 2004). These markers have the prefix **ORPM** and **126 of these SSR were tested**.
- The center for Plant Breeding and Reproduction Research (van der Schoot *et al.*, 2000; Smulders *et al.*, 2001). They have the **WPMS** prefix and **22 of these SSR were tested**.
- SSR with **PTR** prefix were developed by Dayanandan *et al.* (1998) and Rahaman *et al.* (2000) and all the **12 SSR from this source were tested**.
- SSR with **ai**, **bi and bu** prefix were developed from an EST (Expressed Sequence Tag) database by Jorge *et al.* (unpublished), **16 of these SSR were tested**.
- SSRs corresponding to the *P. trichocarpa* AGAMOUS genes, named **PTAG1** and **PTAG2**, developed by Brunner *et al.* (2000), **were also tested**.

All SSR available in our laboratory are described in Appendix 1.

In a first step, SSR were screened on the 2 parents and 6 individuals of the progeny. The PCR conditions used are given in Table 10, the annealing temperature of each SSR primer couple is indicated in the Appendix 1. The amplification products were separated in a 3% high resolution agarose gel, MetaPhor® Agarose (Cambrex BioScience, USA), and visualized by ethidium bromide staining.

Table 10: PCR mix (A) and thermal program (B) for SSR analysis.

A

PCR reaction volume: 12.5µl		
PCR mix	**Vol. (µl)**	**Concentration**
DNA 10ng/µl	1.5	1.2ng/µl
Taq Polymerase 5U/µl	0.05	0.02U/µl
Reaction buffer 10X	1.25	1X
$MgCl_2$ 25mM	0.38	2.26mM
Each nucleotides 2mM	1.25	0.2mM
Each primer 2.5µM	1.25	0.25µM
H_2O	Up to 12.5µl	

B

Thermal program			
T°C	**Time**		**Cycles**
	min	**s**	
94	3	0	1
94	0	30	30
Tm*	0	45	
72	0	30	
72	4	0	1
4	∞	∞	1

* Tm corresponds at the annealing temperature of the SSR primer couple.

All the polymorphic segregating SSRs were analyzed on the 92 progenies in the same condition used for the screening. The number of PCR cycles was adapted according to the PCR efficiency: when the band obtained had a weak intensity, the number of PCR cycles was increased to 42 (Table 10).

When the resolution of MetaPhor® Agarose was not sufficient for the interpretation of the pattern, SSRs were labeled with fluorescent dye-labeled primers (Hex and 6-Fam) and analyzed using an automated capillary sequencer (ABI3100, Applied Biosystem, Italy) which allows more precise separation of the DNA fragments. In order to reduce the cost and time of these analyses, the PCR reactions were performed using tailed primers and the PCR products were multiplexed before the detection with the ABI3100 capillary sequencer. This method (M13-tailed primers) consists in 5'-tailing the forward primer with the M13 sequence 5'-*CACGACGTTGTAAAACGAC*-3. The M13-forward primer is used in combination with a M13 primer that has the same sequence but is dye-labeled at its 5' end (Oetting *et al.*, 1995; Boutin-Ganache *et al.*, 2001; Fukatsu *et al.*, 2005). In this way we use a unique dye-labeled primer for all the reactions (Fig. 25).

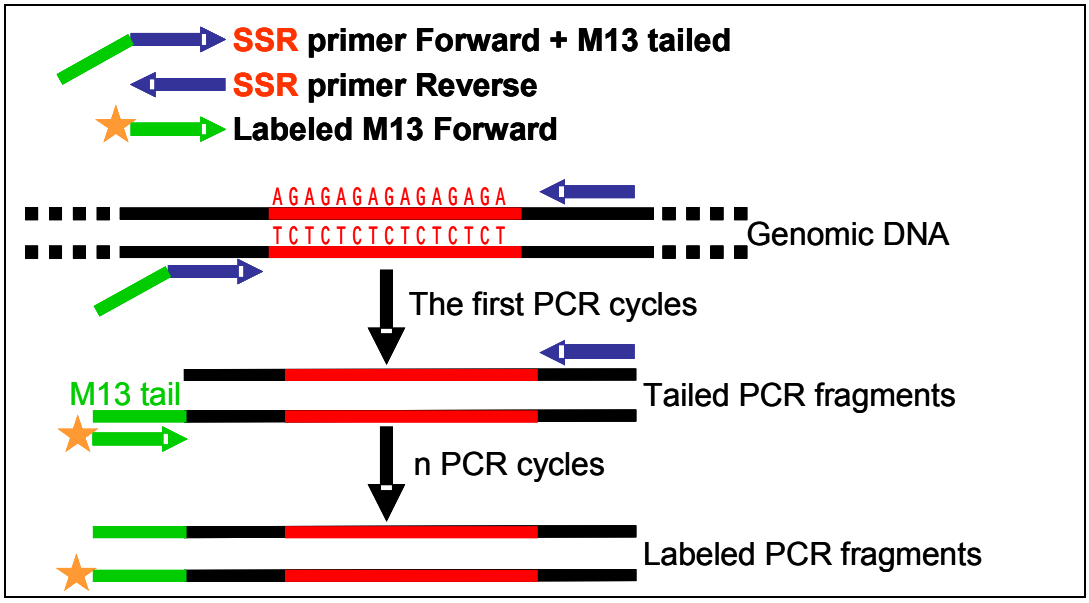


Fig. 25: Description of the M13-tailed primer method.
The amplification of SSRs was performed with 3 primers: a forward primer with the M13 tail, a reverse non-tailed primer, and a dye-labeled M13 primer. During the polymerization reaction the fluorescence reporter was incorporated into the product and a fluorescence signal was emitted. Only the labeled product containing the fluorescent dye could be detected by the DNA sequencer.

Two M13 primers were labeled, one with the Hex dye and the other with the 6-Fam dye in order to perform multiplex analyses. The PCR conditions are given in Table 11. PCR products were analyzed on the ABI3100 sequencer using the standard genotyping module with some modifications of the injection voltage and injection time according to the peak height. Two to 6

primer pairs with different fragments size and/or dye colors were mixed in appropriate ratios (2 folds more PCR product with the dye Hex) and diluted 1:10 with loading buffer containing 98% of deionized formamide and 2% of internal standard, Genescan 500Rox (Applied Biosystems, Italy), then denatured at 95°C for 5 min followed by a rapid cooling on ice.

Table 11: PCR mix(A) and thermal program (B) for SSR analysis with the M13-tailed primer method.

A

PCR reaction volume: 10µl		
PCR mix	**Vol. (µl)**	**Concentration**
DNA 10ng/µl	1.5	1.5ng/µl
Taq Polymerase 5U/µl	0.04	0.02U/µl
Reaction buffer 10X	1	1X
$MgCl_2$ 25mM	0.2	2mM
Each nucleotides 2mM	1	0.2mM
Labeled M13 forward primer 10µM	0.5	0.5µM
M13-tailed forward primer 10µM	0.05	0.05µM
Reverse primer 2.5µM	2	0.5µM
H_2O	Up to 10µl	

B

Thermal program			
T°C	**Time**		**Cycles**
	min	**s**	
94	3	0	1
94	0	20	42
50	0	30	
72	0	30	
72	5	0	1
4	∞	∞	1

In the Appendix 1, the annealing temperature (Tm) of the primers used for the thermal program and the separation method are presented for each SSR analyzed.

2.2.4 SNP analysis

In the frame of a linkage disequilibrium study in *P. nigra* (EC POPYOMICS project), the laboratory of the "Dipartimento di Produzione Vegetale e Tecnologie Agrarie", University of Udine, carried out a search for SNP polymorphisms. Genome regions corresponding to single copy sequences were sequenced in a set of *P. nigra* genotypes, including the parents of our mapping population. The sequences and localization of SNPs were gently provided by M. Morgante and G. Zaina (personal communication). Seven polymorphic SNPs were selected and analyzed by PCR-RFLP, synonym of CAPS (Cleaved Amplified Polymorphic Sequence), or dCAPS (derived Cleaved Amplified Polymorphic Sequence) (Neff *et al.*, 1998). The PCR-RFLP was used when the SNP alter a recognition site for an available restriction enzyme. In this technique, specific primers are used to amplify the template DNA and nucleotide polymorphisms are detected by the loss or gain of a restriction site (Fig. 26, A). When the SNP does not fall within a restriction enzyme recognition site, we used the dCAPS method. It consists in introducing a restriction enzyme recognition site (which includes the SNP into the PCR product) by using a primer containing one or more mismatches to the template DNA. The PCR product, modified in this manner, is then subjected to restriction enzyme digestion and the presence or absence of the SNP is

determined by the restriction pattern (Fig. 26, B). The dCAPS primers were designed with the web-based program, dCAPS Finder 2.0 (Neff *et al.*, 2002).

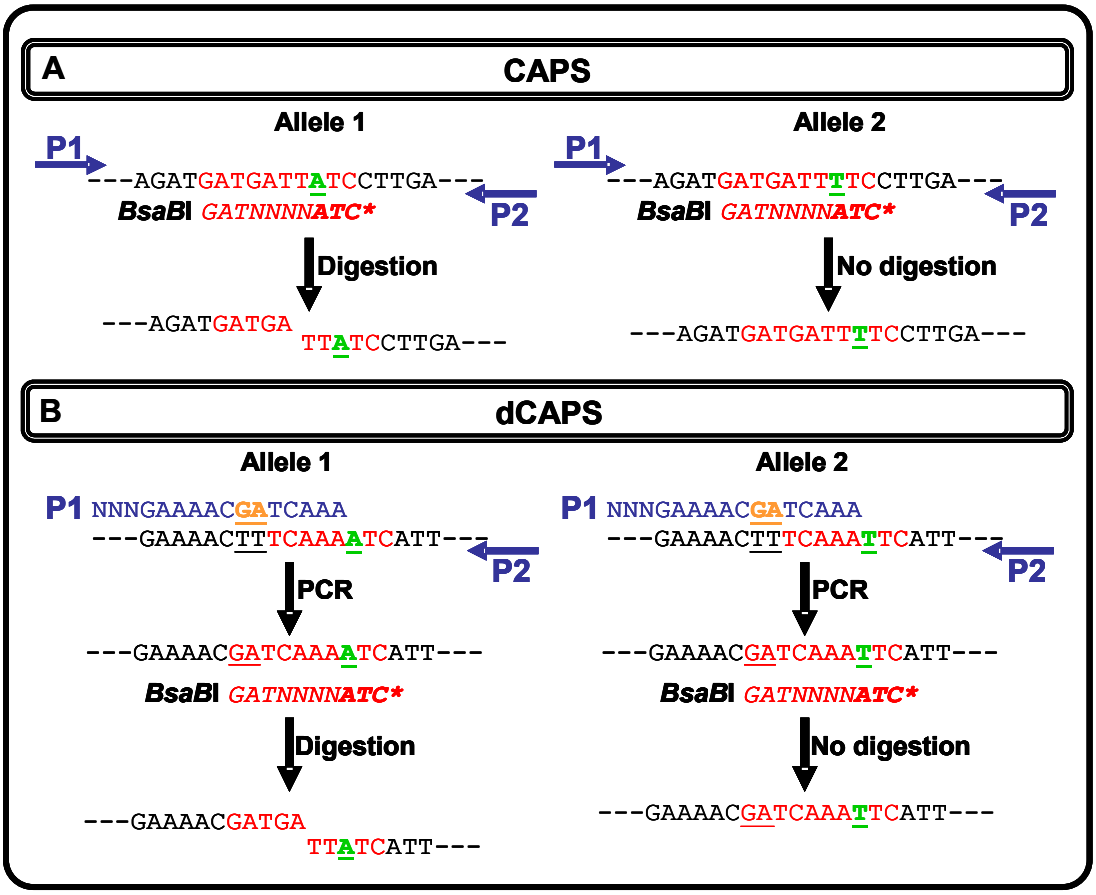


Fig. 26: Description of CAPS and dCAPS method.
A: For the CAPS (Cleaved Amplified Polymorphic Sequence) method, specific primers, P1 and P2 (in blue), are used to amplify template DNA (allele 1 and allele 2) and the SNP (green bold underlined) is detected by the loss or gain of a restriction site (red sequence). Here an example with the restriction enzyme *BsaBI* is showed.
B: dCAPS (derived Cleaved Amplified Polymorphic Sequence) analysis uses a mismatched primer (the mismatch is indicated in orange bold underlined) to create a restriction enzyme recognition site (red sequence). Here an example with the restriction enzyme *BsaBI* is showed.
* the red italic sequence correspond to the *BsaBI* recognition site.

The PCR conditions used are presented in Table 12 and the annealing temperature of each primer couple is indicated in Table 13, with the sequence of the primers. Aliquots (5µl) of the PCR product were digested for 1 h 30 at the temperature indicated by the manufacturer in 10µl total volume with 2 units of the appropriate restriction endonuclease (Fermentas Life Sciences, Italy) (Table 13). After digestion the whole reaction was separated on a 2.5% MetaPhor® agarose gel and visualized by ethidium bromide staining. The characteristic of each SNP with the respective analysis method are presented in Table 14. The available sequences of *P. nigra* with the SNPs were aligned with the *P. trichocarpa* database using BLAST (Basic Local Alignment Search Tool) software (Altschul *et al.*, 1990), to determine the linkage group where these sequences are

located. We expect to map these SNPs on the same linkage groups in *P. nigra* and the expected linkage groups are presented in Table 14.

Table 12: PCR mix (A) and thermal program (B) for SNP analysis.

A

PCR reaction volume: 12.5µl		
PCR mix	**Vol. (µl)**	**Concentration**
DNA 10ng/µl	1.5	1.2ng/µl
Taq Polymerase 5U/µl	0.05	0.02U/µl
Reaction buffer 10X	1.25	1X
$MgCl_2$ 25mM	0.38	2.26mM
Each nucleotides 2mM	1.25	0.2mM
Each primer 2.5µM	1.25	0.25µM
H_2O	Up to 12.5µl	

B

Thermal program			
T°C	**Time**		**Cycles**
	min	**s**	
94	3	0	1
94	0	30	35
Tm*	0	45	
72	0	30	
72	4	0	1
4	∞	∞	1

* Tm corresponds to the annealing temperature of the SNP primer couple.

Table 13: Sequences and annealing temperature of the SNP primers.
The sequence of the forward primer (F) and the reverse primer (R) is indicated for each SNP analyzed with the annealing temperature (Tm) used for the PCR reaction. Letters in bold and underlined are the mismatched nucleotides with template DNA. These primers were used for the dCAPS (derived Cleaved Amplified Polymorphic Sequence) method.

Locus name	**Primers sequence (5'→3')**	**Tm (°C)**
PhyA	F: GCCTTAGATGAGAAAAC**GA**TCAAA R: AAAGAGAAACATCCCCGAAT	55
PhyB1	F: AAGGACTCTTTT**CC**GGCCTTCG R: ATTGCGTAAAAAGGCTTCCC	55
PhyB2	F: AGGTGAGTATTTCTGCTTTG R: ATTAACTTAAAAAAGATTATACAG	52
IAA2	F: TGTTTTACCATAGGTGAGCAAAAT**T**AA R: CCTGACATCATAACAAAGTA	52
I13R	F: AGTAGTCCTAAAATCACAAGC R: AGAAGATCAGATAGAAGGAA	52
A15R	F: TTCATGAGCACCATACG R: AATTCTAGACCGA**G**TGCA	55
H11R	F: AGAATTTTGACTATTTTGTT**G**TA R: AAAAGAAATACTCATCCCAT	55

Table 14: SNP characteristics.
The allelic configuration (A1/A2) of the parents (58-861:female parent, Poli: male parent) is given in the second and third column.

Locus name	58-861 A1/A2	Poli A1/A2	Sequence length (bp)	SNP position (b)	Expected LG[a]	Analysis method	Restriction enzyme	Expected fragment size (bp) ND[b]	D[c]
PhyA	G/A	G/G	572	85	XIII	dCAPS	*Bsa*BI	196	174/22
PhyB1	A/G	A/A	589	303	VIII	dCAPS	*Bsi*YI	123	104/19
PhyB2	A/A	A/T	611	74	X	CAPS	*Bsa*BI	369	298/71
IAA2	A/A	A/T	511	330	II	dCAPS	*Vsp*I	161	137/24
I13R	T/G	T/G	352	245	XIV	CAPS	*Vsp*I	352	245/107
A15R	T/T	T/G	271	132	III	dCAPS	*Apa*LI	130	112/18
H11R	T/T	C/T	410	204	II	dCAPS	*Rsa*I	100	78/22

[a] LG: Linkage Group

[b] ND: Non digested fragment

[c] D: Digested fragment

2.3 Mapping methods

The construction of a genetic map can be divided in 4 steps: i) verification of the Mendelian segregation, ii) detection of the genetic linkage among markers, iii) ordering markers in each linkage group, iv) estimation of the distance among markers. All these steps were performed with different software.

2.3.1 Software

Peak patterns obtained analyzing the PCR products with the ABI3100 sequencer were scored with GeneScan and Genotyper software (Applied Biosystems). The resulting data tables were further processed by the PERL scripts Genomap (http://www.esd.ornl.gov/PGG/scripts.htm) to detect null alleles, anomalous alleles, aneuploidy, discrepancies in repeated samples, segregation distortion, and to infer parental origins of alleles. Data analyzed on the gel were organized in the same way to use Genomap for the complete data set. Discrepancies in the data were corrected by returning to the original traces.

The linkage map was constructed using Mapmaker software version 3.0 (Lander *et al.*, 1987). The data for Mapmaker were prepared with Genomap.

The maps were drawn with the MapChart 2.1 software (Voorrips, 2002).

2.3.2 Map construction

The genetic maps were constructed according to the two-way pseudo-testcross mapping strategy (Ritter *et al.*, 1990; Grattapaglia and Sederoff, 1994). As explained in the introduction, this consists in following the segregation of alleles from each parent separately, generating a map for each of the parents. Therefore, two data matrices were created, one for each parent of the cross. AFLP, SSR, and SNP markers segregating 1:1 in the progeny were used for the construction of the maps of both parents. To detect linkages in repulsion phase, the data set was inverted and added to the original data. Inverted markers are indicated by a "r" and represent markers in repulsion.

2.3.2.1 Test of the Mendelian segregation

For each marker a χ^2 test* (d.f. = 1, $P < 0.01$ and $P < 0.05$) was performed with Genomap to identify deviation from Mendelian ratios. AFLP markers deviating at 1% significance level were excluded for the linkage analysis because of major risk of technical artifact. The other markers (deviating at $0.01 < P < 0.05$) were noted as distorted but conserved in the data set.

* $\chi^2 = \sum ((\text{observed individuals} - \text{expected individuals})^2/\text{expected individuals})$

2.3.2.2 Linkage groups detection

The linkage analysis was performed by Mapmaker with the data type "F_2 backcross" suited for our pseudo-testcross configuration. The "TRIPLE ERROR DETECTION" and the "ERROR DETECTION" features were used to recognize the circumstance when an event was more probably the result of error than of recombination. Initially, markers were grouped by two-point analysis using a LOD (Logarithm of odds) score of 4.0 and a maximum recombination fraction θ of 0.3.

2.3.2.3 Ordering markers in the linkage groups and distance estimation

The most likely order of markers within a linkage group was determined by multipoint analysis. For linkage groups with more than 5 markers, the "THREE POINT" command was used to pre-compute the likelihood of all three point crosses of each group. Then the "ORDER" command was used to select a subset of markers ordered at a minimum LOD of 4.0 with a log-likelihood threshold of 3.0. Additional markers were added by the "TRY" command with a log-likelihood threshold of 2.0. The order of the marker subset was controlled with the "RIPPLE" command,

which compare the likelihoods of the original order to those found when the order of 5 neighboring loci was permuted. New markers were added only if the new order obtained was confirmed with the RIPPLE command. For the linkage groups with less than 5 markers, the "COMPARE" command was used. A framework map was established when marker order was supported by a log-likelihood ratio support of 2.0. Markers that could not be ordered with equal confidence were indicated as accessory and linked to a specific marker on the map. They were placed on the map with the TRY and NEAR commands. Markers showing a segregation distortion at $P < 0.01$ were also incorporated as accessory markers. The marker orders of these groups were equally supported by a log-likelihood of 2.0. Distances between marker loci were calculated from recombination fraction using Kosambi's mapping function (a recombination fraction of 0.3 corresponds to a Kosambi distance of 34.7 cM) with ERROR DETECTION ON and OFF.

2.3.3 Estimation of the genome length

The estimation of the genome length was calculated from partial linkage data according to:

$$G_E = N(N-1)X/K$$

with a confidence interval of:

$$G_E/(1 \pm 1.96/\sqrt{K})$$

where N is the number of markers and N(N-1) is the number of pairwise comparisons, X and K are the maximum distance between 2 adjacent markers in cM and the number of marker couples, respectively, linked at a minimum LOD score (see the introduction and Chakravarti *et al.*, 1991). A minimum LOD score of 4.0 was chosen to estimate the genome length.

We calculated the observed genome length, G_O, simply as:

$$G_O = \sum G_I$$

where G_I is the total genetic distance of linkage group I.

In addition, the observed genome length was calculated using the formula by Nelson *et al.* (1994) which takes into account all markers, linked and unlinked:

$$G_{ON} = G_O + X(L-R)$$

where X is the observed maximum distance between markers, L is the total number of linkage groups, triplets, doublets, and unlinked markers, and R is the haploid number of chromosomes.

2.3.4 Marker distribution analysis

Marker distribution among linkage groups was analyzed by comparing marker density with expectations under the Poisson's distribution hypothesis using the method described by

Remington *et al*. (1999). This test was conducted using all markers, both framework and accessory. Each linkage group *I* was estimated to have a length:

$$L_{IE} = L_I + 2S$$

where L_I is the map distance observed between terminal markers of the linkage group *I*, and S is the average framework marker spacing (S = G_O (observed genome length)/total number of marker of the framework).

Under the assumption of equal marker density for all linkage groups, the expected marker number λ_I in linkage group *I* would be a sample from a Poisson's distribution:

$$\lambda_I = NL_{IE}/\sum_I L_{IE}$$

where N is the total number of markers.

The probabilities $P(N_I \leq \lambda_I)$ and $P(N_I \geq \lambda_I)$ were evaluated under the cumulative Poisson's distribution. As this is a two-tailed test, probabilities less than $\alpha/2$ correspond to deviation from Poisson's expectations of level α.

Clustering of markers on a smaller scale was also examined to identify gaps in the coverage of the current maps. Marker distribution along each linkage group was evaluated by examining markers in windows of variable sizes. The average interval size between 2 markers was calculated (G_O/N-1) for the entire map. Window boundaries were defined by a change in spacing from clustering (interval size lower than the average interval size for the entire map) to dispersion (interval size greater than the average interval). This means that consecutive markers with a distance lower than the average interval form a window (clustered) up to the distance between consecutive markers becomes greater than the average interval, starting a new window (dispersed) and so on. The number of markers in each window was compared to the null expectation for evenly dispersed markers under a cumulative Poisson's distribution using a one-tailed test with $\alpha \leq 0.05$ and $\alpha \leq 0.01$.

2.4 Alignment of maps

The linkage groups were identified thanks to the bridge SSR markers common to the previously published linkage maps of *Populus* spp. from Cervera *et al*. (2001), adopted as a standard map for *Populus*. Markers not previously mapped were searched in the *P. trichocarpa* genome database using BLAST and the sequence of the SSR primers. When we found a good homology of the 2 primers on a linkage group, separated by about 100 to 500 bases, we considered this group as expected for the considered SSR marker.

The availability of the *P. trichocarpa* genome sequence allows us to align our maps with the physical map of *P. trichocarpa*. This map is available on the cMap database (Morgante, personal communication) but the access is still reserved to the partners of the POPYOMICS project. The *P. trichocarpa* physical map was designed with the MapChart software. To simplify the representation, only potential common markers and the first and last marker of each group are indicated. The start base of each SSR was taken as reference. SSR markers not present in the cMap database but localized on the *P. trichocarpa* genome were added to the physical map and the same was done for the SNP and EST based markers. The ratio physical length/genetic length was also estimated (Chagné *et al.*, 2002). It expresses the number of bases per cM. The analysis was carried out at interval scale for the common markers between the *P. trichocarpa* physical map and the *P. nigra* genetic maps. For each two markers interval the ratio between the physical length (bp) of *P. trichocarpa* and the genetic distance (cM) of the *P. nigra* maps was calculated.

3 Results

3.1 Marker analyses

3.1.1 Analysis of AFLP markers

Forty AFLP primer combinations were used to analyze 92 individuals of the mapping progeny. A total of 533 markers were scored as heterozygous in one parent and absent in the other: 296 (55.5%) for the female parent and 237 (44.5%) for the male parent, respectively. The average number of scored markers per primer combination was 7.4 and 5.9 for the female and the male map, respectively. There was a considerable variation in the number of polymorphic AFLP markers revealed by different primer combinations ranging from 5 to 25 (Fig. 27).

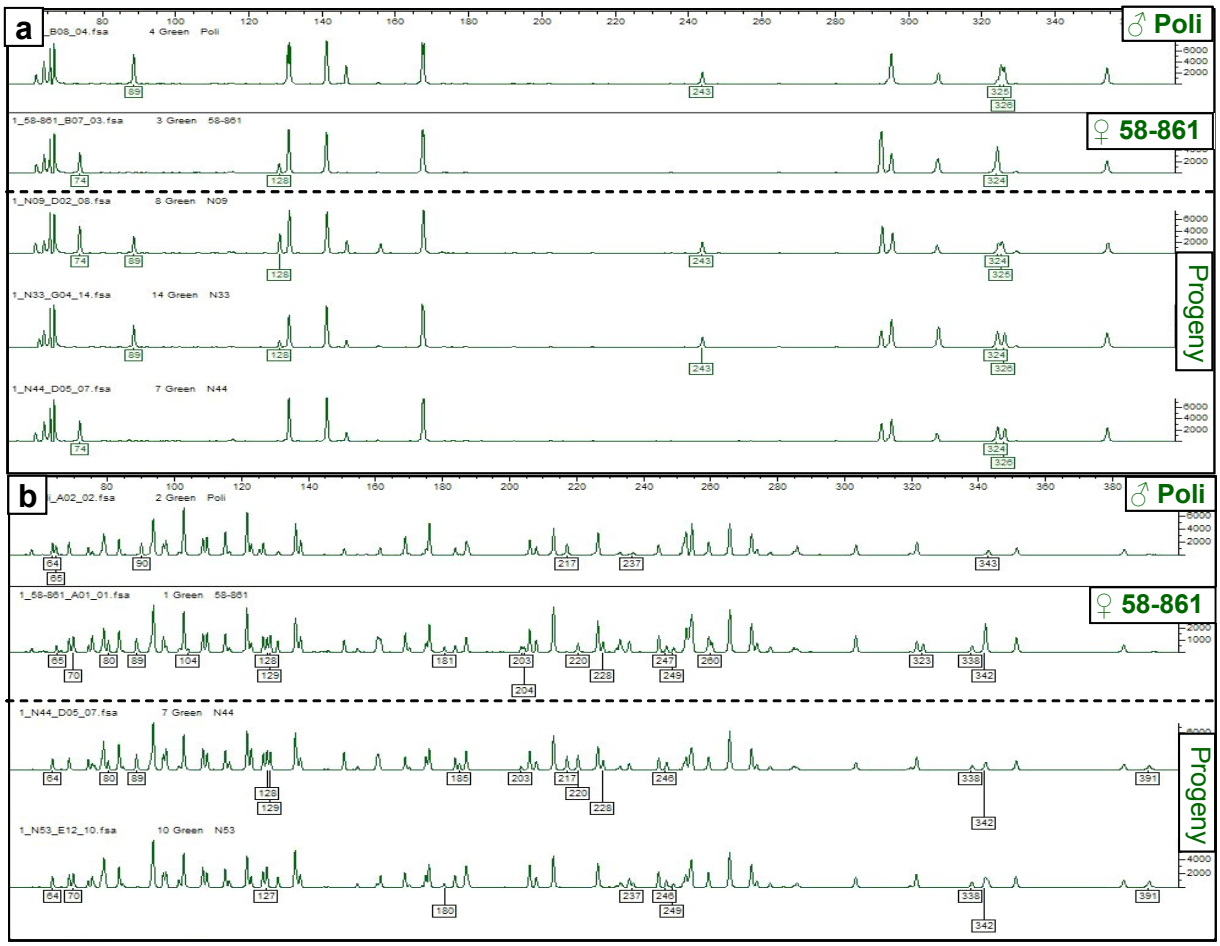


Fig. 27: Example of AFLP results
The first electropherogram represents an AFLP primer combination with a low number of polymorphic markers while the second electropherogram represents AFLP primer combination with a large number of polymorphic markers. Labels indicate the size, in bp, of the polymorphic markers.

3.1.2 Analysis of SSR and SNP markers

SSR and SNP markers are useful as genetic bridges for comparative mapping because they are locus specific and codominantly inherited.

A total number of 330 SSR primer pairs were tested: 152 SSRs from PMGC/GCPM source, 126 from ORPM, 22 from WPMS, 12 from PTR, and 16 originated from EST. These markers were screened on the 2 parents and 6 randomly selected progenies (Fig. 28). Results, sorted by SSR source, are presented in Table 15. Seventy-three (22,1%) of the markers tested did not amplify, 127 (38.8%) did not segregate and 128 (39.1%) generated segregating loci (Fig. 29).

Table 15: Number of SSR tested, segregating and mapped in the *P. nigra* pedigree.

Marker types	Tested	Not amplified	Segregating in F_1	Mapped	
				Female	Male
PMGC/GCPM	152	45	67	47	58
ORPM	126	27	32	26	25
WPMS	22	0	19	11	14
PTR	12	1	3	2	2
EST	16	0	5	3	5
PTAG	2	0	2	2	2
Total	**330**	**73**	**128**	**91**	**106**

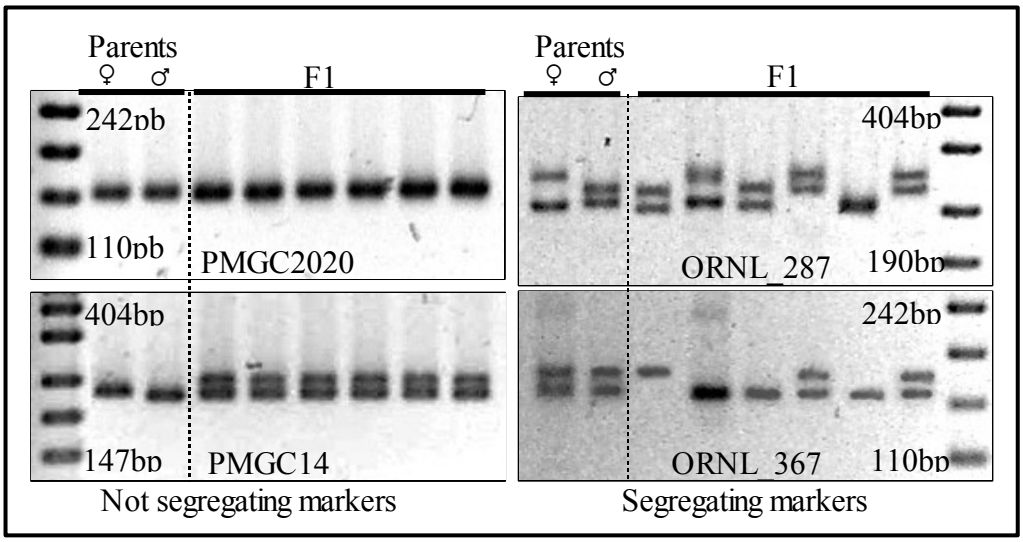


Fig. 28:. SSR screening results on the 2 parents and 6 F_1 progenies.
Monomorphic markers (PMGC2020, PMGC14) are showed on the left part of the figure and polymorphic markers (ORNL_287, ORNL_367) on the right. The amplification products were separated in a 2.5% agarose gel and visualized by ethidium bromide staining.

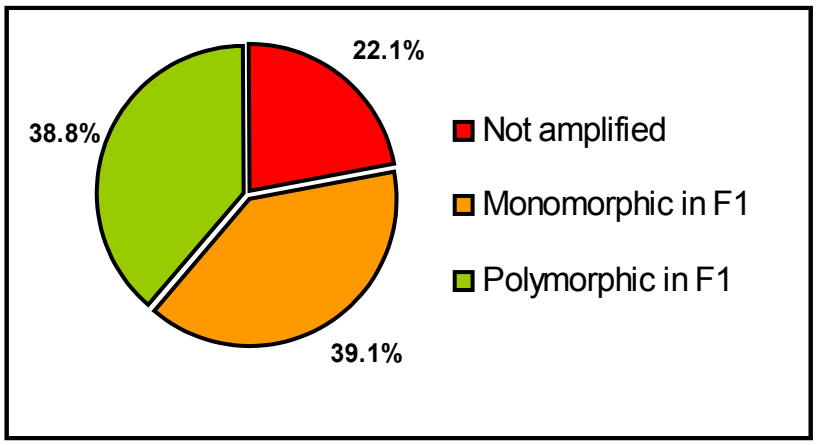


Fig. 29: Results of the screening with SSR markers

Among the 128 segregating SSRs, 13 (10.2%) were maternally informative [1:1], 21 (16.4%) were paternally informative [1:1], 86 (67.2%) were fully informative [1:1:1:1], and 8 (6.2%) segregated [1:2:1].

The sequence of the EST used to develop the ai, bu and bi SSR markers were aligned using BLASTX in order to search for homology with known genes. The results are presented in Table 16.

Table 16: Results of alignment using BLASTX.
Alignment of the EST sequence used to develop SSR markers.

EST corresponding to the SSR marker	BLASTX results Best alignment accession	Description	E Value	Identities (%)
bu810400	gi | 51971563 | dbj | BAD44446.1 |	ClpP protease complex subunit ClpR3 [*Arabidopsis thaliana*]	7e-13	70
bu885452	gi | 19116236 | gb | AAH16608.1 |	Unknown protein [*Mus musculus*]	0.81	60
bu814989	gi | 20965 | emb | CAA40072.1 |	Unnamed protein product [*Populus trichocarpa* x *Populus deltoides*]	3e-59	97
bu818855	gi | 56236462 | gb | AAV84587.1 |	ubiquitin carrier protein [*Populus tomentosa*]	7e-57	97
bu813833	gi | 22669 | gb | CAA49693.1 |	NtpII10 [*Nicotiana tabacum*]	2e-40	78

The heterozygosity levels based on SSR were 41% and 44% for the female and the male respectively.

The species-specific win3 marker (Heinze, 1997), and 4 SSRs (WPMS9, WPMS18, WPMS20, and PMGC14) (Fossati *et al.*, 2003) were analyzed on the *P. nigra* mapping pedigree to check for introgression of *P. deltoides* into *P. nigra.* Win3 gave the expected *P. nigra* specific pattern for all the individuals tested (Fig. 30). WPMS9, WPMS18, and PMGC14 did not show *P. deltoides*-specific alleles, but the WPMS20 allele of 224 bp that is rare in *P. nigra* was found in one parent. The size of alleles found with these SSR markers is presented in Table 17.

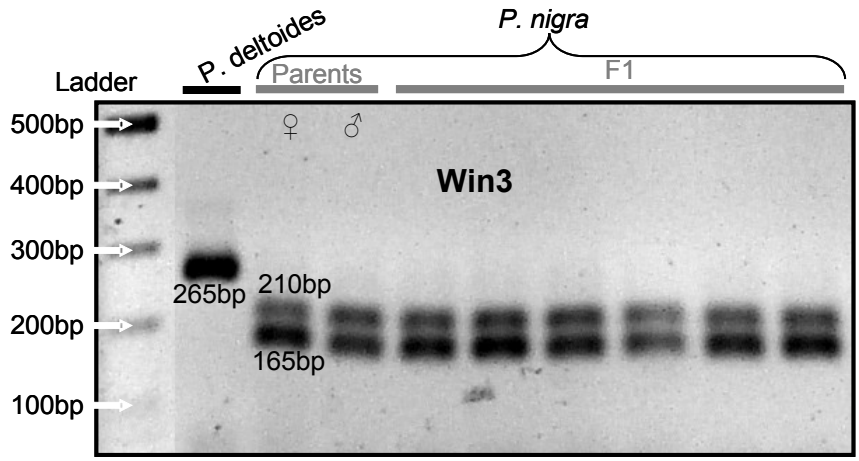


Fig. 30: Amplification pattern of Win3 marker.
Win3 amplification products were separated in a 2% agarose gel and visualized by ethidium bromide staining. The amplification products expected are one fragment of approximately 265bp from *P. deltoides* and two fragments of approximately 165 and 210bp from *P. nigra*.

Table 17: Genotyping of the mapping pedigree parents with 4 SSR markers able to detect *P. deltoides*-specific alleles.

