Construction of a microsatellite based genetic linkage map of almond

By

Iraj Tavassolian

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Discipline of Wine and Horticulture

School of Agriculture, Food and Wine

May 2008

Table of Contents

List of Figures

List of Tables

Abstract

Almond (*Prunus dulcis*) is the most important nut crop in terms of world production. Due to its health benefit and high nutritional value the consumption and world supply of almond is increasing. To remain competitive in the world market, the Australian almond breeding program was established to produce cultivars with better adaptation to Australian conditions. As part of this program an almond mapping population consisting of 93 F_1 progeny derived from a cross between the American cultivar 'Nonpareil' (NP) and the European self-compatible cultivar 'Lauranne' (LA) was produced to construct the genetic linkage maps. The first almond linkage map developed prior to the commencement of this project failed to produce the eight linkage groups similar to the basic chromosome number of almond (x = 8) and many large gaps were also observed on the linkage groups. Therefore, more markers were needed to saturate the maps.

Microsatellite markers are considered one of the best choices for mapping studies. 195 microsatellite markers isolated from *Prunus* species were obtained from published papers or by personal communication. Polymorphism was revealed by three different methods, and in general, polyacrylamide gel electrophoresis (PAGE) compared to the fluorescent labelled marker detection using an automated DNA sequencer or agarose gel electrophoresis, showed the most efficient and cost effective method of genotyping. A subset of 54 markers which produced reliable and easily interpretable polymorphic bands was selected to screen the whole mapping population. Microsatellites originally isolated from almond species showed the highest rate of amplification and polymorphism followed by peach microsatellites and the least informative markers were isolated from cherry. It seems that the level of

transportability and usefulness of microsatellite markers is related to the genetic distance of the closely related species. Almond and peach belong to the same subgenus (Amygdalus) and other *Prunus* species are classified in Prunophora subgenus.

The nut, or kernel, is the commercial part of the almond tree, thus to improve the quality of fruit an understanding of environmental influence, heritability and correlation of traits is required. Pomological and quality characters such as: shell hardness, kernel size, shape, taste, pubescence, colour, and percentage of doubles were measured during three consecutive years (2005-2007) on the total mapping population, but data analysis (ANOVA) was performed only on trees that survived for all three years. Most of the traits showed high broad-sense heritability and kernel shape showed the highest heritability of $H^2 = 0.92$ suggesting high genetic control of this trait. Occasionally larger kernels than either parent were found in the progeny indicating potential for improvement of this trait even with smaller kernel size parent that encompass many desirable characters. High correlation was also found between the in-shell and kernel weight ($r = 0.74$), kernel length / kernel width ($r = 0.67$), kernel weight to kernel length ($r = 0.78$) and kernel width ($r = 0.80$). This correlation estimation pointed out in this study indicates that the improvement of one character may result the progress in another trait. Neither of the parents in the mapping population had bitter or obvious slightly bitter taste but slightly bitter kernels were observed among the progeny. Amygdalin was assumed to be responsible for bitter taste in almond; therefore we measured the amount of amygdalin in sweet and slightly bitter kernel progeny by HPLC. However, the results showed that amygdalin exists in sweet kernels as well. Although the average amount of amygdalin in slightly bitter kernels (20.34 mg kg^{-1} FW) was higher than sweet kernels (3.67 mg kg^{-1} FW), some

sweet kernels had higher amounts of amygdalin suggesting the impact of other components on slightly bitter kernel. The highest variability within the traits was observed in the percentage of double kernel, which showed the highest standard error. Strong environmental effects, particularly low temperature at pre-blossom time is speculated to produce much higher double kernels.

Three genetic linkage maps, one for each parent and an integrated map were constructed by the addition of 54 new microsatellite markers to the previous dataset. All the data was scored and coded according to the coding system necessary by JoinMap3 which was used for map construction. 131 markers including microsatellite, ISSR, RAPD, SCAR and *S*-allele markers were placed on the integrated map covering 590.7 cM with the average density of 4.5 cM/marker. The minimum number of six microsatellite markers was placed on linkage group 8 and the linkage group 1 which is the longest linkage group has 14 microsatellite markers. Comparative mapping study with other *Prunus* maps, especially with the highly saturated reference map showed complete synteny and minor changes in the order of four markers on linkage groups compared with *Prunus* reference map. The conservation of molecular marker order observed in this study supports the idea of looking at *Prunus* genome as a single genetic system and practical application of this similarity would be in cross-transportability of microsatellite markers from well developed linkage maps to the less studied species in *Prunus*. Ten microsatellite loci placed on our map have not been reported before and could be used to improve the density of other *Prunus* maps, especially the reference map.

This study contributed to the better understanding of the mode of inheritance and environmental effect on morphological traits and the effect of amygdalin on kernel taste. The most saturated microsatellite based almond linkage map developed

in this study can serve as a framework for future almond breeding program in Australia and benefit *Prunus* improvement programs internationally.

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or tertiary institution and, to the best of my knowledge contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent for my thesis to be made available for photocopying and loan when deposited in the University library.

Iraj Tavassolian

May, 2008

Abbreviations

A	Amygdalin
AFLP	Amplified fragment length polymorphism
ANOVA	Analysis of variance
BAC	Bacterial artificial chromosome
BC	Before Christ
bp	Base pair
BSA	Bulked segregant analysis
χ^2	Chi-squared
ºC	Degrees celsius
cM	centi-Morgan
CsCl	Cesium choloride
CTAB	Cetyltrimethylammonium bromide
cv.	Cultivar
CAPs	Cleaved amplified polymorphic sequence
DArT	Diversity arrays technology
DETC	Diethylidithiocarbamic acid
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DPSTC	Double pseudo-testcross
EDTA	Ethylene diamine tetraacetic acid
EST	Expressed sequence tag
FAM	6-carboxy-fluorescine
F_1	First familial generation

F_2	Second familial generation
FAM	6-carboxyfluorescein
FAO	Food and agriculture organisation
FISH	Fluorescence in-situ hybridisation
FLB	formamide / bromophenol blue
FW	Fresh weight
g	gram
G	Linkage group
GDR	Genome database for Rosaceae
H	Broad sense inheritability
HCl	Hydrochloric acid
HEX	Hexacholoro-6-flourescine
HEX	Hexachloro-6-carboxy-fluorescein
HPLC	High performance liquid chromatography
ISSR	Inter simple sequence repeat
Kb	Kilobase
LA	Lauranne
LD	Linkage disequilibrium
LOD	Logarithm of odds ratio
LSD	Least significant difference
M	Molar
MAS	Marker assisted selection
min	minute
mg	milligram
µL	Microlitres

mL	Millilitres
mM	Millimolar
mm	Millimetre
NH4Ac	Ammonium acetate
NP	Nonpareil
NSW	New South Wales
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pH	Power of hydrogen
%	Percent
QTL	Quantitative trait loci
®	Registered trademark
r	coefficient of correlation
RAPD	Randomly amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RHS	Royal Horticultural Society
SCAR	Sequence characterised amplified region
SD	Standard deviation
SDS	Sodium dodecylsulphate
SE	Standard error
sec	second
SNP	Single nucleotide polymorphism
sp.	Species
SSR	Simple sequence repeats
TBE	Tris borat- EDTA buffer

TE	Tris-EDTA buffer
TEMED	N,N,N',N'- tetramethylethylenediamone
T × E	‘Texas’ × ‘Earlygold’
U	Enzyme unit
USDA	United state department of agriculture
V	Volts
var.	Variety
v/v	Volume per volume
w/v	Weight per volume
YAC	Yeast artificial chromosome

Acknowledgments

This thesis would not been written without the invaluable assistance provided by many people listed below. If I missed anyone I sincerely apologise. I am deeply grateful to Dr Chris Ford for his peer supervision, constructive comments and endless support and encouragement. I always remember him as a role model in my next career. I express my appreciation to Dr Michelle Wirthensohn for all her patience, expert advice and enthusiasm. She has really brilliant mind and was so helpful to me and my wife. I also extend my gratitude to Prof. Margaret Sedgley who was very inspiring and her constant interest gave me the determination about the project from the start.

I wish to specially thank Dr Shubiao Wu for his important assistance on scoring of the microsatellite markers and comments on a draft of the thesis manuscript. Many thanks also to Prof. Raffaele Testolin, Dr Pere Arús and Dr Werner Howad for providing me the sequence of some unpublished microsatellite markers. My thank goes to Dr Ian Riley for statistical assistance and helpful advice during study in Australia.

I am also grateful to many other friends: Esmail, Abbas and Nasser for joyful conversations and sharing their time with me. Many thanks to the colleagues Dr Tricia Franks, Dr Matthew Hayes, Vanessa Melino, Venetia Joscelyne and Crista Burbidge for lively and lovely atmosphere they provide during my lab work.

I gratefully acknowledge Ministry of Science, Research and Technology of Iran for financial support of this study. Finally I want to express my greatest thanks to my wife, Samaneh Poormohammad, and my parent for their patience, constant

encouragement, support and trust that they placed in me throughout this study. Without their help I would never been able to do this research.

Chapter 1

General introduction and literature review

1.1 General introduction

The improvement of plant characteristics that will increase benefits for farmers and industries is the main objective of plant breeding programs. An understanding of the genetic basis of commercially important characters before any attempt of their manipulation is very important for breeders. This information can be obtained by understanding how the genes of interest are inherited, in which position these are located on the chromosome and how they are regulated. One method for organising genetic information is the creation of a genetic linkage map. This involves placing markers, e.g. a portion of DNA, in order, calculating the genetic distance on the basis of recombination rate and assigning them into the linkage groups to construct a linkage map (Staub et al., 1996). This map gives a theoretical location of markers on the chromosome and the relative location of the gene of interest. The development of saturated linkage maps, with the number of predicted linkage groups in agreement with the basic chromosome number of the species in question, and short interval marker

coverage, has many applications in plant improvement programs (Tanksley et al., 1989). To construct a saturated linkage map, it is important that the markers show a high level of multiplex ratio and reproducibility. When molecular markers and the genes of interest are tightly linked to each other on the genetic linkage map, the gene or trait of interest may be used for early selection at the seedling stage. High-resolution maps with reliable markers also can assist the isolation of economically important genes. Flowering plant genomes are remarkably different in size and arrangement, however, a high degree of synteny (genes remain on corresponding chromosomes) and collinearity (in corresponding order) was observed among many of them. Collinearity and synteny between species has been identified by comparison of one-to-one loci conservation. Comparative mapping by comparison of genes or loci between linkage groups of different species can be used for further understanding of genome arrangement and evolution study.

1.2 Botanical classification of almond

The almond (*Prunus dulcis*) Miller [D.A. Webb] syn. *P. amygdalus* Batsch is classified within the family of Rosaceae. The *Prunus* genus contains many important fruit crops such as peach (*P. persica* (L) Batsch), apricot (*P. armeniaca* L.), sweet cherry (*P. avium* L.) sour cherry (*P. cerasus* L.) and plum (*P. domestica* L.). The chromosome number of *Prunus dulcis* is 2n=16 which is the same as other members of *Prunus* (Kester and Gradziel, 1996). *Prunus* is characterized by species that produce fruit known as a drupe where the seed is enclosed in a hard, lignified endocarp referred to as the stone, and the edible portion is a juicy mesocarp. In contrast to other members of the genus such as peach, apricot and cherry, the seed or kernel is the most valuable

part of the almond tree, which includes an embryo and seed coat and residue of nucellus and endosperm (Kester and Gradziel, 1996; Baird et al., 1998). Almond produces a drupe with a pubescent exocarp, a thin and fleshy mesocarp (hull) that after the period of development becomes dry and dehiscent at maturity, and a distinct hardened endocarp (shell). The dehiscence of the hull distinguishes almond from other *Prunus* species (Kester and Gradziel, 1996).

1.3 Economic importance and uses

The almond represents the largest commercial production of any nut crop. The worldwide production was estimated to have increased from 1,478,313 tonnes in 2000 to 1,766,127 tonnes in 2006 (FAO, 2007). The United States of America is the world's largest producer followed by Spain, Syria, Italy and Iran (Fig 1.1; Fig 1.2). Australia produced less than 1% (12,419 tonnes) of the total world production and the major regions of production are South Australia, Victoria and New South Wales (Fig 1.3) (Australian Bureau of Statistics, 2007). Australia has many advantages in the production of almond, for instance, almond growing areas are among the best reliable and pest free areas in the world and the productivity and growing techniques is similar to the USA. Moreover, Australian almonds are produced counter-seasonal to the majority of world suppliers (Anon, 2007).

Because of its health benefits and high nutritional value almond has been recognised as an important component of many diets and its consumption is growing rapidly. Almond is a good source of energy which contains lipid (52.2%), carbohydrates (20.4%) and 4.45 water-soluble sugar (Kader, 1996). Apart from the excellent source of macronutrients almond has more unique phytonutrients, such as vitamin E (α-

tocopherol), folic and oleic acid (Socias i Company et al., 2008). The main fatty acids in the almond kernel are oleic (70-80% of total fatty acid content) linoleic (10-17%) and palmitic (5.5-6.5%) acids (Nanos et al., 2002). The high level of oleic acid suggested that increased consumption of almonds could be very valuable for the human diet, because it is known that oleic acid reduces low-density lipoprotein cholesterol that inhibits blood circulation (Sabate et al., 1996). Almond is very rich in antioxidants and its oil has been used in facial creams and beauty treatments as it is claimed to prevent skin wrinkles and make radiant skin (Gilbert, 2005).

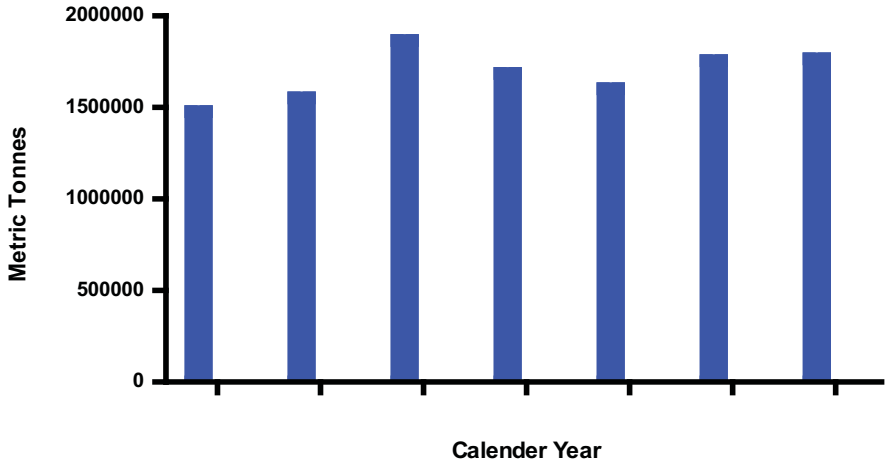


Fig. 1.1 Almond world production during 2000-2006

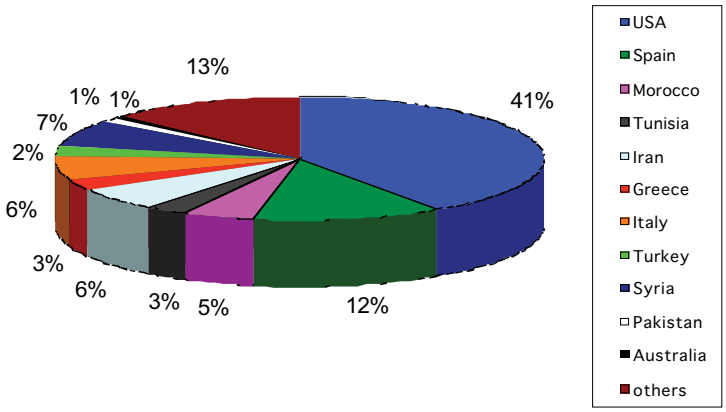


Fig. 1.2 The world top producers of almond

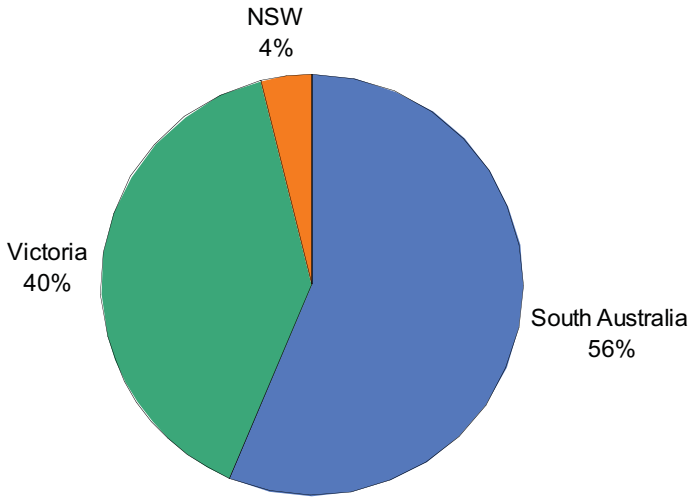


Fig. 1.3 main almond production area in Australia

1.4 Genetic diversity and origin of almond

Almond is one of the oldest fruit tree species cultivated by man: it was mentioned in the Old Testament that the rod of Moses was an almond "Moses striking the rock with sprouting almond rod" (Moldenke and Modenke, 1952). It has been recognised that the almond was first domesticated in central to southern Asia during the third millennium BC (Spiegel-Roy, 1986). The evolution and distribution of almonds can be divided into three stages from the original growing area in mountains of western China and middle Asia to the Americas: Asiatic, Mediterranean and Californian (Fig 1.4) corresponding to the geographical growing area (Kester, 1991; Kester and Gradziel, 1996; Martínez-Gómez et al., 2007). Several hypotheses have been suggested about the origin of cultivated almonds. At first it was proposed that the cultivated almond was generated by selection from *Amygdalus communis* L., resulting in two natural populations (Vavilov, 1930; Denisov, 1988). Wild almond species mostly have bitter kernels due to the accumulation of amygdalin (Conn, 1980). As a result of the genetic similarity between almond and peach, it was proposed that both of them were derived from the same primitive species (Socias i Company, 1998).

1.5 Almond breeding

The domestication of almond was started many centuries ago by selection of superior morphological traits. The most important factor for almond selection was the development of the sweet kernel. The three main objectives of almond breeding programs to meet industry's demand are increasing yield, improvement of kernel quality and reduction of the production costs (Kester and Gradziel., 1996). Due to outcrossing and self incompatibility in almond, it shows high level of heterozygosity, which provide a window of opportunity for genetic studies. Progeny of a cross shows high polymorphiss at many loci (Kester and Gradziel., 1996). Some objectives of almond breeding programs are shown in Table 1.1

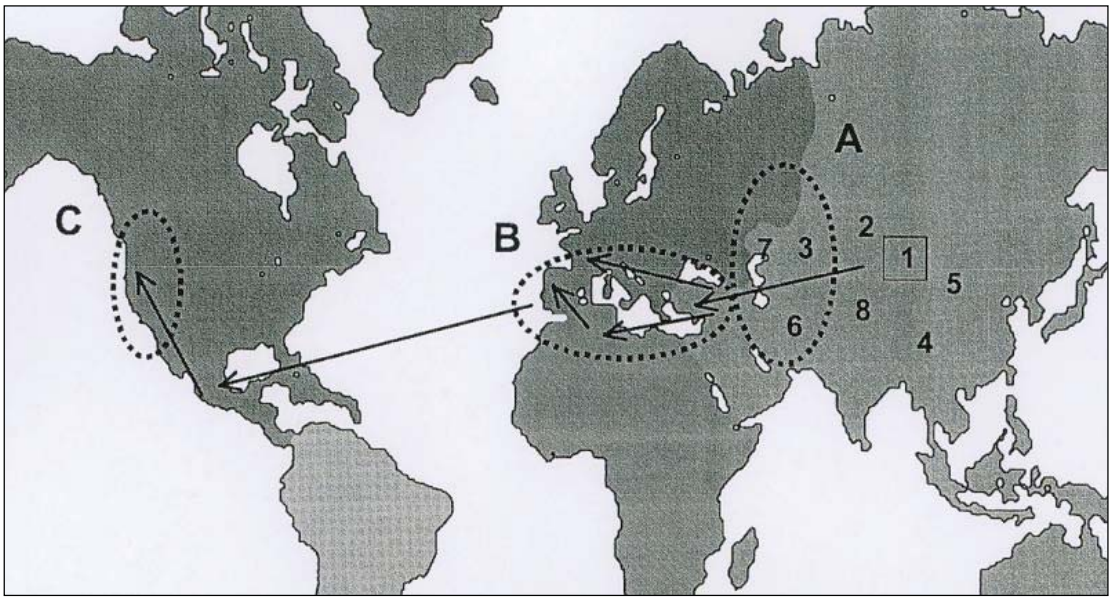


Fig. 1.4 Map of the world showing the origin for almond [*Prunus* dulcis (1)] and different relative *Prunus* species [*P. bucharica*(2), *P. fenzliana* (3), *P. davidiana* (4), *P. persica* (5), *P. scoparia* (6), *P. webbii* (7), and *P. argentea* (8)], the dissemination routes for the cultivated almond [→] and the three main areas for diversification and cultivation of almonds [Asiatic (A), Mediterranean (B) and Californian (C)] (Martínez-Gómez et al., 2007).

Table 1.1. Objectives of Almond Breeding Programs

Category	Trait	Outcome	Programs and Countries
Overcoming Adverse Environmental Conditions			
Spring frosts	late bloom	avoid early and late spring frosts	USA, Ukraine, France Greece, Spain, Italy, Romania Turkey, Bulgaria, Australia
Winter freezing	hardy buds and wood	avoids loss of dormant flower bud; avoid tree damage	Ukraine, Romania Bulgaria
Low winter chilling	low chilling requirements	grow in subtropical area	Israel, Tunisia Australia, Morocco
Moisture stress	drought tolerance	grow with deficient Irrigation	Spain, Italy
Overcoming Adverse Environmental Conditions			
Lack of/ reduced bee populations	self-fertility	eliminate or reduced need for bee pollinisers	France, Greece, Tunisia Spain, Italy, USA
Pathogens (tree or crop)	resistance to fungal and bacterial diseases & insects	eliminate or reduced need for chemical sprays	USA, France Spain, Italy
Adapting to Commercial Conditions			
Reproductive Efficiency/ self-incompatibility	cross- compatibility; blossom timing; density	maximise cross-pollination	USA, France, Ukraine
Self-fertility	self-compatibility autogamy	reduced need for bee pollinisers; single-cultivars orchards	USA, France, Italy Spain, Tunisia Australia, Greece
Tree management	tree size, shape branching, growth habit	efficient orchard management; adjust orchard density; pruning; shaking	USA, Spain
Harvest and handling	time of maturity ease and complete-teness of nut removal, hulling	extend harvest period; efficient and complete harvest	all programs
	shell character very hard shell	kernel protection and storage prevent worm infestation	Spain, France, Italy
	semi soft to soft	higher shelling percentages	USA, Ukraine, France
	well sealed	kernel protection	USA, all programs

Table 1.1 Objectives of Almond Breeding Programs

Category	Trait	Outcome	Programs and Countries
Adapting to Commercial Conditions			
Quality and market use	shell: California-soft shell, blanch-able; Mediterranean hard shell	ability to be marketed at established maximum prices	USA all programs
	kernel: improved characteristics [a]		
Yield	tree: precocity Productivity	early production high yield	all programs
Adapting to Commercial Conditions			
Yield	regularity of bearing	no alternate bearing	
	nuts: control of rejects (worms, gum, blanks, shrivels)	increased kernel yield and reduced need for control measures	all programs all programs
Tree replacement or yield reduction	virus and virus-like infections	"clean stock" programs	USA, Spain, France, Italy
	noninfectious bud failure	low bud failure potential	USA

[a] Quality includes appearance, shape, uniformity, size, colour of testa, thickness of testa, no doubles or creases, flavour, processing characteristics. Adapted from Kester and Gradziel (1996).