Marker	Alleles found		*P. deltoides*-specific allele
	♀ 58-861	♂ Poli	
WPMS9	265 bp	265 bp	234 bp
WPMS18	258/239 bp	244/230 bp	220 bp
WPMS20	244/224 bp	237 bp	224 bp
PMGC14	226 bp	219 bp	193/199 bp

Seven polymorphic SNPs were selected and analyzed with the CAPS and dCAPS technique. The dCAPS is not frequently used and requires careful design of mismatched primers. The set up of dCAPS SNPs gave good results. We obtained the expected pattern for the parents (Fig. 31) and the SNPs analyzed were mapped on the expected linkage groups.

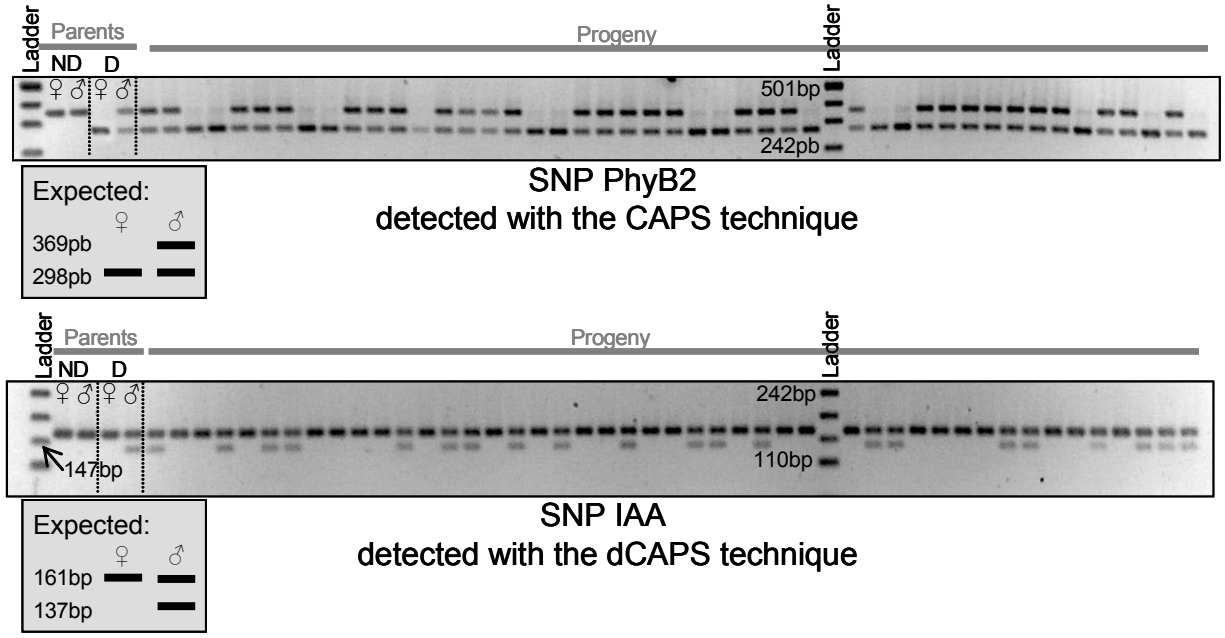


Fig. 31: Example of genotyping of 2 SNP markers.
PhyB2 and IAA revealed by the CAPS (PCR-RFLP) and dCAPS technique respectively. The digestion products were separated on a 2.5% agarose gel and visualized by ethidium bromide staining. The first 4 lanes correspond to the PCR products of the parents without digestion (ND) and after digestion (D). The expected pattern of parents after digestion is represented under the gels.

3.1.3 Sex morphological trait

In 2005, 50 (54%) individuals of the F_1 progeny, which were 4 year old, flowered and their gender was assessed. Thirty-five of them (70%) were male and 15 of them (30%) were female. Female : male ratio deviated from the expected Mendelian ratio 1:1 at 1% significance level.

A summary of the results obtained for each type of marker is presented in Table 18.

Table 18: Presentation of the results obtained for each type of marker

Features	♀ **58-861**	♂ **Poli**
AFLP markers obtained from 40 primer combinations	533	
Average AFLP markers per primer combination before and after $\chi 2$ analysis ($P<0.01$)	7.4 / 7.1	5.9 / 5.6
AFLP markers distorted at p = 0.05	29 (9.8%)	36(15.2%)
AFLP markers distorted at p = 0.01	13 (4.4%)	15 (6.3%)
AFLP markers used for linkage analysis	283	222
AFLP markers mapped	274	206
SSR markers tested	330	
SSR markers polymorphic in F_1	107	115
SSR markers distorted at p = 0.05	12 (14.4%)	16 (20.9%)
SSR markers distorted at p = 0.01	2 (2.2%)	6 (5.7%)
SSR markers used for linkage analysis	93	109
SSR markers mapped	91	106
SNP markers analyzed	7	
SNP markers used for linkage analysis	3	5
SNP markers distorted at p = 0.05	1 (33.3%)	0
SNP markers mapped	3	5
SNP markers distorted at p = 0.01	0	1 (20%)
Morphological marker (sex) used for linkage analysis	1	
Morphological marker (sex) distorted at p = 0.01	1	
Morphological marker (sex) mapped	0	1
Total polymorphic markers before $\chi 2$ analysis	**409**	**360**
Total markers distorted at p = 0.05	**42 (10.3%)**	**52 (14.4%)**
Total markers distorted at p = 0.01	**16 (4.10%)**	**23 (6.61%)**
Total markers used for linkage analysis	**380**	**337**
Total markers mapped	**368**	**316**

3.2 Linkage map construction

3.2.1 Linkage analysis

In order to generate reliable maps, 92 genotypes were analyzed. The map construction was based on the pseudo-testcross strategy, as previously described in the materials and methods, which led to the construction of two parental maps. The computations were made with MAPMAKER 3.0. The map distances were calculated with the error detection function of

Mapmaker both enabled and disabled. All the results presented were obtained with the error detection function disabled, except in Table 19, 20, and 21, which report map distances calculated with both functions.

3.2.1.1 Linkage analysis of the female data

The linkage analysis in the female parent 58-861 was based on 380 markers, including 283 AFLPs, 93 SSRs, 3 SNPs, and one morphological marker, the sex (Table 18). Using a LOD threshold of 4.0, 368 markers were initially assigned to 21 groups, leaving 13 ungrouped markers, including 2 SSRs and the morphological marker. Under these criteria, the linkage group III was separated into 2 groups, but it was possible to join them at a LOD threshold of 3.5. In this way, **368 markers, including 274 AFLPs 91 SSRs and 3 SNPs, were mapped to 20 linkage groups** (Fig. 32). The average distance between 2 markers was 7.58 cM. Linkage groups ranged from 22 to 316 cM in size (Table 19). A highly reliable marker order is needed to detect QTLs. Therefore, a framework map was obtained by excluding the markers ordered with lower reliability. This framework map consisted of 195 markers (143 AFLPs, 50 SSRs and 2 SNPs) (Table 19). The average distance between 2 markers in this map was 10.85 cM.

3.2.1.2 Linkage analysis of the male data

A total of 337 markers, including 222 AFLPs, 109 SSRs, 5 SNPs and one morphological marker, the sex (Table 18), were available for mapping in the male parent. Using a LOD threshold of 4.0, 316 markers were initially assigned to 27 groups, leaving 20 ungrouped markers, including 3 SSRs. The linkage group VI was separated into 2 groups under these criteria, but it was possible to join them at a LOD threshold of 3.5. In this way, **316 markers were mapped to 23 main linkage groups, plus 2 triplets and 1 doublet** (Fig. 32). The average distance between 2 markers was 8.88 cM. Linkage groups ranged from 6.8 to 190.5 cM in size (Table 20). The framework map consisted of 188 markers (126 AFLPs, 61 SSRs and 1 SNP) (Table 20). The average distance between 2 markers in this map was 12.33 cM.

Table 19: Female linkage map data

♀58-861		Map						Framework map					
	Linkage	Markers				Length (cM)		Markers				Length (cM)	
	Group	Tot.	SSR	SNP	AFLP	on*	off**	Tot	SSR	SNP	AFLP	on*	off**
	I	40	7	0	33	240.0	316.0	26	5	0	21	240.4	266.1
	II	27	6	0	21	155.7	214.8	13	4	0	9	159.8	176.5
	III	18	4	0	14	117.2	150.9	12	1	0	11	116.4	134.7
	IV	22	7	0	15	139.1	172.9	11	5	0	6	88.4	92.3
	V	11	5	0	6	73.4	84.0	2	1	0	1	16.9	16.9
	VI	29	9	0	20	182.3	233.9	15	5	0	10	179.4	194.1
	VII	6	1	0	5	44.5	48.8	3	0	0	3	39.6	40.0
	VIII	23	7	1	15	106.1	156.0	12	5	0	7	110.2	121.4
	IX	25	5	0	20	85.8	120.4	11	2	0	9	87.3	96.7
	X	19	5	0	14	135.7	174.1	7	4	0	3	104.1	111.8
	XI	11	2	0	9	80.6	101.6	7	0	0	7	79.3	85.8
	XII	17	5	0	12	74.4	114.1	9	3	0	6	72.7	80.8
	XIII	24	10	1	13	137.6	212.0	14	4	1	9	130.6	165.5
	XIV	14	0	1	13	113.9	135.0	9	0	1	8	103.1	109.6
	XV	10	3	0	7	69.9	76.4	5	2	0	3	65.8	66.6
	XVI	22	3	0	19	86.1	111.7	13	2	0	11	82.9	84.3
	XVII	20	4	0	16	82	145.5	8	2	0	6	67.1	77.2
	XVIII	14	5	0	9	81.5	97.8	7	2	0	5	74.6	80.1
	XIX	11	3	0	8	80.4	100.9	9	3	0	6	82.6	99
	A	5	0	0	5	17.9	22.2	2	0	0	2	16.1	16.1
Total	20	368	91	3	274	2104.1	2789.0	195	50	2	143	1917	2115.5

* Length with the error detection function of Mapmaker enabled

** Length with the error detection function of Mapmaker disabled.

Table 20: Male linkage map data

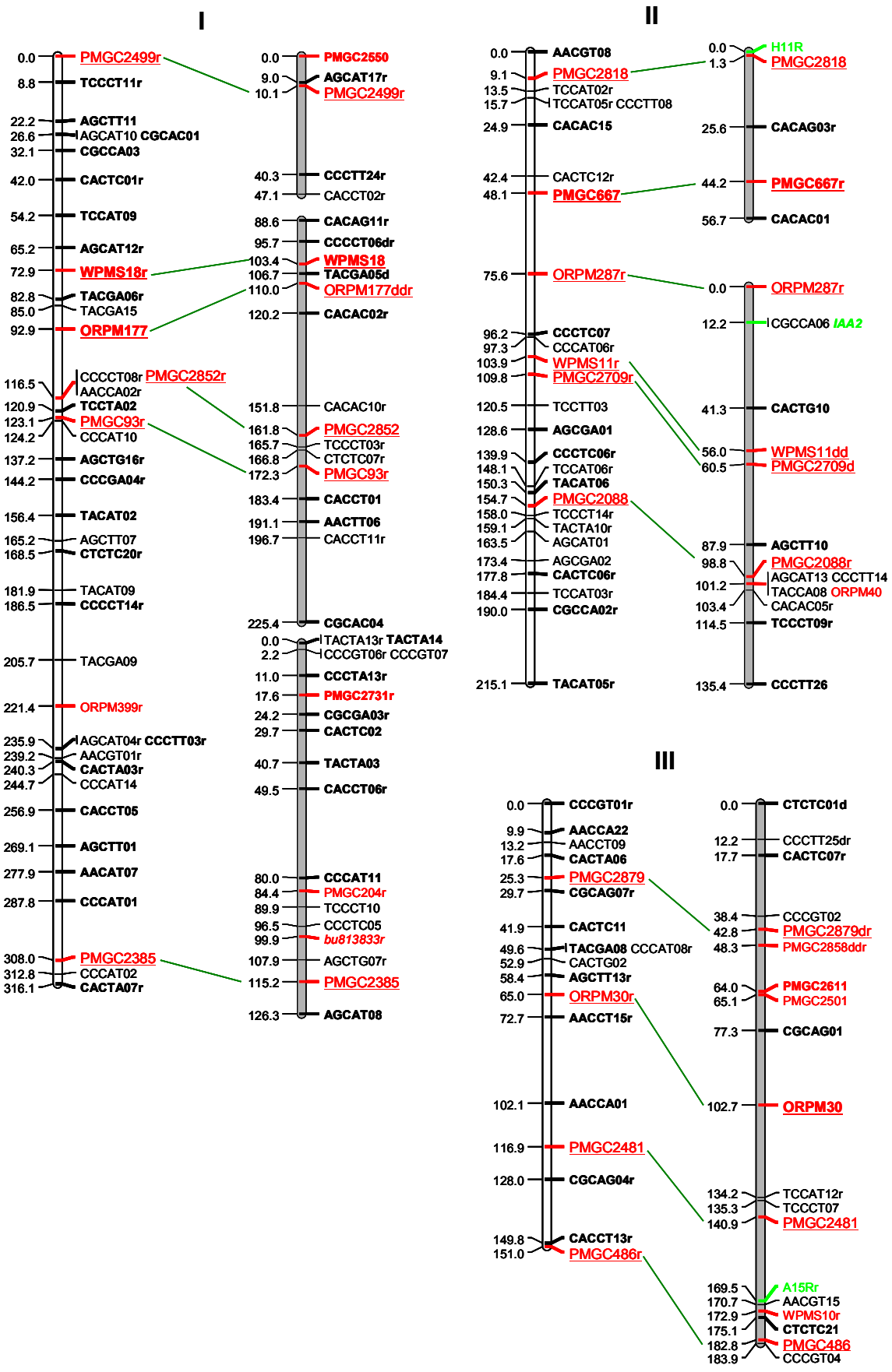
♂Poli		Map						Framework map					
	Linkage	Markers				Length (cM)		Markers				Length (cM)	
	Group	Tot.	SSR	SNP	AFLP	on*	off**	Tot.	SSR	SNP	AFLP	on*	off**
	Ia	5	2	0	3	47.4	47.1	3	1	0	2	40.0	39.9
	Ib	15	4	0	11	118.2	136.9	10	3	0	7	106.0	112.6
	Ic	18	4	0	14	103.6	126.2	11	2	0	9	96.6	102.0
	IIa	15	5	1	9	121.6	135.4	8	3	1	4	120.3	126.7
	IIb	5	2	1	2	51.8	56.7	4	2	0	2	51.2	53.7
	III	19	8	1	10	155	183.7	8	4	0	4	136.9	146.6
	IVa	2	1	0	1	6.8	6.8	2	1	0	1	6.8	6.8
	IVb	11	4	0	7	123.4	131.6	7	1	0	6	73.5	78.9
	Va	19	9	0	10	114.9	137.2	11	4	0	7	110.6	117.4
	Vb	6	1	0	5	45.7	50.7	3	1	0	2	13.1	13.1
	VI	17	7	0	10	173.9	190.5	11	6	0	5	147.5	153.0
	VII	12	4	0	8	64.1	82.5	5	2	0	3	53.4	58.3
	VIII	15	7	0	8	152.2	182.4	9	4	0	5	145.6	152.3
	IX	9	5	0	4	95.9	101.1	6	2	0	4	92.1	95.1
	X	21	6	1	14	125.6	163.5	12	2	0	10	112.3	123.7
	XI	15	6	0	9	122.1	138.5	10	4	0	6	119.1	131.1
	XII	12	3	0	9	102.2	111.1	9	3	0	6	97.0	101.7
	XIII	20	7	0	13	107.8	126.4	8	5	0	3	88.5	92.6
	XIV	12	2	1	9	105.3	113.2	8	1	0	7	100.7	101.7
	XV	9	3	0	6	82.0	87.9	5	1	0	4	75.8	79.4
	XVI	18	5	0	13	92.8	110.5	10	5	0	5	91.0	96.2
	XVII	17	5	0	12	127.6	147.2	11	1	0	10	127.9	137.8
	XVIII	9	3	0	6	75.4	83.6	5	1	0	4	60.7	65.8
	XIX	10	3	1	6	95.8	108.4	6	2	0	4	72.6	75.7
	A	3	0	0	3	17.6	17.7	3	0	0	3	17.6	17.7
	B	3	0	0	3	38.6	39.1	3	0	0	3	38.6	39.1
Total	26	317	106	6	205	2467.3	2815.9	188	61	1	126	2195.4	2318.9

***** Length with the error detection function of Mapmaker enabled

****** Length with the error detection function of Mapmaker disabled.

Fig. 32: Linkage maps of the *P. nigra* pedigree.

The linkage groups on the left (in white) result from the female parent and those on the right (in grey) from the male parent. Framework markers are in bold, AFLP markers in black, SSR markers in red, SNP markers in green and the morphological marker in blue. Markers corresponding to genes are in italic. Markers with a distorted segregation ratio are indicated by "d" ($0.01<P<0.05$) and "dd" ($P>0.01$) suffix. Inverted markers are indicated by a "r" suffix and represent markers in repulsion phase. Markers in common between the 2 maps are underlined and in larger letters. They are indicated with allelic bridges. Linkage groups aligned with the maps of Cervera *et al* (2001) or with the physical map of *P. trichocarpa* are labeled by a roman numeral. The length of the linkage group bars is proportional to the distance in cM (0.5 mm per cM).


I
0.0 PMGC2499r
8.8 TCCCT11r
22.2 AGCTT11
26.6 AGCAT10 CGCAC01
32.1 CGCCA03
42.0 CACTC01r
54.2 TCCAT09
65.2 AGCAT12r
72.9 WPMS18r
82.8 TACGA06r
85.0 TACGA15
92.9 ORPM177
116.5 CCCCT08r PMGC2852r
AACCA02r
120.9 TCCTA02
123.1 PMGC93r
124.2 CCCAT10
137.2 AGCTG16r
144.2 CCCGA04r
156.4 TACAT02
165.2 AGCTT07
168.5 CTCTC20r
181.9 TACAT09
186.5 CCCCT14r
205.7 TACGA09
221.4 ORPM399r
235.9 AGCAT04r CCCTT03r
239.2 AACGT01r
240.3 CACTA03r
244.7 CCCAT14
256.9 CACCT05
269.1 AGCTT01
277.9 AACAT07
287.8 CCCAT01
308.0 PMGC2385
312.8 CCCAT02
316.1 CACTA07r
0.0 PMGC2550
9.0 AGCAT17r
10.1 PMGC2499r
40.3 CCCTT24r
47.1 CACCT02r
88.6 CACAG11r
95.7 CCCCT06dr
103.4 WPMS18
106.7 TACGA05d
110.0 ORPM177ddr
120.2 CACAC02r
151.8 CACAC10r
161.8 PMGC2852
165.7 TCCCT03r
166.8 CTCTC07r
172.3 PMGC93r
183.4 CACCT01
191.1 AACTT06
196.7 CACCT11r
225.4 CGCAC04
0.0 TACTA13r TACTA14
2.2 CCCGT06r CCCGT07
11.0 CCCTA13r
17.6 PMGC2731r
24.2 CGCGA03r
29.7 CACTC02
40.7 TACTA03
49.5 CACCT06r
80.0 CCCAT11
84.4 PMGC204r
89.9 TCCCT10
96.5 CCCTC05
99.9 bu813833r
107.9 AGCTG07r
115.2 PMGC2385
126.3 AGCAT08
II
0.0 AACGT08
9.1 PMGC2818
13.5 TCCAT02r
15.7 TCCAT05r CCCTT08
24.9 CACAC15
42.4 CACTC12r
48.1 PMGC667
75.6 ORPM287r
96.2 CCCTC07
97.3 CCCAT06r
103.9 WPMS11r
109.8 PMGC2709r
120.5 TCCTT03
128.6 AGCGA01
139.9 CCCTC06r
148.1 TCCAT06r
150.3 TACAT06
154.7 PMGC2088
158.0 TCCCT14r
159.1 TACTA10r
163.5 AGCAT01
173.4 AGCGA02
177.8 CACTC06r
184.4 TCCAT03r
190.0 CGCCA02r
215.1 TACAT05r
0.0 H11R
1.3 PMGC2818
25.6 CACAG03r
44.2 PMGC667r
56.7 CACAC01
0.0 ORPM287r
12.2 CGCCA06 IAA2
41.3 CACTG10
56.0 WPMS11dd
60.5 PMGC2709d
87.9 AGCTT10
98.8 PMGC2088r
AGCAT13 CCCTT14
101.2 TACCA08 ORPM40
103.4 CACAC05r
114.5 TCCCT09r
135.4 CCCTT26
III
0.0 CCCGT01r
9.9 AACCA22
13.2 AACCT09
17.6 CACTA06
25.3 PMGC2879
29.7 CGCAG07r
41.9 CACTC11
49.6 TACGA08 CCCAT08r
52.9 CACTG02
58.4 AGCTT13r
65.0 ORPM30r
72.7 AACCT15r
102.1 AACCA01
116.9 PMGC2481
128.0 CGCAG04r
149.8 CACCT13r
151.0 PMGC486r
0.0 CTCTC01d
12.2 CCCTT25dr
17.7 CACTC07r
38.4 CCCGT02
42.8 PMGC2879dr
48.3 PMGC2858ddr
64.0 PMGC2611
65.1 PMGC2501
77.3 CGCAG01
102.7 ORPM30
134.2 TCCAT12r
135.3 TCCCT07
140.9 PMGC2481
169.5 A15Rr
170.7 AACGT15
172.9 WPMS10r
175.1 CTCTC21
182.8 PMGC486
183.9 CCCGT04

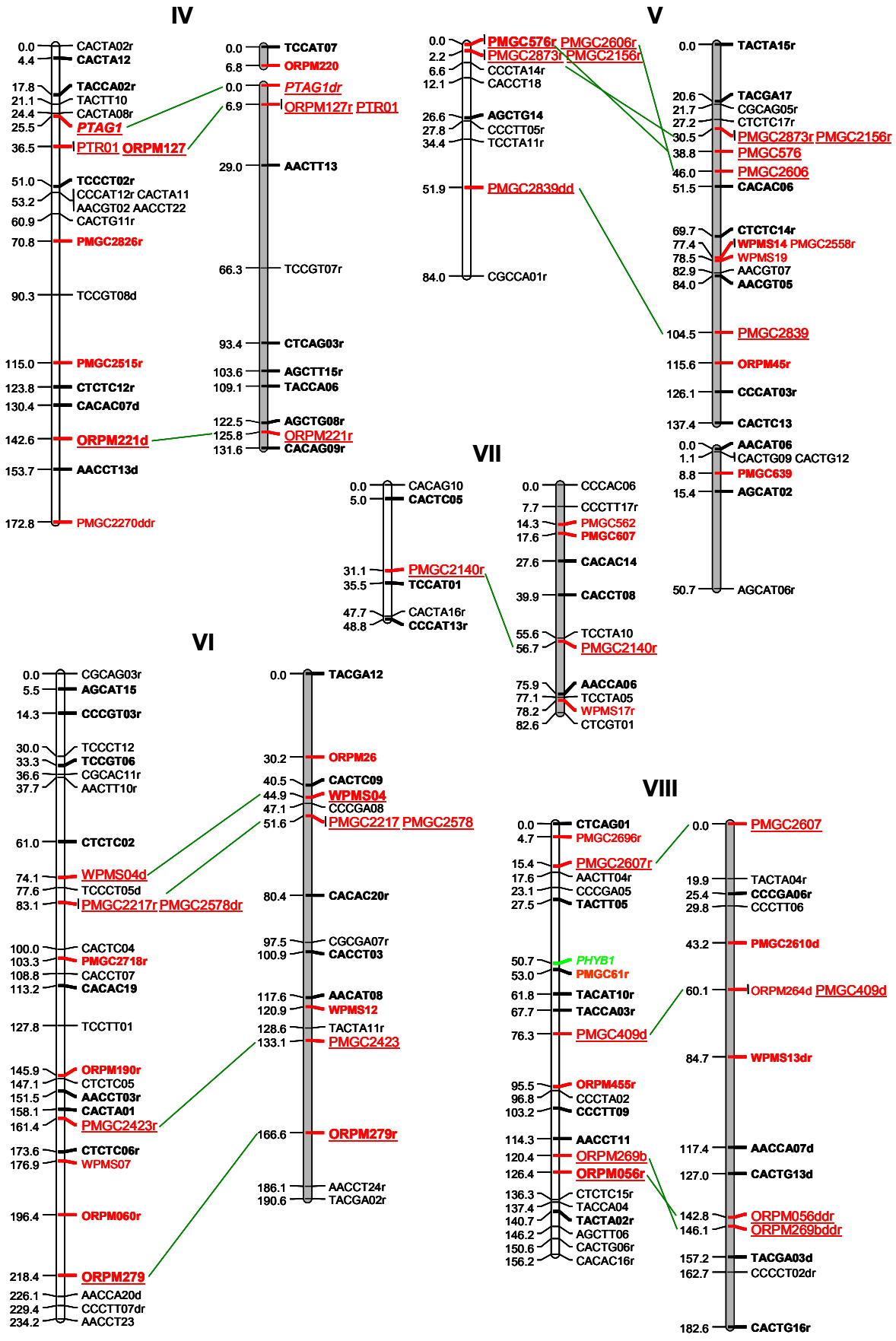


Fig. 32: Continued

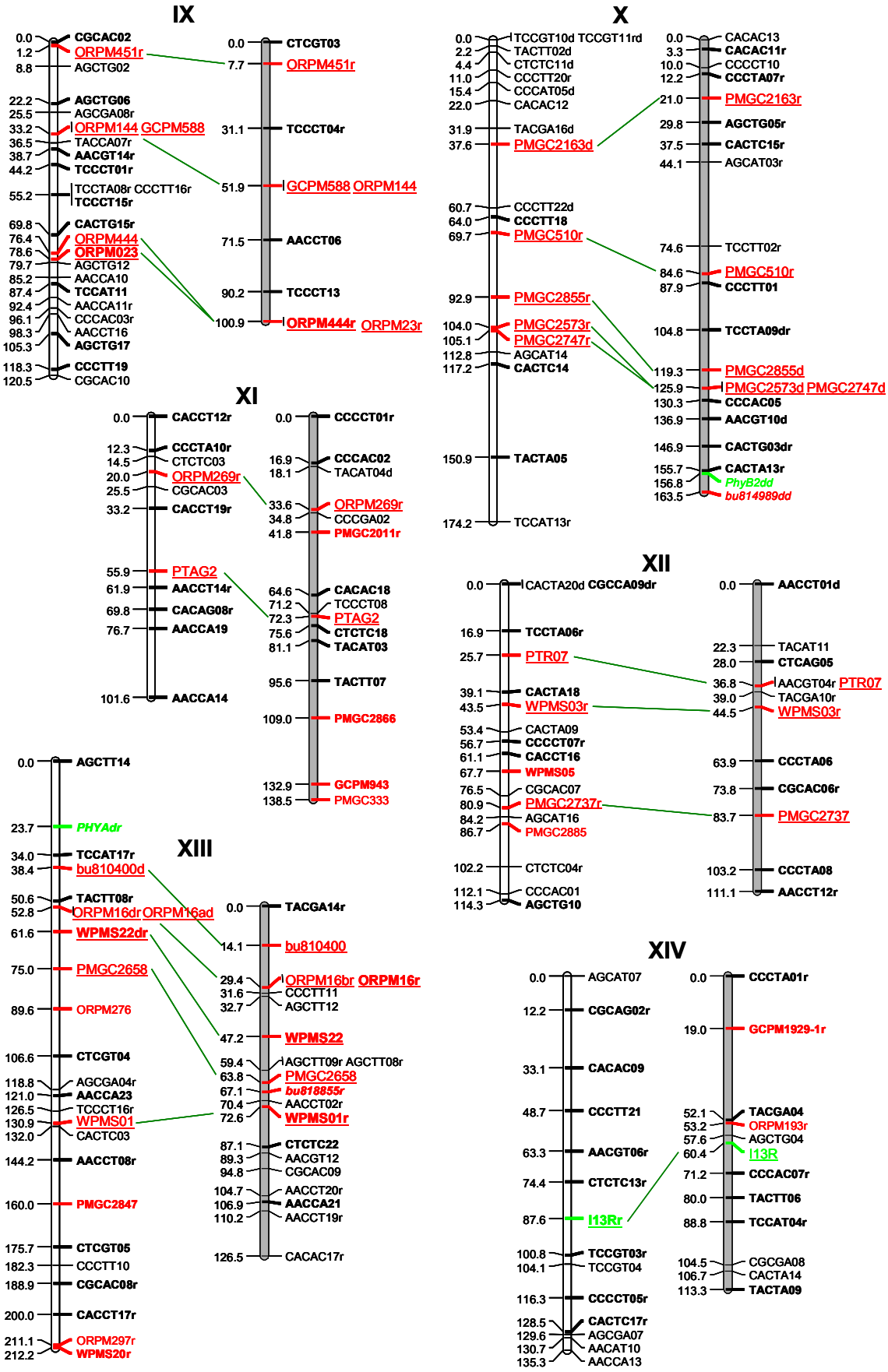


Fig. 32: Continued

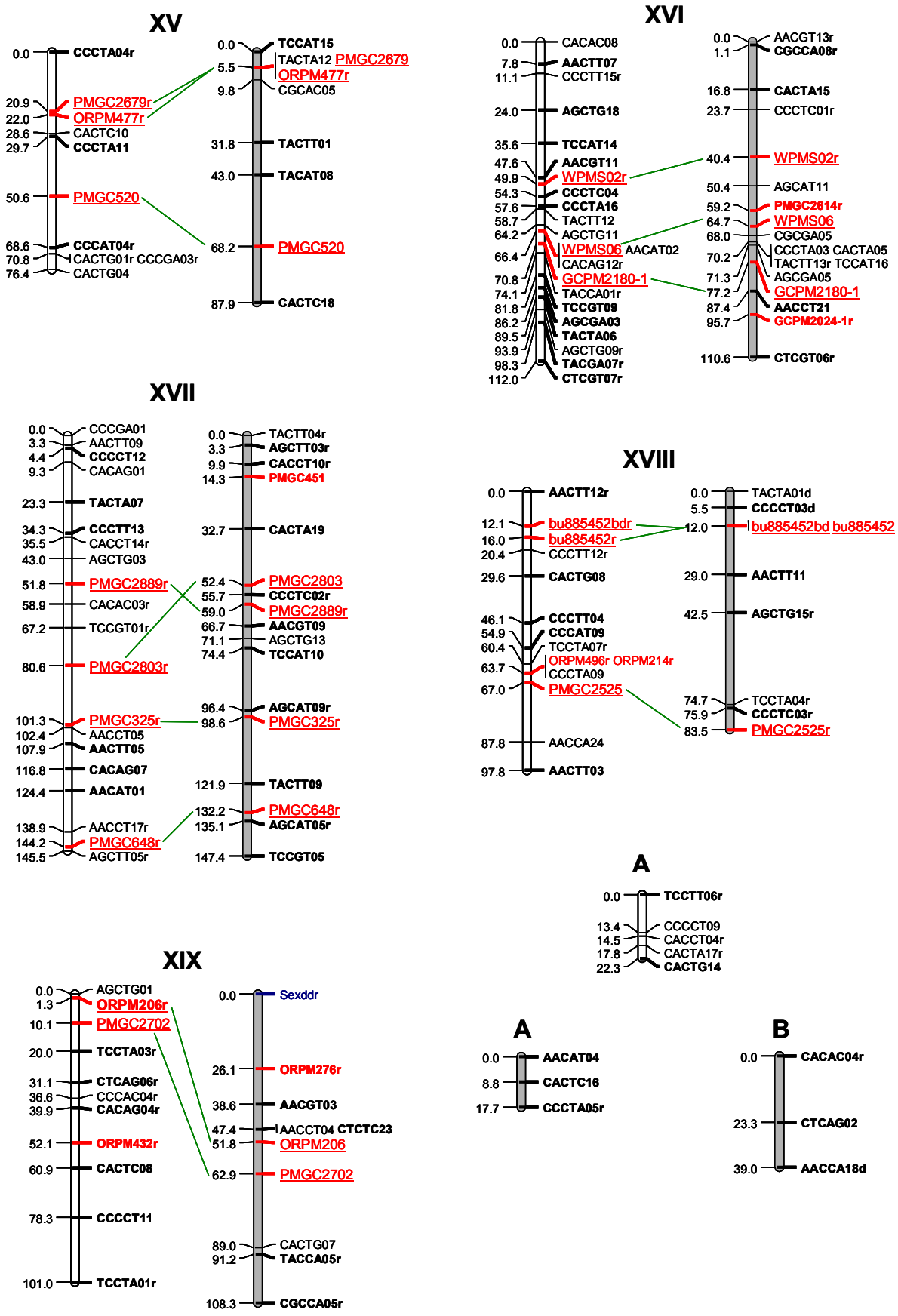


Fig. 32: Continued

3.2.2 Estimated and observed genome length

The estimated and observed genome length of the 2 maps are presented in Table 21. The values of the observed genome length for the 2 parents were in the same range: 2789 cM for the female map and 2815.9 cM for the male map. The observed genome length G_{ON} (see Materials and Methods), which takes in account all linked and unlinked markers, gave similar results between the value 3823.7 cM corresponding to the estimated genome length calculated according to Nelson *et al.* (1994) and the value 3572 cM estimated for the male map. The female map showed a G_{ON} value in the same range (3227.1 cM) but lower than the value obtained for the male map. The estimated genome length calculated according to Chakravarti *et al* (1991) (G_E; see Materials and Methods) was closer to the observed genome length when the number of linkage groups is close to the haploid chromosome number (19) than when the number of linkage groups is higher. Actually, the difference between the genome length observed and estimated was 62.7 cM for the female map composed by 20 linkage groups, whereas the difference was 756.1 cM for the male map, which had 26 linkage groups.

Table 21: Observed and estimated genome length of the 2 parental maps.

Observed genome length	♀ **58-861**		♂ **Poli**	
	OFF	ON	OFF	ON
G_O (cM)	2789	2104.1	2815.9	2467.3
X_o (cM)	33.7	34.4	37.3	34.6
L	32		46	
R	19		19	
G_{ON} (cM)	3227.1	2551.3	3823.0	3401.5
Framework map				
G_O (cM)	2115.5	1917.3	2318.9	2195.4
X_o (cM)	35.1	34.4	33.3	33.4
L	32		46	
R	19		19	
G_{ON} (cM)	2571.8	2364.5	3097.2	3218
Estimated genome lenght	♀ **58-861**		♂ **Poli**	
N	368		317	
X_E (cM)	34.7		34.7	
K	1717		972	
G_E	2726.3		3572	
Lower bound	2603.2		3360.7	
Higher bound	2801.6		3811.6	
Framework map				
N	195		188	
X_E (cM)	34.7		34.7	
K	433		295	
G_E	3028.1		4130.5	
Lower bound	2767.5		3707.5	
Higher bound	3343		4662.6	

X_o, observed maximum distance between 2 markers; L, total number of linkage groups, pairs, and unlinked markers; R, haploid number of chromosomes in poplar; G_O, observed genome length, based on all markers (Map) and framework markers (Framework map); G_{ON}, observed genome length, calculated according to Nelson *et al* (1994); OFF/ON, data column with the error detection function of Mapmaker enabled/ disabled.
N, number of markers; K, observed number of locus pairs with minimum LOD scores of 4.0; X_E, corresponding maximum distance between the locus pairs; G_E, estimated genome length

Far all equations, see Materials and Methods.

3.2.3 Segregation distortion

A χ^2 test (d.f.=1) was performed to test the null hypothesis of a 1:1 segregation ratio of the marker alleles. At 5% significant level, 42 (10.27%) and 52 (14.44%) of the markers analyzed were distorted for the female and the male, respectively (Table 18). A χ^2 test shown that the difference between the 2 parents was not significant. Thirteen (4.39%) and 15 (6.33%) AFLP markers belonging to the female and the male map, respectively, and deviating at 1% significance level were excluded from the linkage analysis because of a major risk of technical artifacts. Therefore, for the female data, 29 distorted markers (0.01<P<0.05) were included in the mapping data set and 28 of them were mapped, whereas 37 distorted markers (P<0.05) were included in the male mapping data set and 34 of them were mapped. These distorted markers were not uniformly distributed along the genome. They were clustered in 5 linkage groups (IV, VI, X, XII, XIII) for the female parent and in 6 linkage groups (I, II, III, VIII, X, XVIII) for male parent as showed in Table 22.

Table 22: Distribution of distorted markers in *P. nigra* maps.

	58-861			Poli		
	30 significantly distorted markers (p<0.05) analyzed			38 significantly distorted markers (p<0.05) analyzed		
LG	**Total markers**	**Distorted**	**%**	**Total markers**	**Distorted**	**%**
I	40	0	0	38	3	7.89
II	27	0	0	20	2	10.00
III	18	0	0	19	4	21.05
IV	22	5	22.73	13	1	7.69
V	11	1	9.09	25	0	0
VI	29	5	17.24	17	0	0
VII	6	0	0.00	12	0	0
VIII	23	1	4.35	15	9	60.00
IX	25	0	0	9	0	0
X	19	8	42.11	21	8	38.10
XI	11	0	0.00	15	1	6.67
XII	17	2	11.76	12	1	8.33
XIII	24	5	20.83	20	0	0
XIV	14	0	0	12	0	0
XV	10	0	0	9	0	0
XVI	22	0	0	18	0	0
XVII	20	0	0	17	0	0
XVIII	14	1	7.14	9	3	33.33
XIX	11	0	0	10	1	10.00
Total	368	28	-	317	34	-

The linkage group X showed distortion regions in the 2 parents but not for the same segments (Fig. 33). In the female map the linkage group IV presented a large region of distortion covering 82.5 cM, corresponding to 47.74% of the group length. In the male map the linkage groups VIII presented also a very large region of distortion covering 119.5 cM, corresponding to 76.34% of the group length (Fig. 33).

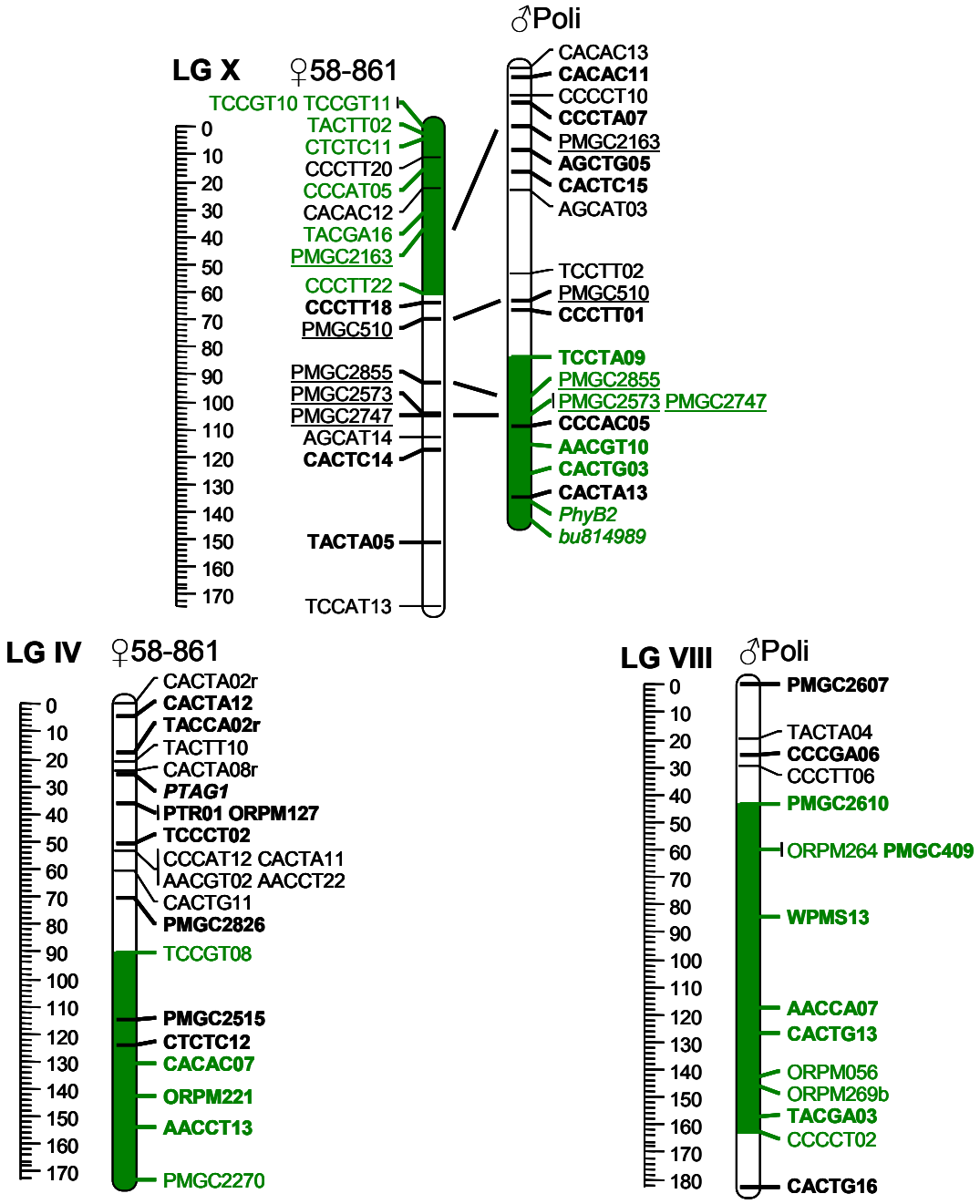


Fig. 33: Distorted region on linkage groups.
Distorted markers are indicated in green as well as the corresponding region of distortion on the linkage groups (LG). Example of the linkage groups X, IV, VIII, showing the largest regions of distortion.

3.2.4 Marker distribution

The distribution of markers among the linkage groups was analyzed to search for any difference in marker density. A two-tailed Poisson test was performed to compare the observed number of markers (N_I) of each linkage group to the expected number ($\lambda_I = 368L_I/3664$, $\lambda_I = 312L_I/3521$ for the female and male map respectively) under the null hypothesis of homogenous marker density among groups. Poisson probabilities observed for the deviation of N_I from λ_I in both directions were greater than 0.025 for all linkage groups for the female and the male map (Table 23 and 24). Thus, we did not detect any significant differences in marker density among linkage groups at a significance level of 0.05.

Table 23: Marker density by linkage group for the female map.

♀ **58-861**		S = 14.38 cM			
Linkage group	Number of markers, N_I	Map length, L_I (cM)	Inferred LG length, L_{IE} (cM)	Expected number of markers λ_I	Poisson two-tailed P-value[a]
I	40	316	344.75	37.71	0.376
II	27	214.8	243.55	26.64	0.498
III	18	150.9	179.65	19.65	0.411
IV	22	172.9	201.65	22.06	0.551
V	11	84	112.75	12.33	0.424
VI	29	233.9	262.65	28.73	0.505
VII	6	48.8	77.55	8.48	0.258
VIII	23	156	184.75	20.21	0.296
IX	25	120.4	149.15	16.32	0.027
X	19	174.1	202.85	22.19	0.292
XI	11	101.6	130.35	14.26	0.239
XII	17	114.1	142.85	15.63	0.397
XIII	24	212	240.75	26.34	0.371
XIV	14	135	163.75	17.91	0.214
XV	10	76.4	105.15	11.50	0.402
XVI	22	111.7	140.45	15.36	0.065
XVII	20	145.5	174.25	19.06	0.445
XVIII	14	97.8	126.55	13.84	0.518
XIX	11	100.9	129.65	14.18	0.245
A	5	22.2	50.95	5.57	0.517
Total	368	2789	3364.1	368	-

S, average framework marker spacing; N_I, number of markers observed in linkage group I; L_I, length of linkage group I; L_{IE}, estimated length of linkage group I; λ_I, expected marker number under Poisson distribution in linkage group I.

[a] Poisson probability of having as many (for $N_I \geq \lambda_I$) or as few (for $N_I < \lambda_I$) markers as the observed number N_I in linkage group I under the null hypothesis that the true marker density is the same for all the linkage groups. As this is a two-tailed test, a p-value of 0.025 correspond to a significance level of 0.05.

Far all the equations, see Materials and Methods.

Table 24: Marker density by linkage group for the male map.

Poli Linkage group	S = 15.05 cM Number of markers, N_I	Map length, L_I (cM)	Inferred LG length, L_{IE} (cM)	Expected number of markers λI	Poisson two-tailed P-value[a]
Ia	5	47.1	77.20	6.84	0.322
Ib	15	136.9	167.00	14.80	0.514
Ic	18	126.2	156.30	13.85	0.327
IIa	15	135.4	165.50	14.66	0.499
IIb	5	56.7	86.80	7.69	0.221
III	19	183.7	213.80	18.94	0.525
IVa	2	6.8	36.90	3.27	0.365
IVb	11	131.6	161.70	14.33	0.233
Va	19	137.2	167.30	14.82	0.168
Vb	6	50.7	80.80	7.16	0.426
VI	17	190.5	220.60	19.55	0.332
VII	12	82.5	112.60	9.98	0.301
VIII	15	182.4	212.50	18.83	0.226
IX	9	101.1	131.20	11.62	0.277
X	21	163.5	193.60	17.15	0.205
XI	15	138.5	168.60	14.94	0.528
XII	12	111.1	141.20	12.51	0.518
XIII	20	126.4	156.50	13.87	0.071
XIV	12	113.2	143.30	12.70	0.496
XV	9	87.9	118.00	10.46	0.402
XVI	18	110.5	140.60	12.46	0.082
XVII	17	147.2	177.30	15.71	0.405
XVIII	9	83.6	113.70	10.07	0.449
XIX	10	108.4	138.50	12.27	0.32
A	3	17.7	47.80	4.23	0.390
B	3	39.1	69.20	6.13	0.140
Total	312	2815.9	3611.7	312.00	-

S, average framework marker spacing; N_I, number of markers observed in linkage group *I*; L_I, length of linkage group *I*; L_{IE}, estimated length of linkage group *I*; λ_I, expected marker number under Poisson distribution in linkage group *I*.

[a] Poisson probability of having as many (for $N_I \geq \lambda_I$) or as few (for $N_I < \lambda_I$) markers as the observed number N_I in linkage group *I* under the null hypothesis that the true marker density is the same for all the linkage groups. As this is a two-tailed test, a p-value of 0.025 correspond to a significance level of 0.05.

Far all the equations, see Materials and Methods.

The tests for marker distribution at a linkage group scale revealed the presence of clustered and dispersed regions:

- in the female map, the marker density was significantly higher than expected in 8 linkage groups: I, II, IV, V, IX, XIII, XVI, XVIII, and significantly lower than expected in 5 linkage groups: II, III, V, X, and XIV (Table 25). The clustered markers represented 13.57% of the mapped markers and only 4.42% of the map length while the dispersed markers represented 4.89% of mapped markers and 12.29% of the map length.

- In the male map, the marker density was significantly higher than expected in 7 linkage groups: IIa, III, Va, VI, VII XIII XVI, and significantly lower than expected in 3 linkage groups: III, IVb, VIII (Table 26). The clustered markers represented 12.50% of the mapped markers and only 3.35% of the map length, while the dispersed markers represented 3.52% of the mapped markers and 8.82% of the map length.

Table 25: Clustered and dispersed regions on each linkage group in the female map.

Regions with clustered markers														
LG	λ_{IJ}	b_{IJ}	$\sum b_{.J}$	**P**	b_I	b_I %	I_{IJ}	E_{IJ}	x_{IJ}	$\sum x_{IJ}$	L_I	L_I %	A x_{IJ}	A L_I
I	1.01	5	9	0.004 **	40	22.5	116.5	123.1	7.7	16.5	316.0	5.2	1.9	8.1
	1.16	4		0.030 *	40		235.9	240.3	8.8				2.9	
II	2.03	6	6	0.018 *	27	22.2	148.1	159.1	15.4	15.4	214.8	7.2	3.1	8.3
IV	0.29	4	4	0.000 **	22	18.2	51.0	53.2	2.2	2.2	172.9	1.3	0.7	8.2
V	1.59	5	5	0.023 *	11	45.5	0.0	6.6	12.1	12.1	84.0	14.4	3.0	8.4
IX	4.67	9	9	0.049 *	25	36.0	69.8	98.3	35.5	35.5	120.4	29.5	4.4	5.0
XIII	0.29	2	2	0.035 *	24	8.3	50.6	52.8	2.2	2.2	212.0	1.0	2.2	9.2
XVI	3.49	10	10	0.003 **	22	45.5	47.6	70.8	26.5	26.5	111.7	23.7	2.9	5.3
XVIII	1.59	5	5	0.023 *	14	35.7	54.9	63.7	12.1	12.1	97.8	12.4	3.0	7.5
Regions with dispersed markers														
II	6.33	2	2	0.049 *	27	7.4	48.1	75.6	48.1	48.1	214.8	22.4	48.1	8.3
III	11.16	5	5	0.034 *	18	27.8	65.0	128.0	84.8	84.8	150.9	56.2	21.2	8.9
V	6.53	2	2	0.042 *	11	18.2	34.4	84.0	49.6	49.6	84.0	59.0	49.6	8.4
X	7.50	2	2	0.020 *	19	10.5	117.2	150.9	57.0	57.0	174.1	32.7	57.0	9.7
XIV	13.26	7	7	0.047 *	14	50.0	0.0	87.6	100.8	100.8	135.0	74.7	16.8	10.4

** significant at α=0.01; * significant at α=0.05;

λ_{IJ} is the expected number of markers in a distance of x_{IJ};

b_{IJ} is the observed number of markers in a distance of x_{IJ};

$\sum b_{.J}$ is the sum of b_{IJ} on linkage group *I*

P is the Poisson one-tailed P-value of $b_{IJ} \leq \lambda_{IJ}$ or $b_{IJ} \geq \lambda_{IJ}$;

b_I is the total number of markers on linkage group *I*;

b_I **%** is the percentage of markers clustered or in lower density than expected;

I_{IJ} is the start position x_{IJ} on linkage group *I*;

E_{IJ} is the end position of x_{IJ} on linkage group *I*;

x_{IJ} is the *J* distance formed by continuous intervals which size is larger or smaller than average on linkage group *I*;

$\sum x_{IJ}$ the sum of x_{IJ} on linkage group *I*;

L_I is the map distance between terminal markers of linkage group *I*;

L_I **%** is the percentage of regions with clustered or dispersed markers on linkage group *I*;

A x_{IJ} is the average interval size of x_{IJ};

A L_I is the average interval size of linkage group *I*

Table 26: Clustered and dispersed regions on each linkage group in the male map.

LG	λ_{IJ}	b_{IJ}	$\sum b_{.J}$	P		b_I	b_I %	I_{IJ}	E_{IJ}	x_{IJ}	$\sum x_{IJ}$	L_I	L_I %	A x_{IJ}	A L_I
Regions with clustered markers															
IIa	0.52	5	5	0.000	**	15	33.3	98.8	101.2	4.6	4.6	135.4	3.4	1.2	9.7
III	1.62	5	5	0.025	*	19	26.3	169.5	183.9	14.4	14.4	183.7	7.8	3.6	10.2
Va	1.60	5	5	0.024	*	19	26.3	69.7	82.9	14.3	14.3	137.2	10.4	3.6	7.6
VI	1.25	4	4	0.038	*	17	23.5	40.5	51.6	11.1	11.1	190.5	5.8	3.7	11.9
VII	0.75	3	3	0.045	*	12	25.0	75.9	82.6	6.7	6.7	82.5	8.1	3.4	7.5
XIII	0.37	3	8	0.006	**	20	40.0	29.4	31.6	3.3	16.5	126.4	13.1	1.7	6.7
	1.48	5		0.018	*	20		59.4	70.4	13.2				3.3	
XVI	3.01	9	9	0.004	**	18	50.0	50.4	71.3	26.8	26.8	110.5	24.3	3.4	6.5
Regions with dispersed markers															
III	7.75	3	3	0.050	*	19	15.8	65.1	102.7	69.1	69.1	183.7	37.6	34.6	10.2
IVb	10.85	4	4	0.017	*	11	36.4	6.9	93.4	96.7	96.7	131.6	73.5	32.2	13.2
VIII	9.28	4	4	0.046	*	15	26.7	60.1	127.0	82.7	82.7	182.4	45.3	27.6	13.0

** significant at a=0.01; * significant at a=0.05;

λ_{IJ} is the expected number of markers in a distance of x_{IJ};

b_{IJ} is the observed number of markers in a distance of x_{IJ};

$\sum b_{.J}$ is the sum of b_{IJ} on linkage group *I*

P is the Poisson one-tailed P-value of $b_{IJ} \leq \lambda_{IJ}$ or $b_{IJ} \geq \lambda_{IJ}$;

b_I is the total number of markers on linkage group *I*;

b_I **%** is the percentage of markers clustered or in lower density than expected;

I_{IJ} is the start position x_{IJ} on linkage group *I*;

E_{IJ} is the end position of x_{IJ} on linkage group *I*;

x_{IJ} is the *J* distance formed by continuous intervals which size is larger or smaller than average on linkage group *I*;

$\sum x_{IJ}$ the sum of x_{IJ} on linkage group *I*;

L_I is the map distance between terminal markers of linkage group *I*;

L_I **%** is the percentage of regions with clustered or dispersed markers on linkage group *I*;

A x_{IJ} is the average interval size of x_{IJ};

A L_I is the average interval size of linkage group *I*

3.3 Map comparisons

The 2 *P. nigra* maps obtained in this work were compared and successively were aligned with the physical map of *P. trichocarpa*. Finally they were aligned with the *P. alba* maps by Paolucci *et al.* (unpublished data) and the *P. nigra* map by Cervera *et al.* (2001) obtained from an inter-specific cross with *P. deltoides*.

3.3.1 Comparison between the two parental maps

The comparison of the 2 *P. nigra* parental maps revealed 69 common markers, including 68 SSRs and 1 SNP. The number of common markers for each linkage groups is presented in Table 27.

Table 27: Common markers of the *P. nigra* parental maps by linkage groups

Linkage Group	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	Other
Common markers	6	6	4	3	5	4	1	4	5	5	2	3	5	1	3	3	4	2	2	0

There was an average of 4 common markers per linkage groups. Two linkage groups presented only 1 common markers (VII and XIV) and 3 linkage groups had 2 common markers (XI, XVIII, and XIX) (Fig. 32). The order of the markers was the same for 63 markers. A non co-linearity occurred within an interval of < 10 cM for 4 common markers. One discrepancy was also found, the SSR marker ORPM276 was mapped on 2 different linkage groups, the XIII for the female map and the XIX for the male map (Fig. 32).

3.3.2 Alignment with the *Populus trichocarpa* physical map

The alignment with the *P. trichocarpa* physical map and the *Populus spp.* published maps (Cervera *et al.*, 2001; Yin *et al.*, 2004b) allowed to identify homologous linkage groups. After the alignment we found the 19 linkage groups corresponding to the number of haploid chromosome of *Populus.* In this way the number of linkage groups in the male map could be reduced to 21 and only 2 triplets remained unaligned. After alignment the female map remained at 20 linkage groups previously identified.

3.3.2.1 Comparison between Populus trichocarpa and Populus nigra maps

The comparison of the *P. trichocarpa* physical map and the 2 *P. nigra* parental maps revealed 102 common markers, 67 of them were common to the female map and 84 to the male map. The

number of common markers between *P. trichocarpa* and each *P. nigra* parental map, ordered on the basis of linkage groups, is presented in Table 28.

Table 28: Common markers with the *P. trichocarpa* physical map for the 2 *P. nigra* parental maps

Linkage Group	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX
♀[a] common markers	4	6	4	4	4	4	1	5	4	4	2	5	7	1	2	2	2	4	2
♂[b] common markers	5	8	9	5	7	5	2	6	4	6	5	3	4	3	2	4	2	2	2

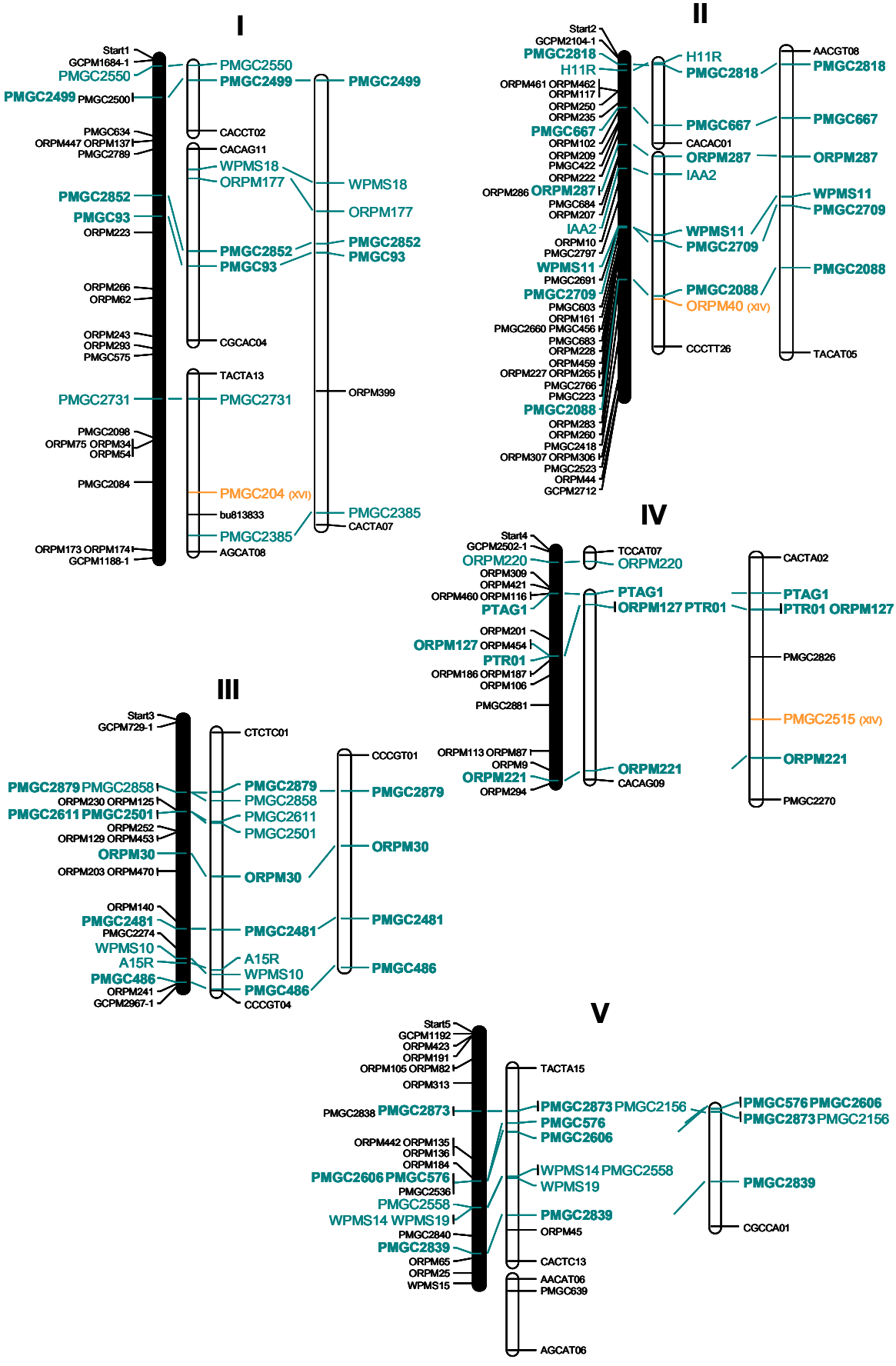
[a] Female 58-861 map

[b] Male Poli map

At least two common markers are needed to align maps. For the male map all groups were identified with reliability because all groups had a minimum of 2 markers common to the *P. trichocarpa* physical map (Fig. 34). For the female map, there were 2 groups, VII and XIV, with only one marker common to *P. trichocarpa* and also one marker common to the male map (Fig. 34). Therefore, for the female map the groups VII and XIV were not identified with a good reliability.

Some discrepancies were also found, 7 markers were not mapped on the same group of *P. trichocarpa* physical map (Fig. 34). The markers ORPM144, ORPM477, and PMGC2803 were common to the 2 parents and mapped on unexpected groups, IX, XV, and XVII, respectively, while they were mapped in *P. trichocarpa* on the VII, XVIII, and IX linkage groups, respectively. One marker, ORPM276, was mapped on 2 different linkage groups, the XIII in the *P. nigra* female map and the XIX in the male map. In *P. trichocarpa*, this marker was mapped on the group XIX like in the *P. nigra* male map. Five marker couples were inverted. Three out of 5 marker couples showed a disorder occurring within an interval lower than 5 cM while the 2 others showed a disorder occurring within an interval < 10 cM.

Fig. 34: Alignment of the *P. trichocarpa* and *P. nigra* maps.
From left to right: the linkage groups of the physical *P. trichocarpa* map (in black), the linkage groups of the male map (in dashed) and the linkage groups of the female map (in white). Markers in common among maps are in blue and in larger letters. Markers common to the 3 maps are in bold. They are indicated with allelic bridges. Markers in orange are mapped on different groups in *P. nigra* and *P. trichocarpa* maps, the linkage group of these markers in *P. trichocarpa* is indicated in parenthesis. Linkage groups names are denoted by a roman numeral. The length of the linkage group bars is proportional to the distance among markers: 0.25 mm per bases for the *P. trichocarpa* physical map, 0.25 mm per cM for the male and female *P. nigra* maps. To simplify the representation, only potential common markers and the first and last marker of each group are indicated.


I
Start1
GCPM1684-1
PMGC2550
PMGC2499
PMGC2500
PMGC634
ORPM447 ORPM137
PMGC2789
PMGC2852
PMGC93
ORPM223
ORPM266
ORPM62
ORPM243
ORPM293
PMGC575
PMGC2731
PMGC2098
ORPM75 ORPM34
ORPM54
PMGC2084
ORPM173 ORPM174
GCPM1188-1
PMGC2550
PMGC2499
CACCT02
CACAG11
WPMS18
ORPM177
PMGC2852
PMGC93
CGCAC04
TACTA13
PMGC2731
PMGC204 (XVI)
bu813833
PMGC2385
AGCAT08
PMGC2499
WPMS18
ORPM177
PMGC2852
PMGC93
ORPM399
PMGC2385
CACTA07
II
Start2
GCPM2104-1
PMGC2818
H11R
ORPM461 ORPM462
ORPM117
ORPM250
ORPM235
PMGC667
ORPM102
ORPM209
PMGC422
ORPM222
ORPM286 ORPM287
PMGC684
ORPM207
IAA2
ORPM10
PMGC2797
WPMS11
PMGC2691
PMGC2709
PMGC603
ORPM161
PMGC2660 PMGC456
PMGC683
ORPM228
ORPM459
ORPM227 ORPM265
PMGC2766
PMGC223
PMGC2088
ORPM283
ORPM260
PMGC2418
ORPM307 ORPM306
PMGC2523
ORPM44
GCPM2712
H11R
PMGC2818
PMGC667
CACAC01
ORPM287
IAA2
WPMS11
PMGC2709
PMGC2088
ORPM40 (XIV)
CCCTT26
AACGT08
PMGC2818
PMGC667
ORPM287
WPMS11
PMGC2709
PMGC2088
TACAT05
III
Start3
GCPM729-1
PMGC2879 PMGC2858
ORPM230 ORPM125
PMGC2611 PMGC2501
ORPM252
ORPM129 ORPM453
ORPM30
ORPM203 ORPM470
ORPM140
PMGC2481
PMGC2274
WPMS10
A15R
PMGC486
ORPM241
GCPM2967-1
CTCTC01
PMGC2879
PMGC2858
PMGC2611
PMGC2501
ORPM30
PMGC2481
A15R
WPMS10
PMGC486
CCCGT04
CCCGT01
PMGC2879
ORPM30
PMGC2481
PMGC486
IV
Start4
GCPM2502-1
ORPM220
ORPM309
ORPM421
ORPM460 ORPM116
PTAG1
ORPM201
ORPM127 ORPM454
PTR01
ORPM186 ORPM187
ORPM106
PMGC2881
ORPM113 ORPM87
ORPM9
ORPM221
ORPM294
TCCAT07
ORPM220
PTAG1
ORPM127 PTR01
ORPM221
CACAG09
CACTA02
PTAG1
PTR01 ORPM127
PMGC2826
PMGC2515 (XIV)
ORPM221
PMGC2270
V
Start5
GCPM1192
ORPM423
ORPM191
ORPM105 ORPM82
ORPM313
PMGC2838 PMGC2873
ORPM442 ORPM135
ORPM136
ORPM184
PMGC2606 PMGC576
PMGC2536
PMGC2558
WPMS14 WPMS19
PMGC2840
PMGC2839
ORPM65
ORPM25
WPMS15
TACTA15
PMGC2873 PMGC2156
PMGC576
PMGC2606
WPMS14 PMGC2558
WPMS19
PMGC2839
ORPM45
CACTC13
AACAT06
PMGC639
AGCAT06
PMGC576 PMGC2606
PMGC2873 PMGC2156
PMGC2839
CGCCA01

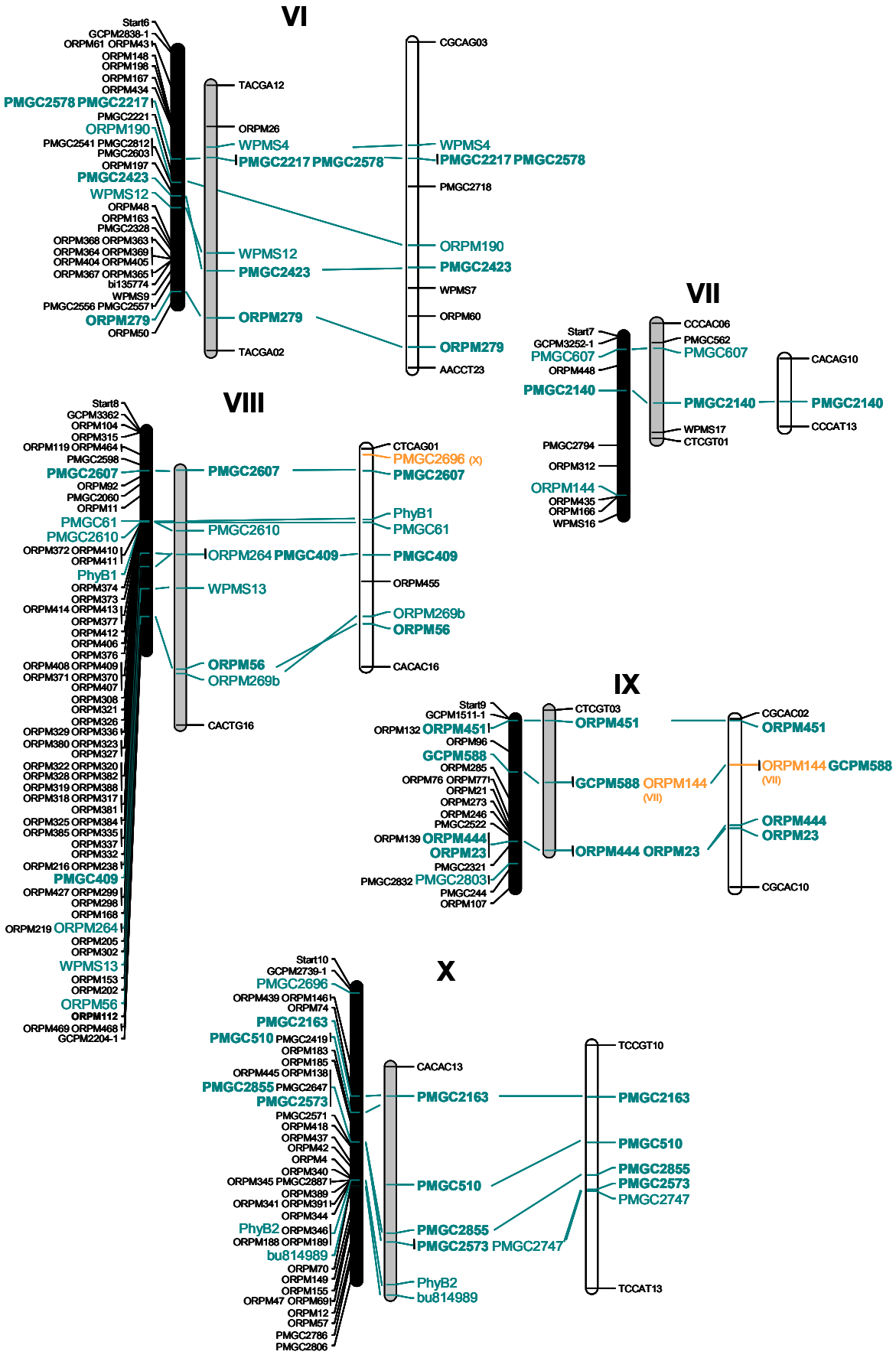


Fig. 34: Continued

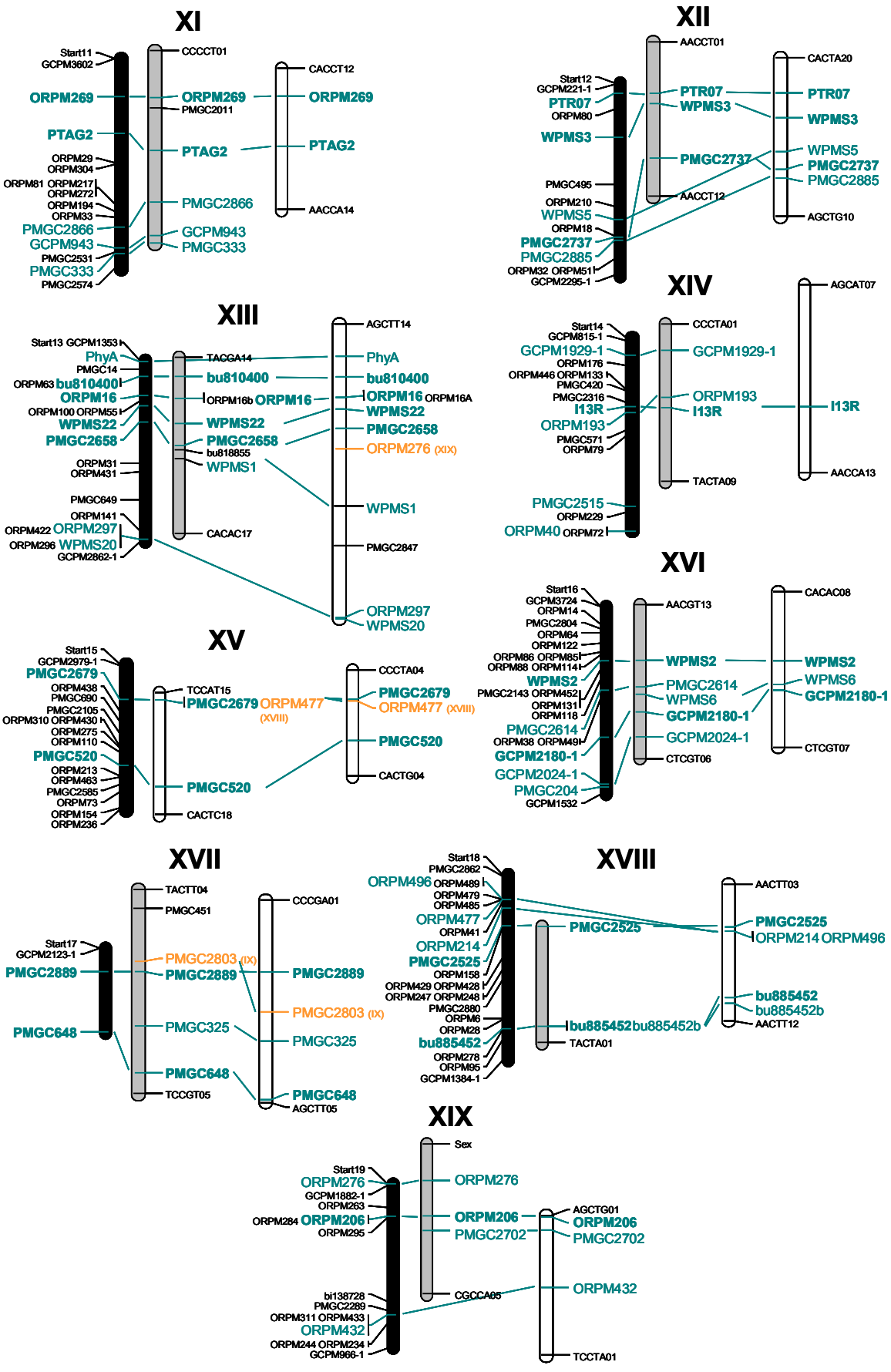


Fig. 34: Continued

3.3.2.2 Correlation between genetic and physical length

Thanks to the availability of the *P. trichocarpa* physical map, the correlation between the genetic and the physical length, and the ratio between the number of bases and cM were estimated for each *P. nigra* map. The results are presented in Fig. 35.