1.6 Genetic markers

A genetic marker is a segment of DNA on a chromosome whose inheritance can be followed. They are usually inherited in a Mendelian manner (Kumar, 1999). Genetic markers include: morphological or visible markers, protein variants or isozymes and DNA markers. Genetic markers have been used for different purposes in plant research such as fingerprinting for classification, genome mapping and marker-assisted selection, (Baird et al., 1998). So-called 'molecular markers' are a class of genetic markers that ultimately represent pieces of DNA on a chromosome, which produce a banding pattern

when DNA is cut or amplified with enzymes. Molecular markers have become widely used for producing genetic (meiotic) maps in the past 15 years.

1.6.1 Morphological markers

Morphological markers have been used for many years for identification and characterization of species or cultivars. These markers can be visualised easily and be used for selection, especially those traits that are controlled by a single locus. In pea, these markers have been used for tagging certain virus resistance genes. For example, the flower mutant, *k*, has been used as a marker for bean yellow mosaic virus resistance, (conferred by the gene *mo)* (Weeden et al., 1994). Despite the simple monitoring, expression of such markers is influenced by epistatic and pleiotropic interactions, and dominant-recessive relationships. Moreover, the numbers of these markers are limited and differentiation between heterozygous dominant and homozygous dominant individuals impossible (Kumar, 1999).

1.6.2 Biochemical markers

Markers that are discriminated based on differences at the expressed protein level are named biochemical markers. They can often be separated by electrophoresis allowing the researcher to identify polymorphisms (Kumar, 1999). One of the most commonly used forms of biochemical markers is isozymes, which are different forms of the same enzyme (Vodenicharova, 1989). Isozymes are reproducible, transferable between labs and have co-dominant expression, thus heterozygotes can be distinguished from homozygotes. Nevertheless their use is limited due to their low variation in a given

species and small number of loci can be analysed by conventional staining methods Martínez-Gómez et al., 2003b).

1.6.3 DNA markers

DNA markers show polymorphism at the DNA level and are abundant on the genome (Kumar, 1999). Depending on the method used for revealing the polymorphism they can be divided into two categories: hybridisation-based polymorphisms and PCR-based polymorphisms. Usually DNA markers have no effect on the phenotype because they are a reflection of the natural variation present in the DNA sequence (Kumar, 1999). Therefore, a detailed genetic linkage map can be constructed utilising only one cross and one mapping population. DNA markers are free of pleiotropic effects, thus allowing any number of markers to be monitored in a single population (Kumar, 1999). Although any DNA marker can be used for gene mapping, the most frequently used are RFLPs, RAPDs, ISSRs, microsatellites and AFLPs.

1.6.3.1 Hybridisation-based polymorphism

RFLPs

Hybridisation-based polymorphism includes restriction fragment length polymorphism (RFLPs) and variable number tandem repeats (VNTR) or minisatellites. RFLPs are DNA fragments that are obtained from DNA digestion, separation of the restricted fragments by agarose gel electrophoresis, transformations of fragments by Southern blot followed by hybridisation with radioactive or fluorescent probes, and can be visualised by staining with ethidium bromide (Wunsch and Hormaza, 2002). There are several reports of the use of RFLP in mapping study in *Prunus* (Rajapakse et al.,

1995; Joobeur et al., 1998; Wang et al., 2000). Although RFLP markers are co-dominant, simply transferable between labs and robust, these markers have limitations including: the development cost of probe-enzyme combinations are relatively high, the quantity of DNA required for generation of the DNA fingerprint is more than other methods, large-scale DNA extraction is laborious and time consuming, and they are complex to use because procedures have not been automated and it often involves using radioisotopes (Baird et al., 1998; Martínez-Gómez et al., 2003b).

1.6.3.2 PCR-based polymorphisms

In the polymerase chain reaction (PCR), a pair of oligonucleotide primers is used to copy each strand of denatured template DNA with the help of *Taq* polymerase enzyme, which adds nucleotides to 3′-end points of the primers. PCR has three steps: template denaturation, primer annealing and enzymatic extension. These steps are repeated several times to amplify the intervening regions that are specified by the primers. As a result, the quantity of the DNA increases exponentially. PCR-based methods allow quick identification of DNA polymorphism by detection of different amplified molecular fragments (co-dominant markers) or by presence or absence of amplified bands (dominant markers) through electrophoresis on agarose or polyacrylamide gels and staining methods (Baird et al., 1998). PCR-based markers are the most common types of marker which have been used for mapping studies.

RAPD

Random Amplified Polymorphism DNA (RAPD) is a fragment of genomic DNA that is amplified by PCR using a decamer primer with arbitrary sequence, where

polymorphism is revealed upon the presence or absence of amplified products (Wunsch and Hormaza, 2002). RAPDs have been used because of the moderately high level of polymorphism that they reveal, the low cost compared with techniques like isozyme analysis and RFLP as well as their simplicity. Although isozymes did not reveal any diversity in red pine, (*Pinus resinosa*) RAPD markers revealed diversity in the population (DeVerno and Mosseler, 1997). Owing to the small size of RAPD primers (10 bp), the amplification pattern may vary between assays with the same materials (Jones et al., 1997). This makes them less reproducible than other techniques. Also, these markers are dominant and heterozygous individuals cannot be distinguished from homozygous dominant individuals (Jones et al., 1997).

Microsatellite

Microsatellites or simple sequence repeats (SSRs) are short tandem repeats of sequence units generally less than 5 base pairs e.g. $(TA)_n$ or $(AAT)_n$ (Bruford and Wayne, 1993). A microsatellite reveals polymorphism because of variation in the number of repeat units. Polymorphism occurs whenever one genome is missing one or some of its motifs or has insertion that modifies the distance between repeats (Kumar, 1999). Studies show that microsatellites primers can be amplified in the same microsatellites region in related taxa (Rabinson and Harris, 1999, Dirlewanger et al., 2004a). Microsatellites have been especially useful for molecular genetic analysis because of their abundance, ability to be tagged in the similar genome region, their high level of polymorphism and their ease of detection via automated systems (Kumar, 1999). These markers are co-dominant and highly reproducible, which make them ideal for genome mapping (Dayanandan et al., 1998). Some difficulties associated with

microsatellites include practical problems for isolation and analysis of microsatellites such as the cost of the production of DNA library and clones containing microsatellites, also stutter bands which are usually less intense than the desired product make the scoring process difficult. If primers have not been designed in earlier research, then the isolation of this marker can be both time consuming and expensive (Powell et al., 1996; Ciofi et al. 1998; Rabinson and Harris, 1999).

ISSR

Inter-simple sequence repeats (ISSR) were developed based on the direct amplification of microsatellite sequences via PCR (Zietkiewicz et al., 1994). Although there is some problem in developing SSR markers whereby flanking sequences must be known to design PCR primers, ISSR primers can be developed without the prior sequence knowledge. The PCR reaction amplifies the sequence between two SSRs, producing a multi-locus marker system useful for fingerprinting, diversity analysis and genome mapping. The efficiency of ISSR is similar to RAPDs when agarose gels and ethidium bromide are used for separation and staining (Henry, 2001). The ISSR markers can be used for saturation of linkage maps. The barley linkage map developed by RFLPs was saturated with 60 ISSRs that distributed to all linkage groups particularly among the clustered RFLPs tip of the chromosomes and areas with low marker density (Becker and Heun, 1995). The disadvantages of ISSRs are they segregate mostly as dominant markers and showed segregation distortion (Wang et al., 1998a).

AFLP

Amplified fragment length polymorphism (AFLP) combine the use of restriction enzymes, similar to the RFLP, with PCR amplification of selective fragments. At first, DNA is double digested by two restriction enzymes, rare cutter such as *PstI* which produce low copy fragments and frequent cutter such as *MseI*, which produce lots of fragments. Then two different short selective nucleotides (adaptors) are ligated to both ends of fragments which act as binding sites for PCR primers, and then primers are designed for these templates for amplification by PCR (Vos et al., 1995; Wunsch and Hormaza, 2002). AFLP markers can detect several loci simultaneously and have been used to construct a high density map especially when a framework map with RFLPs or microsatellites is available (Lambert et al., 2004; Semagn et al., 2006). However, using AFLP is technically difficult and compared with other markers is relatively expensive (Jones et al., 1997).

1.7 Application of the saturated linkage map in plant breeding

To use the polymorphisms as genetic markers, knowledge of their genomic location is required. This information can be obtained by constructing a genetic linkage map (Kumar, 1999). Linkage groups represent the genetic loci along the chromosome. Three steps have been used to construct a genetic linkage map: first, development of a source of probes and primers, which identify polymorphism in the genomic DNA of a segregating population; second, segregation analysis and verification of Mendelian inheritance, the parents should be genetically divergent to show the maximum polymorphism in the progeny although not too far so that they cause sterility. Finally, statistical analysis, after scoring the produced polymorphisms, the results are assigned

to computer programs such as Linkage, MAPMAKER or JoinMap to construct the genetic linkage map (Foolad et al., 1995; Van Ooijen et al., 2001). Marker-assisted selection (MAS) is one of the most valuable applications of genetic linkage maps. Tagging the genes that control the phenotypic trait with a marker, which shows no or very low recombination with the gene, enable the breeder to make an early indirect selection among seedlings and discard those that do not have the marker. However, even if the marker and gene are in linkage disequilibrium, a recombination event can separate them. Unless the marker contains the gene itself, the success of MAS depends on the distance of the marker from the gene (Luby and Shaw, 2001). MAS can be more useful where the trait(s) are simply inherited, the trait is expressed in the mature phase of long juvenile period plant, the cost of screening by conventional methods is much higher and if the locus discovery and marker genotyping are economical (Luby and Shaw, 2001). The SCAR markers linked to the resistance root-knot nematodes developed in Myrobalan plum was proved to be practical for MAS in peach and almond (Lecouls et al., 2004). Nematode resistance (*Ma*) gene is located on linkage group 7 of the *Prunus* reference map (Dirlewanger et al., 2004a).

Map-based cloning has been used to separate the genes that are responsible for Mendelian trait differences (Remington et al., 2001). Map-based cloning has two requirements, first individuals in a population must be divergent in a trait of interest and second, the trait has been mapped on a linkage map adjacent to genetic markers. The availability of a high-resolution linkage map decreases the amount of physical mapping required for chromosome walking (Tanksely et al., 1995). In this method, a genetic linkage map is used to localise the target gene to one or more markers, then chromosome walking is begun, from the flanking marker and series of overlapping

clones [(BACs) bacterial artificial chromosome or (YACs) yeast artificial chromosome], to reach the opposite side of the target gene close to the known marker; i.e. in fact chromosome walking involves passing over the gene of interest (Tanksely et al., 1995). Currently map based cloning and physical mapping of peach as a model for Rosaceae species is underway. The map was composed of 2,138 contigs containing 15,655 BAC clones. In total, 2,633 markers comprised of peach unigene EST set, cDNAs, RFLPs and overgo probes (overgo probe consists of two 24mer oligonucleotides, derived from genomic sequence, which share eight base pairs of complimentary sequence at their 3' ends) are integrated into physical framework. The physical framework was anchored to the *Prunus* genetic map using 152 core genetic probes derived mainly from the *Prunus* reference map 'Texas' × 'Early Gold' (T x E) (Zhebentyayeva et al., 2008).

Comparative mapping allows the transfer of information from map-rich species to map-poor related species (Miller, 1997). The pre requirement of comparative mapping is the availability of co-dominant and transferable markers. The high degree of conserved marker and gene positions make it possible to map genes in a new species by transferring markers isolated and mapped in another one (Testolin, 2003). Microsatellite markers were suggested as the best class of markers for comparative study and many *Prunus* maps share their microsatellites, which enable the investigation of genetic background and homology (Aranzana et al., 2003, Dirlewanger et al., 2004a).

1.8 Morphological traits in Almond

Although among the nut crops almond is considered as a well studied crop, compared with other rosaceous species relatively little information about the

morphological traits and their inheritance is available. To produce improved cultivars, elucidating the role of environmental effects and the mechanism of trait inheritance is very useful. The degree of similarity among relatives (parent and progeny) determines heritability which in turn facilitates the choice of breeding method and improves the efficiency of breeding programs (Falconer and Mackay, 1996). Shell hardness along with, kernel size, shape, colour, taste and the percentage of double kernels are important characters for almond selection.

The preference of the shell hardness in each region and industry is different. In Europe hard shell is acceptable while in America and Australia soft or paper shell which can be removed easily by hand is preferred (Socias i company et al., 2008). It was suggested that shell hardness is controlled by a single gene *D* in which hard shell is dominant *DD* to soft shell *dd* (Grasselly, 1972). Modifier genes can act quantitatively and change the degree of shell hardness (Kester and Gradziel, 1996). Hard shell varieties are more resistance to the attack of insects such as navel orange worm (Sederstrom, 1977) and damage by birds.

Kernel taste is a primary important trait in almond improvement programs. Heppner (1923, 1926) suggested that the kernel bitterness in almond is a monogenic trait in which bitter kernel is recessive (*ss*). This hypothesis was supported by several other studies and showed that major almond cultivars such as ‘Nonpareil’ and ‘Mission’ are heterozygous for this trait (Dicenta et al., 1993; Spiegel-Roy and Kochba, 1981). It was shown that bitter kernel has high levels of the glucoside amygdalin (Kester and Gradziel, 1996). But a few studies have been carried out regarding the contents of amygdalin in slightly bitter kernels. In some cases the amount of amygdalin in sweet kernel was higher than slightly bitter kernel (Dicenta et al., 2002).

Occasionally both ovules in the almond pistil are fertilised and produce double kernels, which is an undesirable trait that decreases the value of the crop. It was suggested that this trait is under strong environmental influence (Spiegel-Roy and Kochba, 1974; Kester et al., 1977). Previous study of pre-blossom temperature indicated the correlation between low temperature and high production of double kernels (Egea and Burgos, 1995). However more studies in different cultivars and environmental conditions are suggested to confirm the role of temperature in double kernel production.

1.9 Development of an Australian almond linkage map, prior to the commencement of this project

1.9.1 Introduction

The long term nature of fruit breeding caused the decline of many tree fruit breeding programs. The application of new technology such as genetic linkage map will speed up breeding programs by decreasing the time, labour and cost. For the breeder, a molecular map may serve as a tool to reach a goal of improved cultivars (Scorza, 2001). The first almond genetic map was constructed using an interspecific F_2 population of a cross between almond ('Texas') and peach ('Earlygold'). This map comprises 246 markers (11 isozymes and 235 RFLPs) and covered the total map distance of 491 cM. Markers were assigned into eight linkage groups, the same as the basic chromosome number of the almond (Joobeur et al., 1998). Later on, 54 RAPD markers and six microsatellites were added to the former almond map and improved the quality of the almond linkage map by filling low-density regions (Joobeur et al., 2000).

Although this map produced eight expected linkage groups, the markers used have limited application in other mapping populations. Isozymes and RFLPs markers are highly transferable across species but they rely on technologies which are considered tedious, labour demanding and costly and RAPD markers are rarely transferable across the same species therefore less suitable for comparative mapping study. The almond × peach map generated in Spain did not produce suitable markers for the Australian almond industry, thus it was decided to construct an Australian almond × almond genetic linkage map (Wirthensohn et al., 2002). A low-density linkage map of the Australian F_1 mapping population derived from the cultivars 'Nonpareil' × 'Lauranne' (NP × LA) was developed by Gregory et al., (2005) using RAPD, ISSR and microsatellites markers which produced three linkage maps, one for each parent and an integrated map. The maps were not saturated, did not match the chromosome number of almond (x = 8), and genome coverage was poor and many large gaps were observed; further details of these maps are presented in the next sections.

The efficiency of the map in breeding programs is directly associated with the amount of map saturation. The closer the markers are to the trait of interest the more useful it becomes. Therefore, the number of markers on the map must be increased to saturate them. The microsatellite marker was chosen over the other DNA based markers to complete the first almond maps reported by Gregory et al., (2005). Microsatellites are robust, co-dominant, very polymorphic and informative, PCR based and easily transferable across related taxa which make it the best choice for comparison and mapping study (Testolin, 2003). Microsatellite markers have been successfully used for the development of a genetic linkage map in almond and many other *Prunus* species (Joobeur et al., 2000; Dirlewanger et al., 2004a). Consequently we intend to use this

marker to produce a reliable and transferable linkage map for almond breeding programs.

1.9.2 Mapping population

1.9.2.1 Introduction

A mapping population consists of the individuals derived from an intraspecific or interspecific cross in which the parents are different in the traits of interest. Genetic tools are utilised to identify factors (or loci) and to determine the recombination distance between loci within the segregating population. The traits to be studied must be polymorphic between the parents and significantly heritable. The more difference in the desirable characters in the parents, the more loci will be found (Schneider, 2005).

In many species, mapping populations are F_2 progeny. However, several reports indicate that in tree species F_1 progenies were successfully used as the mapping population (Grattapaglia and Sedoroff, 1994; Maliepaard et al. 1998; Yamamoto et al., 2002; Grando et al., 2003; Verde et al., 2005). Such species are generally self-incompatible, highly heterozygous and have a long generation time. Although by crossing related species of plant it would be reasonable to expect a higher level of polymorphism than between even very different members of the same species, thereby producing greater likelihood of being able to produce genetic map, it may be difficult to break linkage when the desirable traits are linked to undesirable ones (Schneider, 2005).

1.9.2.2 Australian almond mapping population

The Australian almond improvement program was launched in 1997 to meet the needs of the almond industry. A number of crosses using Australian or overseas elite

cultivars were carried out in the first year of this program. As part of the program, a mapping population from a cross between 'Nonpareil' (NP) and 'Lauranne' (LA) was developed. The almond industry in Australia and California is dominated by the cultivar NP, it has superior nut characteristics such as pale testa colour, paper shell, and good kernel quality. LA is a French cultivar that is currently grown commercially in Europe. It is derived from the cross between 'Ferragnès' and 'Tuono' at INRA-Avignon in 1978. 'Ferragnès' has superb kernel quality for fresh consumption. LA is a hard-shelled cultivar with some desirable traits like late flowering and self-fertility, the latter inherited from Tuono. This family was selected for further analysis due to expected high segregation for important agronomic traits (Godini, 2002; Wirthensohn and Sedgley, 2002; Gregory, 2004).

In 1997, the controlled cross hybridisation between NP and LA was performed at the Waite Agricultural Research Institute (Fig. 1.5), The University of Adelaide, Adelaide, South Australia (034º 53´ 37"S and 138º 33´ 54" E). NP was used as the female parent and kept in a bird-proof cage at the Waite Claremont Orchard under normal irrigation and fertilisation. Two NP clones were chosen for pollination (3-8-4-72 and 3-8-7-72). Woolley et al., (2000) reported that RAPD analysis showed no genetic differences between these clones. At the balloon stage, branches were selected and kept in an insect proof net for 2-5 days to allow the flowers to open before hand pollination. In February 1997, Dr Henri Duval (INRA) kindly provided LA pollen from the original hybrid plant in INRA-Avignon, France. The pollen was stored at -20ºC before use. LA pollen was applied by camel hair paint brush to approximately 300 open NP flowers in the insect proof net. The nuts were allowed to develop until fully mature. Two hundred and forty two nuts from this cross were collected and placed in a closed plastic bag

containing damp vermiculite and stored at 4ºC for 5-7 weeks for stratification. When the seeds had germinated and roots and hypocotyls had emerged 208 viable seeds out of the seedlings were acclimatised in the shade house for four months. The 184 surviving progeny were transplanted on their own roots to the commercial almond orchard owned by Andrew Lacey at Lindsay Point, Victoria, Australia (034º 15´ 27" S, 141º 00´ 00" E). The planting design consisted of two adjacent rows with very close planting distance of 1.0 m between trees and 3 m between successive double rows with micro-sprinkler irrigation between the double rows. The seedlings were maintained under standard orchard management and no pruning has been carried out. From 184 individuals that were transplanted to the orchard, three died and a further 10 were small and weak before the first assessment in 2001.

Fig. 1.5 The controlled cross hybridisation between NP and LA in Waite Campus within the bird cage and branches sealed in insect proof nets

Leaves for DNA isolation were collected from 171 progeny, which were assessed to confirm hybridity by RAPD and *S*-allele analysis. A total of 93 progeny were randomly selected from the 171 to construct the genetic linkage map (Sedgley and Collins, 2002; Gregory, 2004).

1.9.3 Molecular and morphological markers used to develop the first generation of the map

Genomic DNA extractions method, PCR protocols and staining methods are completely explained in Chapter 2. The first generation of the NP x LA map was developed by Gregory, (2004). She screened a total of 119 molecular DNA-based markers: 60 RAPD, 23 ISSR and 36 SSR. All primers were screened on the parents and 10 progeny to see if they would give amplification and polymorphic bands, then, the polymorphic markers were tested on the entire mapping population. Twenty microsatellite primers (Table 1.2), six RAPD primers from Operon Kits A, B and C, and 13 ISSR's that produced reproducible, clear, and polymorphic bands, were selected to construct 3 maps, one for each parent and the integrated map using the markers that produced polymorphic patterns in both parents. From 73 polymorphic markers, 19 were inherited from the female parent (NP) 30 from the male parent (LA) and 24 from both parents. Twelve quantitative traits such as: bloom time, self-compatibility, shell hardness and kernel characteristics, were measured for morphological analysis over three years (2001, 2002, 2003) and χ^2 analyses were performed for deviation from the expected Mendelian ratio. Three loci were placed on the maps. Testa colour was placed at the top of linkage group 4 (LA), kernel taste was placed at the bottom of linkage group B-NP and the locus for self-incompatibility was placed at linkage group six.

Paternal self-incompatibility alleles were *S*7*S*8 and maternal self-compatibility alleles (S3Sf) were scored using allele specific primers, (Gregory, 2004).