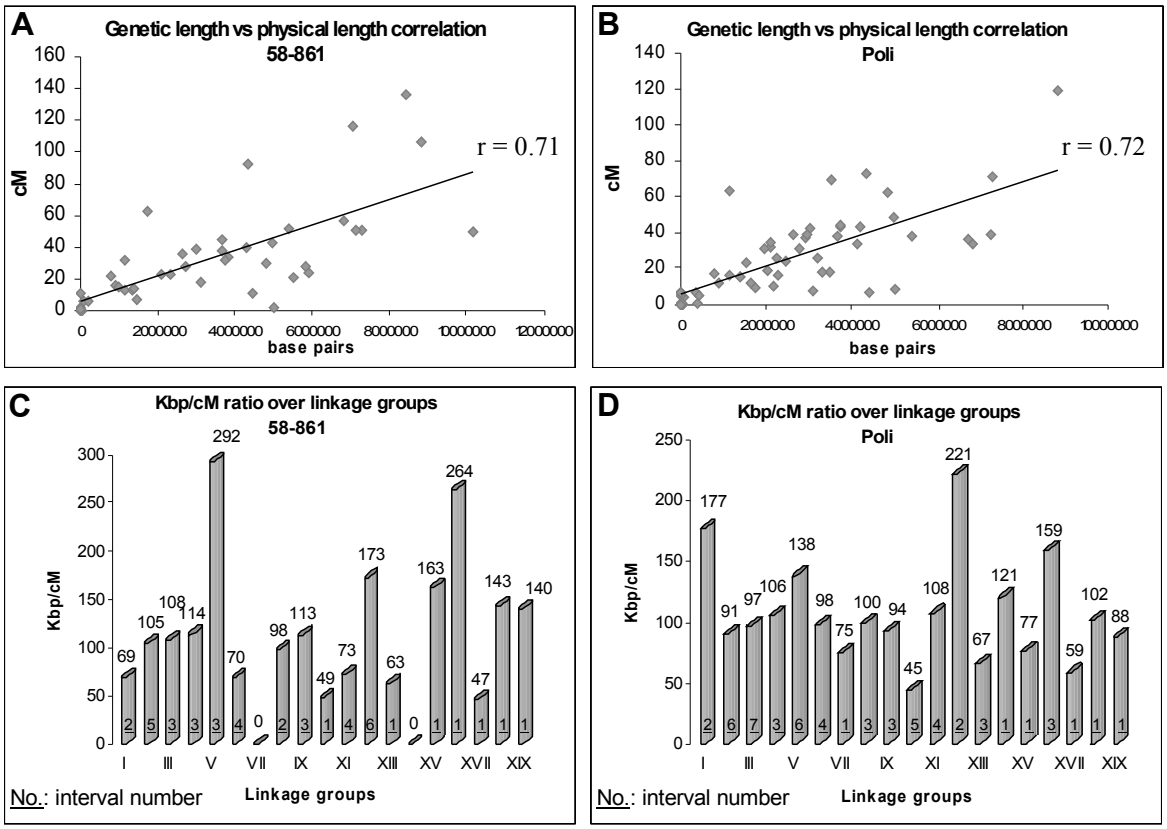


Fig. 35: Correlation and ratio between the genetic length and physical length.
Points of the plots A and B represent the correspondence between the distance in cM for 2 markers mapped in the *P. nigra* female (A) and male (B) maps and the distances in base pairs of these markers in the *P. trichocarpa* physical map. Plots C and D represent, for each linkage group, the ratio between the length in base pairs and the length in cM for the female (C) and the male (D) map. The underlined numerals indicate the number of intervals used to calculate the ratio for each linkage group. A ratio equal to 0 means that there is only one common marker between the *P. nigra* map and the *P. trichocarpa* map therefore the ratio can not be calculated.

The coefficient of linear correlation obtained was 0.71 for the female map and 0.72 for the male map (Fig. 35 A and B). The ratio between the physical length of *P. trichocarpa* and the genetic distance of the *P. nigra* linkage groups was rather variable, from 47 to 292 Kbp/cM for the female map and from 45 to 221 Kbp/cM for the male map. The ratio Kbp/cM was calculated from only one value (only 2 common markers) for 6 linkage groups for both study maps. These linkage groups were XI, XV, XVI,XVII, XVIII, XIX for the female map and VII, XIV, XV, XVII, XVIII, XIX for the male map (Fig. 35 C and D).

3.3.3 Comparison between *Populus alba* and *Populus nigra* maps

The *P. alba* maps of this comparative analysis were gently provided by I. Paolucci (personal communication). They were constructed in our laboratory with the same strategy applied for the *P. nigra* maps. The *P. alba parents* used to obtain the mapping pedigree come from natural Italian populations located in the same regions of the *P. nigra* parents. A mapping pedigree composed of 141 F_1 individuals was used for the *P. alba* map. The same markers were analyzed to produce the *P. nigra* and *P. alba* maps. The comparison of these 4 maps revealed 56 SSR common markers, 11 of them were common to the 4 maps, 26 markers were found on 3 maps and 18 markers were found only on 2 maps. Only three markers were mapped on different linkage groups among the 4 maps. The number of common markers among each parent and the 2 parents of the other species is presented in Table 29.

Table 29: Common markers among the *P. alba* and *P. nigra* maps.

	Linkage Group	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX
Common markers	♀ *P. alba*	2	4	5	3	2	5	3	1	-	4	1	2	1	-	1	-	2	-	3
	♂ *P. alba*	8	1	3	3	1	1	4	1	1	2	1	4	1	-	-	-	1	-	-
	♀ *P. nigra*	6	3	2	3	1	3	1	1	1	5	1	4	1	-	1	-	2	-	1
	♂ *P. nigra*	8	4	5	3	3	3	4	1	1	5	2	2	-	-	1	-	3	-	2

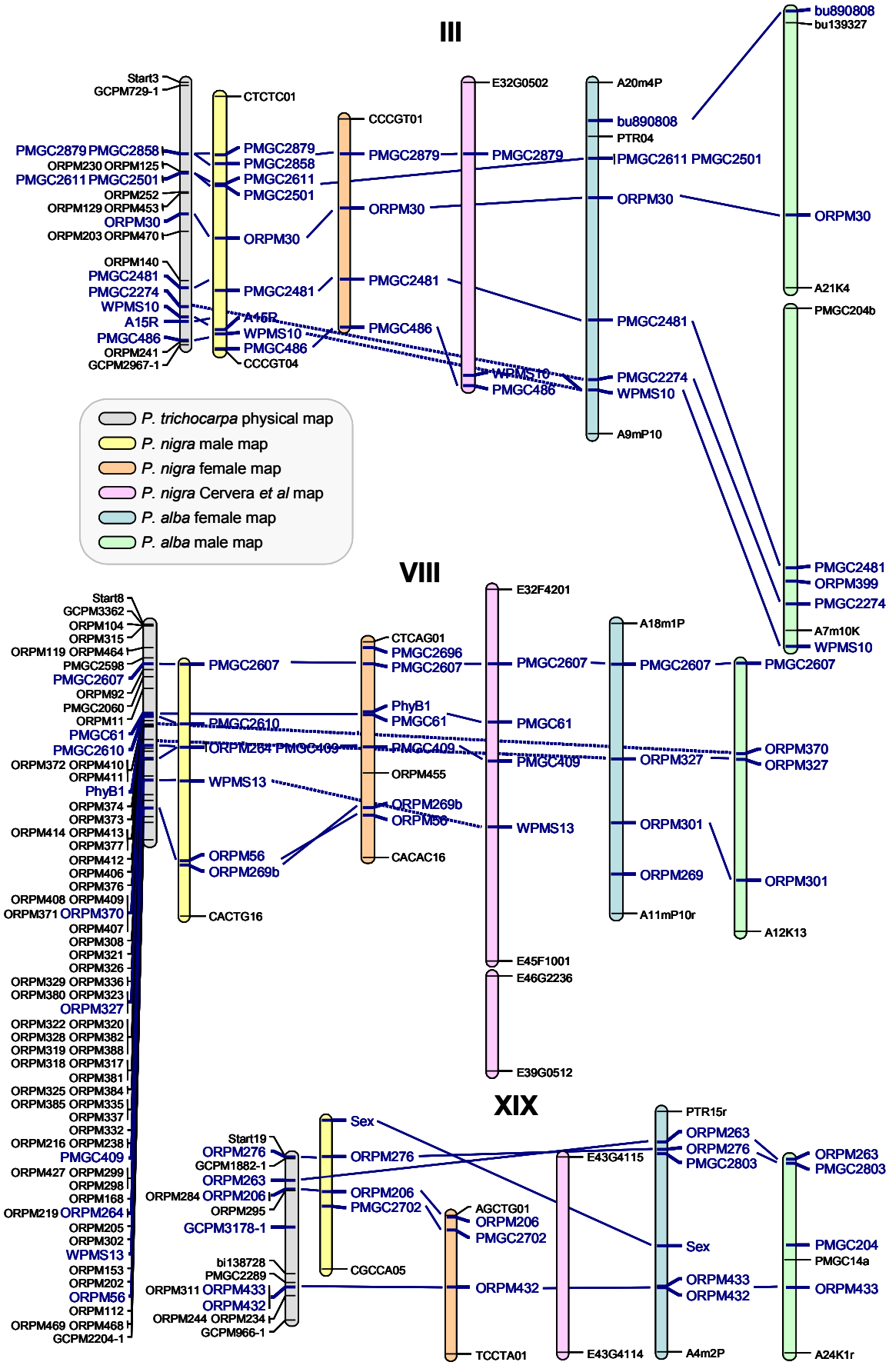
Three linkage groups, XIV, XVI, and XVIII could not be aligned (Table 29). The marker order was conserved in most of the cases. Two linkage groups showed inversions between tightly linked markers, separated by a distance < 10 cM. The ORPM40 marker was mapped on the same linkage group (II) for the *P. alba* female map and the *P. nigra* male map but it was mapped on the linkage group XIV in the *P. trichocarpa* physical map. The sex morphological trait was mapped on the same linkage group, XIX, for the *P. alba* female map and the *P. nigra* male map (Fig. 36). The common markers among the 4 maps, for each linkage group, are presented in detail in Appendix 2.

3.3.4 Comparative mapping among some *Populus species*

Finally, we aligned our *P. nigra* maps with the *P. alba* maps, the *P. nigra* map by Cervera *et al.* (2001) originated from an inter-specific cross with *P. deltoides*, and the *P. trichocarpa* physical map. The coalignment of these 6 maps revealed 166 common markers. Only the PMGC2607 marker (linkage group VIII) was present on the 6 maps (Fig. 36). Seventeen markers were

common to 5 maps, 26 markers were common to 4 maps, 58 markers were common to 3 maps and 64 markers were common to only 2 maps. Discrepancies were found for 17 markers which were mapped on different linkage groups. The marker order was conserved in most of the cases (Fig. 36). The linkage group I was one of the largest group in all the maps. The linkage group IX for the *P. alba* female map and the linkage groups XIV and XV for the *P. nigra* map, by Cervera *et al.* (2001) were not identified. On these 2 groups (XIV and XV), there were few co-dominant makers in the genetic maps analyzed, and we observed a maximum of 3 co-dominant markers. The common markers among the 6 maps are presented in detail in Appendix 2.

Fig. 36: Example of the coalignment of 6 maps for 3 linkage groups.
From left to right: the linkage groups of the physical *P. trichocarpa* map (in grey), the linkage groups of the *P. nigra* male map (yellow), the linkage groups of *P. nigra* female map (orange), the linkage groups of*P. nigra* Cervera *et al* (2001) map (pink), the linkage groups of *P. alba* female map (blue) and the linkage groups of *P. alba* male map (green). The *P. alba* maps come from I. Paolucci (personal communication). Markers in common among maps are in blue and larger letters. They are indicated with allelic bridges. Linkage groups names are denoted by a roman numeral. The length of the linkage group bars is proportional to the distance among markers: 0.25 mm per bases for the *P. trichocarpa* physical map, 0.25 mm per cM for the other genetic maps. To simplify the representation only potential common markers and the first and last marker of each group are indicated.


III
Start3
GCPM729-1
PMGC2879 PMGC2858
ORPM230 ORPM125
PMGC2611 PMGC2501
ORPM252
ORPM129 ORPM453
ORPM30
ORPM203 ORPM470
ORPM140
PMGC2481
PMGC2274
WPMS10
A15R
PMGC486
ORPM241
GCPM2967-1
CTCTC01
PMGC2879
PMGC2858
PMGC2611
PMGC2501
ORPM30
PMGC2481
A46R
WPMS10
PMGC486
CCCGT04
CCCGT01
PMGC2879
ORPM30
PMGC2481
PMGC486
E32G0502
PMGC2879
ORPM30
PMGC2481
WPMS10
PMGC486
A20m4P
bu890808
PTR04
PMGC2611 PMGC2501
ORPM30
PMGC2481
PMGC2274
WPMS10
A9mP10
bu890808
bu139327
ORPM30
A21K4
PMGC204b
PMGC2481
ORPM399
PMGC2274
A7m10K
WPMS10
P. trichocarpa physical map
P. nigra male map
P. nigra female map
P. nigra Cervera et al map
P. alba female map
P. alba male map
VIII
Start8
GCPM3362
ORPM104
ORPM315
ORPM119 ORPM464
PMGC2598
PMGC2607
ORPM92
PMGC2060
ORPM11
PMGC61
PMGC2610
ORPM372 ORPM410
ORPM411
PhyB1
ORPM374
ORPM373
ORPM414 ORPM413
ORPM377
ORPM412
ORPM406
ORPM376
ORPM408 ORPM409
ORPM371 ORPM370
ORPM407
ORPM308
ORPM321
ORPM326
ORPM329 ORPM336
ORPM380 ORPM323
ORPM327
ORPM322 ORPM320
ORPM328 ORPM382
ORPM319 ORPM388
ORPM318 ORPM317
ORPM381
ORPM325 ORPM384
ORPM385 ORPM335
ORPM337
ORPM332
ORPM216 ORPM238
PMGC409
ORPM427 ORPM299
ORPM298
ORPM168
ORPM219 ORPM264
ORPM205
ORPM302
WPMS13
ORPM153
ORPM202
ORPM56
ORPM112
ORPM469 ORPM468
GCPM2204-1
PMGC2607
PMGC2610
ORPM264 PMGC409
WPMS13
ORPM56
ORPM269b
CACTG16
CTCAG01
PMGC2696
PMGC2607
PhyB1
PMGC61
PMGC409
ORPM455
ORPM269b
ORPM56
CACAC16
E32F4201
PMGC2607
PMGC61
PMGC409
WPMS13
E45F1001
E46G2236
E39G0512
A18m1P
PMGC2607
ORPM327
ORPM301
ORPM269
A11mP10r
PMGC2607
ORPM370
ORPM327
ORPM301
A12K13
XIX
Start19
ORPM276
GCPM1882-1
ORPM263
ORPM284 ORPM206
ORPM295
GCPM3178-1
bi138728
PMGC2289
ORPM311 ORPM433
ORPM432
ORPM244 ORPM234
GCPM966-1
Sex
ORPM276
ORPM206
PMGC2702
CGCCA05
AGCTG01
ORPM206
PMGC2702
ORPM432
TCCTA01
E43G4115
E43G4114
PTR15r
ORPM263
ORPM276
PMGC2803
Sex
ORPM433
ORPM432
A4m2P
ORPM263
PMGC2803
PMGC204
PMGC14a
ORPM433
A24K1r

4 Discussion

4.1 Mapping pedigree

4.1.1 Intra-specific cross

Different species generally display much polymorphic variation over a high portion of the genome which is a result of a long evolutionary divergence (Yin *et al.*, 2001). The choice of two highly divergent genotypes of the same species, *P. nigra*, as parents of our mapping pedigree enabled us to obtain results comparable to those obtained from an inter-specific cross. Thirty-nine percent of the tested SSRs markers were found polymorphic in our pedigree whereas 32% were obtained for the inter-specific cross between *P. nigra* and *P. deltoides* by Cervera *et al* (2001). The efficiency of constructing a genetic linkage map in outbred forest trees with the two-way pseudo-testcross strategy depends on the level of heterozygosity of the species (Cervera *et al.*, 2001). In our pedigree the heterozygosity level based on SSR markers was 41% and 44% for our *P. nigra* female and male parent, respectively. In the *P. nigra*, *P. deltoides*, *P. trichocarpa* parents crossed by Cervera *et al* (2001) the heterozygosity level based on SSR markers was 58%, 63%, and 75%, respectively. In spite of a smaller heterozygosity level in our genotypes, we obtained a good number of polymorphic markers (380 and 337 markers for the female and male parent, respectively) useful for the map construction.

Most of the genetic maps published for the genus *Populus* were obtained from inter-specific crosses (Bradshaw *et al.*, 1994; Yin *et al.*, 2001; Cervera *et al.*, 2001; Yin *et al.*, 2002; Zhang *et al.*, 2004). **This *P. nigra* mapping pedigree was produced to generate the first genetic maps from an intra-specific cross for the euroasian species *P. nigra*.**

4.1.2 Populus deltoides introgression

Hybridization in plants is a common event that occurs naturally in the wild. *P. nigra* ability to cross with *P. deltoides* and his cultivated hybrids (*P. x canadensis*) has led to anxieties for the introgression of *P. deltoides* germplasm in *P. nigra* natural populations (Fossati *et al.*, 2003). However, only the crosses with *P. nigra* as father generate hybrid offsprings with *P. deltoides*. In cross-breeding trials with the reciprocal combination *P. nigra x P. deltoides*, the embryos died off at an early stage (Hofmann, 2001). Recent results suggest that the introgression of *P. deltoides* is rare (Fossati *et al.*, 2003; Storme *et al.*, 2003). Nevertheless, five species-specific

markers were used to check for introgression of *P. deltoides* into *P. nigra* parents of our mapping pedigree. Any *P. deltoides*-specific alleles were detected in our pedigree but we found an allele (WPMS20, 224bp) that occurs frequently in *P. deltoides* and rarely in *P. nigra* (Fossati *et al.*, 2003). This allele was present only in one of the two parents. **Thus, no evidence of *P. deltoides* introgression into the *P. nigra* mapping pedigree was found in this study.**

4.2 Marker analysis

An important advantage of using various markers is a better coverage of different genome regions, probably due to distinct target areas of different molecular markers on the genome (Casasoli *et al.*, 2001). The use of AFLP, SSR, and SNP markers on the *P. nigra* genome enabled to generate many polymorphic markers ensuring a good coverage of the genome.

4.2.1 AFLP markers

One main advantage of this technique is the simultaneous recovery of numerous polymorphic markers (5 to 25 in our experiment) in one assay. In order to increase the number of DNA fragments, the number of selective nucleotides at the end of the AFLP primer corresponding to the rare cutter restriction enzyme was reduced from 1 to 0 for the preselective amplification and from 3 to 2 for the selective amplification. In this way, an average of 7 and 6 AFLP markers per primer combinations was obtained for the female and the male parent, respectively. In addition, the reduction of the AFLP reaction volumes allowed to decrease the cost of the analysis by 5 folds. The reproducibility of three popular molecular markers (RAPD, AFLP, and SSR) was examined in a previous study involving several European laboratories. In that comparison, AFLPs showed extremely high reproducibility (Jones *et al.*, 1997).

In this work, the AFLP technique was successfully set up to obtain a large number of markers with reduced time and cost.

4.2.2 SSR markers

SSRs are highly polymorphic, co-dominant markers with a great value for the construction of genetic maps, comparative mapping, population genetic surveys, and paternity analyses. In *Populus*, several hundred SSRs have been identified using various approaches (Tuskan *et al.*, 2004). Since September 2004, *Populus* is the third plant species (after *Arabidopsis* and *Oryza*) and the first tree genome to be sequenced (http://genome.jgi-psf.org/Poptr1/Poptr1.home.html). SSRs will be useful in merging the genetic map in *Populus* with the physical map and the sequence

database. So far about 4166 SSRs are available in *Populus*. The majority of these SSRs were developed from *P. trichocarpa*. They showed a high transferability throughout the genus. Nevertheless, amplification frequency appears to vary in concert with genetic relatedness among taxa (Schlötterer, 2001; Tuskan *et al.*, 2004). A large number of SSR markers (330) were analyzed on the mapping pedigree and 78% of them yielded amplified product. This percentage is in agreement with the amplification frequency (80-99%) within the *Aigeiros* section found by Tuskan *et al.* (2004). Hundred percent of amplification was obtained with SSR markers developed from *P. nigra* (WPMS) whereas 64% only amplified in the *P. alba* mapping pedigree (Paolucci, personal communication). That confirms the better efficiency of amplification with markers developed from closely related species. The published map with the highest number of SSRs is the *P. deltoides* map by Yin *et al.* (2004b) including 105 SSR markers. In the F_1 *P. nigra* pedigree 106 and 91 SSRs were mapped for the male and the female map, respectively. **The *P. nigra* male map includes the highest number of SSRs among the poplar genetic maps published up to now.**

4.2.3 SNP markers

SNPs are co-dominant markers and are present in all parts of the genome. By using assembled ESTs from 14 different cDNA libraries, 556 *Populus* candidate SNPs were identified (Zhang *et al.*, 2005). The development of these markers is interesting because they have widespread application including the construction of high-resolution genetic maps, mapping genes, and analyses of the genetic structure of population (Zhang *et al.*, 2005). For these reasons, sequence specific markers were developed from the sequences and the localization of some SNPs in the *P. nigra* parents, gently provided by M. Morgante and G. Zaina. A fast and low cost genotyping technique was searched to map these markers. The CAPS and dCAPS techniques answered to these exigencies. The CAPS method has been widely used as a tool for rapidly and reliably detecting SNPs that create a restriction site in only one allele per locus. However, the majority of SNPs do not create such restriction site. Therefore, the dCAPS method which is a modification of the CAPS technique was used. It allows the detection of most SNPs by utilizing mismatched PCR primers (Michaels and Amisino, 1998). The difficulty of the dCAPS technique consists in choosing the appropriate mismatches in the primer to conserve the PCR efficiency. All the seven SNPs tested gave amplification products and the expected pattern. Thus the design of the mismatched primers was successful and the SNPs were integrated on the map.

The *P. nigra* maps are the first poplar genetic maps including SNPs (3 SNPs in the female map and the 5 SNPs in the male map).

4.2.4 Gene mapping

Mapping genes is interesting for a better knowledge of the genome structure and function. Genetic maps are also a basis for QTLs mapping. The availability of genes on the maps could allow to identify the gene or genes responsible for variation in the phenotype. ***PhyA*, *PhyB1*, and *PhyB2*, *IAA2* genes were included in the *P. nigra* female and male map, respectively while *PTAG1* and *PTAG2* were present in both maps.**

The SSR markers identified from EST data are closely associated with, or directly in, the coding region of the genes and could provide a platform for the comparisons of genomes, genes searching, and cloning (Zhang *et al.*, 2005). **Two and five EST-based SSRs were mapped for the female and male map, respectively.** Three out of the five EST-based SSRs included in these maps have a known function.

4.2.4.1 Phytochrome genes

Populus has three phytochrome genes, *PhyA*, *PhyB1,* and *PhyB2* (Howe *et al.*, 1998). Both *PhyA* and *PhyB* regulate seed germination and de-etiolation. *PhyA* may play a role in the photoperiodic control of flowering (Howe *et al.*, 1998) and regulates processes which are not affected by other phytochromes (Ingvarsson *et al.*, 2005). *PhyB* also controls aspect of the shade avoidance response and photoperiodically induced flowering and tuberization in other species (Howe *et al.*, 1998). *PhyB2* has been mapped to a linkage group containing QTLs for bud set and bud flush in several replicated experiments using independent mapping populations (Frewen *et al.*, 2000; Chen *et al.*, 2002; Ingvarsson *et al.*, 2005).

4.2.4.2 AGAMOUS orthologous in poplar

The two *P. trichocarpa* genes, *PTAG1* and *PTAG2,* isolated by Brunner *et al.* (2000) were mapped on the *P. nigra* mapping pedigree. They are homologous to the *Arabidopsis* floral homeotic gene *AGAMOUS* (*AG*). *PTAG1* and *PTAG2* are located on separate linkage groups, but their non-coding regions are highly similar, consistent with a phylogenetically recent duplication. Brunner *et al.* (2000) demonstrated the absence of additional genes in the poplar genome with significant *PTAG1*/*PTAG2* homology. They have also shown that *PTAG1* and *PTAG2* exhibit an AG-like floral expression pattern. The high degree of similarity shared by *PTAG1* and *PTAG2* in both sequence

and expression indicates that they are unlikely to be functionally associated with specification of tree gender. Brunner *et al.* (2000) have designed primers for the region flanking the SSRs located in the second intron of *PTAG1* and in the 5' flanking region of *PTAG2*. The primers were used in this study and the *PTAG1* and *PTAG2* genes were mapped on the same linkage groups of *P. trichocarpa*. Thus, these 2 genes appear to be conserved in the *P. nigra* genome.

4.2.4.3 IAA genes

The role of indole acetic acid (IAA/auxin) as an important regulator of growth-related processes such as cell division, elongation and differentiation has been well established over the years. Auxin plays a key role in regulating wood formation through its effects on cambial activity and xylem development. It has been shown to influence diverse aspects of plant growth and development (Moyle *et al.*, 2002). *Aux/IAA* genes form a multigene family. Moyle *et al.* (2002) demonstrate that there are at least eight *Aux/IAA* genes expressed differentially in the tissues of cambial region in hybrid aspen (*P. tremula* L. X *P. tremuloides* Michx). The *IAA2* gene of this multigene family was mapped in this study.

Taking advantage of the SNPs (Morgante and Zaina, personal communication), the SSRs based on EST (Jorge *et al.*, unpublished) and the literature information (Brunner *et al.*, 2000) a male genetic map with nine genes and a female map with four genes were obtained.

4.2.5 Sex morphological trait

Gender, the expression of maleness or femaleness, in dioecious plants has been associated with change in morphology, physiology, ecological position, and commercial importance of several species. Consequently, morphological and/or physiological differences may exist between the sexes, differences which could have a genetic basis. In plant breeding programs, attempt to select for a trait of interest can be complicated by gender since maleness or femaleness can only be determined at the onset of flowering in most dioecious plants species (McLetchie *et al.*, 1994). Several gender associated traits have been identified in Salicaceae (family which contains two genera, *Salix* and *Populus*). Female clones of interamericana *Populus* hybrid (*P. x generosa*) tend to produce greater number of lateral branches (Tschaplinski *et al.*, 1994). Grant and Milton (1979), working with *P. tremuloides*, found that female clones produce higher mean annual growth than male clones. Therefore, the unraveling of sex determination mechanisms would also be extremely useful in plant breeding and improvement (Alstrom-Rapaport *et al.*, 1998). With the advent of the molecular age, the situation is changing rapidly but we are still far from a global

picture of plant sex determination mechanism and their evolution (Semerikov *et al.*, 2003). Little is know about the sex determination of Salicaceae. Two previous studies in *Salix* (Alström-Rapaport *et al.*, 1997; 1998) indicated that the presence of a sex chromosome is unlikely and suggested a multilocus autosomal sex-determination system (Semerikov *et al.*, 2003). Unfortunately there is no published information about sex-determination in *Populus*.

Fifty-four percent of the 4 year old F_1 individual flowered during spring 2005 allowing to determine the gender of 50 genotypes of the *P. nigra* pedigree. The observed preliminary data highly deviated from the 1:1 expected Mendelian segregation (70% of the flowering F_1 individuals were male and only 30% female). Many Salicaceae demonstrate biased sex ratio. For instance, the sex ratio of the basket willow (*Salix viminalis* L.) populations is usually strongly female-biased and controlled crosses suggest that sex biases are not only the result of environmental effects (Alström-Rapaport *et al.*, 1997; Semerikov *et al.*, 2003). It was also observed that the sex ratios of Scandinavian *P. tremula* population are predominantly male biased (Ingvarsson, 2005). A male biased sex ratio in the *P. alba* mapping pedigree was also observed. On a total of 157 *P. alba* F_1 progeny 140 genotypes produced flowers and their gender was checked. Sixty-four percent were male and 36% were female (Sabatti, personal communication). Until now, there is no explanation for the biased sex ratios. In spite of the male biased sex ratio and the missing data sex like a morphological trait was mapped on the *P. nigra* pedigree. It was mapped on the XIX group in the male map and unlinked in the female map. It is unknown whether the sex morphological trait was not placed on the female map because of low quality of the marker (highly distorted and 46% of missing data) or because of a gap in the female map. The second hypothesis could be considered likely because in the male map the sex and ORPM276 marker were placed at the beginning of the linkage group XIX followed by ORPM206 marker whereas in the female map ORPM206 marker began the XIX group (Fig. 32).

Another interesting point is that sex morphological trait was mapped on the linkage group XIX in *P. alba* like in *P. nigra* but on the female map (Paolucci, personal communication). Testolin *et al.* (2001) also mapped the sex determinant in the male linkage map in *Actinidia callosa* (kiwifruit). The alignment of the *P. alba* and *P. nigra* map indicated that sex was mapped in two different regions on the XIX group (Fig. 36). These information indicate that at least two loci could be involved in sex determination in *Populus*. The results, especially for the *P. nigra* maps, could be confirmed by checking the F_1 individuals not flowering in 2005. For a better understanding of the sex determination in *Populus* it would be interesting to search for the presence of genetic markers associated with gender. Gender-related markers would provide knowledge about the sex

determination such as the number of loci and chromosomes involved. Gender-related markers would also be an useful tool to identify gender in *Populus* at early stage (zygote, seedling) within breeding programs.

Although our data are preliminary, this is the first study where sex has been mapped like a morphological trait in *Populus*, showing the loci of the linkage group XIX likely involved in sex determination.

4.3 Linkage map characteristics

4.3.1 Linkage maps

Following the pseudo-testcross strategy, two linkage maps were constructed in an F_1 intra-specific progeny derived from *P. nigra* species of the genus *Populus*, section *Aigeiros*. Like other forest tree species, poplars generally exhibit a high level of heterozygosity, because they are still in wild or semi-wild state. High heterozygosity in these species enables us to detect a large number of polymorphic markers using PCR technology (Grattapaglia and Sederoff, 1994; Barreneche *et al.*, 1998; Yin *et al.*, 2001). Another advantage of the pseudo-testcross strategy lies on the exploitation of a two-generation pedigree, which is currently available in forest tree species (Grattapaglia and Sederoff, 1994; Arcade *et al.*, 2000). The application of this strategy proved to be efficient in a relative short period of time to construct genetic maps of *P. nigra* with a good coverage of the genome.

The 19 chromosomes of *Populus* species are represented in the genome maps of the two *P. nigra* female and male parents by 20 and 26 linkage groups, respectively. One group for each map was separated into two groups (III and VI for the female and male map, respectively). These groups were joined because the SSRs present in the separated groups indicated that they belong to the same linkage group. A small reduction of the LOD threshold (from 4.0 to 3.5) allowed the junction. The most complete genetic map constructed for the genus *Populus* has been reported by Yin *et al.* (2004b). This *P. trichocarpa* x *P. deltoides* map includes 544 markers mapped onto 19 linkage groups equivalent to the *Populus* haploid chromosome number. Three other studies achieved 19 main linkage groups in *Populus*. Wu *et al.* (2000) constructed a *P. deltoides* genetic map based exclusively on AFLP markers which includes 19 major groups and 24 minor groups. Yin *et al.* in (2001) presented a *P. adenopoda* x *P. alba* map based on RAPD markers with 19 major groups plus one triplet and four doublets. In 2002, these authors identified 19 major groups for

the *P. deltoides* and the *P. euramericana* map based on AFLP and RAPD markers. Finally, Cervera *et al*. (2001) reported dense genetic maps for *P. deltoides*, *P. trichocarpa*, and *P. nigra* utilizing AFLP and more than 100 SSR markers. The 19 linkage groups are identified for the three pedigrees but accompanied by minor groups. The *P. nigra* map, which is the unique published, is formed by 34 main groups plus four triplets and two doublets. The maps by Cervera *et al*. (2001) and Yin *et al*. (2004b) provide an excellent starting point for comparative mapping. The linkage group nomenclature proposed by Cervera *et al*. (2001) is adopted as a standard for *Populus*. **In this context, the *P. nigra* map presented in this study amount to the lowest number of linkage groups, 20 for the female map, while the unidentified group is composed of only five markers.**

Although 19 linkage groups were found representing the 19 poplar haploid chromosomes, the results show that some regions of the genome remain uncovered. To obtain the most complete *Populus* map Yin *et al*. (2004b) mapped 544 markers, including 49 (41%) fully informative SSRs out of 105 SSRs mapped. Our maps include 368 and 317 markers for the female and male map, respectively. Thus to reach a better coverage of the genome and obtain the 19 linkage groups, it will be necessary to add some new markers. On the other hand, Cervera *et al*. (2001) mapped 385 markers including 40 SSRs in the *P. nigra* map and they obtained a larger number of linkage groups than in our *P. nigra* maps. In fact, highly polymorphic, multi-allelic markers which detected all the four allelic variants of the mating configuration, such as SSRs, contain more genetic information (Grattapaglia and Sederoff, 1994). The use of a major number of fully informative markers allows to obtain a number of linkage groups closer to the number of chromosome. Among our polymorphic SSRs, 67% were fully informative. This high proportion of fully informative markers could explain the best result obtained in this study with less markers compared to the *P. nigra* map published by Cervera *et al*. (2001). Actually, Cervera *et al*. (2001) published the *P. nigra* map based on 404 markers and composed by 34 linkage groups while in the same study *P. deltoides* and *P. trichocarpa* map amounted to 466 and 364 markers, respectively, and the linkage analysis resulted in 21 and 23 linkage groups for the two species, respectively. The estimated heterozygosity levels of *P. nigra* and *P. trichocarpa* were in the same range but 40 SSRs were mapped for *P. nigra* instead of 59 and 60 for *P. deltoides* and *P. trichocarpa,* respectively. **In addition to their good transferability, SSRs markers demonstrate great efficiency for the construction of genetic maps**.

4.3.2 Genome length

In this study, the *P. nigra* observed genome length, including all the markers, was 2789 cM and 2816 cM for the female and male map, respectively. Values reported in literature vary among studies: Bradshaw *et al.* (1994) observed 1261 cM for the *P. trichocarpa* x *P. deltoides* map; Frewen *et al.* (2000) observed a genome length of 2002 cM for the *P. trichocarpa* map and 1778 cM for the *P. deltoides* map; Wu *et al.* (2000) observed 2927 cM for the *P. deltoides* map; Yin *et al.* (2002) observed 3801 cM for the *P. deltoides* map. Some discrepancies may be partly due to differences in genome coverage and recombination rates between the parents of the cross (Remington *et al.*, 1999). Actually, several parameters, including environmental (e.g., temperature, stress conditions), physiological (e.g., age), and genetic effects, are known to influence the frequency and distribution of crossover events (Lashermes *et al.*, 2001). An additional reason for the differences in observed genome length is the choice of mapping function. Unfortunately, there is no standard for the mapping function and, consequently, it is not possible to compare the different genome lengths observed. We chosen Kosambi function because it is considered closer to the biological reality (Arcade, 1999; Casasoli, 2004) and it is mainly used in plant mapping. The maps presented by Cervera *et al.* (2001) and Yin *et al.* (2004b) are very interesting for comparative mapping but Cervera *et al.* (2001) calculated the length of the genome with Kosambi function, whereas Yin *et al.* (2004b) used Haldane function. Only the maps obtained with the same mapping function can be correctly compared in order to better understand the variations of recombination rate among different genotypes.

Additionally, some overestimates of linkage map length may be attributed to genotyping errors. They occur in every laboratory and they are difficult to detect, unless they lead to Mendelian inconsistencies in the data (Remington *et al.*, 1999; Göring and Terwilliger, 2000a; Göring and Terwilliger, 2000b). The *P. nigra* data set produced in this study was accurately checked for genotyping errors and particularly was paid attention to the distorted markers. Even though this checking genotyping errors, a possibility of unintentional mistakes remains. Double crossing-overs and possibly misscored individuals or loci can be identified by specific commands in various mapping software packages. Therefore, a comparison of map length with and without error detection may give some indication of the level of error in the data set (Yin *et al.*, 2004b). For the current maps, our estimates differed by 685 cM and 349 cM for the female and male map, respectively. These values indicate that genotyping errors are rather low but confirm that they are not absent. Nevertheless, the difference between the estimate with and without error detection decreased considerably for the framework map, where we observed differences of

198 cM and 123 cM for the female and male map, respectively. With a difference of 250.4 cM, Yin *et al*. (2004b) consider that genotyping error is quite low in their map and the estimates of the *Populus* map length are highly accurate. **Thus, the framework maps presented in this study can be considered accurate for the detection of QTLs**.

The results of this study are presented (representation of the maps, segregation distortion, marker distribution, and correlation between genetic and physical length) without error detection. The error detection function is a good tool to check the potential genotyping errors. After the genotyping verification, the errors detected by the software are likely biological events. Therefore, in this case, error detection function could introduce artifacts and bias the dataset. Additionally, it has been shown that there are very important local variations of the recombination rate. For example, there is a restriction of recombination near the centromere (De Vienne, 1997). The lack of information about genotyping errors is another factor that makes difficult to compare map length derived from different studies with differing error rates. Among the published *Populus* maps, only Yin *et al*. (2004b) has given the genome length with and without error detection.

The comparison between the observed and estimated genome length provide an indication about the genome coverage. The estimation of the genome length calculated according to Chakravarti *et al*. (1991) seems to be less efficient when there is a difference between the number of linkage groups and the number of haploid chromosomes. In fact, we obtained 26 linkage groups for the male map representing 2816 cM and the estimated genome length was 3572 cM, whereas 20 linkage groups, representing 2789 cM, were obtained for the female map and the estimated genome length was 2726 cM. The results obtained by Cervera *et al*. (2001) are consistent with this trend. Forty linkage groups representing 2791 cM were observed for their *P. nigra* map and their estimate of the genome length was 3869 cM.

The comparison of genome length among the maps available could help in estimating the most likely genome size of *Populus*. Nevertheless, the discrepancies in the way the different maps are obtained make this comparison difficult.

4.3.3 Marker distribution

The integration of different kind of markers (AFLP, SSR and SNP) allowed to reach a good coverage of the genome in this study. Homogeneous marker density was found among linkage groups. Nevertheless, at linkage group scale clustered and dispersed regions were also found. Tanksley *et al*. (1992) observed the same type of markers distribution in the tomato map. While

the number of markers among linkage groups appeared to be uniformly distributed according to chromosome size, the distribution of markers within chromosome varied dramatically, depending on which part of the linkage group is being examined (Tanksley *et al.*, 1992). The majority of linkage maps of trees revealed clustered and dispersed regions in *Picea abies* (Paglia *et al.*, 1998), *Pinus taeda* (Remington *et al.*, 1999; Sewell *et al.*, 1999), *Castanea sativa* (Casasoli *et al.*, 2001), *Populus* spp. (Cervera *et al.*, 2001; Yin *et al.*, 2004b). The occurrence of sequences that are hot spots for recombination may explain heterogeneities in marker density along the map as the recombination suppression region. For example, genes near the centromere or the telomere recombine at very low frequency (Tanksley *et al.*, 1992; Lashermes *et al.*, 2001; Fu *et al.*, 2002). It has also been shown that meiotic recombination in eukaryotes takes place mostly in genes (Fu *et al.*, 2002).

Dispersed and clustered regions represent hindrances to achieve an even coverage of genetic maps. On the other hand, these heterogeneities in markers density may represent interesting clues for a better understanding of genome structure.

4.3.4 Segregation distortion

Segregation distortion of molecular markers has commonly been observed in mapping population of crops (Bert *et al.*, 1999), forest trees, and fruit trees (Bradshaw and Stettler, 1994; Grattapaglia and Sederoff, 1994; Paglia *et al.*, 1998; Barreneche *et al.*, 1998; Marques *et al.*, 1998). The reasons for segregation distortion in plants are not well understood but are thought related to factors such as chromosome loss, structural rearrangement, genetic load, genetic isolating mechanism, or viability genes. Non-biological factors such as scoring or sampling errors may also result in segregation distortion (Kuang *et al.*, 1999). Bradshaw and Stettler (1994) found that a lethal allele in *Populus* spp. affecting embryo development was the cause of segregation distortion of markers linked to it. Kuang *et al.* (1998) found an allele responsible for seedling death in *Pinus radiata.* As the presence of skewed markers is mainly biological, markers with significant deviation from Mendelian segregation were not excluded because of the interesting patterns of distortion segregation in the genome. On the other hand, it is easier to detect technical artifacts with the SSR pattern than with the AFLP pattern. For this reason it was chosen to exclude AFLPs distorted data at $P < 0.01$. Even if including skewed markers increases the chance of false linkage, we did not found aberrant data with the distorted markers. Actually, they were mapped on the expected linkage groups and only two and four skewed markers remained unlinked for the female and male map, respectively. The proportion of

distorted markers (P=0.05) in this study was 10% and 14% for the female and male map, respectively. Similar proportion of distorted markers were observed in previous genetic linkage maps using inter- or intra-specific crossing population: 16% in *Lolium perenne* (Bert *et al.*, 1999), 15% in *Eucalyptus globulus* and *Eucalyptus tereticornis* (Marques *et al.*, 1998), 22% in *P. tomentosa* and 13% in *P. tomentosa x P. bolleana* (Zhang *et al.*, 2003).

In this study 85% and 89% of skewed markers were found, clustered in five and six linkage groups for the female and male map, respectively. The clustering of distorted markers was also reported in *Populus* (Cervera *et al.*, 2001; Yin *et al.*, 2004b), *Eucalyptus* spp. (Marques *et al.*, 1998), and *Medicago* (Jenczewski *et al.*, 1997). Distorted markers that appear in clusters suggest that these areas contain genes that affect viability (Barreneche *et al.*, 1998; Cervera *et al.*, 2001). Yin *et al.* (2004a) mapped two resistance loci against *Melampsora* rust on linkage groups IV and XIX. Like Cervera *et al.* (2001), they observed that markers co-segregating with the *Melampsora* resistance genes showed significant segregation distortion. Yin *et al.* (2004b) observed extensive regions of distortion covering nearly the entire length of the linkage group IV in *P. trichocarpa* x *P. deltoides*. A region of distortion, covering 48% of the linkage group IV, was also observed in the *P. nigra* female map of this study. However, this distorted region did not correspond to the region where Yin *et al.* (2004a) mapped the *Melampsora* resistance gene. Moreover, a region of distortion for the linkage group XIX was not found in *P. nigra* maps. During the propagation phase, the *P. nigra* pedigree experienced rust attacks, but plant survival was not affected (Sabatti, personal communication). It is likely that no selection pressure was exerted by rust on the mapping pedigree. Therefore, segregation distortion in genome regions close to rust resistance genes was not expected. The size of the distorted regions suggests that multiple loci were involved in causing the segregation distortion. It has been shown that resistance genes typically occur in large clusters in plant genomes. The clustered occurrence of disease-resistance genes appears to play a central role in the generation and maintenance of the tremendous diversity observed in these gene families, as domains are shuffled within and between clusters because of large-scale insertion/deletion events and unequal crossing-over and gene conversion (Young, 2000; Meyers *et al.*, 2003).

There is a variety of other potential selective factors, such as photoperiod, frost tolerance, drought tolerance, and rooting ability, that would have likely segregated in our pedigree and may have resulted in differential survival of progeny. In the case of photoperiod and frost tolerance, selection against the paternal alleles might be expected because the male parent comes from Southern Italy (40°09'N, 16°41'E), whereas the female parent originates from Northern Italy

(45°09'N, 7°01'E) approximately at the same latitude where the pedigree was initially propagated (44°21'N, 8°17'E). Interestingly, the distorted region on the linkage group X of the male parent contains the *PhyB2* gene, which presents segregation distortion at 1% level. Near this gene, Frewen *et al*. (2000) mapped a QTL for the bud burst. However, there are many possible alternative explanations for the patterns of segregation distortion that were observed. Distorted markers dispersed in different linkage groups, suggest that the distortion could be more likely due to genotyping error or chance rather than biological effect.

Skewed markers lead to increased difficulty in both linkage determination and recombination frequency estimation, thereby eventually affecting the map construction (Whitkus, 1998). **However, segregation distortion has some biological explanations and the exclusion of skewed markers could bias the data set and make genome region uncovered. Mapping skewed markers provide a tool to understanding the mechanism of segregation distortion**. Segregation distortion is increasingly recognized as a potentially powerful evolutionary force. Therefore one key to understand the evolutionary importance of non-Mendelian genes is to investigate how often these genetics conflicts arise and how they are resolved in nature. (Taylor and Ingvarsson, 2003).

4.4 Comparative mapping

Comparative mapping relies on mapping orthologous loci in two or more species and comparing their position along homologous linkage groups. With the availability of genetic linkage maps, comparative mapping studies among phylogenetically related species were widely developed in the last decade (Nadeau and Sankoff, 1998). Transferring genetic and molecular information from one species to another, as well as discovering the main processes of genome evolution, are two promising prospects in developing comparative mapping studies. In crops such studies showed that gene content (synteny), marker order (co-linearity), and QTLs position are often conserved among phylogenetically related species (Paterson *et al.*, 2000; Schmidt, 2000). In forest trees, few examples have reported the transferability of genetic and molecular information among closely related species within the same genus: in pine (Sewell *et al.*, 1999; Chagné *et al.*, 2003), eucalypt (Marques *et al.*, 2002), and poplar (Cervera *et al.*, 2001). RFLP, SSR and EST markers were successfully used to identify orthologous loci and to align homologous linkage groups. The high number of SSR markers that were mapped in this study on the *P. nigra* maps enables a comparison with the other *Populus* maps including SSRs.

4.4.1 *Populus nigra* parental maps and *Populus trichocarpa* physical map

The alignment of our *P. nigra* parental maps and the *P. trichocarpa* physical map allowed to identify the homologous linkage groups and control the quality of the generated maps. The alignment is based on 69 markers in common between the two *P. nigra* maps and 102 in total between the two maps and the *P. trichocarpa* map. The comparison of the order of markers between the two *P. nigra* maps indicates a good reliability of the maps, 91% of the common markers were in the same order. The alignment with the *P. trichocarpa* map revealed seven discrepancies. Some of them are likely due to genotyping errors. However, Sterck *et al.* (2005) demonstrate that a large-scale gene-duplication event had occurred for seven poplar species. From their analysis, based on the publicly available EST collection, it is clear that all poplar species share the same large-scale gene-duplication event, suggesting that this event must have occurred in the ancestor of poplar, or at least very early in the evolution of the *Populus* genus. Moreover, duplicated regions of all linkage groups were clearly identified in the whole genome sequence of *P. trichocarpa*, (Tuskan, 2006 personal communication) (Fig. 37). In the present study, the marker ORPM276 is mapped on the linkage group XIII for the female parent and on the linkage group XIX for the male parent as in the *P. trichocarpa* physical map. As showed in Fig. 37, these two groups present duplicated region. Thus, it could be possible that a locus of the group XIII was amplified for the female and a locus of the group XIX for the male. The pattern of the marker ORPM276, does not contradict this hypothesis because one DNA fragment (a,-) was observed for the female parent and another (b,-) for the male parent. Two other discrepancies could be explained in the same way. The marker PMGC2696 was mapped on the linkage group VIII in the female parent, whereas it was found on the linkage group X in the *P. trichocarpa* map. These two groups (X and VIII) are highly duplicated between them, more than 50% of their length (Fig. 37). The marker ORPM40 was mapped on the group II for the male parent and on the group XIV in the *P. trichocarpa*. These two groups, II and XIV, are also highly duplicated. About 80% of the linkage group XIV correspond to the group II. **Therefore, the high level of duplication of the *Populus* genome could explain some discrepancies. It could also explain why some groups are difficult to map.**

The correlation between the genetic and the inferred physical map length for the whole genome of the two parents tends to be linear. At the linkage group level, a high heterogeneity of the ratio between the *P. trichocarpa* physical map length and the genetic distance of the *P. nigra* maps was observed, from 47 to 292 Kbp/cM and 45 to 221 Kbp/cM for the female and the male

parent, respectively. These results could be partially explained by the presence of only two common markers on some groups for which the ratio was calculated on only one interval. On the other hand, we have previously seen that the recombination ratio is variable in the genome. Assuming a random distribution of markers, low levels of meiotic recombination would induce markers physically well separated to cluster on linkage map (Tanksley *et al.*, 1992). Although, while the amount of nuclear DNA in eukaryotes varies greatly, the total length of their genetic map does not (Fu *et al.*, 2002). **The relationship between genetic and physical map distance could be an interesting tool to determine the hot spots or suppression regions for recombination and to have a better understanding of the genome structure.**

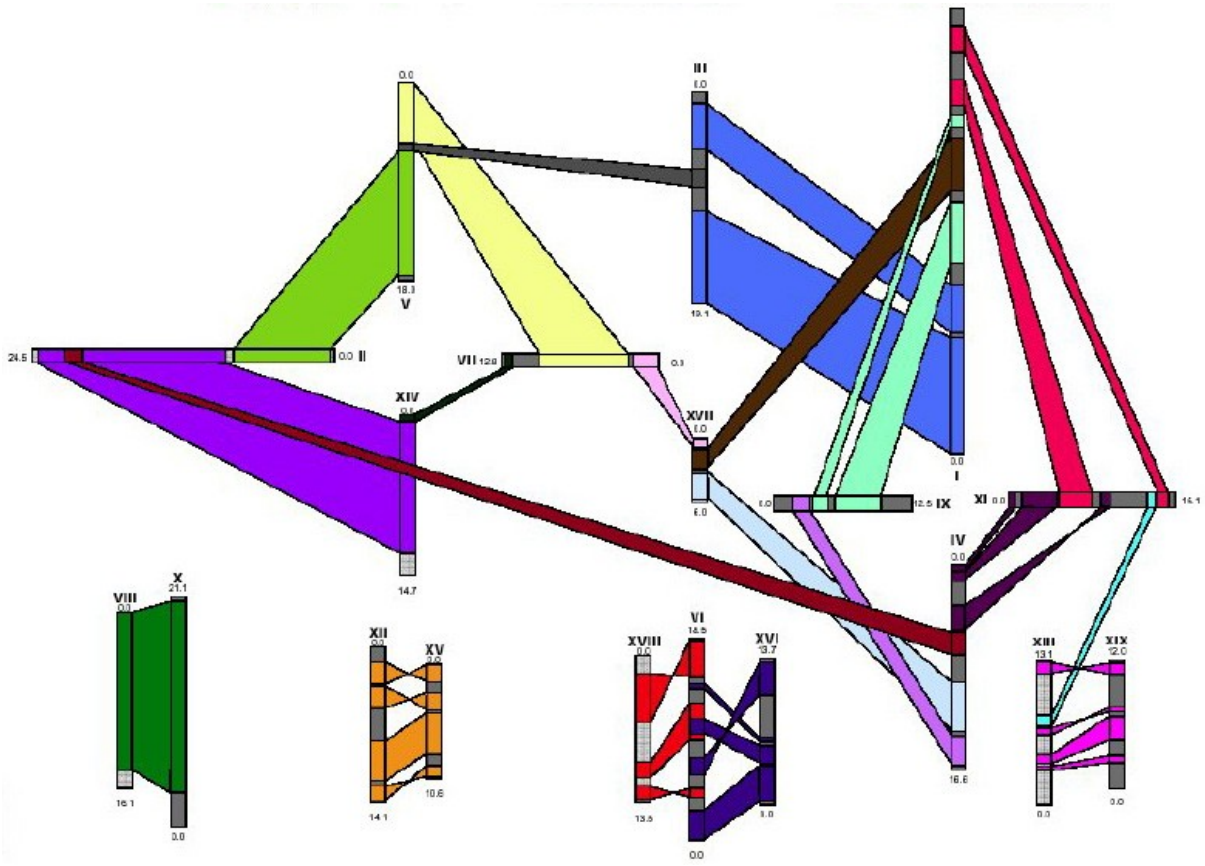


Fig. 37: The *Populus* genome duplications (Tuskan, 2006 personal communication).

4.4.2 Comparison among *Populus nigra*, *Populus alba* and *Populus trichocarpa* maps.

The comparison of the *P. nigra* maps with the *P. alba* maps produced in the same laboratory (Paolucci, personal communication) was interesting because they were constructed with the same strategy for the choice of the parental genotypes with similar origins, and for the kind of markers used. In this way a relatively high number of common SSRs (56) was found. Discrepancies were observed between the two species, and some of them could be due to the duplication of the *Populus* genome. Like in the *P. nigra* male parent, the marker ORPM40 was mapped on the linkage group II in the *P. alba* female parent whereas it is mapped on the group XIV of the *P. trichocarpa* physical map. As we have noticed before, these two groups are highly duplicated. The duplication of the groups I and III (Fig 37) could explain the mapping of the marker ORPM399 in the group I of the *P. nigra* female and in the group III of the *P. alba* male. We observed a very good synteny and co-linearity between *P. nigra* and *P. alba* maps. However, the number of common markers and the genome coverage is too low to understand if there are some chromosomal rearrangements between these two species belonging to different *Populus* sections (*Aigeiros* and *Populus).*

The comparison of the *P. nigra* maps generated in this study with the *P. nigra map* by Cervera *et al*. (2001) reveals 27 common markers and no discrepancy. These results indicate a good reliability of our data. The comparison of the three species, *P. alba*, *P. nigra*, and *P. trichocarpa*, reveals an interesting synteny and co-linearity. An alignment of QTLs could be also done to identify common QTLs in different genetic backgrounds. The result of these comparisons could be summarized in the construction of a poplar consensus map which would be useful in marker-assisted selection. At the moment, it is however difficult to find common markers among all the available maps published for the *Populus* genus. Only one common marker was found among the six maps compared. Some maps should be enriched in SSR markers to make comparative mapping more efficient. On the other hand, as the *P. trichocarpa* genome sequence includes almost all the SSRs used in the *Populus* spp. genetic maps, it may be more efficient to align them with the *P. trichocarpa* physical map.

Discrepancies which could not be explained by the *Populus* genome duplications are likely due to genotyping errors. It should be also possible that they reveal species-specific chromosomal rearrangements. To go thoroughly into these observations it would be necessary to sequence some loci to confirm the conservation or the rearrangement of these loci among the compared species.

In this study, comparative mapping based on the co-alignment of common markers among genetic and physical maps enabled to correlate linkage information from different genetic maps and to validate the accuracy of locus ordering from the different mapping strategies. Comparative mapping also allows the comparison of genome structure within the genus *Populus* and, thus, to study chromosomal evolution by detecting chromosome rearrangements (Cervera *et al.*, 2004). Moreover, the alignment of the genetic map with the poplar genome sequence allows to locate large numbers of candidate genes on the genetic maps and to compare their map position with QTLs.

5 Conclusions and Perspectives

5.1 Genetic mapping

The accurate choice of two *P. nigra* genotypes with divergent phenotypes from natural italian populations, as parents of a controlled intra-specific cross, has originated an efficient mapping pedigree. The first genetic maps of natural italian *P. nigra* were constructed from this pedigree. The double pseudo-testcross strategy, mainly used in forest tree mapping, allowed to construct two highly informative maps in a relatively short time. Actually, these maps include the highest number of interesting and original markers such as SSRs, SNPs, and genes among *Populus* spp. maps published so far. The two *P. nigra* maps obtained are very similar, as well as mapped markers (317 and 368 markers including 91 and 106 SSRs, 3 and 5 SNPs, 4 and 9 genes for the female and male map, respectively), and observed genome length (2789 cM and 2816 cM for the female and male map, respectively). The 19 linkage groups corresponding to the haploid number of *Populus* spp. chromosomes were identified thanks to the availability of the *P. trichocarpa* physical map and the SSRs already mapped by Cervera *et al.* (2001). Moreover, the sex morphological trait was mapped in the male parent. Therefore, for the first time in *Populus* spp., there is an information on the genomic region involved in sex determination.

With all these characteristics, the two maps form a powerful resource for answering to scientific and applicative needs such as:

- a better understanding of the structure and the evolution of the *Populus* genome (comparative mapping);
- the possibility to identify and localize genes or genome regions responsible for traits of economic importance (QTL mapping and map-based cloning);
- providing information which can be used in a program of advanced tree selection, improvement, and breeding (informative maps for the marker-assisted selection).

The information brought by the *P. nigra* maps presented here rises new questions such as how many genes and how many linkage groups are involved in the determination of gender?

The results of this study show that there are a high synteny and co-linearity among the *Populus* species: are there some species-specific chromosomal rearrangements? It could be interesting to enrich the *P. nigra* maps presented here with new SSRs and SNPs in order to increase the density of informative markers and obtain only the 19 linkage groups corresponding to the *Populus* haploid chromosome number. We could pay a particular attention to the linkage group

XIX and try to saturate it. In this way, it could be possible to find gender-related markers and have a better comprehension on the mechanism of sex determination in poplar.

5.2 Applications of comparative mapping

The comparison of the *P. nigra* maps with the *Populus spp.* maps from other studies validates the accuracy of the locus ordering and shows an important synteny and co-linearity. For a more detailed comparison it could be useful to establish standard methods for genome length calculation, such as the choice of mapping function (Haldane or Kosambi) and the way of managing the potential genotyping errors. The comparison of the distance among the common markers could provide important information on the conservation of hot spots and suppression regions for recombination on the genome structure of the species. Therefore, comparative mapping allows two kinds of approach:

- finding the differences among species of the *Populus* genus to understand the evolution of the genome. For this approach it would be necessary to carry out micro-synteny studies at the level of genes and the *P. nigra* maps should be highly enriched in markers;
- paying attention to the genome co-linearity for the construction of a *Populus* consensus map where SSRs constitute anchor points for specific genome region of the different *Populus* species. This consensus map would be used to map traits of economic or scientific interest. For this purpose, the maps presented here are already informative, thanks to their high number of SSRs, SNPs, and genes included. The construction of a *Populus* consensus map is one of the objectives of the European project Popyomics. Indeed, the availability of high-density consensus maps greatly facilitates the construction of new maps and the mapping of specific chromosomal regions (Collard *et al.*, 2005). The *Populus* consensus map would be a very useful tool for both fundamental research and breeding.

5.3 QTL mapping

The construction of relatively dense and accurate framework maps such as those presented in this study facilitate the dissection of complex traits and enable us to study the genetic basis of QTL. Indeed, one of the main application of a genetic map is the mapping of QTLs for a number of traits of importance for yield and disease resistance. Actually, the *P. nigra* mapping pedigree was duplicated and maintained in two different sites, Cavallermaggiore (Cuneo province) in Northern Italy and Montelibretti (Rome province) near Rome, in Central Italy. Measurements of bud set, bud flush, and growth traits (height, circumference, number of sylleptic branches, leaf

area) were carried out in 2003, 2004 and 2005 (Ricciotti, personal communication). The mapping of QTLs for these traits is on going. Their map position could be compared to QTLs mapped in the other *Populus* species analyzed for the same traits (in the frame of the Popyomics project four pedigrees have been analyzed for these traits). QTLs which are co-located could reveal with high probability the regions involved in the expression of the trait. The co-located QTLs could be added to the poplar consensus map.

Another use of the genetic linkage maps consists in the determination of the link between QTLs and the underlying gene or genes (Taylor, 2002). Mapping QTLs is one of the first steps to associate the genes involved in the expression of a particular trait. There are two strategies to attain this purpose: the candidate gene approach and the map-based cloning approach. Candidate gene analysis is based on the hypothesis that known-function genes (the candidate genes) could correspond to loci controlling traits of interest (Pflieger *et al.*, 2001). Some candidate genes (*PhyA, PhyB1, PhyB2*) were already mapped in this study. The mapping of the QTLs for the bud set and bud flush could enable to check if some of the candidate genes are co-located to the QTLs and responsible for the genotypic variation of the trait. To determine the most likely candidate, fine-mapping should be performed to reduce the confidence interval of QTLs (Pflieger *et al.*, 2001). Therefore, if a candidate gene seems to be correlated to a trait, it could be interesting to increase the markers density of the region containing the QTL. The availability of the physical map of *P. trichocarpa* enables to target markers for fine mapping of specific regions. Fine-mapping is also useful for the second approach, the map-based cloning. Saturated linkage maps allowed the molecular cloning of novel genes on the basis of map position alone, without prior knowledge of the encoded protein (Gebhardt *et al.*, 2005). Map-based cloning involves the use of tightly linked markers to isolate target genes by using markers as a "probe" to screen a genomic library (Collard *et al.*, 2005). The maps obtained in this study are in favor of the candidate gene approach but the map-based cloning could be possible in a longer term by mapping additional markers in specific genome regions.

5.4 An application of the *Populus* genome knowledge: marker-assisted selection

Most of the works on genetic and QTL mapping in forest trees conclude with the perspective to provide tools for breeding and marker-assisted selection (MAS). Will molecular tree breeding become a reality in the near future?

Selecting plants that contain appropriate gene combinations in a segregating progeny is a critical component of plant breeding. Moreover, plant breeders typically work with large populations composed by hundreds or even thousands of genotypes. MAS may greatly increase the efficiency of plant breeding compared to conventional breeding methods. Once markers tightly linked to genes or QTLs of interest have been identified, breeders may use specific DNA marker alleles as a diagnostic tool to identify plants carrying the genes or QTLs, prior to field evaluation of a large number of plants (Collard *et al.*, 2005). DNA marker technology has been used in commercial plant breeding programs since the early 1990s and has proved to be useful for rapid and efficient transfer of economically interesting traits into agronomically desirable varieties and hybrids (Sharma *et al.*, 2002). Three main kinds of MAS program are used for early selection of plants to be maintained or used for new crosses. Firstly, there is the simple check for the presence of the allelic form of a gene, which brings a resistance or an economic trait, thanks to the tightly linked markers. Secondly, there is the introgression of a gene or a trans-gene in elite variety by backcross. The plants containing the new gene from the donor parent and the maximum of alleles from the elite recurrent parent are selected. The third MAS program consists in the accumulation in a plant of the alleles with a positive effect for quantitative traits of economic interest. The development of these programs has been possible thanks to the advances in genetics and genomics in crops species: the complete genome sequences of *Oryza* and *Arabidopsis*, the enormous number of ESTs and highly informative markers such as SSRs and SNPs, the genetic maps including QTLs, the physical maps, and the ease of genetic transformation (Fig. 38). Today, the genomic resources of *Populus* are comparable to those available for crop species: the complete genome sequence of *P. trichocarpa*, the ESTs database available for seven different poplar species, the large number of SSR (more than 4000), the tools to develop SNP markers, the numerous genetic maps including QTLs, and the ease of genetic transformation (Fig. 38). However, an important difference between crop species and trees still remains: the generation time to obtain mature plants. On the other hand, poplar improvement programs with traditional breeding methods already exist. The main advantages of MAS are also true for trees: i) time saving from the substitution of complex fields trial with molecular markers; ii) elimination of unreliable phenotypic evaluation associated with fields trials due to environmental effects; iii) selection of genotypes at a seedling stage (less plants to maintain, space saving) (Collard *et al.*, 2005). Many genetic resources for poplar improvement already exist and other ones will become available in a nearest future. Nevertheless, the economic aspect still represents a real barrier to MAS in poplar. The genus *Populus* makes an

important contribution to meeting the global need for paper, timber, and other wood-based products. Besides, new promising opportunities are increasing the interest around cellulosic compounds for the market of energy and biofuels (Farrell *et al.*, 2006). The role of fast-growing trees, like *Populus*, in carbon mitigation aligned to the Kyoto Protocol is also being quantified and may be considerable (Taylor, 2002).Therefore, *Populus* is a forest tree of considerable commercial importance. Molecular tree breeding could really become a reality but the governmental funding agencies and the commercial sector should believe and finance these powerful technologies for poplar improvement.

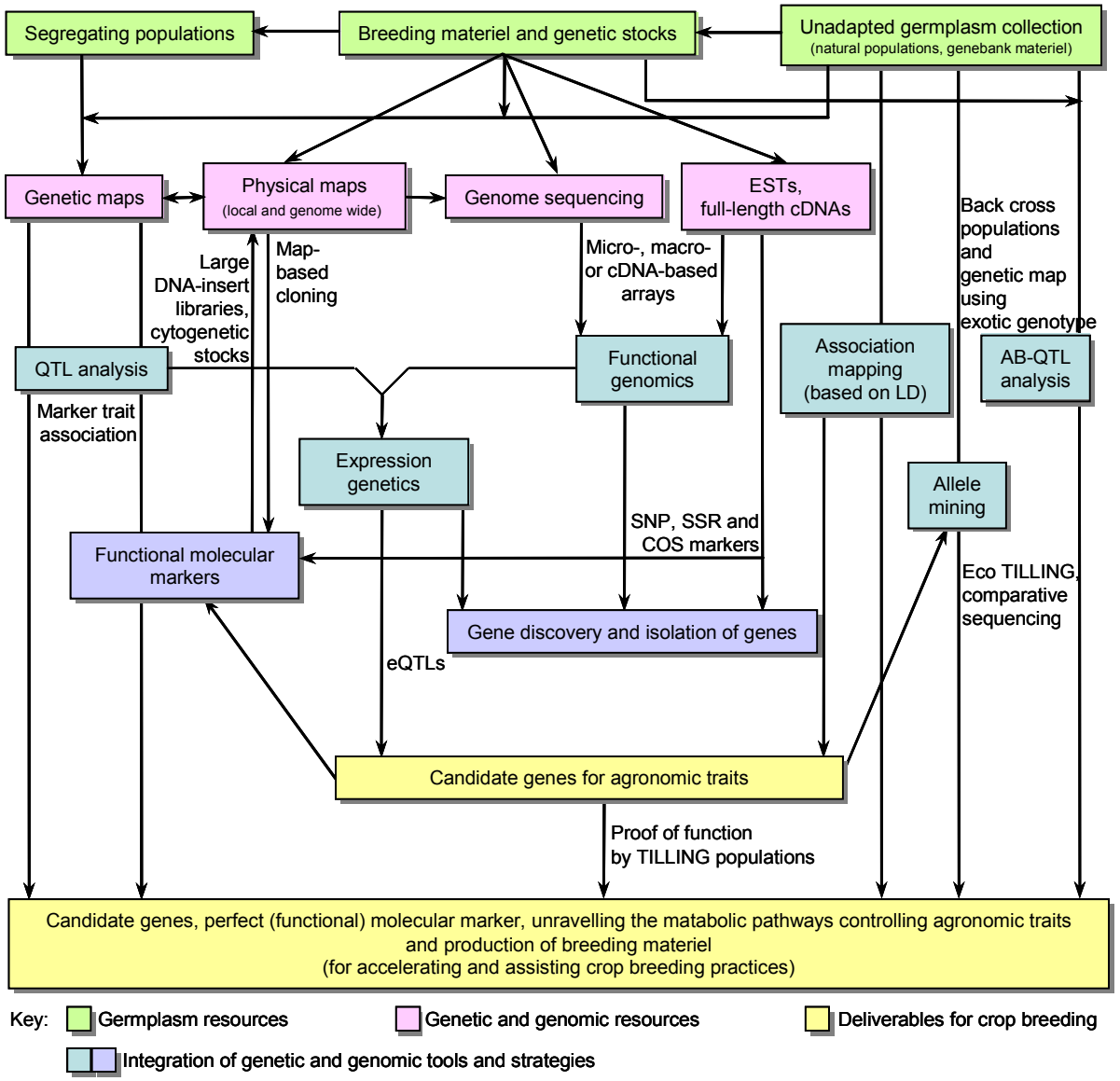


Fig. 38: An integrated view of exploitation of genomic resources for plant improvement via different genetic and genomic strategies. From Varshney *et al.* (2005)
Abbreviations: AB-QTLs, advanced backcross QTL; COS, conserved orthologous set; eQTLs, expression QTLs; ESTs, expressed sequence tags, LD, linkage disequilibrium; QTL, quantitative trait loci; SNP, single nucleotide polymorphism, SSR, simple sequence repeat; TILLING, target induced local lesions in genome.

Bibliographic references

Acheré,V., Faivre-Rampant,P., Jeandroz,S., Besnard,G., Markussen,T., Aragones,A., Fladung,M., Ritter,E., and Favre,J.-M. (2004) A full saturated linkage map of *Picea abies* including AFLP, SSR, ESTP, 5S rDNA and morphological markers. Theoretical and Applied Genetics **108**:1602-1613.

Al Chaarani,G.R., Gentzbittel,L., Sarrafi,A., and Wedzony,M. (2005) Identification of QTLs for germination and seedling development in sunflower (*Helianthus annuus* L.). Plant Science **169**:221-227.

Alström-Rapaport,C., Lascoux,M., and Gullberg,U. (1997) Sex determination and sex-ratio in *Salix viminalis* L. Theoretical and Applied Genetics **94**:493-497.

Alstrom-Rapaport,C., Lascoux,M., Wang,Y.C., Roberts,G., and Tuskan,G.A. (1998) Identification of a RAPD marker linked to sex determination in the basket willow (*Salix viminalis* L.). Journal of Heredity **89**:44-49.

Altschul,S.F., Gish,W., Miller,W., Myers,E.W., and Lipman,D.J. (1990) Basic local alignment search tool. Journal of Molecular Biology **215**:403-410.

Arcade, A. (1999) Détection de locus controlant des caractères quantitatifs dans un plan de croisement factoriel: application au mélèze. Thesis/Dissertation. Université d'Orléans, France, 9-11-1999. 111 pages.

Arcade,A., Anselin,F., Faivre-Rampant,P., Lesage,M.C., Pâques,L.E., and Prat,D. (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. Theoretical and Applied Genetics **100**:299-307.

Arens,P., Coops,H., Jansen,J., and Vosman,B. (1998) Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. Molecular Ecology **7**:11-18.

Barreneche,T., Bodenes,C., Streiff,R., Plomion,C., Roussel,G., Kremer,A., Lexer,C., Steinkellner,H., Glössl,J., Trontin,J.F., Favre,J.M., Fluch,S., and Burg,K. (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.

Barreneche,T., Casasoli,M., Russell,K., Akkak,A., Meddour,H., Plomion,C., Villani,F., and Kremer,A. (2004) Comparative mapping between *Quercus* and *Castanea* using simple-sequence repeats (SSRs). Theoretical and Applied Genetics **108**:558-566.

Barsoum,N. (2001) Regeneration - requirements and promotion measures. In: In situ conservation of *Populus nigra*, F.Lefevre, N.Barsoum, B.Heinze, D.Kajba, P.Rotach, S.M.G.De Vries, and J.Turok, eds (Rome: International Plant Genetic Resources Institute), pp. 16-24.

Benetka,V., Bartàkovà,I., and Mottl,J. (2002) Productivity of *Populus nigra* L. ssp. *nigra* under short-rotation culture in marginal areas. Biomass and Bioenergy **23**:327-336.

Bert,P.F., Charmet,G., Sourdille,P., Balfourier,F., and Hayward,M.D. (1999) A high-density molecular map for ryegrass (*Lolium perenne*) using AFLP markers. Theoretical and Applied Genetics **99**:445-452.

Bisoffi,S., Gemignani,G., Gras,M.A., May,S., and Mughini,G. (1987) Establishment of *Populus nigra* genetic reserves in Italy. Genet. Agr. **41**:105-114.

Bisoffi,S. and Gullberg,U. (1996) Poplar breeding and selection strategies. In: Biology of *Populus* and its implications for management and conservation, R.F.Stettler, H.D.Bradshaw, Jr., P.E.Heilman, and T.M.Hinckley, eds (Ottawa: NRC Research Press, National Council of Canada), pp. 139-158.

Boutin-Ganache,I., Raposo,M., Raymond,M., and Deschepper,C.F. (2001) M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele-sizing methods. BioTechniques **31**:24-28.

Braatne,J.H., Rood,S.B., and Heilman,P.E. (1996) Life history, ecology and conservation of riparian cottonwoods in North America. In: Biology of *Populus* and its implications for management and conservation, R.F.Stettler, H.D.Bradshaw, Jr., P.E.Heilman, and T.M.Hinckley, eds (Ottawa: NRC Research Press, National Council of Canada), pp. 57-85.

Bradshaw,H.D. and Stettler,R.F. (1994) Molecular genetics of growth and development in *Populus*. II. Segregation distortion due to genetic load. Theoretical and Applied Genetics **89**:551-566.

Bradshaw,H.D.Jr., Villar,M., Watson,B.D., Otto,K.G., Stewart,S., and Steller,R.F. (1994) Molecular genetics of growth and development in *Populus*. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. Theoretical and Applied Genetics **89**:167-178.

Brondani,R.P., Brondani,C., and Grattapaglia,D. (2002) Towards a genus-wide reference linkage map for Eucalyptus based exclusively on highly informative microsatellite markers. Molecular Genetics and Genomics **267**:338-347.

Brookes,A.J. (1999) The essence of SNPs. Gene **234**:177-186.

Brunner,A.M., Rottmann,W.H., Sheppard,L.A., Krutovskii,K., DiFazio,S.P., Leonardi,S., and Strauss,S.H. (2000) Structure and expression of duplicate *AGAMOUS* orthologues in poplar. Plant Molecular Biology **44**:619-634.

Bundock,P.C., Cross,M.J., Shapter,F.M., and Henry,R.J. (2005) Robust allele-specific polymerase chain reaction markers developed for single nucleotide polymorphisms in expressed barley sequences. Theoretical and Applied Genetics **112**:358-365.

Casasoli, M. (2004) Carthographie genetique comparée chez les fagacées. Thesis/Dissertation. Université de Bordeaux I, INRA-UMR BIOGECO, France, 14-12-2004. 265 pages.

Casasoli,M., Derory,J., Morera-Dutrey,C., Brendel,O., Porth,I., Guehl,J.-M., Villani,F., and Kremer,A. (2005) Comparison of QTLs for adaptive traits between oak and chestnut based on an EST consensus map. Genetics **172**:533-546.