Another visiting scientist, Dr Mourad Mnejja from Professor Pere Arús laboratory in Spain, tested segregation in progeny of the NP x LA cross of a further 36 microsatellite primer sets chosen from the highly saturated International *Prunus* reference map. They were ordered as fluorescence-labelled primers and analysed on ABI Prism® DNA sequencer. Twenty-one microsatellites showed polymorphism (Table 1.3), 16 primers have been mapped and 5 remain unlinked. Parental and integrated maps were produced by the software JoinMap 3.0 (Van Ooijen and Voorrips, 2001) using the cross-pollination format at a LOD of 3.0-4.0. Map distances were calculated using the Kosambi mapping function and data were presented pictorially with the MapChart 2.1 software (Voorrips, 2002). Figures 1.6 to 1.8 show the final maps constructed by Gregory and Mnejja (unpublished). The parental map of LA with 50 polymorphic loci produced 11 linkage groups with map coverage of 320.9 cM. The maternal map of NP with 59 polymorphic markers coalesced into 8 linkage groups (no linkage group two) with a length of 482.6 cM. The integrated map with 36 polymorphic markers in both parents consisted of 7 linkage groups and spans a total length of 360.9 cM. First generation maps provided the starting point of the current project.

1.9.4 Aims of study and expected outcome

1. To saturated the previous almond map by producing the missing linkage groups and filling the large gaps between the markers

2. To test the transportability of the microsatellite primers from other *Prunus* species to almond

3. To determine the random distribution of microsatellite markers in the almond genome

4. To estimate broad sense heritability, correlation and genetic variation and environmental effects such as temperature and rainfall on important agronomic traits

5. To elucidate the mode of inheritance and amygdalin content effect on almond flavour

6. To compare the synteny and colinearity of the resultant map with other existing *Prunus* maps particularly the *Prunus* reference map (T × E)

7. To develop a tool for future MAS and QTL (quantitative trait loci) mapping in almond

Expected Outcome

Co-dominant microsatellite markers may provide good coverage of the genome with a smaller number of markers than other systems. A microsatellite based map would allow the genome comparison study of the almond with other *Prunus* species and information from well saturated maps such as the location of resistance genes and morphological traits can be predicted by synteny analysis. Investigation of the genotype × environment interaction, inheritance and correlation of morphological traits can lead us to better understanding of the factors controlling these characters. Furthermore, the results can be used to predict the expression of the trait under question and be used for selection. Moreover, the heritability and correlation of almond nut characters can be compared with other nut crops. A saturated map for 'Nonpareil' as the world's most cultivated and standard cultivar for pomological traits will be constructed. The map will be a valuable resource for MAS, physical mapping and map-based cloning of genes, which are responsible for resistance and nut quality.

Table 1.2 Polymorphic SSR markers used by Gregory, (2004).

Origin	Primers	Annealing T°C	References
Almond enriched library of cv. Texas	CPDCT006	62	Mnejja et al. 2005
	CPDCT007	62	
	CPDCT018	47	
	CPDCT019	62	
	CPDCT022	62	
	CPDCT023	62	
	CPDCT024	62	
	CPDCT027	62	
	CPDCT028	62	
	CPDCT034	60	
	CPDCT035	62	
	CPDCT046	60	
	CPDCT046	57	
Peach enriched library	CPPCT005	52	Aranzana et al. 2002
Peach enriched library Of cv. Redhaven	UDP96-001	57	Cipriani et al. 1999
	UDP96-005	57	
	UDP97-401	57	
	UDP98-405	57	
	UDP98-409	57	
Cherry enriched genomic library of cv. Napoleon	PS12e2/ PS12A02	55	Downey and Iezzoni, 2000

Table 1.3 Polymorphic SSR markers screened by Mnejja (2004).

Origin	Primers	Annealing T°C	References
CT/AG enriched genomic library of almond cv. Texas	CPDCT008	62	Mnejja et al., 2005
	CPDCT020	62	
	CPDCT024	62	
	CPDCT025	62	
	CPDCT027	62	
	CPDCT042	62	
	CPDCT045	62	
CT/AG enriched genomic library of Japanese plum cv. "Santa Rosa"	CPSCT006	62	Mnejja et al., 2004
	CPSCT012	62	
	CPSCT018	52	
	CPSCT021	46	
	CPSCT022	62	
	CPSCT024	46	
Peach EST Database	EPDCU2584	57	Dirlewanger et al., 2004a
	EPDCU3083	57	
	EPDCU3392	57	
	EPDCU3454	57	
	EPDCU4658	57	
	EPDCU5183	57	
AG/CT enriched library of cv. Redhaven	UDP98-408	50	Cipriani, 1999
CT enriched genomic library of peach	Pchgms3	54	Sosinski, 2000

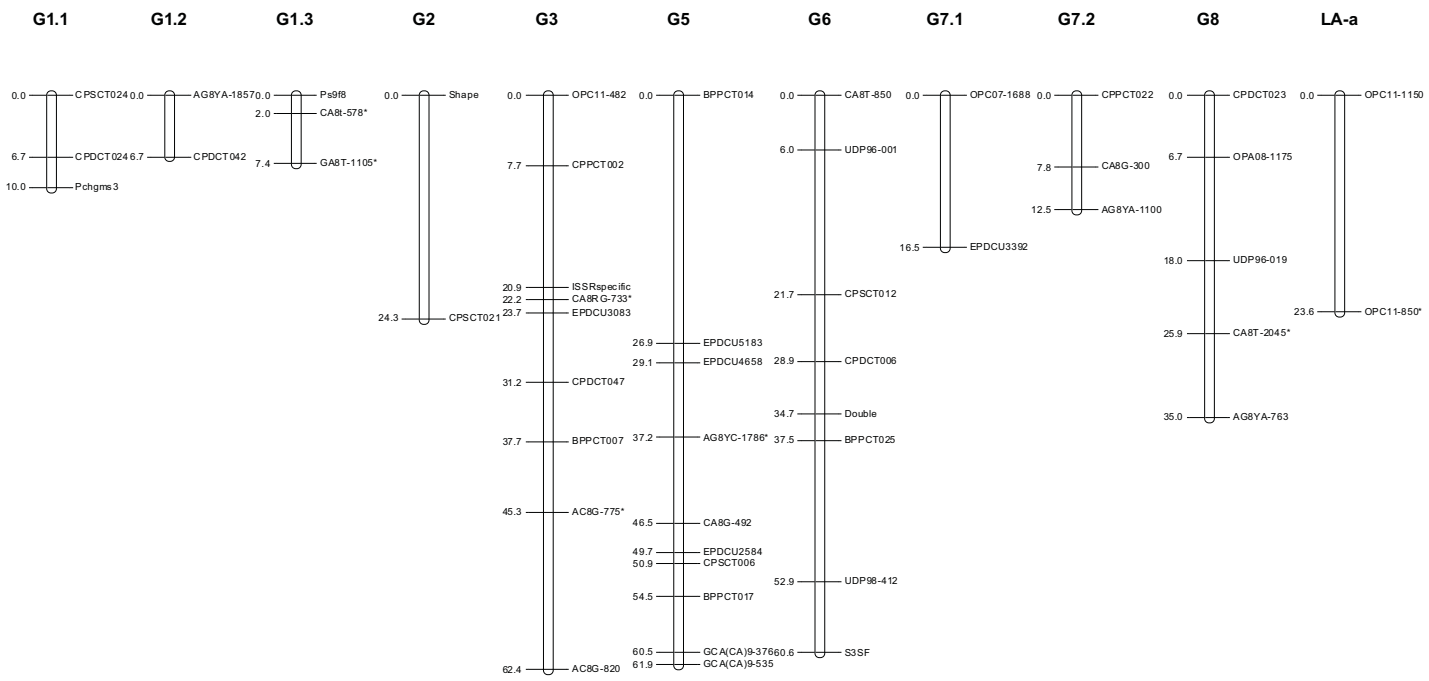

G1.1
0.0 CPSCT024
6.7 CPDCT024
10.0 Pchgms3
G1.2
0.0 AG8YA-1857
6.7 CPDCT042
G1.3
0.0 Ps9f8
2.0 CA8t-578*
7.4 GA8T-1105*
G2
0.0 Shape
24.3 CPSCT021
G3
0.0 OPC11-482
7.7 CPPCT002
20.9 ISSRspecific
22.2 CA8RG-733*
23.7 EPDCU3083
31.2 CPDCT047
37.7 BPPCT007
45.3 AC8G-775*
62.4 AC8G-820
G5
0.0 BPPCT014
26.9 EPDCU5183
29.1 EPDCU4658
37.2 AG8YC-1786*
46.5 CA8G-492
49.7 EPDCU2584
50.9 CPSCT006
54.5 BPPCT017
60.5 GCA(CA)9-376
61.9 GCA(CA)9-535
G6
0.0 CA8T-850
6.0 UDP96-001
21.7 CPSCT012
28.9 CPDCT006
34.7 Double
37.5 BPPCT025
52.9 UDP98-412
60.6 S3SF
G7.1
0.0 OPC07-1688
16.5 EPDCU3392
G7.2
0.0 CPPCT022
7.8 CA8G-300
12.5 AG8YA-1100
G8
0.0 CPDCT023
6.7 OPA08-1175
18.0 UDP96-019
25.9 CA8T-2045*
35.0 AG8YA-763
LA-a
0.0 OPC11-1150
23.6 OPC11-850*

Fig. 1.6 A linkage map of the diploid almond, developed from the markers present in the male parent 'Lauranne' (LA) (Gregory and Mnejja, unpublished data). Linkage groups named according to the convention of the *Prunus* reference map T _ E (Aranzana et al. 2003). Distances between markers are given in
centi Morgans. Asterisks indicate distorted markers

NOTE: This figure is included on page 29 in the print copy of the thesis held in the University of Adelaide Library.

Fig. 1.7 A linkage map of the diploid almond, developed from the markers present in the female parent 'Nonpareil' (NP) (Gregory and Menejja, unpublished data). Linkage groups named according to the convention of the *Prunus* reference map T _ E (Aranzana et al. 2003). Distances between markers are given in centi Morgans. Asterisks indicate distorted markers**.**

NOTE: This figure is included on page 30 in the print copy of the thesis held in the University of Adelaide Library.

Fig. 1.8 Integrated linkage map of the diploid almond cross of LA × NP (Gregory and Mnejja, unpublished data). Linkage groups named according to the convention of the *Prunus* reference map T×E (Aranzana et al. 2003). Distances between markers are given in centi Morgans. Asterisks indicate distorted markers.

Chapter 2

Microsatellite analysis for linkage mapping

2.1 Introduction

Genetic variability is an essential part of any breeding program. Molecular and genetic markers have been used to detect polymorphisms as a pre requirement for construction of genetic linkage maps. Both types are a result of differences between the genotypes, or more accurately, differences in DNA sequences. A few morphological markers have been employed to identify polymorphisms, however, there are some limitations such as difficulty of assessment, variation with environment, cultural practices and time consumption when collecting morphological data in large populations that may be growing in different locations. Molecular markers have no pleiotropic effects, remain unaffected by the environment, can be detected and expressed at all developmental stages, show Mendelian inheritance and there is almost an unlimited number of them (Ahmad et al., 2004).

Among biochemical and molecular markers, isozymes, restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), microsatellite and amplified fragment length polymorphism (AFLP) and their modifications have been extensively utilised for the preparation of molecular linkage maps. Isozymes and RFLPs are the earliest markers used in *Prunus* variability detection

and linkage maps (Vezvaei et al., 1995; Foolad et al., 1995, Joobeur et al., 1998). Since they are time consuming, costly and labour intensive including Southern hybridisation for RFLPs, their use has now been replaced by more convenient PCR-based markers. RAPDs compared with RFLP, can be used with lower investment of time and cost. This is because of their random nature for binding to priming sites with no need to find the sequences of given primer sites to design primers. Joobeur et al., (2000) reported the use of 325 RAPD markers to increase the coverage of an almond intraspecific F_1 population derived from 'Ferragnès' and 'Tuono' which resulted in a 35% increase over markers for a previous map. Gregory et al., (2005) also used RAPD markers to construct a framework of the first generation linkage map which is used in this study. The use of this type of marker is however, limited by its dominant expression, poor reproducibility and low power of discrimination. Dirlewanger and Bodo (1994) reported that from 524 RAPD primers tested only 38 (7%) gave polymorphic bands which segregated in a mapping population. Caparro et al., (1994) reported that only 16% of the 522 RAPD primers assayed gave clear polymorphisms in their population.

Following the invention of RFLP and RAPD markers, a second generation of molecular markers involving microsatellites and AFLPs were developed during the 1990s. AFLP, which is based on PCR amplification of restriction fragments from digested genomic DNA (Vos et al., 1995), has recently been used to generate linkage maps in a variety of plant species such as subtropical grass, *Paspalum notatum*, (Pablo et al., 2001), *Eucalyptus* (Muburg et al., 2003) and lentil (Hamwieh et al., 2005). Due to their high multiplex ratio, AFLP markers were employed to rapidly enrich a previously generated map in *Prunus* (Dirlewanger et al., 1998; Verde et al., 2005). AFLP markers used to saturate a RFLP linkage map of maize exhibited high throughput amplification;

some primer combinations produced 19 polymorphic markers (Castiglioni et al., 1999). However, the disadvantages of AFLP are the complexity of method, dominant nature and requirement of licensing for commercial use (Ahmad et al., 2004).

Microsatellites are short tandem repeats of between 2 and 6 base pairs, extending over regions between and 10 to 60 nucleotides in length (Hemmet et al., 2003). Originally, the term microsatellite has been used for the dinucleotide motif, repeating pattern, CA (GT) (Litt and Luty, 1989; Weber and May, 1989). Various terms have been used to describe tandem repeated sequences, including simple sequences repeat (SSR; Tautz, 1989) and STR (short tandem repeats; Edwards et al. 1991). Microsatellite has now become the most common name for this kind of sequence and they have been discovered in all eukaryotes studied so far.

Microsatellites reveal polymorphism due to differences in the lengths of short tandem repeats at a specific locus of individuals. The microsatellite markers technique has a number of advantages: they are PCR-based so it is quick to detect and requires small amounts of DNA, they are codominant markers, so more informative, evenly distributed throughout the genome, abundant, have high information content (hypervariable), are reproducible and multi-allelic (Powell et al., 1996). It has been shown that microsatellites longer than 20 bp occur every 33kb in plants and every 6kb in mammals (Morgante and Oliviery, 1993). The most common dinucleotide motif in the plant genome is the AT repeat (Powell et al., 1996). Microsatellites can be developed by three methods: first by searching in the sequence databases such as: GenBank, EMBL and GDR; secondly accidentally through the sequencing of some fragments of DNA for gene discovery; thirdly by hybridization of repeat oligonucleotide probes with small insert genomic or cDNA libraries. When repeat

motifs are identified by one of these methods two unique primers for flanking regions of repeats are designed to allow amplification of the motifs through PCR (Valdes et al., 1993; Schubert and Muller-Starck, 2002).

Microsatellites can be categorised into four types: 1. Perfect microsatellite: microsatellite without mutation and interruption anywhere that does not match the repeat pattern, e.g. CTCTCTCTCTCTCTCTCTCTCTCT. 2. Imperfect microsatellite: microsatellite with point mutation that does not fit the repeat structure, e.g. CTCTCTCTCTCTCTCTCACTCTCTCT. 3. Interrupted microsatellite: microsatellite with the small replacement of base pairs that do not fit the repeat structure, e.g.: CTCTCTCTCTCTCTGGGCTCTCTCTCCT. 4. Compound microsatellite: consisting of two or more different tandem repeats next to each other, e.g.: CTCTCTCTCT CTCTGTGTGTGTGTGTGTGTGTGT (Hancock, 2001).

Microsatellite markers have been utilised extensively for many genetic studies in *Prunus* (Cipriani et al., 1999; Downey and Iezzoni, 2000; Testolin et al., 2000; Cantini et al., 2001; Dettori et al., 2001; Jáuregui et al., 2001; Hormaza, 2002; Serrano et al., 2002; Wang et al., 2002a,b; Aranzana et al., 2003; Dirlewanger et al., 2004a; Messina et al., 2004; Mnejja et al., 2004; Mnejja et al., 2005; Pederse, 2006; Vilanova et al., 2006; Xie et al., 2006; Blenda et al., 2007; Sánchez-Pérez et al., 2007b). Microsatellites are usually located in non-coding regions of DNA in which many mutations occur. As a result, the compositions of these sequences are diverse, which allows them to be employed as a highly discriminative system for DNA fingerprinting, in forensics, paternity testing and cultivar identification.

Sosinski et al., (2000) used microsatellites successfully to identify 28 scion cultivars in peach germplasm which exhibited 80% polymorphism level compared to

25% for RFLP probes. Transportability of microsatellites was explored within *Prunus* species such as peach, sweet and sour cherry, European and Myrobolan plum and almond, and within species from other genera of Rosaceae *Malus* (apple) and *Fragaria* (strawberry), and species not belonging to the Rosaceae: C*astanea* (chestnut), *Juglans* (walnut) and *Vitis* (grapevine). Results showed that from forty one markers 75.6% amplified all the *Prunus* species and six microsatellites amplified all the tested species. Among them one microsatellite was placed on a region where a protein was found homologous to the sequences encoding a MADS-box protein in *Malus* × *domestica* (Dirlewanger et al., 2002). Struss et al., (2003) reported using fifteen microsatellite markers for identification of cultivars in sweet cherry and other *Prunus* species. It was suggested that these markers could be employed successfully for Plant Breeders Rights.

Linkage and comparative mapping have been facilitated by the advantages of microsatellites such as: high rate of polymorphism, even distribution throughout the genome, ability to tag distinct regions of genomes and co-dominant mode of inheritance (Rajapakse et al., 1995). In addition, microsatellite markers were used to anchor BAC clones to assist physical mapping in wheat (Mohan et al., 2007) and peach (Zhebentyayeva et al., 2008). The major drawback for using these markers in mapping studies is time and expensive cost for initial isolation and characterization of microsatellite loci when a pre-existing DNA sequence is not available. However, once the locus specific primers become available, the approach becomes efficient and cost effective (Hancock, 2001). Generation of a saturated molecular genetic linkage map just with microsatellites seems to be idealistic, therefore in many studies different kinds of dominant and co-dominant markers have been used.

Beliss et al., (2002) described using additional RFLP, isozyme, resistance genes, RAPDs, cleaved amplified polymorphic sequence (CAPs), and microsatellite markers to expand the linkage map resulting from an almond ('Padre') × dwarf peach (54P455) F2 population. A set of 109 microsatellite primers developed from peach and cherry were used to enrich the previously constructed *Prunus* reference map mostly with RFLPs (Aranzana et al., 2003). The resulting map consisted of eight linkage groups and increased the density and coverage of the genome with transferable polymorphic markers as an anchor map (Aranzana et al., 2003). To improve the coverage of a second generation peach BC_1 progeny, a set of 146 microsatellite and AFLPs was employed; consequently, the addition of more markers resulted in saturation of the map and targeted desired places on the resultant map (Verde et al., 2005).

The choice of molecular marker depends on the purpose of application, convenience and cost involved. Many studies have confirmed high cross transportability of microsatellites among *Prunus* species. Flanking regions of microsatellites are believed to be relatively conserved among related species and thus the expense and time consuming process of developing new microsatellites can be avoided. Considering all aspects, it seems microsatellites are the best choice for this current study. A set of 195 microsatellites was tested to select markers suitable for amplification on parents and progeny of an F_1 almond hybrid population, to saturate the previous linkage map of almond constructed by Gregory et al., (2005).

2.2 Material and Methods

2.2.1 Survey of published and unpublished microsatellites

In total, 195 microsatellite primer pairs isolated previously from bacterial artificial chromosome (BAC) or enriched genomic libraries of several *Prunus* species were screened for polymorphism. Designation, number of microsatellites tested and their origin are presented in Table 2.1. The same name is adopted as in the original papers, for example, the terminology used was as in Mnejja et al., (2004) for CPSCT with a five-letter code (C for Cabrils, PS for *P. salicina*, and CT for the microsatellite repeat selected followed by a three digit number). The names and sequences of the primers, along with annealing temperatures are presented in Appendix A. To saturate the almond linkage map of NP × LA which was initially produced by Gregory et al., (2005), microsatellite primers from *Prunus* species were selected, by looking at the published literature or by personal communication. Markers were chosen according to their genetic informativeness and polymorphism results from previous studies.

2.2.2 Genomic DNA extractions

2.2.2.1 Introduction

The preparation of highly pure DNA is essential for DNA-based molecular techniques. In plant species the extraction of high quality DNA is hampered by secondary metabolites including carbohydrates, proteins, RNA and polyphenolic compounds. Although there are a number of protocols that use different components to isolate DNA from plant tissues, generally all of them follow the same basics. Plant tissue is first

Table 2.1 Microsatellites: designation, number of screened markers and origin

Microsatellite name	Species	No. of Markers	References
UDA	Almond cv. 'Ferragnès'	41	Testoline et al., 2004 and Pers Comm. Dr Testoline, 2005
CPDCT	Almond cv. 'Texas'	8	Mnejja et al., 2005
EPDCU	Almond	8	Dirlewanger et al., 2004a
BPPCT	Peach cv. 'Merrill O'Henery'	28	Dirlewanger et al., 2002
CPPCT	Peach cv. 'Merrill O'Henery	23	Aranzana et al., 2002
UDP	Peach cv. 'Redhaven'	9	Cipriani et al., 1999
Pchgms	Peach cv. 'Nemared'	4	Wang et al., 2002b
Pchgms	peach cv. 'Bicentennial'	2	Sosinski et al., 2000
Pchcms	peach cv. 'Bicentennial'	3	Sosinski et al., 2000
Pchgms6	Peach	1	Pers Comm. Dr Arús, (2006)
AP2M	Peach cv. 'Loring'	1	Pers Comm. Dr Arús, (2006)
MA020	Peach cv. 'Akatsuki'	1	Yamamoto et al., 2002
UDAp	Apricot cv. 'Portici'	28	Messina et al., (2004) Pers Comm Dr Testoline, 2005
PacC	Apricot cv. 'Stark Early Orange'	2	Decroocq et al., (2003)
CPSCT	Japanese plum cv. 'Santa Rosa'	17	Mnenjja et al., 2004
EMPA	sweet cherry cv. 'Napoleon'	3	Clark and Tobutt, 2003
UCD-CH	sweet cherry cv. 'Valerij schkalov'	8	Struss et al., 2003
Ps08e8	Cherry	1	Sosinski et al., 2000
PceGA	sour cherry cv. 'Erdi Botermo'	3	Downey and Iezzoni, 2000
PS5C3	Cherry	1	Sosinski et al., 2000
PMS	Cherry	3	Pers. Comm. Dr Arús, (2006)

ground and homogenised in the presence of extraction buffer, then the DNA is separated from other cellular components by centrifugation with a combination of appropriate chemicals and enzymes. A number of extraction protocols use CTAB (cetyltrimethylammonium bromide) following the procedure reported by Murray and Thompson (1980) and Doyle and Doyle (1991). CTAB is a non-ionic detergent that is major component of the initial extraction buffer. It is used to bind to the DNA and prevent the binding of DNA to proteins while limiting DNA degradation (Gregory, 2004). Other chemicals present in most extraction buffers include PVP (polyvinyl-pyrrolidone), β-mercaptoethanol and DETC (diethylidithiocarbamic acid), utilised to

prevent oxidation of the DNA by polyphenolic compounds. For separating cellular debris from suspended nucleic acids, centrifugation at high speed is used, then to remove contamination by secondary compounds a high concentration salt solution is employed. RNA may be precipitated or digested with RNase or lithium chloride. Finally the DNA pellet is precipitated using either isopropanol or ethanol and resuspended in water or buffer (Tris-HCl, EDTA).