Casasoli,M., Mattioni,C., Cherubini,M., and Villani,F. (2001) A genetic linkage map of European chestnut (*Castanea sativa Mill.*) based on RAPD, ISSR and isozyme markers. Theoretical and Applied Genetics **102**:1190-1199.

Cervera,M.T., Sewell,M.M., Faivre-Rampant,P., Storme,V., and Boerjan,W. (2004) Genome mapping in *Populus*. In: Molecular genetics and breeding of forest trees, S.Kumar and M.Fladung, eds (New York: Food Product Press), pp. 387-410.

Cervera,M.T., Storme,V., Ivens,B., Gusmao,J., Liu,B.H., Hostyn,V., Van Slycken,J., Van Montagu,M., and Boerjan,W. (2001) Dense genetic linkage maps of three *Populus* species (*Populus deltoides, P. nigra and P. trichocarpa*) based on AFLP and microsatellite markers. Genetics **158**:787-809.

Chagné,D., Brown,G., Lalanne,C., Madur,D., Pot,D., Neale,D., and Plomion,C. (2003) Comparative genome and QTL mapping between maritime and loblolly pines. Molecular Breeding **12**:185-195.

Chagné,D., Lalanne,C., Madur,D., Kumar,S., Frigerio,J.-M., Krier,C., Decroocq,S., Savouré,A., Bou-Dagher-Karrat,M., Bertocchi,E., Brach,J., and Plomion,C. (2002) A high density genetic map of maritime pine based on AFLPs. Annals of Forest Science **59**:627-636.

Chakravarti,A., Lasher,L.K., and Reefer,J.E. (1991) A maximum likelihood method for estimating genome length using genetic linkage data. Genetics **128**:175-182.

Chen,T.H.H., Howe,G.T., and Bradshaw,H.D. (2002) Molecular genetic analysis of dormancy-related traits in poplars. Weed Science **50**:232-240.

Collard,B.C.Y., Pang,E.C.K., Jahufer,M.Z.Z., and Brouwer,J.B. (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica **142**:169-196.

Crouzillat,D., Lerceteau,E., Petiard,V., Morera,J., Rodriguez,H., Walker,D., Phillips,W., Ronning,C., Schnell,R., Osei,J., and Fritz,P. (1996) *Theobroma cacao* L.: a genetic linkage map and quantitative trait loci analysis. Theoretical and Applied Genetics **93**:205-214.

Crow,J.F. and Dove,W.F. (1990) Anecdotal, historical and critical commentaries on genetics. Genetics **125**:669-671.

Dayanandan,S., Rajora,O.P., and Bawa,K.S. (1998) Isolation and characterization of microsatellites in trembling aspen (*Populus tremuloides*). Theoretical and Applied Genetics **96**:950-956.

De Vienne,D. (1997) Etablissement de cartes de liaisons genetiques. In: Les marqueurs moléculaires en génétique et biotechnologies végétales, D.De Vienne, ed (Paris) pp. 49-79.

De Vries,S.M.G. and Turok,J. (2001) Introduction. In: *In situ* conservation of *Populus nigra*, F.Lefevre, N.Barsoum, B.Heinze, D.Kajba, P.Rotach, S.M.G.De Vries, and J.Turok, eds (Rome: International Plant Genetic Resources Institute), pp. 5-7.

Devos,K.M. and Gale,M.D. (1997) Comparative genetics in the grasses. Plant Molecular Biology **35**:3-15.

Don,R.H., Cox,P.T., Wainwright,B.J., Baker,K., and Mattick,J.S. (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Research **19**:4008.

Eckenwalder,J.E. (1996) Systematics and evolution of *Populus*. In: Biology of *Populus* and its implications for management and conservation, R.F.Stettler, H.D.Bradshaw, Jr., P.E.Heilman, and T.M.Hinckley, eds (Ottawa: NRC Research Press, National Council of Canada), pp. 7-32.

Ellis,M.H., Rebetzke,G.J., Azanza,F., Richards,R.A., and Spielmeyer,W. (2005) Molecular mapping of gibberellin-responsive dwarfing genes in bread wheat. Theoretical and Applied Genetics **111**:423-430.

Farrell,A.E., Plevin,R.J., Turner,B.T., Jones,A.D., O'Hare,M., and Kammen,D.M. (2006) Ethanol can contribute to energy and environmental goals. Science **311**:506-508.

Fossati,T., Grassi,F., Sala,F., and Castiglione,S. (2003) Molecular analysis of natural populations of *Populus nigra* L. intermingled with cultivated hybrids. Molecular Ecology **12**:2033-2043.

Frewen,B.E., Chen,T.H., Howe,G.T., Davis,J., Rohde,A., Boerjan,W., and Bradshaw,H.D., Jr. (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in populus. Genetics **154**:837-845.

Fu,H., Zheng,Z., and Dooner,H.K. (2002) Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude. Proceedings of the National Academy of Sciences of the United States of America **99**:1082-1087.

Fukatsu,E., Isoda,K., Takahashi,M., Watanabe,A., and Hirao,T. (2005) Development and characterization of simple sequence repeat DNA markers for *Zelkova serrata*. Molecular Ecology Notes **5**:378-380.

Fulton,T.M., Van der Hoeven,R., Eannetta,N.T., and Tanksley,S.D. (2002) Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. The Plant Cell **14**:1457-1467.

Gebhardt,C., Ritter,E., Debener,T., Schachtschabel,U., Walkemeier,B., Uhrig,H., and Salamini,F. (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. Theoretical and Applied Genetics **78**:65-75.

Gebhardt,C., Schmidt,R., and Schneider,K. (2005) Plant genome analysis: The state of the art. International Review of Cytology **247**:223-284.

Gerber,S., Rodolphe,F., Bahrman,N., and Baradat,P. (1993) Seed-protein variation in maritime pine (*Pinus pinaster* Ait.) revealed by two-dimensional electrophoresis: genetic determinism and construction of a linkage map. Theoretical and Applied Genetics **85**:521-528.

Göring,H.H.H. and Terwilliger,J.D. (2000a) Linkage analysis in the presence of errors II: Marker-locus genotyping errors modeled with hypercomplex recombination fractions. American Journal of Human Genetics **66**:1107-1118.

Göring,H.H.H. and Terwilliger,J.D. (2000b) Linkage analysis in the presence of errors III: Marker loci and their map nuisance parameters. American Journal of Human Genetics **66**:1298-1309.

Gosselin,I., Zhou,Y., Bousquet,J., and Isabel,N. (2002) Megagametophyte-derived linkage maps of white spruce *Picea glauca* based on RAPD, SCAR and ESTP markers. Theoretical and Applied Genetics **104**:987-997.

Grant,M.C. and Mitton,J.B. (1979) Elevational gradient in adult sex ratios and sexual differentiation in vegetative growth rates of *Populus tremuloides* Michx. Evolution **33**:914-918.

Grattapaglia,D. and Sederoff,R. (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics **137**:1121-1137.

Haanstra,J.P.W., Meijer-Dekens,F., Van Den Berg,P., Lindhout,P., Wye,C., Verbakel,H., Peleman,J., Odinot,P., Van Heusden,A.W., and Tanksley,S. (1999) An integrated high-density RFLP-AFLP map of tomato based on two *Lycopersicon esculentum* x *L. pennellii* F2 populations. Theoretical and Applied Genetics **99**:254-271.

Haldane,J.B.S. (1919) The combination of linkage values, and the calculation of distance between the loci of linked factors. Belgian Journal of Genetics **8**:299-309.

Heinze,B. (1997) A PCR marker for a *Populus deltoides* allele and its use in studying introgression with native european *Populus nigra*. Belgian Journal of Botany **129**:123-130.

Helentjaris,T. (1987) A genetic linkage map for maize based on RFLPs. Trends in Genetics **3**:217-221.

Helentjaris,T., Slocum,M., Wright,S., Schaefer,A., and Nienhuis,J. (1986) Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theoretical and Applied Genetics **72**:761-769.

Hofmann,M. (2001) Economic aspects of black poplar - Hybrids, yield and utilization of wood-. In: Genetic diversity in river population of European Black Poplar. Implication for riparian eco-sytem management, B.C.Van Dam and S.Bordacs, eds (Budapest, Hungary) pp. 205-212.

Howe,G.T., Bucciaglia,P.A., Hackett,W.P., Furnier,G.R., Cordonnier-Pratt,M.M., and Gardner,G. (1998) Evidence that the phytochrome gene family in black cottonwood has one *PHYA* locus and two *PHYB* loci but lacks members of the *PHYC/F* and *PHYE* subfamilies. Molecular Biology and Evolution **15**:160-175.

Huentelman,M.J., Craig,D.W., Shieh,A.D., Corneveaux,J.J., Hu-Lince,D., Pearson,J.V., and Stephan,D.A. (2005) SNiPer: improved SNP genotype calling for Affymetrix 10K GeneChip microarray data. BMC Genomics **6**:149.

Hulbert,S.H., Ilott,T.W., Legg,E.J., Lincoln,S.E., Lander,E.S., and Michelmore,R.W. (1988) Genetic analysis of the fungus, *Bremia lactucae*, using restriction fragment length polymorphisms. Genetics **120**:947-958.

Imbert,E. and Lefèvre,F. (2003) Dispersal and gene flow of *Populus nigra* (Salicaceae) along a dynamic river system. Journal of Ecology **91**:447-456.

Ingvarsson,P.K. (2005) Nucleotide polymorphism and linkage disequilibrium within and among natural populations of European aspen (*Populus tremula* L., salicaceae). Genetics **169**:945-953.

Ingvarsson,P.K., Garcia,M.V., Hall,D., Luquez,V., and Jansson,S. (2005) Clinal variation in *phyB2*, a candidate gene for day-lengh induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). Genetics **DOI**:10.1534.

Jenczewski,E., Bonnin,I., Prosperi,J.M., Gherardi,M., Huguet,T., and Olivieri,I. (1997) Insight on segregation distortions in two intraspecific crosses between annual species of *Medicago* (Leguminosae). Theoretical and Applied Genetics **94**:682-691.

Jermstad,K.D., Bassoni,D.L., Wheeler,N.C., and Neale,D.B. (1998) A sex-averaged genetic linkage map in coastal Douglas-fir *Pseudotsuga menziesii* [Mirb.] Franco var *menziesii* based on RFLP and RAPD markers. Theoretical and Applied Genetics **97**:762-770.

Jones,C.J., Edwards,K.J., Castiglione,S., Winfield,M.O., Sala,F., van de Wiel,C., Bredemeijer,G., Vosman,B., Matthes,M., Daly,A., Brettschneider,R., Bettini,P., Buiatti,M., Maestri,E., Malcevschi,A., Marmiroli,N., Aert,R., Volckaert,G., Rueda,J., Linacero,R., Vazquez,A., and Karp,A. (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Molecular Breeding **3**:381-390.

Kenis,K. and Keulemans,J. (2005) Genetic linkage maps of two apple cultivars (*Malus x domestica Borkh.*) based on AFLP and microsatellite markers. Molecular Breeding **15**:205-219.

Kosambi,D.D. (1944) The estimation of map distance from recombination values. Ann. Eugen. **12**:172-175.

Krutovsky,K.V., Troggio,M., Brown,G.R., Jermstad,K.D., and Neale,D.B. (2004) Comparative mapping in the Pinaceae. Genetics **168**:447-461.

Kuang,H., Bongarten,B., Richardson,T., Carson,S., and Wilcox,P. (1999) Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. I. Genetic map with distorted markers. Theoretical and Applied Genetics **98**:697-703.

Kuang,H., Bongarten,B.C., Richardson,T.E., and Carson,S.D. (1998) An allele responsible for seedling death in *Pinus radiata* D. Don. Theoretical and Applied Genetics **96**:640-644.

Kumar,L.S. (1999) DNA markers in plant improvement: An overview. Biotechnology Advances **17**:143-182.

Kurata,N., Nagamura,Y., Yamamoto,K., Harushima,Y., Sue,N., Wu,J., Antonio,B.A., Shomura,A., Shimizu,T., Lin,S.Y., Inoue,T., Fukuda,A., Shimano,T., Kuboki,Y., Toyama,T., Miyamoto,Y., Kirihara,T., Hayasaka,K., and Miyao,A. (1994) A 300 kilobase interval genetic map of rice including 883 expressed sequences. Nature Genetics **8**:365-372.

Lagercrantz,U. and Lydiate,D.J. (1996) Comparative Genome Mapping in Brassica. Genetics **144**:1903-1910.

Lander,E.S., Green,P., Abrahamson,J., Barlow,A., Daly,M.J., Lincoln,S.E., and Newburg,L. (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1**:174-181.

Lashermes,P., Combes,M.C., Prakash,N.S., Trouslot,P., Lorieux,M., and Charrier,A. (2001) Genetic linkage map of *Coffea canephora*: Effect of segregation distortion and analysis of recombination rate in male and female meioses. Genome **44**:589-596.

Laureysens,I., Pellis,A., Willems,J., and Ceulemans,R. (2005) Growth and production of a short rotation coppice culture of poplar. III. Second rotation results. Biomass and Bioenergy **29**:10-21.

Lespinasse, D. (1999) Cartographie génétique de l'Hévéa (*Hevea* spp) et détermination de la résistance au champignon pathogène *Microcyclus ulei*. Thesis/Dissertation. Université Montpellier II, France, 5-3-1999. 110 pages.

Lespinasse,D., Rodier-Goud,M., Grivet,L., Leconte,A., Legnate,H., and Seguin,M. (2000) A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. Theoretical and Applied Genetics **100**:127-138.

Liu,Q., Thorland,E.C., Heit,J.A., and Sommer,S.S. (1997) Overlapping PCR for Bidirectional PCR Amplification of Specific Alleles: A Rapid One-Tube Method for Simultaneously Differentiating Homozygotes and Heterozygotes. Genome Research **7**:389-398.

Liu,Z. and Furnier,G.R. (1993) Comparison of allozyme, RFLP, and RAPD markers for revealing genetic variation within and between trembling aspen and bigtooth aspen. Theoretical and Applied Genetics **87**:97-105.

Marques,C.M., Brondani,R.P.V., Grattapaglia,D., and Sederoff,R. (2002) Conservation and synteny of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species. Theoretical and Applied Genetics **105**:474-478.

Marques,C.M., Whetten,R., O'Malley,D.M., Liu,B.H., Sederoff,R., Araújo,J.A., and Ferreira,J.G. (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. Theoretical and Applied Genetics **96**:727-737.

McLetchie,D.N., Tuskan,G.A., and Dietrichson,J. (1994) Gender determination in *Populus*. Journal of Agricultural Science **18**:57-66.

Menz,M.A., Klein,R.R., Mullet,J.E., Obert,J.A., Unruh,N.C., and Klein,P.E. (2002) A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP®, RFLP and SSR markers. Plant Molecular Biology **48**:483-499.

Meyers,B.C., Kozik,A., Griego,A., Kuang,H., and Michelmore,R.W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell **15**:809-834.

Michaels,S.D. and Amisino,R.M. (1998) A robust method for the detecting single-nucleotide changes as polymorphic markers by PCR. The Plant Journal **14**:381-385.

Mohan,M., Nair,S., Bhagwat,A., Krishna,T.G., Bhatia,C.R., Yano,M., and Sasaki,T. (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding **3**:87-103.

Mottl,J. and Uradnieek,L. (2003) Poplars and their leaves (X-Ray patterns of poplar leaves). Acta Pruhoniciana **74**:31-59.

Moyle,R., Schrader,J., Stenberg,A., Olsson,O., Saxena,S., Sandberg,G., and Bhalerao,R.P. (2002) Environmental and auxin regulation of wood formation involves members of the *Aux/IAA* gene family in hybrid aspen. Plant Journal **31**:675-685.

Myburg,A.A., Griffin,A.R., Sederoff,R.R., and Whetten,R.W. (2003) Comparative genetic linkage maps of Eucalyptus grandis, *Eucalyptus globulus* and their F(1) hybrid based on a double pseudo-backcross mapping approach. Theoretical and Applied Genetics **107**:1028-1042.

Nadeau,J.H. and Sankoff,D. (1998) Counting on comparative maps. Trends in Genetics **14**:495-501.

Neff,M.M., Neff,J.D., Chory,J., and Pepper,A.E. (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. The Plant Journal **14**:387-392.

Neff,M.M., Turk,E., and Kalishman,M. (2002) Web-based primer design for single nucleotide polymorphism analysis. Trends in Genetics **18**:613-615.

Nelson,C.D., Kubisiak,T.L., Stine,M., and Nance,W.L. (1994) A genetic linkage map of longleaf pine (*Pinus palustris* Mill.) based on random amplified polymorphic DNAs. Journal of Heridity **85**:433-439.

Oetting,W.S., Lee,H.K., Flanders,D.J., Wiesner,G.L., Sellers,T.A., and King,R.A. (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. Genomics **30**:450-458.

Otto, W. T. (1885) Flora von Deutschland Österreich und der Schweiz. http://nl.wikipedia.org/wiki/Afbeelding:Illustration_Populus_nigra0.jpg.

Paglia,G.P., Olivieri,A.M., and Morgante,M. (1998) Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). Molecular Genetics and Genomics **258**:466-478.

Paracchini,S., Arredi,B., Chalk,R., and Tyler-Smith,C. (2002) Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. Nucleic Acids Research **30**:e27.

Paris,M., Potter,R.H., Lance,R.C.M., Li,C.D., and Jones,M.G.K. (2003) Typing Mlo alleles for powdery mildew resistance in barley by single nucleotide polymorphism analysis using MALDI-ToF mass spectrometry. Australian Journal of Agricultural Research **54**:1343-1349.

Paterson,A.H., Bowers,J.E., Burow,M.D., Draye,X., Elsik,C.G., Jiang,C.X., Katsar,C.S., Lan,T.H., Lin,Y.R., Ming,R., and Wright,R.J. (2000) Comparative genomics of plant chromosomes. Plant Cell **12**:1523-1539.

Pelgas,B., Bousquet,J., Beauseigle,S., and Isabel,N. (2005) A composite linkage map from two crosses for the species complex *Picea mariana* x *Picea rubens* and analysis of synteny with other Pinaceae. Theoretical and Applied Genetics **111**:1466-1488.

Pflieger,S., Lefebvre,V., and Causse,M. (2001) The candidate gene approach in plant genetics: A review. Molecular Breeding **7**:275-291.

Rahman,M.H., Dayanandan,S., and Rajora,O.P. (2000) Microsatellite DNA markers in *Populus tremuloides*. Genome **43**:293-297.

Rao,G.U., Ben Chaim,A., Borovsky,Y., and Paran,I. (2003) Mapping of yield-related QTLs in pepper in an interspecific cross of *Capsicum annuum* and *C. frutescens*. Theoretical and Applied Genetics **106**:1457-1466.

Remington,D.L., Whetten,R.W., Liu,B.H., and O'Malley,D.M. (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. Theoretical and Applied Genetics **98**:1279-1292.

Ritter,E., Gebhardt,C., and Salamini,F. (1990) Estimation of recombination frequencies and construction of RFLP linkage maps in plants from crosses between heterozygous parents. Genetics **125**:645-654.

Rong,J., Pierce,G., Waghmare,V., Rogers,C., Desai,A., Chee,P., May,O., Gannaway,J., Wendel,J., Wilkins,T., and Paterson,A. (2005) Genetic mapping and comparative analysis of seven mutants related to seed fiber development in cotton. Theoretical and Applied Genetics **111**:1137-1146.

Rostoks,N., Borevitz,J.O., Hedley,P.E., Russell,J., Mudie,S., Morris,J., Cardle,L., Marshall,D.F., and Waugh,R. (2005) Single-feature polymorphism discovery in the barley transcriptome. Genome Biology **6**:R54.

Scalfi,M., Troggio,M., Piovani,P., Leonardi,S., Magnaschi,G., Vendramin,G.G., and 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.

Schiex, T., Givry, S., Chabrier, P., and Bouchez, M. Carthagene. [0.999R]. (1995). Computer Program. http://www.inra.fr/bia/T/CarthaGene/.

Schlötterer,C. (2001) Genealogical inference of closely related species based on microsatellites. Genetical Research **78**:209-212.

Schmidt,R. (2000) Synteny: Recent advances and future prospects. Current Opinion in Plant Biology **3**:97-102.

Semerikov,V., Alström-Rapaport,C., Lascoux,M., Lagercrantz,U., Tsarouhas,V., and Rönnberg,W. (2003) Genetic mapping of sex-linked markers in *Salix viminalis* L. Heredity **91**:293-299.

Sewell,M.M., Sherman,B.K., and Neale,D.B. (1999) A consensus map for loblolly pine (*Pinus taeda* L.). I. Construction and integration of individual linkage maps from two outbred three-generation pedigrees. Genetics **151**:321-330.

Sharma,H.C., Crouch,J.H., Sharma,K.K., Seetharama,N., and Hash,C.T. (2002) Applications of biotechnology for crop improvement: Prospects and constraints. Plant Science **163**:381-395.

Sibov,S.T., Lopes De Souza Jr,C., Garcia,A.A.F., Garcia,A.F., Silva,A.R., Mangolin,C.A., Benchimol,L.L., and De Souza,A.P. (2003) Molecular mapping in tropical maize (*Zea mays* L.) using microsatellite markers. 1. Map construction and localization of loci showing distorted segregation. Hereditas **139**:96-106.

Smulders,M.J.M., van der Schoot,J., Arens,P., and Vosman,B. (2001) Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). Molecular Ecology Notes **1**:188-190.

Song,X., Wang,K., Guo,W., Zhang,J., and Zhang,T. (2005) A comparison of genetic maps constructed from haploid and BC1 mapping populations from the same crossing between *Gossypium hirsutum* L. and *Gossypium barbadense* L. Genome **48**:378-390.

Stam,P. (1993) Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. The Plant Journal **3**:739-744.

Sterck,L., Rombauts,S., Jansson,S., Sterky,F., Rouze,P., and Van de Peer,Y. (2005) EST data suggest that poplar is an ancient polyploid. New Phytologist **167**:165-170.

Storme,V., Vanden Broeck,A., Ivens,B., Halfmaerten,D., Van Slycken,J., Castiglione,S., Grassi,F., Fossati,T., Cottrell,J.E., Tabbener,H.E., Lefevre,F., Saintagne,C., Fluch,S., Krystufek,V., Burg,K., Bordacs,S., Borovics,A., Gebhardt,K., Vornam,B., Pohl,A., Alba,N., Agundez,D., Maestro,C., Notivol,E., Bovenschen,J., Van Dam,B.C., Van Der,S.J., Vosman,B., Boerjan,W., and Smulders,M.J. (2003) Ex-situ conservation of Black poplar in Europe: genetic diversity in nine gene bank collections and their value for nature development. Theoretical and Applied Genetics **108**:969-981.

Su,X.H., Zhang,Q.W., Zheng,X.W., Zhang,X.H., and Harris,S. (1998) Construction of *Populus deltoides* Marsh x *P. cathayana* Rehd. molecular linkage map. Scientia Silvae Sinica **34**:29-37.

Sugita,T., Kinoshita,T., Kawano,T., Yamaguchi,K., Nagata,R., Todoroki,A., Shimizu,A., Chen,L., Kawasaki,S., and Yuji,K. (2005) Rapid construction of a linkage map using high-efficiency genome scanning/AFLP and RAPD, based on an intraspecific, doubled-haploid population of *Capsicum annuum*. Breeding Science **55**:287-295.

Sun,X., Ding,H., Hung,K., and Guo,B. (2000) A new MALDI-TOF based mini-sequencing assay for genotyping of SNPS. Nucleic Acids Research **28**:e68.

Tanksley,S.D., Ganal,M.W., Prince,J.P., De Vicente,M.C., Bonierbale,M.W., Broun,P., Fulton,T.M., Giovannoni,J.J., Grandillo,S., Martin,G.B., Messeguer,R., Miller,J.C., Miller,L., Paterson,A.H., Pineda,O., Roder,M.S., Wing,R.A., Wu,W., and Young,N.D. (1992) High density molecular linkage maps of the tomato and potato genomes. Genetics **132**:1141-1160.

Taylor,D.R. and Ingvarsson,P.K. (2003) Common features of segregation distortion in plants and animals. Genetica **117**:27-35.

Taylor,G. (2002) *Populus*: *Arabidopsis* for forestry. Do we need a model tree? Annals of Botany **90**:681-689.

Testolin,R., Huang,W.G., Lain,O., Messina,R., Vecchione,A., and Cipriani,G. (2001) A kiwifruit (*Actinidia* spp.) linkage map based on microsatellites and integrated with AFLP markers. Theoretical and Applied Genetics **103**:30-36.

Toplu,F. (2005) Breeding and conservation of black poplar (*Populus nigra*) gene resources in Turkey. Unasylva **56**:26-30.

Tschaplinski,T.J., Tuskan,G.A., and Gunderson,C.A. (1994) Water-stress tolerance of black and eastern cottonwood clones and four hybrid progeny. I. Growth, water relations, and gas exchange. Canadian Journal of Forest Research **24**:364-371.

Tsuro,M., Kubo,N., Hirai,M., Suwabe,K., and Matsumoto,S. (2005) Construction of a molecular linkage map of radish (*Raphanus sativus* L.), based on AFLP and Brassica-SSR markers. Breeding Science **55**:107-111.

Tuskan,G.A., Gunter,L.E., Yang,Z.K., Yin,T., Sewell,M.M., and DiFazio,S.P. (2004) Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. Canadian Journal of Forest Research **34**:85-93.

Van Buuren,M.L., Salvi,S., Morgante,M., Serhani,B., and Tuberosa,R. (2002) Comparative genomic mapping between a 754 kb region flanking *DREB1A* in *Arabidopsis thaliana* and maize. Plant Molecular Biology **48**:741-750.

van der Schoot,J., Pospiskova,M., Vosman,B., and Smulders,M.J.M. (2000) Development and characterization of microsatelite markers in black poplar (*Populus nigra* L.*nigra*). Theoretical and Applied Genetics **101**:317-322.

Vanden Broeck,A. (2003) EURFORGEN Technical Guidelines for genetic conservation and use for European black poplar (*Populus nigra* L.). International Plant Genetic Resources Institute:1-6.

Varshney,R.K., Graner,A., and Sorrells,M.E. (2005) Genomics-assisted breeding for crop improvement. Trends in Plant Science **10**:621-630.

Voorrips,R.E. (2002) MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. The Journal of Heredity **93**:77-78.

Vos,P., Hogers,R., Bleeker,M., Reijans,M., Van de Lee,T., Hornes,M., Frijters,A., Pot,J., Peleman,J., Kuiper,M., and Zabeau,M. (1995) AFLP: a new Technique for DNA fingerprinting. Nucleic Acids Research **23**:4407-4414.

Whitkus,R. (1998) Genetics of adaptive radiation in Hawaiian and Cook islands species of Tetramolopium (Asteraceae). II. Genetic linkage map and its implications for interspecific breeding barriers. Genetics **150**:1209-1216.

Winfield,M.O., Arnold,G.M., Cooper,F., Le Ray,M., White,J., Karp,A., and Edwards,K.J. (1998) A study of genetic diversity in *Populus nigra* subsp. *Populus nigra betulifolia* in the Upper Severn area of the UK using AFLP markers. Molecular Ecology **7**:3-10.

Wu,R.L., Han,Y.F., Hu,J.J., Fang,J.J., Li,L., Li,M.L., and Zeng,Z.B. (2000) An integrated genetic map of *Populus deltoides* based on amplified fragment length polymorphisms. Theoretical and Applied Genetics **100**:1249-1256.

Wullschleger,S.D., Tuskan,G.A., and DiFazio,S.P. (2002) Genomics and the tree physiologist. Tree Physiology **22**:1273-1276.

Yin,T., Huang,M., Wang,M., Zhu,L.H., Zeng,Z.B., and Wu,R. (2001) Preliminary interspecific genetic maps of the *populus* genome constructed from RAPD markers. Genome **44**:602-609.

Yin,T., Zhang,X., Huang,M., Wang,M., Zhuge,Q., Tu,S., Zhu,L.H., and Wu,R. (2002) Molecular linkage maps of the *Populus* genome. Genome **45**:541-555.

Yin,T.M., DiFazio,S.P., Gunter,L.E., Jawdy,S.S., Boerjan,W., and Tuskan,G.A. (2004a) Genetic and physical mapping of *Melampsora* rust resistance genes in *Populus* and characterization of linkage desequilibrium and flanking genomic sequence. New Phytologist **164**:95-105.

Yin,T.M., DiFazio,S.P., Gunter,L.E., Riemenschneider,D., and Tuskan,G.A. (2004b) Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. Theoretical and Applied Genetics **109**:451-463.

Young,N.D. (2000) The genetic architecture of resistance. Current Opinion in Plant Biology **3**:285-290.

Zhang,B., Zhou,Y., Zhang,L., Zhuge,Q., Wang,M.-X., and Huang,M.-R. (2005) Identification and validation of single nucleotide polymorphisms in poplar using publicly expressed sequence tags. Journal of Integrative Plant Biology **47**:1493-1499.

Zhang,D., Zhang,Z., Yang,K., and Li,B. (2004) Genetic mapping in (*Populus tomentosa* x *Populus bolleana*) and *P. tomentosa* Carr. using AFLP markers. Theoretical and Applied Genetics **108**:657-662.

Appendix 1

A. Description of the available SSR

For each SSR are presented the primer sequences, the annealing temperature* (Tm), the length expected, the motif of the SSR, the linkage group where the SSR is expected and if it was tested on the *P. nigra* pedigree.

* Tm = 2(A+T)+4(C+G)-5, in the table is indicated the lowest Tm of the primer couple. It is the Tm used for the PCR reaction

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP257	ai164591	CCACCCAAACCATGCCCTTTATC	GAATCCCCTAAATCGCGCTCAG	63	162	AGG	XIV	yes
POP258	bi135774	CCTTCTTTGGAACCTCACAAAC	TTGCAGACCATGCAGATAAG	53	169	GA	VI	yes
POP261	bi138728	TCGCCTCTTATTTGATCGCC	TGCACACATTCCTCTGCCTC	55	188	CT	XIX	yes
POP262	bi139308	ACAATAGACAGCAGGCATGG	GAACGAGAATGTTGGAGGGG	55	173	CAG	VIII?	yes
POP263	bi139327	GCAGAAACCAGCTTCTTGGAC	AATTCCCGACAAGGCTCGAC	57	140	GA		yes
POP259	bu810400	CAAAAGCTGAAGCTGTGGTAAC	GGCAATGAACATGTCGTCAAAG	59	116	GA	XIII	yes
POP260	bu810907	TTCTTGTTCTCTTCGCAGCGCC	CCAACCCAATGCTTTTCCAGAACC	63	197	CAG		yes
POP264	bu813610	GTATGTCTCTCACTTCACACAC	CTGGTGAGCATAGAAGCAGG	57	108	TCC	VIII	yes
POP265	bu813833	GAGGAGCCCTTCTTGTTTAC	TGCATAAATTGGAGCAGCAC	53	176	CAG	Scaf	yes
POP266	bu814260	CAGCCAAATATACACAGCCC	ACACACCCCACTTCAACTAC	55	138	CA	XIII	yes
POP267	bu814989	GGCAGCAACGAGAGAGAGAAAAG	CACCTTTGAACCAGGTAATACG	59	134	AG	X	yes
POP268	bu818855	AAAGGCAAACCCTCCTCCTC	CCTCATCAACCTCTTCCTTGCC	57	197	TC	XIII	yes
POP272	bu831219	TTCGGCAGCTCCCATCCAAAAC	GTTCCTCTTCCAAACTGCTTCACC	63	161	AAG		yes
POP269	bu867968	AGGGTGCAATGGACCATGTC	GGCTCTCTTATTCCACACCG	57	300	TA		yes
POP270	bu885452	TCCTCTGAAAGACTGCGAAG	CTGGTAAATACATCACAGCCTC	55	175	CTT	XVIII	yes
POP271	bu890808	CCTCCTCAATAATTCAATGGCTGC	GGTAGTAAGAAGTCGAGGTAGG	61	172	CTG		yes
POP331	GCPM1027-1	ATATCCTGCCACAAAGATCA	TATCTTCACTTGAGGGGATG	51	211	AT	VII	yes
POP332	GCPM1240-1	TAATTATCAAACTGCGATGC	CACCCTCTCCAATTATACCA	49	216	AT		no
POP333	GCPM1374-1	ACCTTACCGGTTTTGGTG	CCATTGTCTCTGCCACTTAT	49	223	ACAT		no
POP320	GCPM1381	CAATGTCAAGTGCTCAGAAA	GTATTGGGTGAAGGTTGAGA	51	224	AAT	XVII	no
POP334	GCPM1418-1	GTACACTCAGCTGTGGTCCT	TGGGAGTCATGAAATCTACC	53	193	CT		no
POP335	GCPM1438-1	GACTCATGCCTTTATTTTCG	CAACTTCTACCTCCGATTTTT	51	215	GAAA		no

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP336	GCPM1547-1	TCTTCTGCCTTTGCTCTTAC	CAACATGCGGATTAGTAGGT	53	151	AT		no
POP337	GCPM1717-1	TGGCATCTCACAGACAAGTA	TATCTCGTTTCCACCATTTC	51	151	AT		no
POP338	GCPM1896-1	TTGTCAATTGGGTTCCTAAG	AGAGTTTTGGGTGTTTTAGC	51	106	TA	VII	yes
POP339	GCPM1899-1	TTTCATGAACATTTTATGGTT	TAAATTGTCCCCAAGTCAAC	47	165	AT		no
POP340	GCPM1929-1	ACCTTTGTTCGAGAATCCTT	TCCAAGATGATGGGAAATAG	51	146	TG	XIV	yes
POP321	GCPM2024-1	GTTGCCTACTCGTTTGTCTC	GGAAAATATGGTGGTGCTAA	51	229	CT	XVI	yes
POP341	GCPM2050-1	ACATGAAACCTGGATGAGAG	TTGTGATTTGTGGGAGTACA	51	190	AT		no
POP322	GCPM2180-1	GATAAAGAAGAGTGAGGCCA	TTTTTAGCATGACCCAACTA	49	123	AT	XVI	yes
POP342	GCPM2267-2	TGACTTGGCTAATCTTTTCTTT	CAGGTTTAAAGACAAATAAGGTTT	53	216	TTTA	VII	yes
POP343	GCPM2288-1	TGGAGGAAGATGAAAAGAGA	TACCCCTAACCACAATTCAG	51	163	TA	VII	yes
POP344	GCPM2582-1	TCTTTGTTGCTTTGGTTTCT	TGAACAAGCTCAACATGCTA	49	230	TA	XIV	yes
POP323	GCPM2662	CAAATTTTTGCCTCCTTCTA	ATCAGTTGGCATCATTCTTC	49	119	CA	XVII	no
POP324	GCPM2834	ACTTTTCTTGTTGTCGTTGG	AAAGTGTTGAAATGAGTTGACA	51	156	CT	XVII	no
POP345	GCPM2860-1	CAATAAAAATACGGACGGAA	TGTGAAAATAAAGGCCAAAT	47	128	TCT	VII	yes
POP346	GCPM2995-1	CATTCACTTTCACTTCCCAT	TCTACGTGGTAATCCCTGTC	51	166	AG	VII	yes
POP347	GCPM3096-1	GGCCCTTTTAATTATACGAAG	CGTAAACGGTAGCAAAAGAT	51	217	TA	XIV	yes
POP325	GCPM3178-1	AAAAAGGTTTGAAATGCTCA	CTAGCACAACCAGGAAGAAC	47	220	AT	XIX	yes
POP348	GCPM3208-1	GATTGAGCTGAAGAGTGGAG	CTATCCACACACAAAACACG	53	154	TC	VII	yes
POP349	GCPM336-1	AACCTTATGGAACCTAAGCC	AGAGAAATTAAAGGGGAAAAA	49	191	CTTT	XIV	yes
POP350	GCPM350-1	CCTCAAACAAATCAAACTCC	TTGATGTGTAACTAACCCGA	51	205	CTC	VII	yes
POP351	GCPM3503-1	TTCGTCAACGTAAAGAAAGAC	AAAATACTCTATTTCACCCTTAAAA	53	123	AT		no
POP352	GCPM3646-1	TGTTTCATGCTTGCAATTTA	TCTTCTTCCTTTCCATTTCA	47	220	AT	VII	yes
POP353	GCPM3893-1	TTCCACAAAACTTATACCCG	GAAAGAGCCCTTGATAGGTT	51	184	AT	VII	yes
POP354	GCPM4002-1	GAGAAAAATCTCAGTGAGCG	AAACACACCAGGGAATTAGA	51	164	TAA	XIV	yes
POP355	GCPM416-1	GATTGTAGCATTTTGTGGTG	TTTATGAAAAAGAAAAGGAATGA	51	213	AT	XIV	yes
POP356	GCPM505-1	GTCGAAAAGATCAAAAATGC	CCCTTCAACAACAAATCAGT	49	227	GCT		no
POP326	GCPM588	GTATACGTCACTGTCAAAGCC	ACGCTGTGATATATTGTCCC	53	220	CT	IX	yes
POP357	GCPM598-1	TCACAGAACACATCTTTCCA	AAATCTCAAACCTCAAGCAA	49	214	AT		no
POP327	GCPM641	CTAGGTTCCGGAAATAACATT	TCAAGCTTTTGGGTTATCAT	49	152	AGA	XVII	no
POP358	GCPM748-1	AACAAGGCTCATCTCAACAT	AACCAAAGGGGCTAAAATAG	51	230	TA	XIV	yes
POP328	GCPM943	CAGTACTCTCTACCATGCCC	CCATGCTACATTGTATTGGT	51	205	TC	XI	yes
POP177	ORPM104	GCATGTGTGGGGATCAGAAT	CGGTCCTAGCTAGCTCCTCTT	55	178	[AAT]4		yes
POP275	ORPM105	AGCCTGTGATACCAGAACACC	TTTTCTTAGCCTCTGCTCTGC	57	190	[AAT]6	V	yes
POP276	ORPM106	TCCAAAGTCCTGCCTACACTG	GCAGCCTTATCCGAGTTTCA	55	243	[AT]4*	IV	yes
POP277	ORPM107	AATCTGGTGGCTTGCCTCT	TTGAGGAACACGTGCAGACT	53	190	[TAAA]4	IX	yes
POP153	ORPM11	CGATTAATCTCTTCTACTAGGCCATT	GCTGCTGCTAATGGATGTGA	55	185	C[AT]4AC[AT]4		yes
POP278	ORPM112	CATGAATTGGGGGTTTCAAG	TCAGCTCAGAACTAGACCAACC	53	168	[TCT]CT[TCT]3	VIII	no

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP178	ORPM114	ATGAGCCACTGGTAGGGATG	CAAGTCAGACTGGGGAAGTGA	57	200	[CT]4*		yes
POP279	ORPM118	GCAAGAATGACCAGCACAGA	ATTGGAACGCGATGAAGTTT	51	257	[AT]4TATAA	XVI	yes
POP179	ORPM121	GAATTCCTGCAGGTTTTTCG	ATAGCATCCTGCTGGCAATG	53	197	[TG]6		yes
POP180	ORPM127	TCAATGAGGGGTGCCATAAT	CTTTCCACTTTTGGCCCTTT	53	200	[TG]8	IV	yes
POP181	ORPM128	TCCATTCTCGAGATTTTGTGC	CGTTCCCTCCATTGGTTATC	55	195	[AG]4*		yes
POP359	ORPM133	TGGGACATGCTCCATGGTAT	CTTTCCCTTTCCCTCTCTCC	55	209	[AG]4*	XIV	yes
POP182	ORPM137	CCGTGCATCTGCTCACTTTA	GCATTTGCAGATGAAATTGGT	53	185	[AT]7		yes
POP154	ORPM14	GGGCTGCAGCAGATATTGA	CCAAAGGAACCCAAAGAAGA	53	159	[GCTC]4*	XVI	yes
POP280	ORPM144	TCAACCATGCACAAGTTTCC	CTACTGGTTCCAGCCAAACC	53	187	[AT]4	VII	yes
POP183	ORPM147	GTCCAAGACACCTGCGTGTA	TTGAACCCATCACCAATTCA	51	243	[CGTT]4		yes
POP184	ORPM149	GTCTCTGCCACATGATCCAA	CCCGAAATGGATCAAACAAG	53	216	[AT]4?CT]4		yes
POP155	ORPM15	CGTGAGTTTTGAGGCCATTT	CATGGAAAGGATCACCCACT	53	257	[AT]14		yes
POP281	ORPM154	TCAAGGTAAGCCAACACAAGC	TCCTTCAAAGGGTCAAGCAT	53	211	[AT]4	XV	yes
POP156	ORPM16	GCAGAAACCACTGCTAGATGC	GCTTTGAGGAGGTGTGAGGA	57	238	[CTT]15	XIII	yes
POP185	ORPM166	TCATTGGAGCACAAGACACC	GGAGAAGCCTGTTTCCTCAA	55	200	[CT]5		yes
POP186	ORPM167	TGCACTATTTACTCGCAGTCTCTC	AAGCTTTTCCGAAACCGAAG	53	178	[CT]4		yes
POP187	ORPM173	TCTAAACATCCGCCACGTAA	ACGACGACGTGAGTGAGTTG	53	195	[CT]4*		yes
POP282	ORPM176	TGCAAGGTGACACGAACACT	TGAAGTTGAGAATGTAACGGAGA	55	202	[AT]4*	XIV	yes
POP188	ORPM177	TGCAGTAAACACAAAGTGTCGTC	CCATGGACTGCAAATGGTTA	53	226	[TA]7TTA		yes
POP283	ORPM184	CGAGAACAGCCAAGCACTCT	TGCATTTACGCTATCGATTCA	53	186	[TTG]4	V	yes
POP189	ORPM186	GGCTAGGAATACCCTGGAGAA	AAGCCATCTCGACTATACACCA	59	234	[TTTA]5		yes
POP190	ORPM188	TCGCATCTCACTCCTTGGAT	TGAGCTAAACCACCTCTGTTGA	55	156	[GA]4*		yes
POP191	ORPM190	CCCTGGTTTTCTCTTCTTGG	CCAGATTGGACTTGGGATTC	55	209	[TG]7*	VI	yes
POP360	ORPM193	CCGCTGGATTTGTTTGTTTT	TGAGCAGAAAGATGCGAAGA	51	181	[ATTTT]4	XIV	yes
POP284	ORPM194	AAAGCCGAGTTCAAACATGA	GGTGATGCACTCTTTTGTCG	51	222	[TA]6	XI	yes
POP285	ORPM197	GTCAGTTTGCCCTCTTCGTC	TGAGGGCGTCTCCTCTTTTA	55	191	[GA]4	VI	no
POP157	ORPM20	ATGGGTGGTAATTGCAGCAT	GCCAGGAGATGAAGAGTAGCA	53	202	[CTTT]4[CTT]2TT		yes
POP192	ORPM202	TCGCAAAAGATTCTCCCAGT	TTCAAATCCCGGTAATGCTC	53	190	[TAA]5	VIII	yes
POP193	ORPM203	CCACCAGGCATGAGATATGA	TCAAACCGAAAGGTCAACAA	51	209	[TA]4(A/T-rich region)	III	yes
POP194	ORPM206	CCGTGGCCATTGACTCTTTA	GAACCCATTTGGTGCAAGAT	53	196	[GCT]7	XIX	yes
POP195	ORPM207	TGCATATTTCACGTGCCTTT	CAAAGTGAGGAAGCGTCAGA	51	199	[TC]8	II	yes
POP196	ORPM209	TAGCAGATTTGGTGGCACTC	AAAATCGACCACTGCCAAAG	53	198	[TC]4(T-rich region)		yes
POP158	ORPM21	GGCTGCAGCACCAGAATAAT	TGCATCCAAAATTTTCCTCTTT	53	206	[AG]4*	IX	yes
POP286	ORPM210	TGACCATTTTGTTGGGACAG	TAAGGGGCTCAGTTATGCAC	53	195	[CT]4(T-rich region)	XII	no
POP287	ORPM213	TGCAGGATTCAAGAAAAATTTAGA	AGGAGTCATGGGGCTCTTCT	57	211	[TC]4	XV	yes
POP197	ORPM214	TTTTCACAAGCCTCGAAGGA	TGGAAGACCCGAACTTTTTC	53	166	[TC]11		yes
POP288	ORPM217	ATTGGCCACATACGGCTTAC	AGGCAAAATCAGGATCTCCA	53	201	[TC]6	XI	yes

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP289	ORPM219	AAACCATAGGCATCAAAGCA	TTCATGCAAGGCATCAATTC	51	196	[GT]4	VIII	no
POP198	ORPM220	AGCTAGCCTGTCGTCAAGGA	CAAGGAAGCATTCTCGCAAT	53	190	[TTTA]6		yes
POP290	ORPM221	TGGAGGCTGTCTGTTTTGTG	AGATTTGAGCGACTCCGAAA	53	211	[AG]17	IV	yes
POP291	ORPM229	TGGCAACAATGTTTCTGCAT	GTTGCAGCCTTGCAGAGTTA	51	223	[AG]4	XIV	yes
POP159	ORPM23	ATTCCATTTGGCAATCAAGG	CCCTGAAAGTCACGTCTTCG	51	197	[AT]6?AG]6*	IX	yes
POP199	ORPM232	TCTTCCAGCTTCACTCATGC	AATGCTGGTGGCTGAGAGTT	55	259	[AT]9		yes
POP200	ORPM244	GATTAGGTCCATTCGCTCCA	GCAAAAGTATCCAGGCTTGC	55	186	[GT]4		yes
POP201	ORPM257	GGCAGAAGTTGGAAATTCTATGA	TGGTTTAAATATCGGGATCGAG	57	201	[TA]4		yes
POP160	ORPM26	GCTGCAGTCAAATTCCAAAA	CGAGCGTCTTCTTCATGGAT	51	213	[CA]8	VI	yes
POP202	ORPM260	TTCTAGTCCTGGCATAGCTTCA	CAGAGATTTGAATCGCAGCA	53	220	[AAT]10	II	yes
POP292	ORPM263	AGCACATCTTTCGAGCATGA	TGTAGCAATTTGCCAAAATCA	51	243	[AT]8	XIX	yes
POP203	ORPM264	AAACCATAGGCATCAAAGCA	TTCATGCAAGGCATCAATTC	51	197	[GT]4TpolyA		yes
POP204	ORPM268	TTGCTGGGTACCCTATCTCA	AGCGTATTTGAAGCGATTTGA	53	200	[AG]4	VIII VI	yes
POP293	ORPM269	CGCTTCAAATACGCTTTATGC	ACGTGGGTTCATTTTGACCA	53	264	[AT]4?TA]15	XI	yes
POP205	ORPM276	GCAGGAGAAAACACCAGGAA	TCGCGAAAGAGAAGAAAAGC	53	205	[TA]6	XIX	yes
POP206	ORPM277	CTTTGGATTGCTTGCGTTTT	TTACCATTGCTGCCATTTCA	51	201	[GA]4		yes
POP207	ORPM278	CAATATATTATTTTTATCCCTCACTTT	GAAAATGGCGAGACTCAACC	55	194	[TA]5		yes
POP208	ORPM279	TCAAATCAAACCACAAAAACACA	TGAGACGAACATATCCTTCACC	55	197	[AT]18	VI	yes
POP161	ORPM28	GGATCGACTTCCAACCCATA	AATTCCCAGATGAAGGCTCA	53	204	[AT]7*	XVIII	yes
POP273	ORPM28	GGATCGACTTCCAACCCATA	AATTCCCAGATGAAGGCTCA	53	204	[AT]7*	XVIII	yes
POP209	ORPM285	GACAGGCGAGCATGTACAAA	GATCAACCCTGTGTTCAGCA	55	198	[AG]4		yes
POP210	ORPM286	TCAGGCAGAAGGGTAGAGGA	CCTGACCCTGCTTGCTTATC	57	164	[GT]4[GA]8	II	yes
POP211	ORPM287	GATAAGCAAGCAGGGTCAGG	GGGAAGGGAGAAAAGGATTG	55	264	[AT]18		yes
POP294	ORPM295	ACAGCGAGCAAAAGCAAACT	GGGAGGCACGTGATATCCT	53	203	[CA]4	XIX	yes
POP212	ORPM297	CCCAGTATGTCGTGCTTCAA	GGATCCCTTTGAGTTCACCA	55	222	[GTTCTG]4+[GA]4		yes
POP162	ORPM30	ATGTCCACACCCAGATGACA	CCGGCTTCATTAAGAGTTGG	55	224	[TC]9	I III	yes
POP213	ORPM301	CAAAGATGGTGACTGGATGC	AGCCTATTGCTTCCGATCCT	55	201	[CT]5		yes
POP214	ORPM303	CCTCGAAACAGAGTCCCAAA	TGTAGATGAGGCTGCTGCTG	55	205	[AC]4		yes
POP295	ORPM304	AGGCTTTCAGGCTTGGTTTT	GTGGGGCTTGCCTCTTTACT	53	196	[AG]4	XI	yes
POP296	ORPM310	TTGAACAACATTGCTGGCTA	CTAGGAGACGGAGGGGACTC	51	238	[CA]4	XV	no
POP215	ORPM312	GTGGGGATCAATCCAAAAGA	CCCATATCAAACCATTTGAAAAA	53	194	[CCT]6	VII V	yes
POP297	ORPM313	GTGTGCAGAGTTGACGTGGT	GGAGGCGAAAATGAAGATGA	53	212	[TG]4	V	yes
POP298	ORPM32	CCAGCATTAGAGGACTCCTGA	TCTGGACACCCTTTGACTCC	57	198	[AT]9	XII	no
POP216	ORPM327	ATAGACGACCGCGTTTTCAC	GGAATCGAAACCCTAACATGG	55	200	[TC]6		yes
POP299	ORPM33	CGGCTCCTTATTCCTTTTGA	TGCACTTTTTCTGGAACATCC	53	198	[AG]7	XI	no
POP217	ORPM330	AATACCCTCTTCAATTCTAAATTTCTT	ATTACCATCGAGGGCTGAAA	53	168	[AT]6		yes
POP218	ORPM339	ACGGTGACCTGGGAAAAATA	CAGGTCACGGGTTTGAAAGT	53	247	[AT]8		yes

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP219	ORPM340	CACGTGGATTGATGTTTCTGA	ATCGCTTTCAGCTGGATGAT	53	195	[CTT]4		yes
POP220	ORPM344	GGAGATTGTCGGAGAATGGA	TGGACGTTACGATAGGAGTGG	55	229	[TC]8	X	yes
POP221	ORPM345	CCAATCCCAGATTCGTAACAA	CCTCCCGTTAATCAAACCAA	53	345	[AT]8	X	yes
POP222	ORPM346	AAAGAATATCAATCTCACGCTAAGTT	TCGATTCCTCGTGCATAAAA	51	251	[TAT]4?TA]4		yes
POP223	ORPM347	CGTCGGTAAAATCGTCGGTA	CCTGAGCATGGAGGAGAAGA	55	188	[TTC]6		yes
POP224	ORPM354	TCACAGCCTCCTATGTGCTC	TGCTGCCTCCTCTTCTTTCT	55	205	[GA]6		yes
POP225	ORPM355	ACCGACAGATTGAACCCATT	GAATGGTGTTGAAGTGACCAAA	53	204	[TTC]5		yes
POP226	ORPM361	TGGATATGGATCCGTTGGTT	TTTTACAGTTCAAGTTTCCATTCC	53	275	[AT]7		yes
POP227	ORPM363	CAAAGCCAACTCAATCAATCA	GCACAACAAACACATTCTTAAAGG	53	265	[ATA]4	VI	yes
POP228	ORPM367	AAACATGTATCAAAATAATGTGTTCG	TTCTTTGAAATGTCCAATTTTCTTC	59	173	[AAT]5		yes
POP300	ORPM370	TTCCCCTTGTTTTGATCCAC	CCCTGGCTAGATTTCGTTTG	53	187	[ATA]4	VIII	no
POP229	ORPM372	AGCTCTTCTGCTGGTGCTGT	GAGGGAGGGAGGGTAAAAGA	57	190	[TCTT]5		yes
POP230	ORPM379	TTGGTCCTCATTCTTTTGATTG	TCATGGAAAATGAGCTGGAA	51	284	[ATT]4		yes
POP163	ORPM38	CCAATATGTGTTGTCGAACATTG	TCTCCCCCTCTCCCTCTCTA	59	257	[ATA]9*		yes
POP231	ORPM389	GTTTCCCGGTCTTGTTCTTG	CTGCTCAATCCCATCAATCA	53	187	[TAA]4		yes
POP232	ORPM390	TGTAAATGCCATTCGGATCA	GAGGAAGGGGAAAAGAGGAG	51	215	[CTT]4		yes
POP233	ORPM393	TGGATTTCGACTCGTTTTCA	TGCCTTTGATTTCTTTTTCTCA	51	227	[GAT]4		yes
POP234	ORPM399	TTAACCCGTTAAACCCGAGA	GGGTACTTTCCCATGCCTTT	53	276	[ATA]4		yes
POP361	ORPM40	CAAACATTTTCCTTGGCAGAA	TTATTGGGTTTGGGTTTTCG	51	194	[AG]16	XIV	yes
POP235	ORPM415	CTCGGTGCAAATATCGGTTC	AGATCGATGGTCCTTTCCTG	55	225	[GGCG]4		yes
POP301	ORPM421	AAATGATGTTGCGATTTCCA	TCCCATCTCAACTACTCCAACA	49	203	[TA]7	IV	yes
POP302	ORPM427	GTAGGAGTGGATTGGCTTCC	GGTTTTTGCATTGGCAATTT	49	175	[TA]4	XVIII	no
POP303	ORPM43	TCTCGGGGAGGATGAGTATG	TTTAAACGTTGGTGGGGTTG	53	176	[ATAA]4	VI	no
POP274	ORPM430	CCTTGGAAAAACCCCAAAAT	CAGCTCGACTCATTGCAAAA	51	202	[AT]9	XV	yes
POP304	ORPM432	TTTGCAGGTTCATCAAGTGAG	ACCCCAGTAGCTTTGGCATT	55	196	[ATA]4	XIX	yes
POP236	ORPM433	CCATCAGTTTCGAGGAGATTC	GCGTGATGTCAAGCAAGGTA	55	203	[TA]5		yes
POP305	ORPM435	TAACCCACAACACGCCTACA	TGTCTTGCACAAGCAGGTAAA	55	215	[AGC]4	VII	yes
POP306	ORPM438	GTCTTCTGGCTTGCCACTTT	AACAAATGCTGCACCTAGCA	53	206	[CT]3...[TC]5	XV	no
POP237	ORPM441	GGGCTGCAGTCGTTCTTTG	TGAAATCAAACCAGCAGATCA	53	229	[AC]11		yes
POP307	ORPM444	CCCAAACAAGGGCAAAAATA	CACGGCAATCACATTACCAA	51	205	[TA]6[AT]4[AG]13[GA]4	IX	yes
POP238	ORPM446	GGGCTGCAGACAAATTAAGG	TGGGACATGCTCCATGGTAT	55	249	[CT]3...[CT]4		yes
POP308	ORPM446	GGGCTGCAGACAAATTAAGG	TGGGACATGCTCCATGGTAT	55	249	[CT]3...[CT]4	XIV	yes
POP309	ORPM448	CGGGCTGCAGATTTTGTTTA	TCTGCAACTCCAACAAATGG	53	258	[GAA]4	VII	yes
POP164	ORPM45	CATTCAAGGCAGTCGTACCC	TTGAATCCTCCGTGTGAATG	53	215	[TA]4...[AT]5	V	yes
POP310	ORPM451	ATGGACGTTCTTGGCATCTC	TTGCCTCGCACACTACTGAC	55	198	[TA]5	IX	yes
POP239	ORPM455	GAGTTAAACCCACCCTGCAA	GCCGAAGTTGACGATAGCTC	55	195	[GA]5		yes
POP240	ORPM477	ATTCCAGAAACCCTTGGAAA	TGGCTTAGCAAAACCCAAAA	51	190	[TC]7		yes

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP311	ORPM479	GAGCTGCAAACATAGGCACA	GCCCAGTTTGGAATTAGAGG	55	206	[AAT]6	XVIII	no
POP241	ORPM482	TTTCCCAACTGTGGATTTGC	GGAATCTCCAAGGAGACTGGA	53	223	[AC]6		yes
POP242	ORPM488	CTCCAGCCGCTTCTATCCTT	TGTCGTGGGAAAGAACCAGT	55	200	[TTA]6		yes
POP165	ORPM49	AAAGGGCTTTGGACGATTTT	GATTTATGAGCCTGCCCAAC	51	195	[GA]6		yes
POP243	ORPM496	CAGCAGTGCAAGCTCCTAAA	GGCCACTGACAGAGACCAAG	55	185	[GGA]4		yes
POP312	ORPM50	AAGAATTTGGGGCGGTTTAC	GCCTCAAAGGGAATTCTCAA	53	198	[A]7[TA]4[A]6	VI	no
POP166	ORPM55	AGGTTTGTGCGTAGCTTGGT	ATTTTCGCGAGACAAACTGC	53	205	[CT]4*		yes
POP167	ORPM56	CCATGCATACATTTTTAGCTCCT	AGAGGGCACTGTCCATTCAT	55	176	[AT]9...[CT]4	VIII	yes
POP168	ORPM59	TGCTAGTAACTGCGCATTGG	GATGTTTTTCGCACGCATTA	51	213	[AT]6	XIV	yes
POP313	ORPM6	CCAAATTGCTCTCGTGTTCC	ACTCAGCAGTGCTCCATGTG	55	241	[AT]38	XVIII	no
POP169	ORPM60	ATAGCGCCAGAAGCAAAAAC	AAGCAGAAAGTCGTAGGTTCG	53	212	[AAT]5		yes
POP170	ORPM62	CGGAGTCAGCTTGAGGTAGC	CGGCAATATTGAGGAGAATGA	55	203	[AT]4?ATTTT]3		yes
POP314	ORPM64	AAAGGCCTCTGCTTCGCTAT	TTGCAGACATGATCCCAATG	53	222	[CA]4	XVI	yes
POP315	ORPM65	CCAACATTCCTTCGATCTTGA	CAAAATACTGGGCACCCTTG	55	215	[CT]4	V	yes
POP171	ORPM66	AGCCTCCAAACACCATGAAC	ACAGTGGTGTGGATCCTGCT	55	213	[GAAA]4*		yes
POP172	ORPM76	CGTTTTTCCTAAAGCAAACAGA	TCCATCTGCTGCACATTGTT	53	196	[TG]4		yes
POP316	ORPM79	GAAGCTGAAAACAACAACAAACA	GGGTTTTTAACATAATAAAAGCTTGG	57	160	[AAT]4	XIV	yes
POP152	ORPM8	CGATAACGTTGATATCGAATTCCT	CCTCATGGAGTGGAAGTGCT	57	266	[CAT]6		yes
POP317	ORPM80	GCTGCAGCCTCATTTACACC	TGACTGCTTCACTCCTTTGG	55	180	[CA]4	XII	no
POP173	ORPM86	CCACATCCATAGCTCTGCAAC	GTACTACCTCGCCTGCCAAC	59	204	[CTT]5		yes
POP174	ORPM87	GGGTCCCTATTTTTGGCTTG	CTTGGGTGCTCTCTGTCCAT	55	195	[CT]4		yes
POP318	ORPM88	GCCACCCCAGAGTCTCTTCT	TTCTCACCCTCCTTCACTCC	57	250	[AG]4	XVI	yes
POP175	ORPM91	CTCAAGCTCTCTGCCGTTG	GTGGAATGCCCGTAAAAGC	53	197	[CTT]4		yes
POP176	ORPM92	TGACTCTGAGTGGTCGAGGA	GAGAATCCAACCCACTTCCA	55	207	[TGC]5		yes
POP319	ORPM95	GATGGTTTGGTGGTGGCTAA	TTCCGAATAACCAGGACACA	53	187	[TG]4	XVIII	no
POP051	PMGC108	TGCAGGTGATGTCATCACCG	AACCGAATCCATGCGTCACC	57	330	CTT		yes
POP021	PMGC14	TTCAGAATGTGCATGATGG	GTGATGATCTCACCGTTTG	49	210	CTT	XIII	yes
POP030	PMGC2011	TCTACGAGGAAAGGGAAGGG	CTTTATAATGCATCATAAAGTTCC	57	105	GA		yes
POP100	PMGC2015	TTTTGGCATTCAAAGACTTGGC	AGTTGATTCCATGTCGTGTCC	57	160	GA		yes
POP010	PMGC2020	TAAGGCTCTGTTTGTTAGTCAG	GAGATCTAATAAAGAAGGTCTTC	57	150	GA	IV IX	yes
POP060	PMGC2030	TCCACAACTCTTGGCTAACC	GGACTACAATGTGCGTGACC	55	85	GA	XVII	yes
POP096	PMGC204	GAAGATAAATTCTCCAGCTC	TAACTTTCCCCGCATGT	45	225	CTT		yes
POP061	PMGC2055	TCAATTATTTAAGCTACTCGCTC	GCAATGTGCCATAAAATGCGTC	57	75	GA		yes
POP031	PMGC2060	CTCTCAAATGCTGATTTACCG	TCTTCAGTTGCAGTATTCAAAG	55	185	GA		yes
POP101	PMGC2084	CCCCCACCACTAGATTCAGC	GAGTGGTGATGATGGTTGCC	57	190	GA		yes
POP062	PMGC2088	TCACAAAAGGTTAACGACTTCG	CAGTACTCAGCTGCAGGTCC	57	180	GA		yes
POP012	PMGC2098	CACAGTGCCAAAAACAGAGTGG	TCTACTTCATTGTTATTCATGTTAC	59	265	GA	I	yes

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP063	PMGC2105	ATTTCTCTAGGAAACAACAACC	CCTTAAGATGTTGCTGAACTC	55	135	GA		yes
POP032	PMGC2140	GCTGTCAGAATCAAACACTTC	AAGCAGATAACTAAGACATGCC	55	180	GA	VII	yes
POP064	PMGC2143	TCATCATCCATTACTCAACTTG	GCGTAAGAAGCTATTATCGTC	55	160	GA		yes
POP033	PMGC2156	GATCTCTCTTACATCACTCATC	GAATGTCTTTACTCCATTGTTGG	57	135	GA	V	yes
POP102	PMGC2163	CAATCGAAGGTAAGGTTAGTG	CGTTGGACATAGATCACACG	55	220	GA		yes
POP034	PMGC2217	ATTAGCTTCTTCTAAAGCAGC	TGACTGACTGTCTGTCTTCG	53	160	GA	VI	yes
POP002	PMGC223	CGATGAGGTTGAAGAAGTCG	ATATATGTACCGGCACGCCAC	55	170	CTT	II	yes
POP035	PMGC2235	GCCAAAATAGTAAGTGTGATGG	CACACATTCTCTCATTCAAAGC	57	145	GA		yes
POP013	PMGC2270	CAAAAAACATGCAGAAATCTTCAG	TTCACAGCTTATATAGCACTACG	59	105	GA	?	yes
POP103	PMGC2274	GGGGCTAAAATACTTGATGGG	ATCTTCTTCATCATATTTATGTTC	55	135	GA		yes
POP014	PMGC2289	GTCTATCTGTCTGATGTCACC	AAATCTCACATTATAAAAGATTTAG	55	265	GA	XIX (XII)	yes
POP104	PMGC2315	CTGTGGTATTTGTGCAATGTG	CAACAGAGCAAACTTGAGTCG	55	143	GA		yes
POP105	PMGC2316	TACAGGTCAACGCAGTTGAC	TTAGAAGTCAGTGGGATTAGG	55	132	GA		yes
POP106	PMGC2321	ATTCACATGGCTACTACCATG	AAAACTTGGACATCTTGCGTG	55	109	GA		yes
POP036	PMGC2328	CAAAGGTGAAGTTACAGTCAC	CCATTAGGCCATTATAGACAC	55	105	GA		yes
POP107	PMGC2385	ATTCTTCACCTGGGCAATATG	CTTGGCTGTAAATGACGAGTC	55	140	GA	I	yes
POP065	PMGC2392	AAGAGAGATAGCATCACCAAG	TATGTCGAGGAAATCCTTAGC	55	192	GA	XI	yes
POP066	PMGC2408	TAGGTCACTAGAGTGGCGTG	CGAAAATGGTAGCTCTAATGCC	57	137	GA		yes
POP037	PMGC2418	AATTTTCTCTCTTTACCGCCAG	TGATCCCTCAATGTCTTTACAG	57	152	GA		yes
POP067	PMGC2419	TTTCCCTGTCATCGGCACTG	CATTGGAGACAGCTAATCAGC	57	112	GA		yes
POP068	PMGC2420	GACACCACTTCTAAAGATGGC	ACATGCCTTAGGCTAGTTGC	55	204	GA		yes
POP069	PMGC2423	AAACGGACAATGATAGTCTTTC	GGATTAGCATGTCCTGAGTC	55	132	GA		yes
POP003	PMGC244	CTTAGTTGAAAGTCTCTTAAC	GAATAACAGTTGGTTTTTCAG	51	91	CTT	IX	yes
POP108	PMGC2481	CAAAAGAAGGGTAGAGTCTAC	TTCTTCGGTGTGTGTTATTGC	55	225	GA		yes
POP109	PMGC2499	AGAGGGTTTTCAATAACATACC	TATTGGAACTCTCGTCGACC	55	229	GA		yes
POP110	PMGC2500	AATGTCGACCACTCCACGC	AGAGGGTTTTCAATAACATACC	55	138	GA		yes
POP070	PMGC2501	CACAGGACGTTTTGGAGCAG	AATTCGGACAGTCAGTCACC	55	283	GA	III	yes
POP011	PMGC2515	GAAAAGGGATTGTTAATAAACCC	CCAAAATCATAAAAGACAGGGC	57	223	GA	XIV	yes
POP111	PMGC2516	TAGTTCATTATCCTTGGGCTG	GTCAGTCGCCTTGCATTGC	55	150	GA		yes
POP015	PMGC2522	TCTGTTAATTTCTCAGCTGTTG	TGCTTTACTAAACTTTTTACTGC	55	175	GA	IX	yes
POP016	PMGC2523	TTTTTCTAATTAAAGTCTACAAAC	CAAATCATGCCGGATTTATAGC	53	171	GA	II	yes
POP071	PMGC2525	CGAGTCACAAGCTCCCAATAG	GCAGGCTGTCCTATCTGCG	57	188	GA	XVIII	yes
POP112	PMGC2531	TAAGAGAATTGGGAGAGCAAC	TTTTATCTTTTCCAGTTGTCTAC	55	140	GA	XI	yes
POP072	PMGC2536	GCTGTCTAACATGCCATTGC	CATTTCTTTATCATCACCTTAAC	55	182	GA		yes
POP073	PMGC2541	CATTATAGTCCTGATTGATCTTC	CACTCAAGATCGAGTCTATGG	57	205	GA		yes
POP113	PMGC2550	AGGTTACAAACTTTGTTGTAGC	GAACAAACTCTCACTGTGGTC	55	118	GA		yes
POP038	PMGC2556	ACAAAATGGTCCCCATCTTTC	GCCTACCAATACTAAGAGCC	55	129	GA	VI	yes

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP039	PMGC2558	CCAGAGAAAGAGAGTGCTTC	AATGCAGATGTCGTTGTTTGC	55	155	GA		yes
POP074	PMGC2571	TCTCGCAGATTCATGTAACCC	GACTGTATGTTGACCATGCCC	57	100	GA	X	yes
POP040	PMGC2573	TGCCTTTCTGATCAATATCGC	CTCGTTAATTAGAGTCGAATTAG	55	113	GA	X	yes
POP017	PMGC2574	TGTCTGTTCTTATTTTCCTCTG	AGTTGTATCTCAACTAGAATTAG	55	158	GA	XI	yes
POP041	PMGC2578	GAGAACTCGGTGACTGACTG	CAGCAACATCCACATATTAGC	55	194	GA		yes
POP114	PMGC2585	ACTGCTGTGTATTGCCCTAG	TAGTTGAAGTTGGAGCACAAC	55	132	GA	XV	yes
POP075	PMGC2598	TGAACTGGTCATCATTTGACG	CATATATACAAGTCGTTACCATC	55	146	GA		yes
POP076	PMGC2599	ACAGTACGCAGAAAGCTTGG	TTCTGTTTCGGAGATGTTCAC	55	152	GA		yes
POP077	PMGC2603	CATACTTTCAATTTTCTTACTGC	AAGACTCACAAATGCATCTTGC	55	166	GA		yes
POP078	PMGC2606	AATTTACATTTCTTTATCATCACC	GCTGTCTAACATGCCATTGC	55	188	GA		yes
POP042	PMGC2607	TTAAAGGGTGGTCTGCAAGC	CTTCTTGCACCTCGTTTTGAG	55	177	GA	VIII	yes
POP115	PMGC2610	AACACGCAAGAACATACATAAG	GATTAACATGTTTCGCTACGC	55	114	GA	VIII	yes
POP043	PMGC2611	TGACGATTACAGTTTTTGATCG	CTCCTAATTCCTGACAACCAC	55	277	GA		yes
POP079	PMGC2614	TATACAAAATGTCACCTAAAGAG	CGACATATGGTAGGCATATTTC	55	248	GA		yes
POP044	PMGC2647	CTCGTTAATTAGAGTCGAATTAG	TTGTTATCCACTGCCAGTGC	55	129	GA	X	yes
POP080	PMGC2658	GCCCTTGAATACCATGAGCG	ACCTTCAGTAGATCAGGTTAGTG	57	251	GA	XIII	yes
POP081	PMGC2660	GTTCTATGTGTAGGAGATATCC	TAACAATATGCTTCATAGCACAG	57	127	GA		yes
POP116	PMGC2675	CACACCGACAAATTATGAGTG	TTTTAGAGTGAATTTTCCTGCG	55	?	GA	V	yes
POP117	PMGC2679	GGAATCCGTTTAGGGATCTG	CGTCTGGAGAACGTGATTAG	55	118	GA		yes
POP018	PMGC2691	ATTTTGAATTTGAATTATGTTGTTG	TTTCAGAGTATTTTAGGGTGTC	55	101	GA	II	yes
POP118	PMGC2696	ACACACACACCAGGCTTCTC	TTCTTCATGCAGGAAGGAGC	55	142	GA		yes
POP119	PMGC2699	TTTCCTCCATGTATATCAAACC	AACCCTAATTTCAAGAATTGGG	55	156	GA		yes
POP120	PMGC2702	GCTTATCTCTCCTTCCACCG	GCATCAGCACATCAGAGTTC	55	255	GA		yes
POP121	PMGC2709	ATTGTAATTATTGAACACATGCC	GTGCAGTTCAGAGTATTGTTG	55	210	GA		yes
POP122	PMGC2718	ATCTACCAAACTACATTATCTTG	ACAACTATAAATATAGGCTGCC	55	100	GA		yes
POP082	PMGC2730	GGCTTAATATGGGTCAGGTTC	GAAAACCAAAGAGTCTTCACAG	57	161	GA	VII	yes
POP123	PMGC2731	CGTATAGTACTTGAAGAATCCC	CTGGTCAACAGCTACTGCAC	57	211	GA		yes
POP124	PMGC2737	AGATTAACCTTAGTTTGCTTGG	AGAAGTTAAGGTAACGCTAGG	55	226	GA		yes
POP083	PMGC2765	GAGATAGCATCACCTAGAGG	GATATGTCAAGGAAATCCTTAG	55	163	GA		yes
POP125	PMGC2766	AATAGGATGAGGCTTCAATAC	CATTTCAATCTTTTGTGCCGC	53	231	GA	II	yes
POP126	PMGC2775	ATACCGCCGCGTCCACTC	TGGCGCTGGTTTCTTAGCTG	55	176	GA		yes
POP127	PMGC2786	CTTGAATTTCAATAGAATCGCAAG	TAGCAAAGGAGAGGGTTTCC	55	109	GA	I X	yes
POP128	PMGC2794	GTCACGCTAAGCCAGATCTC	CTAATTTAACTTGCTTGGTCTG	55	151	GA	VII	yes
POP129	PMGC2803	AGGCCTCACTATTTTCAGCC	GCTTTGCTAAGCAGGCTTTC	55	178	GA	IX	yes
POP084	PMGC2804	AAAGTTTTTCATTTTCAATCCTTG	TAATCGCCTATACACAGGCG	55	129	GA	XVI	yes
POP085	PMGC2806	CTTTGATTGCTGATGAATGGC	GGTCCTATCTTCCAAATTCTG	55	178	GA		yes
POP130	PMGC2812	TGCATTATAGTCCTGATTGATC	GATCACAAATGCATCTTGCAAG	55	137	GA		yes