2.2.2.2 Extraction method

Former PhD student Davina Gregory extracted DNA of the mapping population as explained here and kept it concentrated at –20ºC, this DNA was used for PCR amplification in the present study. Healthy young leaves of the F_1 mapping population were collected in spring 2000, wrapped in aluminium foil and kept under ice before transfer to -80ºC. To achieve the best protocol for genomic DNA isolation from *Prunus dulcis* leaves with respect to DNA yield and quality the following methods were compared: Alijanabi and Martinez (1997); Messeguer et al. (1994); Steenkamp et al. (1994); and modified Lamboy and Alpha (1998). Assessment of the DNA content and quality was determined by ultraviolet irradiation spectrometry, in which the level of ultraviolet absorbance is used as an indication of both yield and purity. The method A, modified Lamboy and Alpha (1998), was the most time consuming, but produced the best results with respect to yield and absence of contamination (Gregory, 2004). Brief details of this method are provided below to explain the process by which genomic DNA was prepared in the earlier work of Gregory (2004).

2.2.2.3 Modified Lamboy and Alpha (1998) protocol macrotechnique

This technique was very similar to the Lamboy and Alpha (1998) method except that the amount of leaf tissue was increased ten times. Midrib, petiole and stem were removed from 1.5 to 2 g of leaf tissue, then placed in a pre-chilled mortar (-20ºC) and ground to a fine powder covered with liquid nitrogen. The fine powder was transferred to a sterile 50 mL tube on ice containing 10 mL of extraction buffer [250 mM Tris-HCl (pH 8.0) 250 mM NaCl, 50 mM EDTA, 4.0% PVP-40 (w/v), 1.5 % SDS (w/v)] with 50 mM DETC, 100 mM sodium ascorbate and 1% β-mercaptoethanol added quickly just before use and mixed thoroughly. The tissue-buffer mixture was held on ice until all samples were prepared. Nucleic acid extraction was initiated by incubation in a water bath at 37ºC for 30 min with gentle inversion once every 10 min. Then samples were taken from the water bath and an equal volume (12 ml) of cold (4ºC) chloroform: isoamyl alcohol [24:1 (v/v)] was added, the tube mixed gently for five min on a spinning wheel and then centrifuged at 11,000 rpm for 15 min at 4ºC. The upper aqueous layer was transferred to a clean sterile tube and a 0.7 volume of 5 M NaCl was added and mixed thoroughly. Then the sample was centrifuged at 11,000 rpm for 30 min at 4ºC. The supernatant was decanted into a new tube, leaving the contaminant pellet. A 0.66 volume of cold isopropanol (-20ºC) was added with gentle inversion and placed at 20ºC for 30 min to precipitate the nucleic acids. DNA was recovered by centrifuging at 11,000 rpm for 15 min at 4ºC. The upper aqueous layer was gently decanted and the pellet containing the DNA washed with 10 mL of cold wash buffer [76% ethanol, 10 mM NH_4Ac] overnight at 4ºC, or until white colour was visible. The wash buffer was discarded and the pellet rinsed with cold 70% ethanol (-20ºC) for 5 min to remove excess salts. The pellet was allowed to air dry at room temperature for 10 min and transferred to a sterile 1.5 ml Eppendorf tube and dissolved in 600 µl of TE

buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8). Three µl heat-treated RNAse A (10 mg/mL) was added to each sample and incubated at 37ºC for 30 min to digest contaminating RNA. After incubation, to precipitate proteins and RNAse, 240 µl 7.5 M NH_4Ac was added and placed at 4ºC for 30 min followed by centrifugation at 11,000 rpm for 15 min at 4ºC. Supernatant containing DNA was carefully removed with a 1 mL micropipette and transferred equally to two sterile Eppendorf tubes. DNA was precipitated by adding 2 volumes of cold absolute ethanol (-20ºC) to each tube and allowed to stand for 30 min at -20ºC; the solution was gently inverted to concentrate the DNA. DNA was recovered by centrifuging at 11,500 rpm for 15 min at 4ºC, the ethanol discarded and the white pellet washed with 400 µl cold wash buffer (-20ºC) for 15 min. The pellet was washed with 200 µl of cold 70% ethanol (-20ºC) for 5 min. The pellet was left to air dry at room temperature, the DNA dissolved in 200 µl of TE buffer and the corresponding samples recombined to give 400 µl in total (Lamboy and Alpha, 1998). Comparison of DNA extracted by modified Lamboy and Alpha (1998) showed that there was no sign of DNA degradation. All in all, the technique of upscaled Lamboy and Alpha (1998) produced the best high quality intact DNA and was used to extract DNA from the mapping population (Gregory, 2004).

2.2.3 DNA quantification and PCR amplification

Before any application of the mapping population DNA, which was extracted by Gregory et al., (2005), all the DNA samples were quantified using a Nanodrop spectrophotometer ND-1000® at the absorption ratio A_{260}/A_{280} for quantification and quality control, then diluted to a final stock concentration of 20 $ng.\mu l^{-1}$ with TE buffer (10mM Tris, 1 mM EDTA pH 8.0) and kept at -20°C for subsequent PCR amplification.

40 ng was used as template for PCR. Standard and 6-FAM and HEX fluorescent primers were synthesised by Sigma-Proligo® and NED fluorescent primers by Applied Biosystem®. Primers were diluted to a 20 µM working solution for a 1µM dilution in the reaction mixture.

The PCR was performed containing a final concentration of 1 × PCR reaction buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTP mix, 0.2 µM of each primer, 40 ng of template DNA and 1 unit of *Taq* polymerase (Bioline®) and sterile MQ water to a total volume of 20 µl. PCR amplification was performed in a thermocycler manufactured by Bio-Rad My Cycler® Thermal cycler and PTC-100® Programmable Thermal controller, MJ Research, Inc. In general, the PCR protocol was 95°C for 5 min, 34 cycles of: 30 seconds at 95°C, 30 seconds at primer annealing temperature, 30 seconds at 72°C, 7 min at 72°C which acts as a final extension step. The PCR was optimised on a small number of DNA samples across a range of annealing temperatures. All PCR plates contained a negative control with no template DNA. PCR product was held at 4°C before analysis.

2.2.4 Segregation analysis and polymorphism detection

2.2.4.1 Fluorescent labelled microsatellite detection

For automated microsatellite genotyping, the forward primers were labelled with 6-carboxy-fluorescine (FAM), hexacholoro-6-flourescine (HEX), or NED. Up to 5 fluorescently labelled (HEX, FAM, or NED) markers typically can be used for multiplexing. The names of nine forward primers that were labelled at the 5' end with a fluorescent label are shown in Table 2.2. To optimize PCR concentrations for the DNA sequencer, initially, equal concentrations for all primer pairs were used, then

adjustments were carried out until all the peak heights were relatively equal. The dilution factor should be adjusted to the instrument sensitivity. Due to differences in signal intensity of the three dyes, FAM-labelled markers were diluted by adding 48 µl MQ sterile water to 2µl PCR products, NED and HEX-labelled markers were diluted by adding 45µl MQ sterile water to 1 µl PCR products prior to electrophoresis.

Table 2.2 Microsatellite name and fluorescine label

Microsatellite Primers	Fluorescine label
UDA-002 UDA-009 UDA-014 BPPCT013	FAM
UDA-005 UDA-012 UDA-013	NED
UDA-008 UDA-018	HEX

Only markers whose products do not overlap in size (regardless of which dye they contain) were mixed together. The ABI Semi-Automated DNA sequencer 3730® consists of a capillary electrophoresis instrument and software for data collection and genotyping. Data were collected and analysed using GeneMapper 3.7® (Applied Biosystems, USA) or Peak Scanner 1.0 and manual editing of allele size was carried out for verification and accuracy.

2.2.4.2 Agarose gel electrophoresis

Microsatellite markers were tested on parents and some selected progeny of the mapping population on 3% agarose gels (LE, Promega, Australia) in 1x TBE buffer (~5V/cm). 1 X TBE contains 10.8 g of Tris base, 5.5 g of boric acid and 4 ml of 0.5 M EDTA (pH 8.0) per litre (ref). TBE was made and stored as a 5 x stock solution and

diluted before use, both gel solution and the electrophoresis buffer were made from the same concentrated stock solution (Sambrook and Russell, 2001). 8 µl of PCR products was added in a well on the gel together with 2 µl of loading buffer (50% (v/v) glycerol, 0.075 M EDTA, 0.2% (W/V) SDS, 0.1% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol FF). Gels were run at 180V for 2.5 hours, by which time the bromophenol blue dye had migrated approximately 2/3 of length of the gel. Electrophoresis with Metaphor® agarose (FMC BioProducts Inc.,USA) gels was used also to separate alleles of microsatellite markers. Following electrophoresis, PCR products were stained with ethidium bromide and visualised under UV light before image capture using a Bio-Rad ChemiDoc®XRS system (Bio-RAD, California).

2.2.4.3 PAGE gel electrophoresis

Five µl of FLB (formamide loading buffer: 95% formamide, 20 µM EDTA, 0.05% Bromophenol blue and 0.05% Xylene cyanol) was added to each tube of PCR product and denatured for 5 min at 95°C before being immediately chilled on ice. Ready to mix polyacrylamide gel (60 mL) (SequaGel® 8, National Diagnostics, USA) consisted of 8% polyacrylamide, 7 M Urea, TEMED and Bis-acrylamide (19:1, w/w) plus 15 mL TBE and 600 µL of fresly prepared 10% w/v ammonium persulphate was mixed completely in a total volume of 75mL for casting the polyacrylamide gel. To detect the DNA fragments, 6 µl of the PCR product-dye mixture was loaded on an 8% polyacrylamide gel at the thickness of 1.5 mm and the gel was run at a constant power of 200 V for 3 hours. DMW- pUC19/HpaII DNA Ladder (GeneWorks®) was used as standard. The gel was stained with ethidium bromide solution (0.5 µg/ml) for 20 min and destained in MQ water and the image captured by the Bio-Rad ChemiDoc®XRS

system. Although reproducibility and reliability of microsatellite amplification are well documented, occasionally, center parts of the gels become hotter than the sides and produced faint bands. Therefore, to ensure reproducibility, all PCR reactions and gel runs were conducted twice or more. All the microsatellite data was also scored indipendantly by Dr Shubia Wu, another member of almond breeding project.

2.3 Results

Screening of the 195 microsatellites produced 161 amplified, and 107 polymorphic markers. Of the polymorphic markers, almost half produced very complicated or faint bands not suitable for reliable scoring therefore, 54 markers that produced reproducible and easily scorable fragments were selected for further screening on the whole mapping population. Remarkable differences in the rate of polymorphism were observed according to the origin of the markers. One marker UDAp-479 was multiallelic and detected two loci. A summary of results of microsatellite primers used in this study are described in Table 2.3.

2.3.1 Microsatellites isolated from almond

Amplification was successful in parents and progeny for most markers initially developed in almond. Of the 57 almond microsatellites, 52 amplified (91.2%) and of these 32 were polymorphic (56.1%). Among the almond isolated microsatellites UDA set gave the highest rate of polymorphism and scorable loci. Loci UDA-(005, 008, 009, 012, 013, 014, 022, 025, 035, 043, 046, 051 and 053) showed clear gel profile patterns suitable to be used for scoring on the entire population. Alleles appeared in the range of 100 bp to 300 bp. Marker UDA-043 exhibited a deletion in one progeny. None of the almond ESTs detected polymorphism and only marker CPDCT032 was used for

screening on the entire population. Optimization carried out on almond microsatellite primers did not dramatically change polymorphism detection of CPDCT marker series.

Table 2.3 Summary of microsatellite marker analysis results

Designation	Species	No. tested	Amplification rate (%)	Polymorphism rate (%)	References
UDA	Almond	41	(38) 92.7%	(25) 61%	Pers. Comm and Testolin et al., 2004
CPDCT	Almond	8	(7) 87.5%	(5) 62.5%	Mnejja et al., 2005
EPDC	Almond	8	(7) 87.5%	0	Pers. Comm. Dr Arús 2006
BPPCT	Peach	28	(22) 78.6%	(16) 57.1%	Dirlewanger et al., 2002
Pchcms	Peach	3	(2) 66.7%	(2) 66.6%	Sosinski et al., 2000
Pchgms	Peach	7	(5) 71.4%	(4) 57.1%	Sosinski et al., 2000 Wang et al., 2002b
AP2M	Peach	1	(1) 100%	0	Pers. Comm. Dr Arús, 2006
MA020	Peach	1	(1) 100%	(1) 100%	Yamamoto et al., 2002
PMS	Peach	3	(1) 33.3%	(1) 33.3%	Cantini et al., 2001
UDP	Peach	9	(8) 88.9%	(5) 55.6%	Cipriani et al., 1999
CPPCT	Peach	23	(15) 65.21%	(12) 44.4%	Aranzana et al., 2002
CPSCT	Japanese plum	17	(11) 64.7%	(8) 47%	Mnenjja et al., 2004
UDAp	Apricot	28	(22) 78.6%	(13) 46.4%	Messina et al., 2004 Pers. Comm. Testolin
Pac	Apricot	2	(2) 100%	0	Decroocq et al., 2003
EMPA	Sweet cherry	3	3 (100%)	(2) 66.7%	Clark and Tobutt, 2003
PS5C3	Sweet cherry	1	0	0	Sosinski et al., 2000
UCD-CH	Sweet cherry	8	(5) 62.5%	(2) 25%	Struss et al., 2003
PceGA	Sour cherry	3	(2) 66.7%	(2) 66.7%	Downey and Iezzoni, 2000
Ps08e8	Cherry	1	(1) 100%	(1) 100%	Sosinski et al., 2000

2.3.2 Microsatellites isolated from peach

Microsatellites isolated from peach showed the highest available primer information. In total, 75 peach microsatellites were tested. 54 (72%) gave amplification and 41 (54.7%) showed polymorphic pattern. BPPCT and CPPCT series showed higher polymorphism rate and readable profiles. Out of 41 polymorphic peach microsatellite 24 primer pairs including: BPPCT- (001, 002, 004, 010, 013, 027, 029, 037, 040); CPPCT- (003, 006, 008, 009, 021, 023, 033 and 034); UDP (97-403, 98-024, 98-409); Pchgms- (6 and 29); Pchcms5 and MA020 were used for scoring.

2.3.3 Microsatellites isolated from other *Prunus* species

Among the 17 microsatellites isolated from Japanese plum, almost half of them revealed polymorphic patterns but only three of these CPSCT- (033, 042 and 044) had a single locus and scorable gel profile. Of the 30 microsatellites isolated in apricot 12 (40%) primer pairs gave polymorphism. One marker (UDAp-479) detected two loci and the other markers were single locus. Loci UDAp- (407, 418, 461, 463, 473, 479, 483 and 493) that showed the most clear and simple marker pattern were chosen for scoring. Around half (8) of the cherry primers amplify in the mapping population and four loci: EMPA004; UCD-CH (15, 19) and PceGA34 displayed scorable patterns. Fig 2.1 shows a profile of polymorphism detected by microsatellite EMPA004 in some progeny of the segregating mapping population (NP × LA) which showed null allele. Sometimes due to mutation at primer binding site or differential amplification (short allele amplify more efficiently during the PCR), an allele can not be detected which is called null allele. Although the microsatellites originated from other *Prunus* species showed high levels of polymorphism, the segregating marker pattern sometimes was very complex or faint

and not suitable for scoring. In addition to desired microsatellite amplifications in the expected size range, fragments with higher molecular weight or faint minor bands and heteroduplex bands were also observed. Three markers (BPPCT040, CPSCT044 and EMPA004) showed null alleles. "Microsatellite null allele is any allele that could not be amplified to the detected level via PCR" (Dakin and Avise, 2004). Null alleles made unexpected genotypic patterns and sometimes resulted in difficulty of marker scoring.

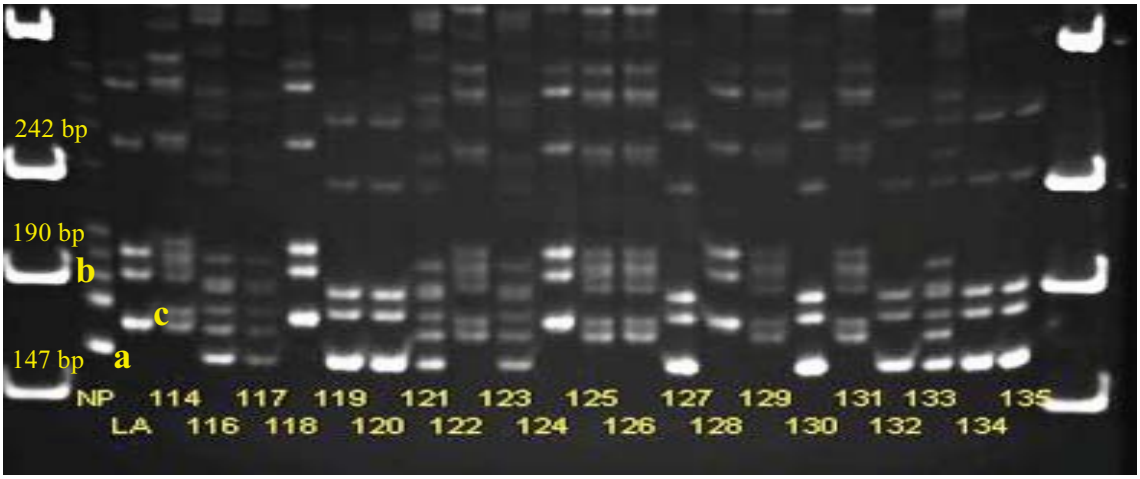


Fig. 2.1 Microsatellite polymorphism in some progeny amplified by marker EMPA004

The majority of scorable microsatellites with reported repeat sequences were perfect except for loci CPSCT044, EMPA004, UDA-005, UDA-009, UDA-013, UDA-022 and UDAp-407 which had interrupted or compound repeat motifs. In fluorescent labelled markers the highest peak which gave most reliable scoring was selected. Dilution of the PCR product before loading onto the instrument reduced the number of 'pull-up' peaks caused by the other fluorescent labelled primers, when loaded in the same lane. Fig 2.2 shows the profiles of polymorphism detected by fluorescent labelled microsatellite BPPCT013. From nine fluorescent labelled primer pairs five loci showed scorable polymorphism. The time needed for microsatellite analysis was reduced by using fluorescent labelled primers. Microsatellite loci isolated in other *Prunus* species

gave correct PCR amplification close to the reported size in the mapping population. As expected, the number of alleles resolved on PAGE gels was higher than Metaphor® or standard agarose gels. However, for the fragments with a difference more than 5 bp Metaphor® agarose was used successfully.

2.4 Discussion

De novo discovery of microsatellite markers for a small scale mapping population is unreasonable due to initial expensive cost, time consumption and technical difficulty. To avoid these problems the available primer sequences from *Prunus* species were used for polymorphism detection. Results of the study showed a high degree of cross-transportability of microsatellites from different *Prunus* species. More than 76% of markers amplified and half of them (50.26%) showed polymorphism. This is in agreement with the previous reports by Testolin et al., (2000), Verde et al., (2005) and Sánchez-Pérez et al., (2007). The results confirmed microsatellites as a powerful tool for synteny studies in *Prunus*. For the *Prunus* derived microsatellites the smallest percentage of markers that did not produce any products was obtained from almond (8.8%) followed by apricot (21%), peach (26%), Japanese plum (35.5%) and half of the cherry markers. Some of these markers have never been used on almond before. The highest polymorphic and scorable marker was expected for the UDA primer set since it had been developed from almond and for the first time was tested on this population. All the apricot microsatellite primers were for the first time tested in this mapping population. While some modifications such as increasing the DNA template or decreasing the annealing temperature increased the level of polymorphism for most microsatellites already tested, a few of them still could not be amplified.

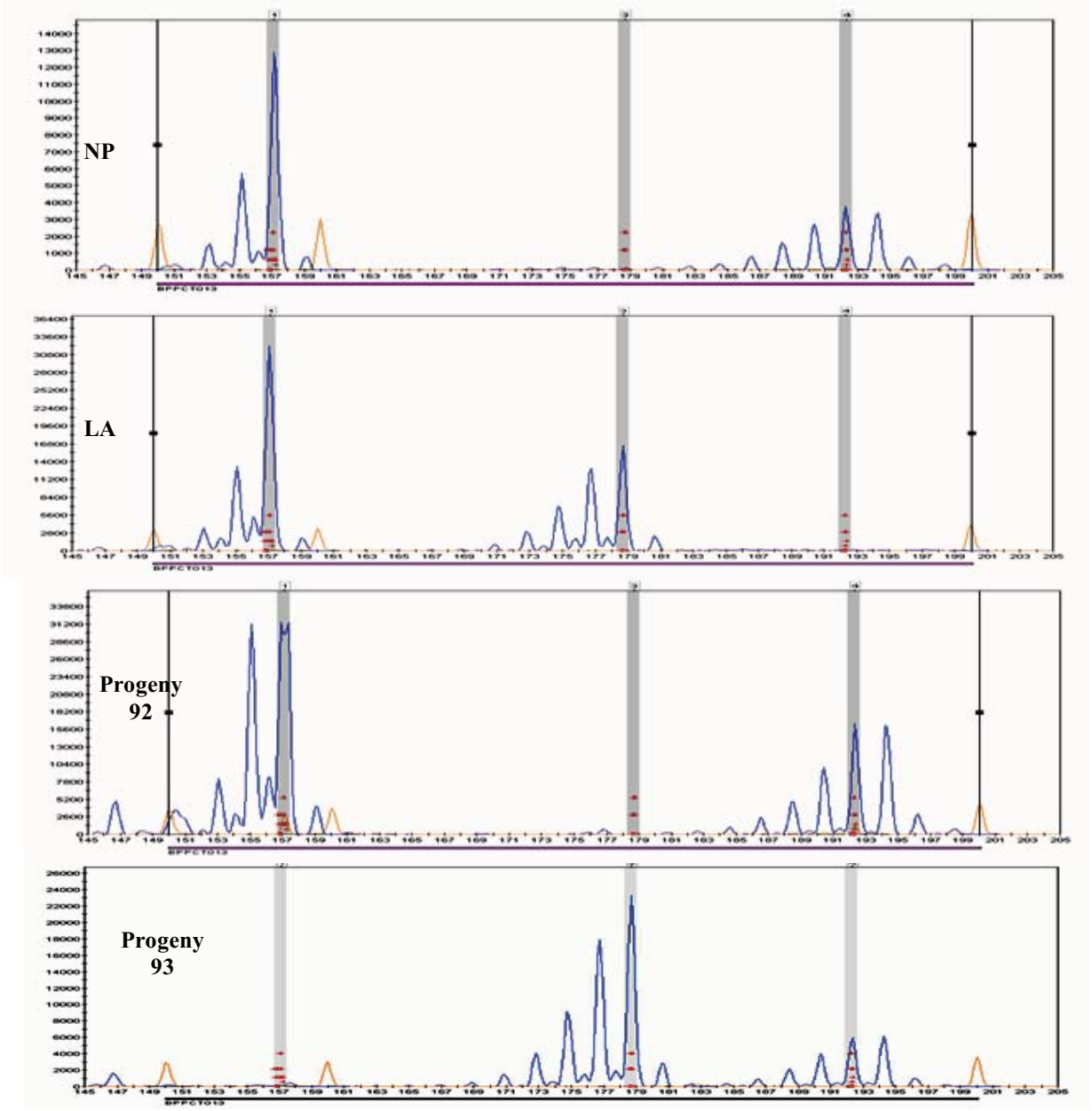


Fig. 2.2 Electrophorogramme of fluorescent labelled alleles detected by loci BPPCT013

These results suggest a strong homology of the genomes between these species and probably between Prunophora and Amygdalus sub-genera. Peach and almond are two species of the genus *Prunus* subgenus Amygdalus subfamily of Prunoideae (Martínez-Gómez et al., 2003a). A decrease in the amount of polymorphic microsatellite markers would be expected as genetic distance increases from the species (Martínez-Gómez et

al., 2003a). Scorable microsatellites were lower in this study compared with reports for some *Prunus* species which indicate that the value of each primer is species and cultivar specific.

The levels of transportability and polymorphism observed here were lower than in the study by Aranzana et al., (2003) which found 80% polymorphism in a study of the F2 almond × peach progeny. This is predictable because the Aranzana et al., (2003) map is based on an interspecies cross, thus more likely to reveal more polymorphism compared with our study, using an intraspecific cross. However, polymorphisms obtained here are higher than for some other plants. For example, 26% of the barley microsatellite primers showed amplification in oat (Li et al., 2000) and only 10% of soybean microsatellites provide useful markers for cowpeas, broad beans or lupins (Peakall et al., 1998).

The range of amplified band sizes in almond was also similar to those reported in the literature using the same primer pairs. Weber and May (1989), reported that the most polymorphic, and therefore the most informative microsatellites are uninterrupted arrays, however, some polymorphic loci contained interruptions. This study revealed only a weak correlation between polymorphism and microsatellite repeat length which is in agreement with the finding by Ramsay et al., (2000) for barley but disagrees with the study by Struss et al., (2003) that could find no correlation between the number of microsatellite repeats and the number of alleles detected. Interspecific allelic differences during evolution might be more complex than simple changes in repeat number. Products amplified in different species might include mutation, rearrangements and duplications in the flanking region and/or changes in the number of repeats (Peakall et al., 1998).

It was recommended that to increase transportability of the marker among taxa, microsatellites isolated from cDNA or the EST library might be used, due to their flanking sequence location in exons which are more likely to be conserved (Winter et al., 2002). Although most of the microsatellites isolated from the EST library were amplified, none of them showed polymorphism. Microsatellites developed from a cDNA library are often reported to be less polymorphic as a result of repeat shortness and less mutation rate (Dendini et al., 2007).

Only three microsatellites displayed null alleles and one detected two loci in mapping population. These findings are similar to the observations of Aranzana et al., (2003); Lambert et al., (2004) and Verde et al., (2005) confirming the high level of codominance among microsatellites and their mostly single locus nature. Several reasons for potential causes of null alleles have been suggested. Each microsatellite primer is designed to amplify complementary sequence in a genomic or cDNA library from specific cultivar, therefore when a primer pair is used in the same or different species it may cause poor primer annealing because of point mutation or indels in one or both priming sites (Kwok et al., 1990). Furthermore, alleles with shorter length often amplify much better than larger ones, longer motifs can produce differential amplification called partial nulls that can be visualised by loading more sample or changing the photography conditions (Wattier et al., 1998). The quality and quantity of template DNA is another source of null alleles. In some cases microsatellite alleles amplify in some progeny but not others thus these poor samples seem to be homozygous for the null allele rather than heterozygous (Dakin and Avise, 2004).

During the PCR amplification of microsatellite loci, minor products that are 1-4 repeat units shorter than the main alleles are produced. This product is referred to as

"stutter" and is believed to be PCR artefacts. Stutter peaks might be caused by polymerase slippage during elongation. Stutter bands can be a severe problem in analysing heterozygous individuals with similar alleles. Dinucleotide microsatellites are highly prone to stuttering, trinucleotides less so, and the larger motif sizes stutter very little, if at all (Goldstein and Schlötterer, 2001).

Capillary electrophoresis also has been used to determine length polymorphism of microsatellite markers, but this method requires a sophisticated instrument and fluorescently tagged primers, which are expensive. Fluorescent labelling is the most efficient methods for visualizing microsatellite but they are expensive and need special facilities, the use of these fluorescent dye makes the high throughput approach rather expensive. For this reason, PAGE was preferable to the fluorescent-labelled detection method.

This work confirmed the possibility of cross-species transfer for several microsatellite markers and the value of markers developed in one species for study within the genus and also provides a cost effective method for construction of a consensus map. This marker type is extremely useful due to its ease of use and high amount of information generated. These findings showed that with a large number of microsatellites in common with other *Prunus* maps, microsatellites would enable researchers to achieve further synteny studies.

Chapter 3

Pomological trait analysis of an F_1 pseudo-testcross population of almond

3.1 Introduction

Almond was once considered as a separate genus, *Amygdalus* due to its morphological differences compared with other *Prunus* species (Srinivasan, 2005). Almond is the first fruit tree to bloom because of its lowest winter chilling requirement among deciduous fruit tree (Kester and Gradziel, 1996). Almond growth cycle follows its original Mediterranean or desert climate habitat, where plants are dormant during the winter and blooming and when temperatures become mild vegetative growth begins in late winter or early spring (Kester and Gradziel, 1996). During the early blooming period frost has an extremely negative impact on yield. Rain in early spring also interferes with flower pollination by reducing visits from the main pollinator vector in almond, the honeybee (*Apis mellifera*) (Jackson, 1996). Consequently, control of

blooming time has been an important objective of most almond breeding programs to avoid frost injury and crop reduction (Asíns et al., 1994; Ballester et al., 2001; Silva et al., 2005). Genetic diversity of almond is very high and there is wide variability for many traits within the species. The differences observed are much higher than in other closely related species e.g. peach. This heterogeneity probably results from the ancient origin of almond, the propagation method, growth habit in many diverse areas and local adaptations to different microclimates (Kester and Gradziel, 1996).

Most cultivars of almond have gametophytic self-incompatibility, which makes them obligately outcrossing and in consequence highly heterozygous (Ballester et al., 1998). This wide genetic variation has provided a very valuable genetic pool for selective breeding purposes. To improve the efficiency of breeding programs, an understanding of the genetic basis, heritability and correlation between the morphological traits and environmental effects is necessary.

When the soft layer of mesocarp is eaten it is called a fruit and when a seed is the edible part it is called a nut. The term nut refers to the kernel in the shell, but kernel refers to produce free of the shell. One of the main objectives in almond breeding is to identify and develop cultivars with superior nut and kernel quality; therefore many studies have concentrated on these traits. In breeding strategies not only the physical aspects of nut and kernel quality, but also organoleptic factors must be considered. Almond nuts of different cultivars vary widely in a number of nut and kernel traits such as shell hardness, kernel size, shape, taste, pubescence, colour, and percentage of doubles.

3.1.1 Shell hardness

Shell hardness is associated with the lignification of endocarp during nut development. It is categorised into five classes: stone (very hard), hard, semi-soft / semi-hard, soft, and paper. Shell hardness is expressed as the weight of kernel divided by the weight of in-shell nut, multiplied by 100 (Kester and Gradziel, 1996). Grasselly (1972) studied a few cultivars and suggested that this trait is controlled by a single gene with hard-shell *D* dominant over soft-shell *d*. However, this finding was not confirmed by the data of another research group, who considered it as a quantitative character (Socias i Company, 1998).

There is a wide range of shell hardness in cultivars that are adapted and well suited for different applications and consumers. In general, most hard-shelled cultivars are grown in the Mediterranean region, due to better adaptation to non-irrigated culture, resistance to birds and some pests, and better storing ability because of the slower rate at which they become rancid (Socias i Company, 1998). In contrast, in California and Australia soft-shelled cultivars are preferred (Socias i Company, 1998). Paper shells have a soft pithy shell that can be easily removed and soft-shells have firm shells that can be broken by hand pressure. Paper and soft-shell cultivars are susceptible to damage by pests such as Navel orangeworm, *Paramyelois transitella* (Soderstrom, 1977), Carob moth, *Ectomyelois ceratoniae* (Pound and Collins, 2003) and birds. Bird damage to paper and soft-shell nuts is one of the biggest problems in almond orchards (Fig. 3.1). Almonds are very attractive for birds and sometimes become their major dietary component (Gathercole**,** 2002). The shell consists of outer and inner layers which are connected by channels of vascular tissues and protect the edible part (kernel). If the outer and inner layers remain attached to the nut it makes a heavier shell but often in

paper shell the outer layer separates from the inner layer which produces nuts with poor seal (Kester and Gradziel, 1996; Godini, 2002).

Fig. 3.1 Birds are very problematic, mostly in paper / soft-shell cultivars

3.1.2 Kernel size

Since the kernel is the commercial part of the tree the shape and weight are very important both for consumer and industry. The kernel shape is a function of kernel length, width and thickness; these traits generally show high genetic heritability and correlations. Production intensity and environmental conditions such as soil moisture and temperature during the developmental phase of the nut have strong influences on average size of the kernel (Kester and Gradziel, 1996). Industry needs different sizes of kernel for manufacturing, processing and marketing operations. Kernel size is graded by differences in width; the kernels are passed through a graduated series of round hole

screens of given diameter, each screening size is subsequently converted to a weight classification (Kester, 1965). Nut weight shows variability between individual nuts of the same tree and same cultivar in different locations. However, the inheritance of nut weight was found to be high (r = 0.89) (Chandrababu and Sharma, 1999) which suggests a high probability of genetic components controlling the trait. In-shell weight in hard and very hard-shelled almonds is not a major trait for the breeder because it mostly related to the shell and not to the kernel (Sánchez-Pérez et al., 2007a).

3.1.3 Testa colour and smoothness

Development of the outer and inner integuments of the ovule produces the seedcoat or testa. Testa cells enlarge significantly during the first a few weeks after flowering and became filled with darkly staining contents. The testa provides the developing kernel with metabolites, with many vascular tissues connecting the pericarp to the nucellus, endosperm and the embryo, and also protects it from damage. Over the ripening period it forms a dry hairy skin over the embryo and changes colour from white to brown (Hawker and Buttrose, 1980). Testa is a rich source of the antioxidant 2, 2-diphenyl-1-picrylhydrazyl, (Sang et al., 2002), which has the ability to capture free radicals. Testa colour is among the most important features in determining consumer preference to distinguish the quality of kernel. Kernel testa shows a variety of different colours (Fig 3.2. this figure does not represent colour variability of the NP x LA mapping population) and is considered to be under genetic control (Kester and Asay, 1975).

Fig. 3.2 Variation of testa colour in almond

3.1.4 Kernel taste

Kernel sweetness is considered as a major trait in almond breeding programs. Sweet kernels were selected for domestication among the wild populations, which usually have a bitter kernel (Kester et al., 1990). Heppner (1923) suggested that sweet kernel is a result of a mutation of an originally bitter almond. Although regular selection decreased the frequency of alleles responsible for the bitter flavour, some cultivars still carry these alleles and as a consequence, if they cross with each other can produce individuals with bitter kernels. Almond kernels may be sweet, slightly bitter or bitter. Sweet kernels have an almond or nutty fragrance and taste and for slightly bitter kernels although there is no common definition, they have a marzipan-like taste (Wirthensohn et al., 2008). Generally sweet and slightly bitter kernels categorised in the same group in breeding process.

Dicenta and García, (1993) suggested that slightly bitter kernel must be heterozygous (*Ss*) for taste character and occasionally the recessive allele (*s*) produce a slightly bitter flavour. A monogenic character with Mendelian inheritance for bitter was

first proposed by Heppner (1923; 1926), with bitter as a recessive allele *ss* and sweet as *Ss* or *SS*. However, further studies have suggested a more complex genetic control. Spiegel-Roy and Kochba (1974) showed that three genes could be responsible in the expression of this trait. The bitter taste of the kernel is suggested to be due to the presence of high levels of cyanogenic diglucoside amygdalin, which is hydrolysed by a suite of β-glucosidase enzymes to produce benzaldehyde and hydrogen cyanide. Hydrogen cyanide is toxic and causes the bitter taste (Kester and Gradziel, 1996). The mean amount of amygdalin in bitter kernel was found 500-times more than the mean value of sweet kernels and 90 times higher than slightly bitter kernels. However, sometimes the amount of amygdalin in sweet kernel was greater than slightly bitter kernels (Wirthensohn et al., 2008). The association of amygdalin content (diglucoside) with bitter kernel taste is well studied, but few studies have been carried out on the relationship between amygdalin and slightly bitter kernel (Dicenta et al., 2002), and the role of amygdalin in slightly bitter kernels still remain unclear.

3.1.5 Percentage of doubles

There are two ovules in the flower of *Prunus* species. Usually, following pollination the secondary ovule can not develop and aborts, but if both ovules in each carpels of the almond flower develop and are fertilised, the production of double kernels is seen (Fig 3.9) (Pimienta and Polito, 1982). Almond cultivars producing a high rate of double kernels are not preferred commercially. They are less profitable for growers and receive a lower grade because of misshapen appearance. Therefore factors affecting the development and fertilisation of both ovules are important for industry (Gregory, 2004).

However, the double kernel phenotype can be useful for genetic studies. In some cases one of the double kernels grows weakly and remains dwarf, these kinds of seedlings showed a high percentage of mortality. Molecular studies have shown an aneuploid chromosome in the dwarf seedlings (Martínez-Gómez et al., 2002). The aneuploid chromosome could be used for genetic studies such as locating the specific genes on chromosome by comparison with normal almond karyotype, isolating and sequencing genes on distinctive chromosome and selective transfer of specific chromosome (Riley and Law, 1984). It seems that the percentage of double kernels is frequent in some cultivars and very low in others such as 'Nonpareil' (Spiegel-Roy and Kochba, 1974). In addition, environmental condition has an immense influence on the production of double kernels. Low heritability and high variability of double kernels were found in many studies (Spiegel-Roy and Kochba, 1974; Kester et al., 1977; Dicenta et al., 1993; Arteaga and Socias i Company, 2001). In fruit breeding programs, in which seed-to-seed selection takes a long time and is therefore an expensive undertaking, understanding the similarity of performance of progeny compared with their parents is very crucial. Most of the important nut characters are controlled by many genes and their expression is associated with the interaction of genotype and environment that is expressed as heritability. In turn, these characters can be assessed by the data produced from the population of a controlled cross. Many of the above-mentioned characters such as shell hardness, kernel size, shape, taste, pubescence, colour and percentage of doubles, were studied by a PhD student at the University of Adelaide, Davina Gregory during 2001-2003 (Gregory, 2004). It is assumed that the age of the tree, and its developmental stage, have a significant influence on certain traits that makes it difficult to accurately measure them. For the first time morphological traits was measured when the trees were

just three years old and the results of her study highlighted the existence of a lack of homogeneity in the juvenile phase of the tree. She observed some trees were not as mature as other trees in each of the years. For example, in 2001 one tree just had three nuts, all of which were doubles; hence that observation may be insufficient for precise and accurate evaluation of pomological traits. In almond, tree is considered mature after six year. The aim of this study was to investigate the genotype by environment interaction, inheritance and correlation of nine morphological traits for better knowledge of the factors controlling these characters. Moreover, this information would be useful for future incorporation into linkage map.

3.2 Materials and Methods

Morphological traits were measured for three consecutive years (2005-2007) on the total population, but for data analysis only trees (106 progeny) that survived for 3 years were included in the analysis. Meteorological data including monthly average rainfall and mean maximum and minimum air temperature was collected from the nearest weather station (Renmark station # 240480, South Australia) in 2004-2007 and presented in Table 3.6.

3.2.1 Plant material and pomological traits measurement

The mapping population utilised in this study was a double pseudo testcross population from the cross between two almond cultivars ‘Nonpareil’ × ‘Lauranne’, detailed descriptions of this cross are presented in Chapter 1. The population was maintained under standard cultural practices for irrigation and pest control without any pruning. All seedlings studied were grown on their own roots, closely planted (0.5 × 3

m spacing). Almonds are harvested by hand once their hull has opened and the green hull is immediately removed in order to prevent mould growth and the nuts were then left to dry at ambient temperature. More attention was paid to avoid mixing of nuts from neighbouring trees, because of the close spacing. A sample of 30 typical sound nuts was randomly selected for kernel measurements of each tree and 10 nuts were used for assessment. The number of double kernels was based on the 30 nuts. Nuts that showed obvious defects were removed prior to nut assessment. A digital calliper was used for all kernel size measurements (Mitutoyo, Tokyo, Japan). The method of measurement for characters was similar to the method used by Gregory (2004).

3.2.1.1 Shell hardness and kernel weight assessment

The average in-shell weights and kernel weights were obtained using ten randomly selected nuts (not including doubles). The shell (endocarp) was removed and the average weight from the ten kernels was calculated for kernel weight. A manual nut cracker was used for kernel examination and shell hardness measurements. A pan balance was used for weighing the kernel samples. Shell hardness (thickness) was calculated as the percentage of the in-shell to kernel proportions of the total nut weight by the following equation (Rugini, 1986), and scored into classes based on the percentage values as shown in Table 3.1:

Shell hardness (%) = individual kernel weight / whole nut weight × 100

3.2.1.2 Kernel size and shape

Kernel length was measured as the distance from the basal end to the kernel apex; kernel width was determined as the distance across the kernel at its widest point.

Kernel thickness was measured as the maximum distance from front to back between the thickest points for ten kernels. Kernel shape was calculated by the kernel width / length ratio of the mean of ten nuts, classes are represented in Table 3.1. For size measurements, all dimensions of the kernels were recorded.

3.2.1.3 Testa characteristics

Both testa colour and testa pubescence were scored subjectively and individually, using both parents as comparisons. Testa colour was scored as light 'Nonpareil' and dark 'Lauranne' for years 2005-2007 using RHS colour charts (RHS, 1966) and then grouped into one of these two categories (Table 3.1). Testa pubescences of the seed coat were scored as pubescent or smooth (Table 3.1).

3.2.1.4 Kernel taste and amygdalin quantification

Kernel flavour determination was carried out by tasting the kernels by two different evaluators over each of the three years of the study. An average of 5 kernels from each individual tree was used for assessment. Taste was classified as sweet or slightly bitter (Table 3.1). Amygdalin was extracted and quantified by HPLC using a modification of the method of Dicenta et al.. (2002) by Dr Tricia Franks, who is working on an almond flavour project in the laboratory. For each genotype, 8 to 10 kernels were pooled together, weighed, then frozen in liquid nitrogen and ground to a homogeneous and fine powder using an IKA®A11 basic analysis mill (IKA® Works). Still frozen, ground tissue (0.2 g FW^{-1}) was resuspended in 10 ml of methanol by vortexing, and then heated in a water bath at 80°C for 10 min. The extracts were filtered (0.45 µm Millex-HV unit, Millipore) and analysed by HPLC using Agilent 1100 series apparatus. Methanol extracts (7.5 µl) were injected by autosampler and separated at 25°C through a ZORBAX Eclipse XDB-C18 5 µm column (Agilent) with dimensions: 4.6 mm

x 150 mm. Mobile phases were water (A) and acetonitrile (B) and conditions were: 20% B at 200 µl min-1 for 15 min; 20% B to 100% B and increasing from 200 µl min-1 to 400 µl min-1 over 12 min; 100% B at 400 µl min-1 for 2 min; 100% B to 20% B at 400 µl min-1 over 3 min; 20% B at 400 µl min-1 for 8 min. Wavelengths from 190 nm to 400 nm were monitored. Generally, one extract was prepared for each genotype and a single replicate was analysed by HPLC (Franks et al., 2008).

3.2.1.5 Percentage of doubles

The percentage of double kernels for each individual tree was determined from the samples of nuts containing two kernels divided by the total of 30 nuts per tree. The data was transformed according to Phippen and Ockendon, (1990) by logarithmic transformation of the percentage to improve the normality of the distribution as required prior to analysis of variance (ANOVA).

3.2.2 Statistical analysis

Data was analysed using an analysis of variance (ANOVA) with GenStat (version 8.2, Rothamsted Experimental Station). Means for each year were compared using least significant difference (LSD) methods at a confidence level of $P < 0.05$ to identify significant differences. Pearson's correlation coefficient was also calculated to test the relationships between the traits. Student's paired *t*-test was used to compare the significant differences of the two traits. For traits such as kernel taste, colour and smoothness, segregation in the F_1 was tested for goodness of fit against the expected segregation ratios for one gene by chi-square analysis (χ^2).

Chi-square test (χ^2) for expected frequency 3:1 calculated by

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} + \frac{(O_j - E_j)^2}{Ej}$$

Where O_i *and* O_j are an observed frequency for first and second trait; E_i *and Ej* are an expected frequency, asserted by the null hypothesis for first and second trait. Broad sense heritability (H) estimates (i.e., the proportion of total phenotypic variation that is due to all genetic causes) were obtained from the analysis of variance based on the linear model:

$Y_{ij} = \mu + y_i + g_j + \varepsilon_{ij}$

Where Y_{ij} is the phenotypic value of j_{th} progeny (j=1, 2, 3,..., 106) in ith year (i=1, 2, 3); μ is the mean value of the trait; y_i is the effect of the ith year on the trait; g_j is the genotypic effect of progeny j; and ε_{ij} is the year × genotype interaction. Broad-sense heritability (H) estimates were calculated using the following equations:

$$\sigma_g^2 = \frac{(MS_g - MS_{residual})}{y}$$

$$H_{(\text{broad sense})} = \frac{\sigma_g^2}{\sigma^2 / y + \sigma_g^2}$$

Where σ_g^2 is the genetic variance; MS_g is the estimated mean square of genotypes, and $MS_{residual}$ is the estimated mean square of residual error, σ^2, which is a measure of variability due to genotype × year interaction (Wang et al., 2000).

Table 3.1 Descriptor list for ten traits scored over the three years on the F_1 progeny, * in AD indicates trait is listed in the almond descriptors list (Gülcan, 1985; Gregory, 2004). KSH calculated according to the width/length ratio.