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP045	PMGC2818	AAGCTTCATCGTCCTGCTTG	CGTATCAATTCACGACTCTCG	55	131	GA		yes
POP046	PMGC2826	GCTTCTTTAGCGACATGCATC	GTCAGAACTGTGACAGTAACC	57	237	GA	IV	yes
POP131	PMGC2832	CAAGCTTGGCTTTGCTAAGC	AGGCCTCACTATTTTCAGCC	55	185	GA		yes
POP086	PMGC2838	TGGTCGGCGAGAGTGACCAG	ATTGATCTCTCTTTACATCACTC	57	182	GA		yes
POP047	PMGC2839	AACCCATAGCAAGAAGCTAG	CAATTACCGAAGAGGATTACTG	53	197	GA	V	yes
POP019	PMGC2840	TATATTGAGATCTCATCTAACAG	ACAAAATTCAATTGTGTTGTAATC	55	228	GA	V	yes
POP132	PMGC2847	AATAGGTAACCAGTAATTGTCAG	TAGGCTAGTTGCCAGGACC	55	316	GA		yes
POP048	PMGC2852	ATAATCTCCCTAGCTTAATTCC	GAATAACATGGATAATGTGTTTG	55	113	GA	I	yes
POP087	PMGC2855	GGTATCTTGTTATCCACTGCC	TTTTCCTCGTTAATTAGAGTCG	55	144	GA	X	yes
POP088	PMGC2858	CTTACCATCTTTATCCTAATGC	TTTCAAAAAATAAAAAGCAGCGC	55	100	GA		yes
POP089	PMGC2861	GTTTTTCATGCGACATTGAGG	TTTGATACACAAGTTCATTTGTG	55	227	GA		yes
POP020	PMGC2862	TTTGTAACTAATGAAGATTTGTAC	ATTTTTGTTCTTTTAACCAAAATTC	55	?	GA	XVIII	yes
POP090	PMGC2866	ATTGTTCAAAATCCTCAGGTTC	TAGCATAGTAGCTAGCTAGTG	55	203	GA		yes
POP091	PMGC2873	TGGTTGGAATGTCTTTACTCC	ATACATTGATCTCTCTTTACATC	55	?	GA		yes
POP049	PMGC2879	TTGATTCGAGCCTCACGAGC	AAACTCCAACATTTTAAGGACC	55	152	GA		yes
POP050	PMGC2881	CCTCACTTTCAAATTGAAGCC	AACACATAAATCTTGAAAGGAAC	55	171	GA		yes
POP133	PMGC2885	CATGATCAAATTGGATTTGAATG	AAAGATGAACATGGCTAGCTC	55	317	GA	XII	yes
POP134	PMGC2889	CCCAAGATCCGATTTTTGGG	CACAATGTACAAATCGCTGTC	55	207	GA	XVII	yes
POP097	PMGC325	CGATTTATGACAGACAGCTTG	GTACCGTTGAGGTGGCTAG	55	295	CTT		yes
POP023	PMGC333	CTTAGTGGTGAAGTATTC	GAGTGGGTGCTGATTCATCC	45	110	CTT	XI	yes
POP004	PMGC409	ACGTATATGAAGTTCTTGATTGC	GACAGATCATTATGATTACTACAG	57	150	GA	VIII	yes
POP024	PMGC420	ATGGATGAGAAATGCTTGTG	ACTGGCACACTCTTTAACTGG	51	105	GA	XIV	yes
POP052	PMGC422	AACCTCGAATTAAGAATAACCC	GTCTCGGTTAAGGTATTGTCGC	55	168	GA	II	yes
POP025	PMGC433	GCAGCATTGTAGAATAATAAAAG	AAGGGGTCTATTATCCACG	51	215	GA	XVI	yes
POP005	PMGC451	AATTACAACCACTTTAGCATATTC	TGCCGACACATCACACATACC	57	210	GA	?	yes
POP053	PMGC456	TGTAGGAGATATCCACGTGG	AACAATATGCTTCATAGCACAG	55	115	GA		yes
POP026	PMGC486	AGAAGTTGTTGAACCCGATGGG	GCTACAAACTTTGTTGTACCC	55	150	GA	III	yes
POP054	PMGC510	AGTCCTGGTCCTGGATTGG	CTACATTAATTTCCCTGTCATC	55	130	GA		yes
POP027	PMGC520	TAACTCACTAGAAAAACCTTTG	TTGCTAGCTAGCTTGTTAG	49	120	GA	XV	yes
POP055	PMGC562	TTTTGGGAGGGGAGTCGAG	ACAACTCTCAACTTCCTAATC	53	185	GA	VII	yes
POP028	PMGC571	CTGGTACCGATGGAGAAGAC	CAAACCAACAACTCACCGTAC	57	180	GA	XIV	yes
POP098	PMGC573	GTCATAATCGCCTATACACAG	GATTGTGAACTCGATCTAAAGG	55	175	GA		yes
POP056	PMGC575	TAAATTCATGTAGATTGACG	CTTACTATTTCATGGTTGTC	47	145	GA	I	yes
POP057	PMGC576	GCTGTCTAACATGCCATTGC	AATTTACATTTCTTTATCATCACC	55	185	GA		yes
POP006	PMGC607	TATTTCTACAACATACCAAAACG	CATTACTCAAGCACATGCACGC	55	140	GA	VII	yes
POP022	PMGC61	GATCCCTCTGCACCGTTTAC	ACCCTAAATTTGCTGACAAC	51	360	CTT	VIII	yes
POP029	PMGC639	AACAAATTTGGCCTGCAGGG	TCAAAATATTATCACTAAACGCG	55	120	GA		yes

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP099	PMGC648	GAAGAATAGGATTACATG	ATAAACTCTCTCCTGTTGATTC	43	235	GA		yes
POP007	PMGC649	CATCCATGATATCAAACCAAATTAG	TGTAATCCAAACATAAAATCCCAAG	61	115	GA	XIII	yes
POP058	PMGC667	CATTCGTTCAGTAGTTAAGGC	GGTTAAGCTACCTCTGCTAC	55	220	GA	II	yes
POP059	PMGC683	CCAGCAATGATTGATTGCTCC	GAGCTTTAACTGTCCAGTAGC	57	260	GA		yes
POP008	PMGC684	GAAATTGAATATCTCTCACTTACC	TAATACGTGAAAAGTCAGGTTTTG	59	210	GA	II	yes
POP009	PMGC690	AAAGGAACTTGTTGATGTCAAG	CTTATTTTTCAATCTCGTTTGAGC	55	130	GA	XV	yes
POP001	PMGC93	ATCATGCGTTCGGCTACAGC	CTCAAACTCCAACTGTTATAAC	55	350	CTT	I	yes
POP329	PTAG1	CTTGTAATTAAGAGCAAGCCA	ATGTTAAACTACCTCAAACATATCC	53	209	TA	IV	yes
POP330	PTAG2	CGAATATAGTGGATGGTTATTG	CGAATCTGAGTAGGAGAGATG	55	237	TA	XI	yes
POP245	PTR01	AGCGCGTGCGGATTGCCATT	TTAGTTTCCCGTCACCTCCTGTTAT	59		[GGT]5n45[AGG]9	IV	yes
POP246	PTR02	AAGAAGAACTCGAAGATGAAGAACT	ACTGACAAAACCCCTAATCTAACAA	63		[TGG]8		yes
POP247	PTR03	CACTCGTGTTGTCCTTTTCTTTTCT	AGGATCCCTTCCCTTTAGTAT	55		[TC]11		yes
POP248	PTR04	AATGTCGAGGCCTTTCTAAATGTCT	GCTTGAGCAACAAACACACCAGATG	65		[TC]17		yes
POP249	PTR05	CTTCTCGAGTATAAATATAAAACACCA	TCACATCACCCTCTCAGTTTCGC	65		[TG]7		yes
POP250	PTR06	AGAAAAGCAGATTGAGAAAAGAC	CTAGTATAGAGAAAGAAGAAGCAGAAA	57		[AT]8		yes
POP251	PTR07	ATTTGATGCCTCTTCCTTCCAGT	TATTTTCATTTTCCCTTTGCTTT	53		[CT]5AT[CT]6	XII	yes
POP252	PTR08	TAGGCTAGCAGCTACTACAGTAACA	TTAAGTGCGCGTATCCCAAAGA	59		[A]11[CT]8		yes
POP253	PTR11	ATGATTGAGCTCTCTCAAGGTTGCT	TTTGCAACCATGCTATCTACTTCAA	63		[GT]3G[GT]2GGT		yes
POP254	PTR12	AATAACCATCCCTCCAATAACCTAC	TATTTTGCACCTAAATGGCTGTTCT	63		[AAAG]3A6n7[AAAG]2		yes
POP255	PTR14	TCCGTTTTTGCATCTCAAGAATCAC	ATACTCGCTTTATAACACCATTGTC	63		[TGG]5		yes
POP256	PTR15	CGTGATTGAAGGCGCACTAACCAT	CTTTGTTCTCAGTGGCTGCCTATT	65		[GA]3AA[GA]5n10[GA]5		yes
POP135	WPMS01	AACCACTATGCCACCTTCTT	AACTAACTCCATTCATTGCTAAA	53	150	GA		yes
POP136	WPMS02	AGAAATACCCCTGCTAATC	AATGTTTTTGGTCCGTGAAT	49	200	GA	XVI	yes
POP137	WPMS03	TTTACATAGCATTTAGCCTTTAGA	TTATGATTTGGGGGTGTTATGGTA	57	250	GT	XII	yes
POP138	WPMS04	TACACGGGTCTTTTATTCTCT	TGCCGACATCCTGCGTTCC	53	275	GT		yes
POP092	WPMS05	TTCTTTTTCAACTGCCTAACTT	TGATCCAATAACAGACAGAACA	53	280	GT	XII XV	yes
POP244	WPMS06	GTATAACGATGACCCCACGAAGAC	TATAAATAAAGGCATGACCAGACA	59	200	GT		yes
POP139	WPMS07	ACTAAGGAGAATTGTTGACTAC	TATCTGGTTTCCTCTTATGTG	53	230	GT	VI	yes
POP140	WPMS08	TAACATGTCCCAGCGTATTG	TTTTTAGAGTGTGCATTTAGGAA	53	225	GT		yes
POP093	WPMS09	CTGCTTGCTACCGTGGAACA	AAGCAATTTGGGTCTGAGTATCTG	57	275	GT		yes
POP094	WPMS10	GATGAGAAACAGTGAATAGTAAAGA	GATTCCCAACAAGCCAAGATAAAA	61	250	GT		yes
POP141	WPMS11	TAAAGATGATGGACTGAAAAGGTA	TAAAGGAGAATATAAGTGACAGTT	57	230	GT	II	yes
POP095	WPMS12	TTTTTCGTATTCTTATCTATCC	CACTACTCTGACAAAACCATC	51	170	GT	VI	yes
POP142	WPMS13	GATCCTGAACAATGTCGTACTTC	ACGATAACCTGCGAGAAATGT	55	141	GT		yes
POP143	WPMS14	CAGCCGCAGCCACTGAGAAATC	GCCTGCTGAGAAGACTGCCTTGAC	65	245	CGT	V	yes
POP144	WPMS15	CAACAAACCATCAATGAAGAAGAC	AGAGGGTGTTGGGGGTGACTA	61	193	CCT	V	yes
POP145	WPMS16	CTCGTACTATTTCCGATGATGACC	AGATTATTAGGTGGGCCAAGGACT	65	145	GTC	VII	yes

Laboratory code	Locus name	Left Primer (F)	Right Primer (R)	Tm PCR	Length (bp)	Motif	LG	Tested
POP146	WPMS17	ACATCCGCCAATGCTTCGGTGTTT	GTGACGGTGGTGGCGGATTTTCTT	67	140	CAC	VII V	yes
POP147	WPMS18	CTTCACATAGGACATAGCAGCATC	CACCAGAGTCATCACCAGTTATTG	65	245	GTG		yes
POP148	WPMS19	AGCCACAGCAAATTCAGATGATGC	CCTGCTGAGAAGACTGCCTTGACA	65	204	CAG	V	yes
POP149	WPMS20	GTGCGCACATCTATGACTATCG	ATCTTGTAATTCTCCGGGCATCT	61	252	TTCTGG	XIII	yes
POP150	WPMS21	TGCTGATGCAAAAGATTTAG	TTGGAACTTCAACATTCAGAT	49	242	GCT	II	yes
POP151	WPMS22	ACATGCTACGTGTTTGGAATG	ATCGTATGGATGTAATTGTCTTA	55	129	TGA	XIII	yes

B. SSR analyzed on the F_1

For each SSR analyzed on the F_1 progenies the annealing temperature (Tm), the number of cycles used for the PCR and the separation method of the PCR product are presented. Gel indicates that the amplification products were separated in a 3% high resolution agarose gel, MetaPhor® Agarose (Cambrex BioScience, USA) and ABI3100 indicates that the amplification products were separated with the ABI3100 capillary sequencer (Applied Biosystems, Italy). In the table is also indicated the length of the alleles found for each parent and their allelic configuration.

Lab name	Locus name	Alleles 58-861 (bp)	Alleles Poli (bp)	Allelic configuration		PCR		Separation method
				58-861	Poli	Tm	N Cycles	
POP259	bu810400	102/110	110/133	a,b	b,c	55	42	ABI3100
POP265	bu813833	149	149/152	a,-	a,b	55	42	ABI3100
POP267	bu814989	129	137/145	a,a	b,c	55	42	ABI3100
POP268	bu818855	191	191/196	a,a	a,b	57	30	Gel
POP270	bu885452	177/195	177/190	a,b	c,d	55	30	Gel
POP270b	bu885452	242/260	242/260	a,b	a,b	55	30	Gel
POP340	GCPM_1929-1	154	156/161	a,a	b,c	55	42	ABI3100
POP350	GCPM_350-1	205/211	211	a,b	b,-	55	42	ABI3100
POP321	GCPM2024-1	196	202/214	a,a	b,c	55	42	ABI3100
POP322	GCPM2180-1	129/137	142	a,b	c,-	55	42	ABI3100
POP326	GCPM588	220/222	218/222	a,b	b,c	55	42	ABI3100
POP328	GCPM943	180	180/186	a,a	a,b	55	42	ABI3100
POP156	ORNL_016	226/235	219	a,b	c,-	55	42	ABI3100
POP156A	ORNL_016A	200/207	207	a,b	a,-	55	42	ABI3100
POP156B	ORNL_016B	229	222	a,a	b,-	55	42	ABI3100
POP159	ORNL_023	212/195	195/182	a,b	a,c	51	30	Gel
POP160	ORNL_026	204	213/215	a,a	b,c	55	42	ABI3100
POP162	ORNL_030	233/237	218/237	a,b	b,c	55	42	ABI3100
POP164	ORNL_045	236	248/234	a,a	a,b	53	30	Gel
POP167	ORNL_056	194/181	219/194	a,b	c,d	55	30	Gel
POP169	ORNL_060	199/187	187	a,b	a,a	50	30	Gel
POP180	ORNL_127	156/160	156/159	a,b	a,c	55	42	ABI3100
POP188	ORNL_177	235/241	227	a,b	c,-	55	42	ABI3100
POP191	ORNL_190	192/200	200	a,b	b,b	55	42	ABI3100
POP194	ORNL_206	203/193	213/193	a,b	b,c	53	30	Gel
POP197	ORNL_214	178/164	178	a,b	b,b	53	30	Gel
POP198	ORNL_220	225	225	a,a	a,b	53	30	Gel

Lab name	Locus name	Alleles 58-861 (bp)	Alleles Poli (bp)	Allelic configuration		PCR		Separation method
				58-861	Poli	Tm	N Cycles	
POP259	bu810400	102/110	110/133	a,b	b,c	55	42	ABI3100
POP203	ORNL_264	-	206	-,-	a,-	51	30	Gel
POP205	ORNL_276	227	222	a,-	b,-	45	35	Gel
POP206	ORNL_277	178/190	178/190	a,b	a,b	55	42	ABI3100
POP208	ORNL_279	198	189/196	a,-	b,c	55	42	ABI3100
POP211	ORNL_287	286/244	266/249	a,b	c,d	55	30	Gel
POP212	ORNL_297	246/229/214	225	a,b	c,c	55	30	Gel
POP234	ORNL_399	186/193	188	a,b	c,c	55	42	ABI3100
POP239	ORNL_455	218/220	218	a,b	a,a	55	30	Gel
POP240	ORNL_477	223/211	223/203	a,b	b,c	51	35	Gel
POP243	ORNL_496	218/224	220	a,b	c,c	55	30	Gel
POP361	ORPM_40	192	189/192	a,a	a,b	55	42	ABI3100
POP280	ORPM144b	220/222	218/222	a,b	b,c	55	42	ABI3100
POP360	ORPM193	210	210/204	a,a	a,b	45	35	Gel
POP290	ORPM221	209/232	213/228	a,b	c,d	55	42	ABI3100
POP293	ORPM269	229	232	a,-	b,-	55	42	ABI3100
POP304	ORPM432	195	-	a,-	-,-	55	42	ABI3100
POP307	ORPM444	186/195	175/186	a,b	a,c	55	42	ABI3100
POP310	ORPM451	219/223	213/215	a,b	c,d	55	42	ABI3100
POP030	PMGC2011	97	102	a,a	a,b	57	30	Gel
POP096	PMGC204	296	292	a,-	b,-	45	35	Gel
POP062	PMGC2088	174/183	168/172	a,b	c,d	55	42	ABI3100
POP032	PMGC2140	143/168	162	a,b	c,-	55	42	ABI3100
POP033	PMGC2156	102	104/106	a,-	b,c	55	42	ABI3100
POP102	PMGC2163	275/270	246/241	a,b	c,d	55	30	Gel
POP034	PMGC2217	148/158	148/154	a,b	a,c	55	42	ABI3100
POP013	PMGC2270	129,7/110,2	110.2	a,b	a,-	59	30	Gel
POP107	PMGC2385	?	159/132	a,b	c,d	55	30	Gel
POP069	PMGC2423	171/141	162/119	a,b	c,d	55	40	Gel
POP108	PMGC2481	?	240	a,b	b,c	55	30	Gel
POP109	PMGC2499	235	231/246	a,-	b,c	55	40	Gel
POP070	PMGC2501	238	242	a,a	a,b	55	40	Gel
POP011	PMGC2515	199/204	199	a,b	a,a	57	30	Gel
POP071	PMGC2525	249/201	226/201	a,b	a,c	57	40	Gel
POP113	PMGC2550	154	135/145	a,a	b,c	55	30	Gel
POP039	PMGC2558	130	128/132	a,a	b,c	55	42	ABI3100
POP040	PMGC2573	120/108	100/90	a,b	c,d	55	30	Gel
POP041	PMGC2578	174/184	174/180	a,b	a,c	55	42	ABI3100
POP078	PMGC2606	331/225	357/309	a,b	c,d	55	40	Gel
POP042	PMGC2607	151/175	161/171	a,b	c,d	55	30	Gel

Lab name	Locus name	Alleles 58-861 (bp)	Alleles Poli (bp)	Allelic configuration		PCR		Separation method
				58-861	Poli	Tm	N Cycles	
POP259	bu810400	102/110	110/133	a,b	b,c	55	42	ABI3100
POP115	PMGC2610	92	98/92	a,a	b,c	55	30	Gel
POP043	PMGC2611	254	248/254	a,a	a,b	55	42	ABI3100
POP079	PMGC2614	344	380/340	a,a	a,b	55	30	Gel
POP080	PMGC2658	300/270	312/250	a,b	c,d	57	40	Gel
POP117	PMGC2679	99/101	99/103	a,b	a,c	55	42	ABI3100
POP118	PMGC2696	152/132	132	a,b	a,a	55	40	Gel
POP120	PMGC2702	260/232	252/218	a,b	c,d	55	30	Gel
POP121	PMGC2709	198/202	196/208	a,b	c,d	55	42	ABI3100
POP122	PMGC2718	103/88	88	a,b	a,a	55	30	Gel
POP123	PMGC2731	185/196	201/242	a,a	b,-	57	30	Gel
POP124	PMGC2737	188	170/167	a,b	c,d	55	30	Gel
POP044	PMGC2747	125/135	119/107	a,b	c,d	55	30	Gel
POP129	PMGC2803	182/188	186/190	a,b	c,d	55	30	Gel
POP045	PMGC2818	131/133	120/126	a,b	c,d	55	42	ABI3100
POP046	PMGC2826	207/209	209	a,b	b,b	55	42	ABI3100
POP047	PMGC2839	193/202	191/195	a,b	c,d	55	42	ABI3100
POP132	PMGC2847	180	-	a,-	-,-	55	42	ABI3100
POP048	PMGC2852	99/101	99/101	a,b	a,b	55	42	ABI3100
POP087	PMGC2855	172/161	155/142	a,b	c,d	55	30	Gel
POP088	PMGC2858	122/118	122	a,-	a,b	55	30	Gel
POP090	PMGC2866	213	234/213	a,a	a,b	55	30	Gel
POP091	PMGC2873	118	120/122	a,-	b,c	55	30	Gel
POP049	PMGC2879	176/159	185/168	a,b	a,c	55	30	Gel
POP133	PMGC2885	310/320	308	a,b	c,c	55	42	ABI3100
POP134	PMGC2889	199/201	199/207	a,b	a,c	55	42	ABI3100
POP097	PMGC325	323/302	312	a,b	c,-	55	40	Gel
POP023	PMGC333	100	97/100	a,a	a,b	55	42	ABI3100
POP004	PMGC409	168/174	178/186	a,b	c,d	55	42	ABI3100
POP005	PMGC451	182	166/184	a,a	b,c	55	42	ABI3100
POP026	PMGC486	168/155	159/137	a,b	c,d	55	30	Gel
POP054	PMGC510	141/147	143/151	a,b	c,d	55	35	Gel
POP027	PMGC520	101/114	103/113	a,b	c,d	55	42	ABI3100
POP055	PMGC562	239	241/246	a,a	b,c	55	42	ABI3100
POP057	PMGC576	226/165	259/x	a,b	c,d	55	40	Gel
POP006	PMGC607	137	157	a,a	b,-	55	42	ABI3100
POP022	PMGC61	388/395	375/390	a,b	c,d	45	30	Gel
POP029	PMGC639	95	95/99	a,a	a,b	55	35	Gel
POP099	PMGC648	164/176	200	a,-	b,-	45	30	Gel
POP058	PMGC667	204	249/222	a,b	c,d	55	30	Gel

Lab name	Locus name	Alleles 58-861 (bp)	Alleles Poli (bp)	Allelic configuration		PCR		Separation method
				58-861	Poli	Tm	N Cycles	
POP259	bu810400	102/110	110/133	a,b	b,c	55	42	ABI3100
POP001	PMGC93	350/356	353/359	a,b	c,d	55	42	ABI3100
POP329	PTAG1	266	248/256	a,-	b,c	55	42	ABI3100
POP330	PTAG2	254	242/254	a,-	a,b	55	42	ABI3100
POP245	PTR1	256/261	256/261	a,b	a,b	55	42	ABI3100
POP251	PTR7	272/238	282/262	a,b	b,c	53	30	Gel
POP092	WPMS05	267/279	282	a,b	c,c	55	42	ABI3100
POP135	WPMS1	175/139	150/139/132	a,b	a,c	53	30	Gel
POP094	WPMS10	249	227	a,a	b,-	55	30	Gel
POP141	WPMS11	229	177/189	a,-	b,c	55	42	ABI3100
POP095	WPMS12	164	164/170	a,a	a,b	45	30	Gel
POP142	WPMS13	121	127/110	a,a	a,b	55	30	Gel
POP143	WPMS14	249	235/240	a,a	b,c	55	42	ABI3100
POP146	WPMS17	150	160/150/139	a,a	a,b	60	30	Gel
POP147	WPMS18	233/238	227/230	a,b	c,d	55	42	ABI3100
POP148	WPMS19	215	220/204	a,a	b,c	65	30	Gel
POP136	WPMS2	255/208	255	a,b	b,c	50	30	Gel
POP149	WPMS20	221/239	234	a,b	c,c	55	42	ABI3100
POP151	WPMS22	139/104	143/124	a,b	a,d	55	30	Gel
POP137	WPMS3	283	306/274	a,b	c,d	50	40	Gel
POP138	WPMS4	285	335/282	a,b	c,d	55	30	Gel
POP244	WPMS6	203/218	201/214	a,b	c,d	59	30	Gel
POP139	WPMS7	270/243	222	a,b	c,c	55	30	Gel

Appendix 2

A. Common markers between *Populus nigra* and *Populus alba* maps

The common markers among the 4 maps are listed below by linkage group. The markers in red were not mapped on the same linkage group. The marker in blue was mapped on the same group in *P. nigra* and *P. alba* maps but on another group in the *P. trichocarpa* physical map. The grey filling pattern indicates groups without any common markers.

	P. nigra		*P. alba*	
Linkage group	♀ 58-861	♂ Poli	♀ 14P11	♂ 6K3
I	PMGC2499	PMGC2499	PMGC2499	PMGC2499
	WPMS18	WPMS18		WPMS18
	ORPM177	ORPM177	ORPM177	ORPM177
	PMGC2852	PMGC2852		PMGC2852
	PMGC93	PMGC93		PMGC93
		PMGC2731		PMGC2731
	ORPM399			
		bu813833		bu813833
	PMGC2385	PMGC2385		PMGC2385
II	PMGC2818	PMGC2818	PMGC2818	
			bu813833	
	PMGC2709	PMGC2709	PMGC2709	
	PMGC2088	PMGC2088	PMGC2088	PMGC2088
		ORPM40	ORPM40	
III		PMGC2611	PMGC2611	
		PMGC2501	PMGC2501	
	ORPM30	ORPM30	ORPM30	ORPM30
	PMGC2481	PMGC2481	PMGC2481	PMGC2481
		WPMS10	WPMS10	WPMS10
				ORPM399
IV	PTAG1	PTAG1	PTAG1	PTAG1
	PTR01	PTR01	PTR01	
	PMGC2826		PMGC2826	PMGC2826
		ORPM221		ORPM221
V		WPMS14	WPMS14	
	PMGC2839	PMGC2839	PMGC2839	
		PMGC639		PMGC639
VI	PMGC2578	PMGC2578	PMGC2578	
		ORPM26	ORPM26	
	ORPM190		ORPM190	
		WPMS12	WPMS12	
	ORPM60		ORPM60	ORPM60

Linkage group	*P. nigra* ♀ 58-861	♂ Poli	*P. alba* ♀ 14P11	♂ 6K3
VII		PMGC562		PMGC562
		PMGC607	PMGC607	PMGC607
VII	PMGC2140	PMGC2140	PMGC2140	PMGC2140
		WPMS17	WPMS17	WPMS17
VIII	PMGC2607	PMGC2607	PMGC2607	PMGC2607
IX	ORPM451	ORPM451		ORPM451
X	PMGC2163	PMGC2163		PMGC2163
	PMGC510	PMGC510	PMGC510	PMGC510
	PMGC2855	PMGC2855	PMGC2855	
	PMGC2573	PMGC2573	PMGC2573	
	PMGC2747	PMGC2747	PMGC2747	
XI	PTAG2	PTAG2		PTAG2
		PMGC333	PMGC333	
XII	PTR07	PTR07	PTR7	PTR7
	WPMS5			WPMS5
	PMGC2737	PMGC2737	PMGC2737	PMGC2737
	PMGC2885			PMGC2885
XIII	ORPM276			
	PMGC2847		PMGC2847	PMGC2847
XIV				
XV	PMGC2679	PMGC2679	PMGC2679	
XVI				
XVII	PMGC2803	PMGC2803		
		PMGC451	PMGC451	
	PMGC325	PMGC325	PMGC325	
	PMGC648	PMGC648		PMGC648
XVIII				
XIX		Sex	Sex	
		ORPM276	ORPM276	
			PMGC2803	PMGC2803
	ORPM432		ORPM432	

B. Common markers among *Populus nigra*, *Populus alba* and *Populus trichocarpa* maps

Below are reported, by linage group, the common markers among 6 maps: *P. nigra* female and male maps from an intra-specific cross, *P. nigra* map by Cervera *et al* (2001) from an inter-specific cross with *P. deltoides*, and *P. alba* maps by I. Paolucci (personal communication). The markers in red were not mapped on the same linkage group. When there are no common markers, cells are in grey.

Linkage group	*P. trichocarpa*	*P. nigra* ♀ 58-861	*P. nigra* ♂ Poli	Cervera *et al*	*P. alba* ♀ 14P11	*P. alba* ♂ 6K3
I	PMGC2550		PMGC2550			
	PMGC2499	PMGC2499	PMGC2499		PMGC2499	PMGC2499
	PMGC2500				PMGC2500	PMGC2500
		WPMS18	WPMS18			WPMS18
		ORPM177	ORPM177		ORPM177	ORPM177
					PMGC2696	
	PMGC2852	PMGC2852	PMGC2852			PMGC2852
					bu813610	bu813610
	PMGC93	PMGC93	PMGC93	PMGC93		PMGC93
	PMGC575					
			PMGC2731			PMGC2731
			PMGC204			
		ORPM399				
			bu813833			bu813833
		PMGC2385	PMGC2385			PMGC2385
	ORPM173				ORPM173	
II	PMGC2818	PMGC2818	PMGC2818	PMGC2818	PMGC2818	
	H11R		H11R			
	PMGC667	PMGC667	PMGC667	PMGC667		
	ORPM287	ORPM287	ORPM287			
	IAA		IAA			
	WPMS11	WPMS11	WPMS11			
					bu813833	
	PMGC2709	PMGC2709	PMGC2709		PMGC2709	
	PMGC2088	PMGC2088	PMGC2088		PMGC2088	PMGC2088
	PMGC2523				PMGC2523	PMGC2523
			ORPM40		ORPM40	
III					bu890808	bu890808
	PMGC2879	PMGC2879	PMGC2879	PMGC2879		
	PMGC2858		PMGC2858			
	PMGC2611		PMGC2611		PMGC2611	
	PMGC2501		PMGC2501		PMGC2501	
	ORPM30	ORPM30	ORPM30		ORPM30	ORPM30
	PMGC2481	PMGC2481	PMGC2481		PMGC2481	PMGC2481
III	WPMS10		WPMS10	WPMS10	WPMS10	WPMS10

Linkage group	*P. trichocarpa*	*P. nigra* ♀ 58-861	♂ Poli	Cervera *et al*	*P. alba* ♀ 14P11	♂ 6K3
I	PMGC2550		PMGC2550			
	A15R		A15R			
						ORPM399
					PMGC2274	PMGC2274
	PMGC486	PMGC486	PMGC486	PMGC486		
IV	ORPM220		ORPM220			
	PTAG1	PTAG1	PTAG1		PTAG1	PTAG1
	ORPM127	ORPM127	ORPM127			
	PTR01	PTR01	PTR01		PTR01	
		PMGC2826		PMGC2826	PMGC2826	PMGC2826
	PMGC2881			PMGC2881		
		PMGC2515				
	ORPM221		ORPM221			ORPM221
						PMGC2020
		PMGC2270		PMGC2270		
V	PMGC2838			PMGC2838		
	PMGC2873	PMGC2873	PMGC2873	PMGC2873		
	PMGC2606	PMGC2606	PMGC2606			
	PMGC576	PMGC576	PMGC576			
		PMGC2156	PMGC2156	PMGC2156		
	PMGC2558		PMGC2558			
	WPMS14		WPMS14		WPMS14	
	WPMS19		WPMS19			
	PMGC2839	PMGC2839	PMGC2839	PMGC2839	PMGC2839	
			PMGC639			PMGC639
VI	PMGC2578	PMGC2578	PMGC2578	PMGC2578	PMGC2578	
			ORPM26		ORPM26	
		WPMS4	WPMS4			
	PMGC2217	PMGC2217	PMGC2217	PMGC2217		
	ORPM190	ORPM190			ORPM190	
	ORPM197				ORPM197	
	PMGC2423	PMGC2423	PMGC2423	PMGC2423		
	WPMS12		WPMS12	WPMS12	WPMS12	
	PMGC2328			PMGC2328	PMGC2328	
	WPMS9			WPMS9		
		ORPM60			ORPM60	ORPM60
	PMGC2556				PMGC2556	PMGC2556
	ORPM279	ORPM279	ORPM279			
VII			PMGC562			PMGC562
	PMGC607		PMGC607	PMGC607	PMGC607	PMGC607
				PMGC575		
	PMGC2140	PMGC2140	PMGC2140		PMGC2140	PMGC2140
	ORPM144					
			WPMS17	WPMS17	WPMS17	WPMS17
	WPMS16			WPMS16		
VIII		PMGC2696				
	PMGC2607	PMGC2607	PMGC2607	PMGC2607	PMGC2607	PMGC2607
	PMGC61	PMGC61		PMGC61		
	PMGC2610		PMGC2610			
	PhyB1	PhyB1				
VIII	ORPM370					ORPM370
	ORPM327				ORPM327	ORPM327

Linkage group	*P. trichocarpa*	*P. nigra* ♀ 58-861	♂ Poli	Cervera *et al*	*P. alba* ♀ 14P11	♂ 6K3
I	PMGC2550		PMGC2550			
					ORPM301	ORPM301
	PMGC409	PMGC409	PMGC409	PMGC409		
	ORPM264		ORPM264			
	WPMS13		WPMS13	WPMS13		
	ORPM56	ORPM56	ORPM56			
		ORPM269b	ORPM269b			
					ORPM269	
IX	ORPM451	ORPM451	ORPM451			ORPM451
				PMGC2020		
	GCPM588	GCPM588	GCPM588			
		ORPM144	ORPM144			
	ORPM444	ORPM444	ORPM444			
	ORPM23	ORPM23	ORPM23			
	PMGC2803					
X	PMGC2696					
	PMGC2163	PMGC2163	PMGC2163			PMGC2163
	PMGC510	PMGC510	PMGC510		PMGC510	PMGC510
	PMGC2855	PMGC2855	PMGC2855		PMGC2855	
	PMGC2573	PMGC2573	PMGC2573		PMGC2573	
		PMGC2747	PMGC2747		PMGC2747	
	ORPM389				ORPM389	ORPM389
	PhyB2		PhyB2			
	bu814989		bu814989			
XI	ORPM269	ORPM269	ORPM269			
			PMGC2011	PMGC2011		
	PTAG2	PTAG2	PTAG2			PTAG2
	ORPM217				ORPM217	
	PMGC2866		PMGC2866	PMGC2866		
	GCPM943		GCPM943			
	PMGC2531				PMGC2531	
	PMGC333		PMGC333	PMGC333	PMGC333	
XII	PTR07	PTR07	PTR07		PTR7	PTR7
	WPMS3	WPMS3	WPMS3			
	ORPM210				ORPM210	
	WPMS5	WPMS5		WPMS5		WPMS5
					PMGC108	PMGC108
					GCPM3178-1	
	PMGC2737	PMGC2737	PMGC2737		PMGC2737	PMGC2737
	PMGC2885	PMGC2885		PMGC2885		PMGC2885
XIII	PhyA	PhyA				
	PMGC14			PMGC14	PMGC14	PMGC14
	bu810400	bu810400	bu810400			
	ORPM16	ORPM16	ORPM16			
	ORPM55				ORPM55	
	WPMS22	WPMS22	WPMS22			
	PMGC2658	PMGC2658	PMGC2658			
		ORPM276				
		WPMS1	WPMS1			
XIII					bu814260	bu814260
		PMGC2847			PMGC2847	PMGC2847

Linkage group	*P. trichocarpa*	*P. nigra* ♀ 58-861	*P. nigra* ♂ Poli	Cervera *et al*	*P. alba* ♀ 14P11	*P. alba* ♂ 6K3
I	PMGC2550		PMGC2550			
	ORPM297	ORPM297				
	WPMS20	WPMS20				
				PMGC420		
XIV	GCPM1929-1		GCPM1929-1			
	PMGC420					
	I13R	I13R	I13R			
	ORPM193		ORPM193			
	PMGC571				PMGC571	PMGC571
	GCPM748-1				GCPM748-1	GCPM748-1
	PMGC2515					
	ORPM40					
XV	PMGC2679	PMGC2679	PMGC2679		PMGC2679	
	ORPM438					ORPM438
		ORPM477	ORPM477			
	PMGC520	PMGC520	PMGC520			
XVI	GCPM1418-1					GCPM1418-1
		WPMS2	WPMS2			
	PMGC2614		PMGC2614			
		WPMS6	WPMS6			
	GCPM598-1				GCPM598-1	
				PMGC433		PMGC433
	GCPM2180-1	GCPM2180-1	GCPM2180-1			
	GCPM2024-1		GCPM2024-1			
	PMGC204					
XVII	PMGC2889	PMGC2889	PMGC2889			
		PMGC2803	PMGC2803			
			PMGC451		PMGC451	
	GCPM641				GCPM641	GCPM641
		PMGC325	PMGC325		PMGC325	
	PMGC648	PMGC648	PMGC648	PMGC648		PMGC648
XVIII	ORPM496	ORPM496				
	ORPM477					
	ORPM214	ORPM214				
	PMGC2525	PMGC2525	PMGC2525			
	ORPM6				ORPM6	ORPM6
	bu885452	bu885452	bu885452			
XIX			Sex		Sex	
	ORPM276		ORPM276		ORPM276	
	ORPM263				ORPM263	ORPM263
					PMGC2803	PMGC2803
	ORPM206	ORPM206	ORPM206			
	GCPM3178-1					
		PMGC2702	PMGC2702			
						PMGC204
	ORPM433				ORPM433	ORPM433
	ORPM432	ORPM432			ORPM432	