Code	Trait	Descriptor	AD
DBL	Percentage Doubles	1. Low (0-5%) 2. High (> 5%)	*
KWE	Kernel Weight/Size	1. Very Small (<0.9 g) 2. Small (1.0-1.1 g) 3. Medium (1.2-1.4 g) 4. Large (1.5-1.8 g) 5. Very Large (> 1.9 g)	*
SHH	Shell Hardness	1. Paper (>55%) 2. Soft (45-55%) 3. Semihard (35-44%) 4. Hard (25-34%) 5. Stone (<25%)	*
KSH	Kernel Shape	1. Very Narrow (0.4) 2. Narrow (0.40-0.48) 3. Medium (0.49-0.55) 4. Broad (0.56-0.65) 5. Very Broad (> 0.65)	*
KTH	Kernel Thickness	1. Very Thin (< 6.0 mm) 2. Thin (6.1-6.9 mm) 3. Medium (7.0-7.9 mm) 4. Thick (8.0-8.9 mm) 5. Very Thick (>9.0 mm)	
TCO	Testa Colour	1. Pale (RHS 164b) 2. Dark (RHS 164a/ 165b)	*
KTA	Kernel Taste	1. Sweet 2. Slightly Bitter	*
TPU	Testa Pubescent	1. Smooth 2. Pubescent	*

3.3 Results

3.3.1 Shell hardness

The shell hardness of individuals was categorised from stone type to papery. The mean of shell hardness for all nuts analysed during three years varied from 32.7% (hard-shell) as an average of all collected nuts in 2005 to 38.7% (semi-hard) in 2007 (Table 3.2; Fig 3.3A). No stone shell was found in 2007 and the overall trend showed an increase toward the semi-hard and soft-shell category over the three years. LSD test showed differences among three years. Around 80% of the individuals were classified as hard or semi-hard shell (Fig. 3.6 A). Heritability estimation was 0.752 (Table 3.4) and a significant negative correlation with in-shell weight (r = -0.36) was calculated (Table 3.5).

3.3.2 In-shell weight

The average value of in-shell weight for each tree ranged between a minimum of 1.33 g in 2007 and a maximum of 5.79 g in 2006 (Table 3.2), representing wide variation among the progeny. The mean value of all nuts in 2006 (3.91 g) was much higher than 2005 and 2007 (2.86 g) and average in-shell weight did not change in 2005 and 2007 (Fig 3.3 C). The LSD test detected significant differences between year 2006 and other years. A large positive correlation coefficient (r = 0.74) was found among in-shell weight and kernel weight (Fig 3.4) and negative correlation was calculated with shell hardness r = -0.36 (Table 3.5). This trait showed heritability of 0.805.

3.3.3 Kernel weight

The weight of almond kernel was generally light and mean value for all nuts ranged between 0.9 g in 2005 to 1.4 g in 2006 (Fig 3.3 D). The minimum of kernel weight (0.5 g) was in 2005 and the greatest among all trees (2.05 g) in 2007 (Table 3.2). Kernel weight change was found significantly different over all three years, and it was significantly correlated to both length (r = 0.78) and width (r = 0.80). A low correlation was observed with shell hardness (0.33). Heritability of H^2 = 0.764 was estimated for kernel weight.

3.3.4 Kernel length

The largest variation was in 2007, with a minimum of 16.8 mm and a maximum of 27.7 mm recorded. Average kernel length was the highest in 2006 at 23.9 mm. LSD test showed no differences between 2005 and 2007 but both were significantly different from the 2006 value. The trait showed a high heritability of 0.85. A positive strong correlation was observed with kernel weight (r = 0.78), kernel width (r = 0.67) and a poor correlation with kernel thickness r = 0.23 (Table 3.5).

3.3.5 Kernel thickness

The trait is divided into five categories, the average value varied significantly among the years. The highest mean value (8.87 mm) was in 2006 (Table 3.2) and no thin or very thin kernels were observed in this year. By and large the trend showed a shift from thin towards thick kernels. Thick kernel (Table 3.1) was the largest group in 2006 and 2007 but medium kernel was the largest group in 2005 (Fig 3.6 B). LSD test showed differences across years. Kernel thickness showed an intermediate correlation with kernel width (r = 0.405) and the inheritability of 0.72.

3.3.6 Kernel width

The kernel width ranged from a minimum of 9.9 mm in 2007 to a maximum of 18.0 mm in 2006. The kernel width exhibited the highest average in 2006 at 14.7 mm. Estimate of broad sense heritability from the analysis of variance showed high heritability of 0.83 (Table 3.4). The highest correlation among the kernel width and other morphological traits was obtained between kernel width and kernel weight (0.80) and lower correlation $H^2 = 0.67$ with kernel length (Table 3.5, Fig 3.5).

3.3.7 Kernel shape

Shape is a subjective factor depending on kernel width and length. The shape of the kernels can be classified into five categories (Table 3.1) but in our experiments the results showed the presence of only three of these classes: medium, broad and very broad. Distribution categories are presented in Fig 3.6 C. Kernel shape appeared generally more consistent compared with other traits. The shape trend showed slight increase from broad kernel toward very broad over the three years. Overall distribution of the kernels was normal, and the largest group was medium in all three years. The highest heritability estimated from all traits belongs to the kernel shape $H^2 = 0.916$ (Table 3.4). As expected the negative and significant correlation was obtained between kernel shape and kernel length ($r = -0.417$) and positive correlation with kernel length $r = 0.388$ (Table 3.5).

3.3.8 Testa colour and pubescence

Testa colour was divided into two categories: pale and dark. Although when kernel colour was compared with the RHS colour chart the gradual change of colour

was observed, for comparison with previous data obtained by Gregory (2004) which classified the colour into two groups the same assessment was carried out (Fig 3.7A). In all three years this trait corresponds to a 3:1 ratio as tested by X^2 analysis (Table 3.4).

Testa pubescence did not segregate in a 3:1 ratio in any years (Table 3.3). In 2007 the numbers of smooth kernels increased around 10%. In 2005 and 2006 the percentage of progeny in each category was almost identical (Fig 3.7 B).

3.3.9 Kernel taste and amygdalin quantification

The number of sweet kernels identified by the tasting panel was higher in 2006 than the other two years; the number of sweet individuals was approximately four times the number of slightly bitter individuals in 2006 and 2007 (Fig 3.7 C). The χ^2 showed a significant deviation from a 3:1 segregation ratio in all three years (Table 3.3). Average value of amygdalin in the slightly bitter kernels was 20.34, and for sweet kernels 3.67 mg amygdalin kg^{-1} FW. The amygdalin peak was recorded at a wavelength of 210 nm with a retention time (about 10.4 min) consistent with the amygdalin reference compound (purchased from Sigma). Amygdalin was quantified using a 4-point calibration curve (0.25 mg L^{-1}, 0.5 mg L^{-1}, 1.0 mg L^{-1}, 2.0 mg L^{-1}; r^2 = 0.999). Amygdalin content was variable and observed in both sweet and slightly bitter kernels (Fig 3.12, 3.13). The highest amygdalin contents were observed in the slightly bitter with 70 mg kg^{-1} FW. The amygdalin level in sweet kernels varied between 0 to 14.5 mg kg^{-1} FW and for slightly bitter it was between 5.5 to 70 mg kg^{-1} FW. The mean content of amygdalin was higher in slightly bitters than sweet kernels.

Table 3.2 Morphological characters measured on F_1 hybrid population of cross between 'Nonpareil' × 'Lauranne', mean value for data collected for three years (2005-2007).

2005

Traits	Mean	SD	SE	Lowest value	Highest value	Conf. inter 95%
Double (%)	12.66	12.11	2.2	0	46.7	4.52
Shell Hardness [a] %	32.67	5.18	0.50	20	55.54	0.99
In-shell Wt. (g)	2.86	0.56	0.05	1.52	4.60	0.11
Kernel Wt. (g)	0.92	0.17	0.02	0.5	1.48	0.03
Kernel Shape [b]	0.6	0.05	$4.7\ e^{-3}$	0.49	0.73	$9.34\ e^{-3}$
Kernel Thickness [c]	7.34	0.78	0.08	5.33	9.34	0.15

2006

Traits	Mean	SD	SE	Lowest value	Highest value	Conf. inter 95%
Double (%)	12.46	9.89	1.60	0	50	3.25
Shell Hardness [a] %	37.57	7.45	0.73	22.65	63.17	1.46
In-shell Wt. (g)	3.91	0.94	0.09	1.76	5.79	0.18
Kernel Wt. (g)	1.43	0.27	0.03	0.78	2.05	0.05
Kernel Shape [b]	0.62	0.05	$5.19e^{-3}$	0.49	0.76	0.0103
Kernel Thickness [c]	8.78	0.72	0.071	6.97	10.69	0.14

2007

Traits	Mean	SD	SE	Lowest value	Highest value	Conf. inter 95%
Double (%)	24.87	26.62	3.76	0	100	7.57
Shell Hardness [a] %	38.66	6.01	0.59	25.32	57.19	1.17
In-shell Wt. (g)	2.86	0.7	0.07	1.33	5.58	0.14
Kernel Wt. (g)	1.07	0.22	0.02	0.57	1.85	0.04
Kernel Shape [b]	0.62	0.056	$5.45e^{-3}$	0.5	0.80	0.011
Kernel Thickness [c]	8.11	0.74	0.073	5.87	10.96	0.144

[a] Shelling percentage (kernel wt./shell wt.). Paper-shell: 55-65%+, Soft-shell: 45-55%, Semihard shell: 35-45%, Hard-shell: 25-35% and Stone-shell: 20-25% (Rugini, 1986)

[b] Kernel shape (width/length) ratio: very narrow (<0.40), narrow (0.40-0.48), medium (0.49-0.55), broad (0.56-0.65), very broad (>0.65)

[c] mm: very thin (<6), thin (6-6.9), medium (7-7.9), thick (8-8.9), very thick (>=9) (Gregory, 2004).

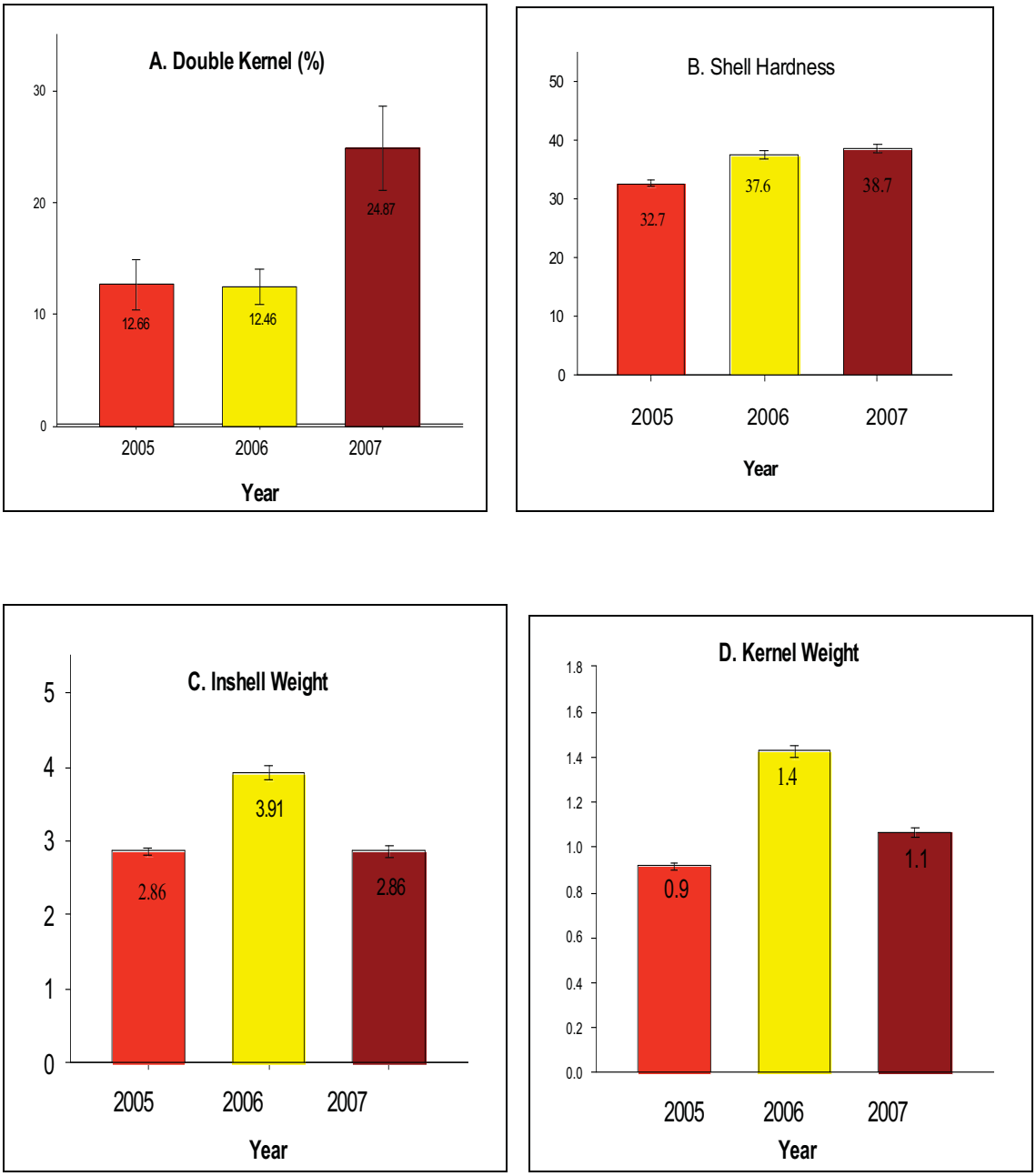


Fig. 3.3 Average of traits over three year's measurements (2005-2007). A. Percentage of double kernels, B. Shell hardness; C. In-shell weight and D. Kernel weight. Bars indicate standard errors of means.

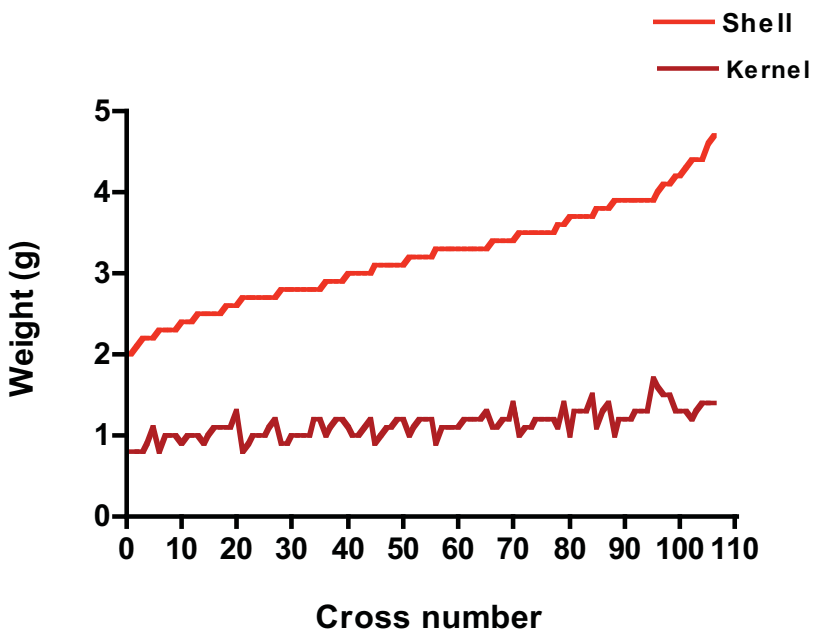


Fig. 3.4 The graphs depict the correlation between in-shell and kernel weight among the mapping population of cross 'Nonpareil' × 'Lauranne' in the average value of three years.

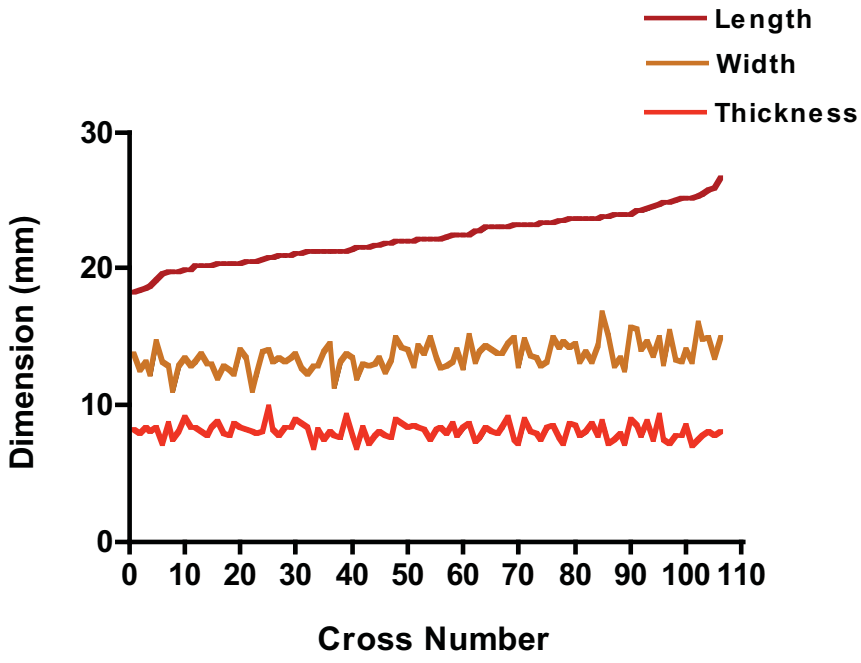


Fig. 3.5 The graph shows the correlation between kernel width and length and low correlation between kernel thickness and the other kernel size traits in a segregating population.

3.3.10 Double kernel

A significant variation was observed in double kernels over three years of the study, although only small differences were observed between years 2005 and 2006. The trait showed the highest standard deviation and error among all traits measured (Fig 3.3 A). The highest individual value obtained was 100% in 2007 and the lowest value among all trees was 0% observed in all years. In all three years around 55 % of individuals showed 0% doubles, with the next largest group containing approximately a fifth of all individuals in the class of 0.1-10% (Fig 3.8). Chi-squared analysis was used to determine if the trait was inherited in a 1:1 or 3:1 ratio. The results of this analysis differed for each year studied. In 2005, it followed 3:1 Mendelian inheritance pattern but not in other years and in 2007 it corresponded to a simple inheritance pattern of 1:1 (Table 3.3).

3.4 Discussion

Studies of economically important characters in fruit trees have been carried out usually using trees grown at a single location over a period of several years. The data analysis of these studies has allowed the estimation of genetic components, the effect of environment and the interaction of genotype × environment, which can help the breeder decide how many years are required for evaluation to distinguish superior genotypes (Yao and Mahlenbacher, 2000). As stated by these authors, 'for traits to be useful they must be consistent from year to year and tree to tree, this consistency can be predicted by the estimation of heritability'. Almost all the quantitative traits evaluated for three years were highly heritable. Furthermore, valuable correlations were obtained between a numbers of traits that affect kernel quality.

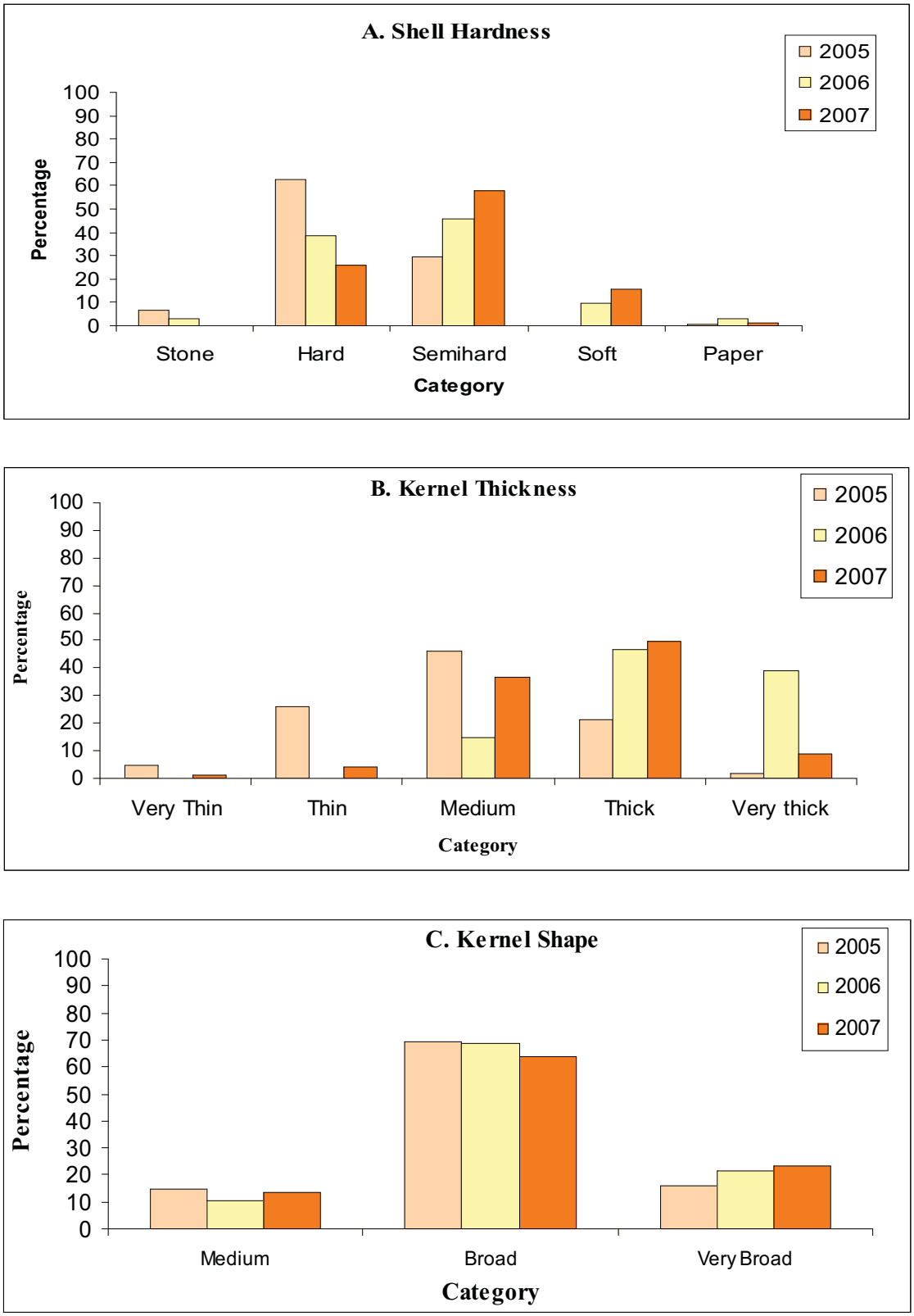
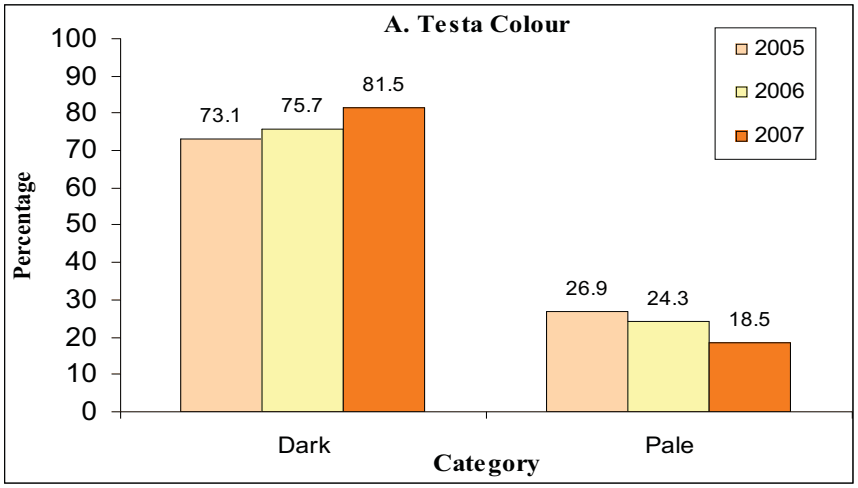
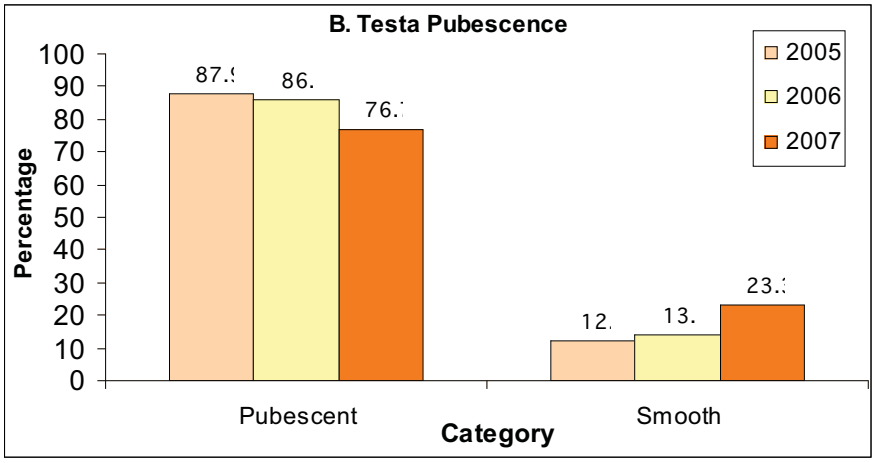


Fig. 3.6 A. B. C Percentage of progeny traits over three years for A. Shell Hardness, B. Kernel Thickness and C. Kernel Shape





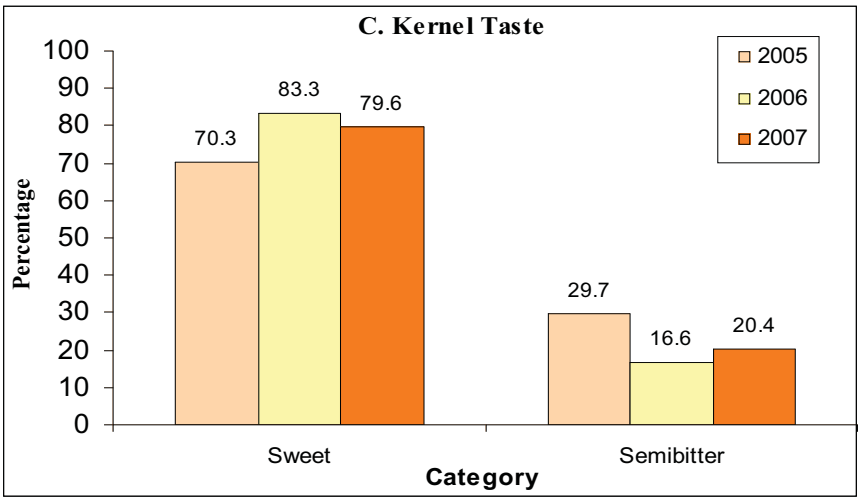


Fig. 3.7 A. B. C Distribution of percentage frequency of the population NP × LA grouped into two classes for A. Testa colour, B. Testa pubescence and C. Kernel taste

Table 3.3 Chi-square test (χ^2) for expected frequency 3:1 or 1:1 of kernel taste, colour, smoothness and double kernels in the population NP × LA. N= number of total observation.

Traits	Year	N	Observed		Expected		χ^2	Significance P<0.05
Kernel Taste			Sweet	Slightly bitter	Sweet	Slightly bitter		
	2005	1053	735	318	789.75	268.25	13.02	ns
	2006	1074	880	180	795	265	36.35	ns
	2007	695	563	132	521.25	173.75	13.38	ns
Kernel Colour			Dark	Pale	Dark	Pale		
	2005	1054	778	276	790.5	263.5	0.79	<.00.1
	2006	1050	800	250	787.5	262.5	0.79	<.00.1
	2007	1050	787	263	787.5	262.5	0.0013	<.00.1
Kernel Smoothness			Smooth	Pubescent	Smooth	Pubescent		
	2005	1074	938	117	791.25	263.75	108.87	ns
	2006	1060	920	140	795	265	78.61	ns
	2007	1047	821	226	785.25	261.75	6.51	ns
Double Kernels 3:1			Single	Double	Single	Double		
	2005	106	78	28	79.5	26.5	0.113	<00.1
	2006	106	68	38	79.5	26.5	6.65	ns
	2007	106	56	50	79.5	26.5	27.79	ns
Double kernels 1:1			Single	Double	Single	Double		
	2005	106	78	28	53	53	23.58	ns
	2006	106	68	38	53	53	8.49	ns
	2007	106	56	50	53	53	0.34	<00.1

Table 3.4 Broad sense heritability of traits in the population NP × LA for three years (2005-2007)

Trait estimate	Heritability
Shell Hardness	0.752
In-shell Weight	0.805
Kernel Length	0.850
Kernel Shape	0.916
Kernel Weight	0.764
Kernel Thickness	0.720
Kernel Width	0.830

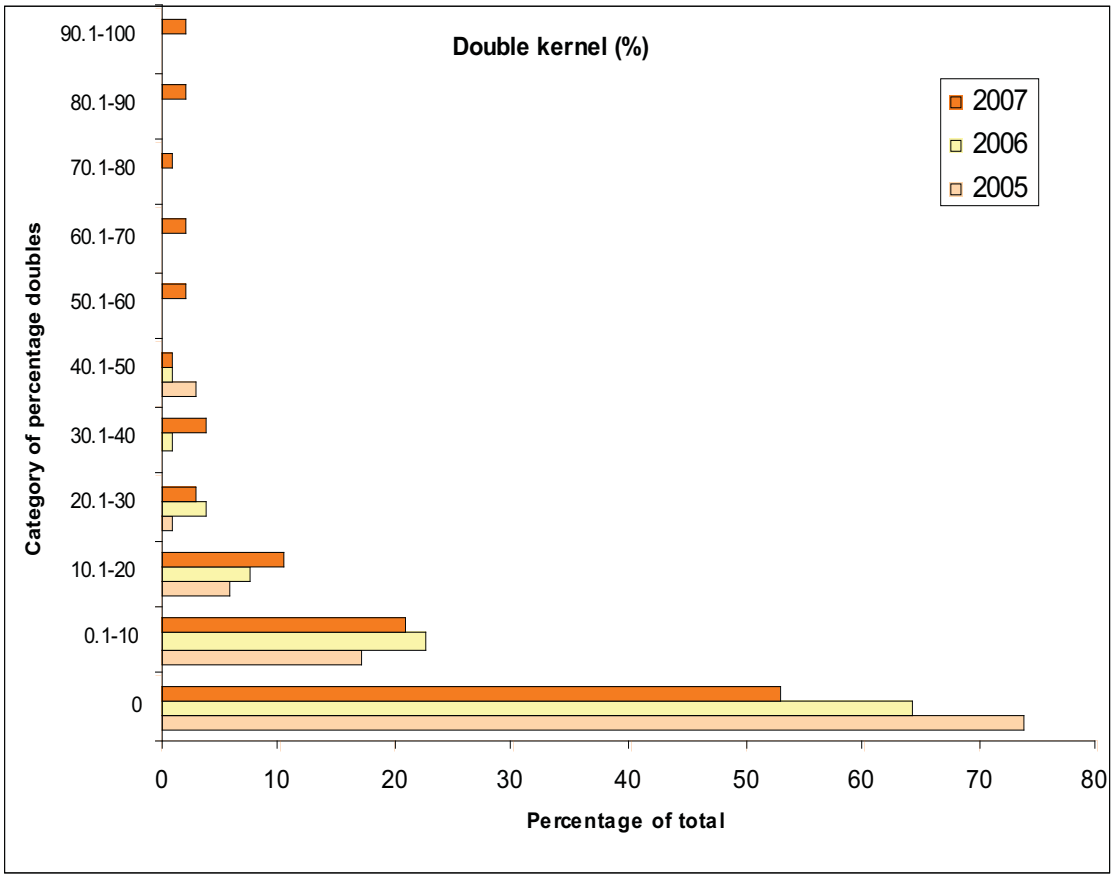


Fig. 3.8 Frequency of total progeny showing 0-100% doubles in 10% increment categories over three years (2005-2007)

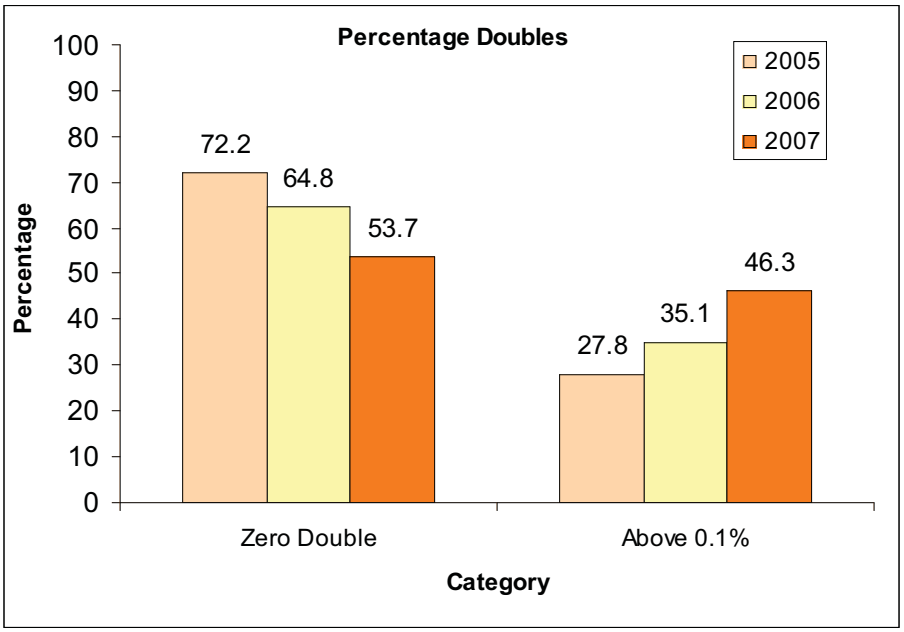


Fig. 3.9 Percentage of double kernels in two categories over three years (2005-2007)

Table 3.5 Pearson correlation coefficients between different pomological traits in the mapping population of 'Nonpareil' x 'Lauranne'

Pomological trait	Pearson correlation coefficient
Shell hardness/ In-shell weight	-0.36***
In-shell weight/Kernel weight	0.74***
Kernel weight/ Kernel thickness	0.643***
Kernel weight/ Kernel length	0.78***
Kernel weight/ Kernel width	0.80***
Kernel length/ Kernel width	0.67***
Kernel length/ Kernel thickness	0.23***
Kernel width/ Kernel thickness	0.405***
Kernel weight/ Shell hardness	0.33***
Kernel shape/ Kernel length	-0.417***
Kernel shape/ Kernel width	0.388***
Testa smoothness/ Kernel taste	0.04**
Kernel taste/ Testa colour	0.09**
Testa colour/ Testa smoothness	0.33***
Kernel taste/ Amygdalin	0.63***

Fig. 3.10 Double Kernel produced in almond. Compared with single kernel, (illustrated on right) double kernels are deformed and misshapen.

Fig. 3.11 The variation of morphological traits in nuts and kernels of the segregating population of NP × LA

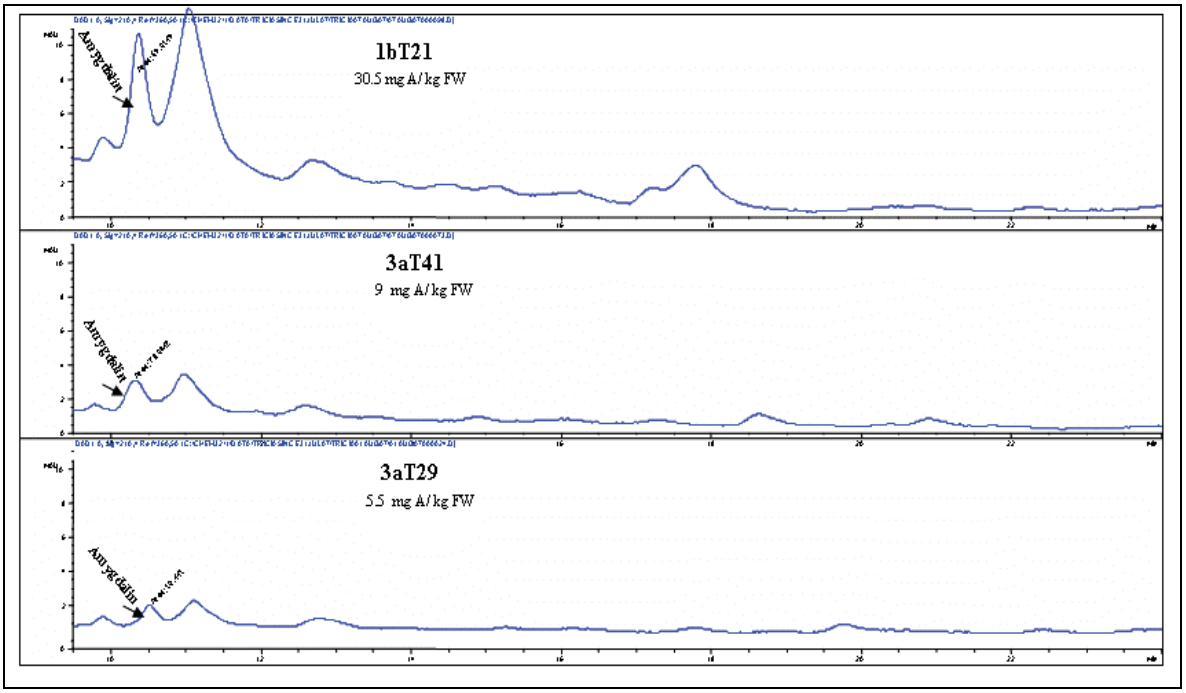


Fig. 3.12 Chromatogram of amygdalin in slightly bitter kernels

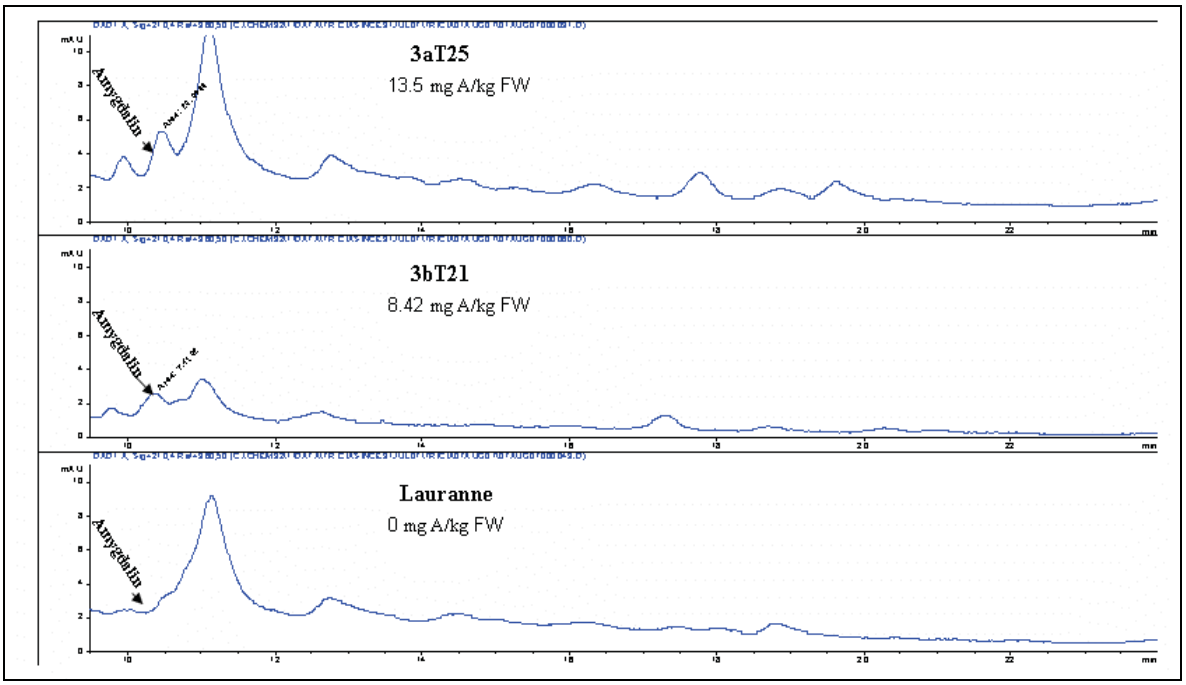


Fig. 3.13 Chromatogram of amygdalin in sweet kernels

Table 3.6 Weather data for Renmark Aero weather station number 024084, average air temperature (A) and average rainfall data (B) for 2004-2007. Collected by Bureau of Meteorology, Australia

A. Temperature Data

Year	Measurement and Unit	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Doc	Annual
2004	Mean	30.3	34.9	29.3	25.5	19.6	17.0	15.8	19.0	21.0	27.5	27.9	30.5	24.9
2005	Maximum Air	32.0	30.3	28.6	28.0	21.9	17.5	16.5	19.1	20.5	23.9	28.5	33.1	25.0
2006	Temperature	36.5	31.7	30.6	22.2	18.1	16.0	15.7	19.7	24.0	26.9	30.1	31.5	25.3
2007	(ºC)	32.4	35.0	28.9	25.8	21.6	14.9	16.6	20.5	23.4	26.4			
2004	Mean	13.5	16.1	11.0	9.2	3.6	6.2	4.1	4.8	6.1	9.3	12.5	14.8	9.3
2005	Minimum Air	15.5	13.8	11.5	10.6	6.2	6.0	4.2	4.8	6.8	10.0	12.1	13.9	9.6
2006	Temperature	19.5	14.6	13.5	7.3	5.0	0.9	3.6	3.1	6.6	7.4	11.7	13.6	8.9
2007	(ºC)	17.2	17.4	14.3	10.4	9.4	2.4	3.1	4.0	6.2	8.6			

B. Rainfall Data

Year	Measurement and Unit	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Doc	Annual
2004	Total Monthly	3.0	1.2	3.4	0.6	6.2	43.2	14.4	27.4	14.2	2.0	22.6	51.4	189.6
2005	Precipitation	36.4	4.6	3.8	5.2	0.4	60.4	26.8	20.6	38.4	74.8	21.6	8.0	301.0
2006	(mm)	3.6	5.6	14.0	34.6	7.6	6.8	35.6	1.2	6.6	0.0	0.0	0.0	115.6
2007		37.6	2.2	27.8	43.6	19.4	8.6	19.6	1.6	4.2	12.8	42.6	20	240

3.4.1 Shell hardness

Shell hardness was estimated as the percentage of kernel to the in-shell weight. This trait varies greatly from 20% (stone) to 63% (paper). The segregating population is a result of cross between genetically paper-shelled (NP = 56.4%) to hard-shelled (LA= 33.3%) almonds. The results indicate that nearly 85% of progeny showed semi hard or hard shell. Grasselly, (1972) proposed that this trait is determined by major gene in which hard-shell is dominant *DD* and soft-shell recessive *dd* (Socias i Company, 1998). According to this simple pattern of inheritance the population should produce only hard-shelled kernels. Alternatively, the presence of modifier genes can lead to recovery of a semi hard shell phenotype. However, in the population nearly 10-15% of the progeny showed soft or paper shells and around half of the progeny showed soft shell. This result is consistent with the earlier evaluation from this population by Gregory (2004). The results indicate that the determination of shell hardness seems to be more complicated and this trait may be controlled with other genes that show more complex inheritance. In hazelnut, Thompson, (1977) suggested that heredity of shell hardness was controlled by additive genes. Monogenic inheritance of this trait has not been confirmed in nut crops, and the trait has been considered as a quantitative trait (Kester et al., 1977; Dicenta et al., 1993; Socias i Company, 1998; Sánchez-Pérez, 2007a).

The results obtained showed heritability of $H^2 = 0.75$, which can be compared with the previous studies of shell hardness $H^2 = 0.82$ (Spiegel-Roy and Kochba, 1981) and $H^2 = 0.55$ (Kester, 1977) in almond and 0.91 in walnut families (Hansche et al., 1972). Gregory (2004) found heritability of 0.348 for this population by the regression method. Our result showed a higher value, which is comparable to the estimation of the other studies. Heritability obtained by regression, where the parents are involved, is

lower than those obtained by variance, where only the progeny are considered. Moreover, the age of trees could be responsible to some extent for the discrepancy. Shell consists of two centric layers, which in some cultivars such as 'Ferragnès' may be separated during the shelling removal procedure and produce small gaps. As a result, the de-shelling equipment needs to be readjusted and it is probable that kernels will be broken, reducing the value of the crop (Socias i Company et al., 2008). Suitable shell hardness reflects the conditions of cultivation environment and the target market, and has implications for the equipment available for nut processing. Uniformity of shell type is necessary due to the needs for initial calibration of the cracking machines. Preference of the shell type in different regions of the world is not same. For example, in the Mediterranean regions hard-shelled cultivars are generally preferred because of their better tolerance to drought, resistance to birds, pests and better storing ability. In California and the new regions of almond culture, however, soft-shelled cultivars with high kernel percentage are preferred (Socias i Company, 1998).

3.4.2 In-shell weight

In-shell weight in almond includes the weight of the kernel and the endocarp (shell). Depending on the shell thickness, the shell can constitute between 35 and 65% of the in-shell weight. Therefore, small papery-shelled nuts can have the same kernel weight as large hard-shelled nuts. Thinner-shelled nuts showed higher damage during nut development by pests and diseases that were observed in the course of this study. The result indicates that in-shell weight varies from one year to another. However, the results obtained in this study are not conclusive, as only three year studies were made. Sánchez-Pérez et al., (2007a) reported similar findings that in-shell weight showed high

variations among years. The frequency of smaller size nuts was higher in 2005 and 2007 compared with 2006. The increase of yield level makes smaller nuts so the acceptable nut size for the target market must be considered. Nevertheless, in spite of the effect of year, the heritability of this trait was high, H^2= 0.80. The high heritability of in-shell weight is similar with data obtained by other studies. Kester et al., (1977) found a heritability of 0.81 for nut weight in the populations of several almond crosses, Spiegel-Roy and Kochba, (1981) obtained high heritability for this trait H^2 = 0.87. Estimations of broad sense heritability in other nut trees agreed with this value, being highly heritable in walnut H^2 = 0.85 (Hansche et al.., 1972) and hazelnut (H^2 = 0.84) (Thompson, 1977). In-shell weight comprises of shell percentage and kernel weight. The commercial part of almond is kernel, thus for breeding program in-shell weight is not as important of kernel weight.

3.4.3 Kernel weight

Kernel weight is considered to be a quantitative trait and as pointed out by Spiegel-Roy and Kochba, (1981), it must be controlled by additive genes. Kernel weight was quite different between year 2006 (heaviest mean value of 1.4 g) with year 2005 and 2007. The trend was comparable with in-shell weight results which indicated that the values for the two traits were consistent over years. The correlation estimates between kernel and in-shell weight were relatively high (r = 0.74). The strong genetic correlation between nut and kernel weight was also found in a previous study in almond (Gregory, 2004) and in macadamia (Hardner et al., 2001). Considering all the information of the previous study by Gregory, (2004) and the results of present study, the difference observed between year 2006 and other two years can be to some extent

explained by the alternate bearing cycle that is expressed in almond (García et al., 1994; Murua, 1994). Alternate bearing is a seasonal cycle with one year heavy 'on' and next year with light 'off' crop. This occurs because of the low or high flower condition, fruit set, or yield (Monselise and Goldschmidt, 1982). In the 'off' year (2006) larger nuts and kernels were produced. The meteorological influence could overlap the influence of the alternate bearing cycle (Murua, 1994). It was shown that almond kernels reach their highest growth around 15 weeks after anthesis (Hawker and Buttrose, 1980). Cooler temperatures in October and November could reduce the stress and respiration rate, which allows more nutrients and metabolites to be stored in the seed and increase the kernel weight. Traynor, (1966) indicated that water stress during flowering can influence pollination and reduce the number of nuts per tree, but the extreme water stress required for this effect is unlikely to happen in commercial almond orchards. The data of weather conditions showed a large decrease in monthly rainfall in August 2006, but the same reduction also occurred in August 2007, hence this factor alone could not be accountable for change in kernel weight.

Heritability of 0.764 was estimated for kernel weight which is approximately equal to the estimate previously reported by Kester, (1977) $H^2 = 0.64$ and by Spiegel-Roy and Kochba (1981) $H^2 = 0.79$ for this trait. Also Hansche et al., (1972) observed heritability of 0.87 in walnut families. Previous work by Gregory, (2004) on the same population showed substantial low heritability of 0.031 for this trait. The low level of heritability was in disagreement with the results of the present study and was furthermore not substantiated by data presented by other authors (Kester, 1977; Spiegel-Roy and Kochba, 1981). It is not readily apparent why the data of Gregory (2004) differ so widely from those reported here and in other research.

3.4.4 Kernel size and shape

Dimensions of the kernel size (except kernel thickness) showed generally similar values over three years of the study. Kernel width to length ratio was used for shape calculation. Kernel shape usually maintains in a cultivar and is defined as one of the characteristics for cultivar identification (Gülcan, 1985). The linear dimensions of almond nuts and kernels are established within 10-12 weeks after flowering (Hawker and Buttrose, 1980). When linear growth was completed, accumulation of metabolites in the kernel continues until near harvest. Linear dimensions, plus the accumulation of storing material such as protein, lipid, starch and sucrose determine the kernel weight (Kester, 1965; Hawker and Buttrose, 1980). Nearly constant kernel shape and high correlation between kernel weight and kernel dimension were observed in this study which is in agreement with findings of Kester (1965).

The highest heritability was estimated for kernel shape ($H^2 = 0.916$). In addition, all the kernel size components showed high heritability except kernel thickness which varies more and showed the lowest heritability. High heritability observed for shape corresponded to the estimation of the ratios between dimensions involved in shape measurement (kernel width / kernel length). Many studies in almond found high heritability value for these traits. Initially, Kester (1977) estimated heritability of kernel dimensions in a number of crosses of almond and reported the values of heritability for kernel length $H^2 = 0.77$, kernel thickness (0.71) and kernel width (0.62). Spiegel-Roy and Kochba, (1981) found heritability of 0.73, 0.76 and 0.50 for kernel length, width and kernel thickness, respectively in almond. The highest heritability in almond was estimated by Chandrababu and Sharma, (1999) who provide and estimate of $H^2 = 0.96$ for kernel length, 0.91 for kernel width and 0.90 for kernel thickness. Positive correlation coefficients among the kernel size dimensions were observed, varying from

r = 0.67 for kernel length / kernel width to r = 0.23 between kernel length and kernel thickness. As mentioned earlier (Table 3.5), high correlation was found between kernel size and kernel weight. Similar phenotypic correlations obtained for these traits were calculated when investigating hazelnut families (Yao and Mehlenbacher, 2000). High correlation indicates that improvement of one trait culminated in the improvement of another. Moreover, the evaluation of one trait of a highly correlated pair would be sufficient to provide an estimation of both traits, resulting the increasing the efficiency of measurement. The association could mean that the large positive correlations between these traits may be a result of the partial linkage among genes affecting these traits. The result obtained by Ledbetter and Palmquist, (2006) showed that kernel thickness was not related to width or length. Poor correlation among estimates of kernel thickness with kernel weight compared to either kernel width or kernel length may be explained by the similarity of kernel thickness in large or small kernels; it is apparent that kernel thickness does not necessarily change in a proportional rate as the kernel length does.

3.4.5 Testa colour and pubescence

Testa colour was scored as pale or dark and most of the individuals in the population had a dark kernel colour. The visual appeal of kernels, a critical quality factor for consumers, is mostly determined by the kernel colour (Socias i Company et al., 2008). The inheritance of the trait followed a 1:3 inheritance ratio that was confirmed by χ^2 analysis and was consistent over the consecutive three years of the study. Heritability values of H^2 = 0.3, 0.51 and 0.66 for kernel colour were estimated by Arteaga and Socias i Company, (2001); Kester, (1977) and Spiegel-Roy and Kochba,

(1981), respectively. They found this trait inconsistent because of environment effects. Testa colour in the harvested kernel is determined by several factors including maturity at harvest time, the conditions and duration of storage. Testa colour may be controlled by a major gene, with epistasis effects and modifier genes potentially capable of changing its expression. Sánchez-Pérez et al., (2007a) assigned testa colour intensity into five rating scales from 1 = very light brown to 5 = very dark brown which could show much better differences among the progeny. Another option to increase the accuracy of testa colour assessment may be the measurement of quantitative factors in testa such as luminosity, chroma and hue angle. Light golden kernel colour is preferred in the market place because unusually dark kernels are assumed to have turned rancid or lost their freshness (Sang et al.., 2002).

Testa pubescence was scored as smooth or pubescent, and higher pubescent kernel was observed in all years. The trait did not segregate to a 3:1 inheritance pattern in the progeny as tested by χ^2 analysis. This result is in agreement with those obtained by Gregory, (2004). Very limited studies have been carried out on testa pubescence despite the importance for consumer preference. Low heritability of 0.30 and 0.39 was estimated for testa pubescence by Arteaga and Socias i Company, (2001) and Kester, (1977) respectively. This trait was rated by naked eye, so inconsistencies of measurement may be the factor limiting the significance of their heritability (Kester et al.. 1977). This low heritability might also suggest that strong environmental factors can modify the expression of the trait. Low correlation coefficient between testa colour / testa smoothness (r = 0.33) was calculated. The basis for this correlation is not quite clear but it is possibly due to the nature of sensory evaluation for these two traits. During the assessment the light kernel seemed less pubescent or smooth. Gregory,

(2004) evaluated the testa pubescence with light and electron microscopy; the results showed that the phenotypically pubescent testa has distinct large expanded cells arising from the epidermal surface which are uniform in shape but smooth testa showed less uniform and smaller expanded cells. In addition, both expanded cells with high density and collapsed cells were observed in pubescence kernels but only collapsed cells observed in smooth kernels.

3.4.6 Kernel taste and amygdalin quantification

Kernel taste was classified as sweet or slightly bitter. The analysis of χ^2 did not show the 3:1 segregation ratio corresponding to a simple Mendelian pattern, in contrast with the earlier work on the same population of almond (Gregory, 2004). In the work of Gregory, (2004), a 3:1 ratio of sweet to slightly bitter kernels was reported in the first year of study (2001), but not in the two following years. According to the monogenic inheritance of almond taste, Dicenta and García (1993) suggested that the slightly bitter taste was produced due to the presence of the recessive allele (*Ss*) which may express a small degree of bitterness. However, the results described in the present work are different from their hypothesis and in agreement with finding by Spiegel-Roy and Kochba, (1974) and Ledbetter and Pyntea (2000), who suggested that a more complex mode of inheritance of kernel taste must be involved. It seems that the ability to produce slightly bitter kernels is a quantitative trait that can be partially inherited and its expression can be modified by different factors. 'Nonpareil' rarely produces a few slightly bitter kernels which can be considered heterozygous for the sweet kernel (Spiegel-Roy and Kochba, 1974). Moreover, it was reported that the male parent does not affect the taste of the kernel (Dicenta et al., 2000), therefore, it can be speculated

that the maternal parent, 'Nonpareil', should be heterozygous for this trait. More study regarding the effect of modifier genes in the expression of slightly bitter taste is required. Although the fact that there is a high amount of amygdalin in bitter kernels is well known, for slightly bitter individuals the relationship between bitterness and amount of amygdalin is unclear (Dicenta et al., 2002). In the present work the mean value of amygdalin in sweet kernels (3.67 mg amygdalin kg^{-1} FW) was lower than in slightly bitter kernels (20.34 mg kg^{-1} FW) but some sweet individuals had higher amounts of amygdalin compared to slightly bitter. Similar results were found by Dicenta et al., (2002) who found considerable variation in the concentrations of amygdalin in the sweet, slightly bitter and bitter genotypes. From these results it can be concluded that other compounds may be affecting the expression of slightly bitter taste of the kernel.

Warner and Creller, (1997) hypothesised that the sweet kernel (*sk*) allele might reduce the synthesis of cyanogenic glucoside in the seed or prevent the transport of cyanogenic glucosides from vegetative parts of the plant into the developing seed. On the other hand, sweet kernel individuals may have normal amounts of seed cyanogenic glucosides, but the *sk* allele may code for defective cyanoglycoside catabolic enzyme, resulting in a seed that is acyanogenic (Warner and Creller, 1997). A moderate to high correlation coefficient was found in this study (0.63) between amygdalin content and taste. It was found that all the progeny with amygdalin levels of more than 20 mg kg^{-1} FW had slightly bitter taste. This is a complex trait that needs more investigation to elucidate, but if some threshold, in our case 20 mg kg^{-1} FW, for the amount of amygdalin were to be considered it may help the assessment by chemical analysis. However, it is worth to mention that the measurement of amygdalin using HPLC was

performed only for kernels collected during one growing season (2007), and the amounts of amygdalin in some samples were so small that we were working at the limits of detection. More replication between years would be more effective and allow better validation of the results.

3.4.7 Percentage of double kernels

Double kernels' variance was highly significant and affected greatly by year effects, which showed a high standard error. The analysis of χ^2 showed the trait was corresponding to the 3:1 in 2005 and 1:1 segregation ratio in 2007. There was no consistency in the inheritance pattern, indicating the dual nature of transmission of this character as a qualitative and quantitative trait. However, double kernel is a complex character that was considered as a quantitative trait and subject to many external influences and therefore, the result of segregation ratio in this study might be biased by strong environmental effects (Spiegel-Roy and Kochba, (1981); Dicenta et al.. (1993); Socias i Company, (1998) and Arteaga and Socias i Company (2001)).

The parents of the segregating population produced a very low percentage of double kernels, but in the progeny a high mean value was found, with double kernels close to 50% in 2007. Several studies reported the influence of environmental conditions during ovule development and particularly temperature during pre-blossom on expression of this trait. Egea and Burgos, (1995) in a 10-year study of the temperatures in months leading up to bloom, found a negative correlation between high pre-blossom temperatures and double-kernelled seed production. Spiegel-Roy and Kochba, (1974) mentioned that low temperatures at blooming time increased the production of double kernels. Trees of the mapping population flowered during August,

and the meteorological data showed that the average minimum temperature during the previous months of blooming dropped sharply i.e. 0.9°C in June and 3.6°C in July 2006 (Table 3.6). The nuts of these trees were collected in February 2007 when the highest percentage of double kernels over the three years was observed. The result obtained in this study is in complete agreement with the previous finding and indicates the key role of environment effects in determining this phenotype. A small increase in pre-blossom temperatures brings about faster embryo development and therefore an early degradation of the secondary ovule before fertilisation (Egea and Burgos, 1995). Another factor affecting the production of double kernel is the size of the primary and secondary ovule. Egea and Burgos, (2000) reported that in single kernel cultivars the size of the secondary ovules were much smaller than primary ones. 'Lauranne' is an offspring of 'Tuono' and 'Ferragnès'. 'Tuono' is a self-compatible cultivar with a high percentage of double kernels (Sánchez-Pérez et al., 2007b). Although during this study no double was observed in 'Lauranne', this cultivar was known to have double kernels. The high percentage of double kernel in segregating population also can be attributed to the genetic composition of the parent.

3.5 Conclusion

Understanding of the genetic nature, inheritance and environmental effects on the traits can assist in choosing correct parents for future breeding programs. Strong correlations which were observed in the present study could enhance trait selection and allow many traits that co-segregate to be improved simultaneously. High heritability indicates that the gains of genetic improvement are quick and feasible. Different

heritability and morphological variations observed in this study compared to the previous work on this population showed that tree age has an impact on expression of the traits. Tree age is rarely mentioned in the almond literature. It appears that at the early stage trees are more subject to environmental variation than mature trees, which are more phenotypically stable. When inheritability and the correlation of similar traits were compared in almond with the other nut crops such as hazelnut, walnut and macadamia the results reflect the wider relationship between nut crops as stated by Hansche et al., (1972). We used a segregating population grown in a commercial orchard under field conditions. To obtain better understanding of the impact of environment and genetic factors, it would be better to perform a study under controlled environmental conditions.

Apart from the use of inheritability and correlations as tool for selection, the results of this study have other practical implications for almond industry and breeding programs. For example, the effect of the pre-blossom air temperature showed that if growers want to obtain a profitable yield they must avoid establishing their orchards on regions with low pre-blossom temperature even with cultivars that show a low percentage of doubles. Although the effect of alternate bearing in almond which was observed in this study is much lower than some other tree fruits such as pistachio or apple, a study of factors that influence this character such as cultural practices or environmental conditions would benefit the almond industry. To manipulate a trait effectively, different aspects or components that affect it must be elucidated. The results of amygdalin analysis showed that the amount of this metabolite alone cannot be responsible for kernel taste determination, which shed some light on the investigation between the relationships of other less studied components with this trait. Finally, the

continuous distribution of the presented data showed these traits can be used for QTL (quantitative trait loci) mapping and ultimately be assigned into a linkage map for future use as marker assisted selection or gene discovery through map-based cloning.

Chapter 4

Construction of intraspecific genetic linkage maps of almond using RAPD, ISSR, microsatellites and self-compatibility allele

4.1 Introduction

The aim of plant improvement programs is to decrease production costs, generate more profit for industry and increase the efficiency of global food production in a competitive market. Conventional plant breeding has made an enormous contribution to this end; however, selection for desired characters may take several years especially in a fruit tree like almond with a long juvenile period, large plant size and self-incompatibility. Therefore, any methods that accelerate or reduce the length of the selection process can save much cost, time and labour. New technological advances can assist tree breeders to achieve these goals by providing methodologies that shorten the time required to determine the best parents for a cross, and to select the most

appropriate progeny resulting from that cross. Genetic linkage maps are a valuable resource which can be utilised as a framework for a number of plant breeding programs; For instance, marker assisted selection (MAS), map based cloning and physical and comparative mapping.

4.1.1 Marker assisted selection

One of the most important objectives to construct a saturated linkage map is MAS. This method is based on presence or absence of markers that tightly link to the gene of interest. A genetic linkage map tagged with agronomically important genes enables a breeder to assess the genetic potential of seedlings as early as the first leaves emerge without actually measuring the phenotypic trait, which reduces the requirements of growing area, maintenance and generation cycles to produce new cultivars (Gardiner et al., 1994). The closer the linkage between a marker and a cosegregating gene, the more precise the indirect selection. Otherwise, the degree of selection error will be larger. This error can be further reduced by using two markers flanking the trait of interest. It has been considered that a maximum distance of 10 cM between the gene and each marker is acceptable because the chance of double crossing over event in this distance is low (Tartarini and Sansavini, 2003). The close association (2 cM) of microsatellite marker PceGA25 with kernel bitterness on almond map derived from cultivar ‘R1000” and the cultivar ‘Desmayo Largueta’ was confirmed in different sweet and bitter almond cultivars. It showed the efficiency of marker assisted selection for kernel bitterness in almond (Sánchez-Pérez et al., 2006a).

4.1.2 Map-based cloning and physical mapping

Map-based cloning and physical mapping are two approaches for genome discovery. High resolution maps decrease the amount of chromosome walking and provide useful information to isolate and clone candidate genes (Wang et al, 2001) Linkage maps are based on recombination frequency , whereas physical maps are based on the actual physical location (in base pairs) of genes and markers. A bacterial artificial chromosome (BAC) library of peach cv. 'Jingyu' was constructed and screened with two markers linked to white flesh and nectarine genes. The results showed ten positive clones which were suitable for map based cloning and physical mapping of peach (Wang et al, 2001).

Currently, physical mapping of the entire peach genome as a model for Rosaceae and tree fruits is underway and the latest sequence and/or map information of the BAC clones, expressed sequence tags (ESTs) and their contigs are publicly available in the genome database for Rosaceae (GDR). Furthermore, peach bacterial artificial chromosome of contiguous set of clones (BAC contigs) anchored to mapped markers can be accessed directly from the *Prunus* reference map ('Texas' × 'Early Gold') and putative genetic and physical map positions of the sequence in question can be compared with mapped peach ESTs (Jung et al, 2004). A physical map of the apple genome from a total of 74,281 BAC clones was constructed and the reliability of contig assembly was evaluated by checking consensus maps of contigs and DNA markers. This map can be used for marker development in targeted chromosome regions, high-resolution mapping and isolation of genes / QTL, comparative genomic analyses of plant chromosomes, and large-scale genome sequencing (Han et al, 2007).

4.1.3 Comparative mapping

Information presented on a linkage map allows predicting the location of other markers and traits on closely related species. Regions that are densely mapped in one species may be filled by markers from the related species. When two or more loci from one species carry similar loci in another species this is known as synteny between them, and if more than two genes or markers are homologous and the relative order of the genes is conserved between them, this is referred to as a conserved linkage (Miller, 1997). Comparative mapping showed the role of conserved regions by demonstrating that such regions are important to the plant and have therefore been conserved through evolution (Kumar, 1999). Moreover, it is interesting to find out what is the function of conserved regions by comparative mapping and positional cloning of the genes in well studied species and compare them with the less studied counterparts (Kumar, 1999).

4.1.4 Construction of genetic linkage maps

The process of genetic mapping requires three critical steps: 1) development of an appropriate mapping population; 2) genotyping of the population including parents and progeny using polymorphic genetic markers, and 3) linkage analysis (Collard et al., 2005).

4.1.4.1 Mapping population

A suitable segregating population derived from sexual reproduction must have sufficient genetic variation with many morphologically different traits to reveal polymorphism. The types of mapping populations vary and each of them has advantages and disadvantages (Paterson, 1996). Generally F_2, backcross (BC) recombinant inbred lines (RILs), double haploids (DH) and near isogenic lines (NILs) have been employed in mapping programs for most of the plant species (Semagn et al., 2006). However, in

woody plants such as almond, production of pure inbred lines is prevented due to long reproductive cycle and self-incompatibility (Socias i Company, 1997). Therefore, a two-way (double) pseudo-testcross (DPSTC) strategy has been employed in these plants due to their crossbred nature (Ritter et al., 1990; Grattapaglia and Sederoff, 1994). A DPSTC mapping population is generated by crossing two distinct heterozygous individuals / varieties (Myburg, et al., 2003). The DPSTC strategy has been successfully used in a number of linkage map constructions such as: *Eucalyptus* (Grattapaglia and Sederoff, 1994); apple (Conner et al., 1997); rose (Debener and Mattiesch, 1999); sour cherry (Wang et al., 2000); chestnut (Casasoli et al., 2001); olive (Wu et al., 2004); apricot (Dondini et al., 2007), passion fruit (Oliveir et al., 2008).

4.1.4.2 Population genotyping

The next step in the development of a linkage map is the identification of genetic polymorphism across the entire mapping population, which is known as genotyping (Collard et al, 2005). Initially, isozymes were used for this purpose but these have been supplanted by the use of DNA-based markers. Many types of DNA markers are available now, but highly informative microsatellites, due to their co-dominant inheritance, reliability and transferability, are the ideal choice for linkage maps (Rossetto, 2001). The use of codominant markers segregating for both parents (anchor markers) can help to compare and integrate maps.

4.1.4.3 Linkage analysis

The final step of genetic mapping involves the calculation of pairwise recombination frequencies between markers, establishment of the linkage groups,

estimation of map distance and determination of the map order by statistical programs (Semagn et al., 2006). During meiosis, each chromosome duplicates and non-sister chromatids may exchange segments (crossing over) and produce recombinant gametes. The proportion of recombinant genotypes can be used to calculate recombination fractions (r), which are related to the distance between loci. Generally the closer the markers are the less chance of recombination and vice versa (Collard et al., 2005). Mapping functions convert the recombination fraction (r) into mapping units (cM). One unit of map distance (1 cM) is equivalent to 1 percent crossing over (Paterson et al., 1988).

Following the scoring of each marker, the data must be converted to a special format to enter into a mapping program for analysis. Several computer packages such as LINKAGE (Suiter et al., 1983), MapMaker (Lander et al., 1987), GMENDEL (Echt et al., 1992), MapManager QTX (Manly et al., 2001) and JoinMap (Van Ooijen et al., 2001) have been developed for data analysis to assist map construction. Computer programs use different algorithms to estimate recombination frequencies of all markers usually by calculation of maximum likelihood or logarithm of odd ratios (LOD) to assign markers into linkage groups. The LOD score, to the base 10 (log), hypothesises the likelihood of no linkage existing ($r = 0.5$ null hypothesis) to the linkage existing ($r < 0.5$ alternative hypothesis) based on expected and observed marker class frequencies in the progeny of a cross (Weising et al., 2005). LOD scores of ≥ 3, which indicates that the linkage between two markers is 1000 times more likely (i.e. 1000:1) than no linkage (null hypothesis), are considered to be strong evidence for linkage and have been used to construct linkage maps (Stam, 1993). When the LOD score is determined then the potential linkage between loci and map distances can be tested with statistical

procedures and data pictorially represented in linkage groups as the linear arrangement of genetic markers (Weising et al., 2005). Between markers at a genetic distance less than 10 cM, the chance of double crossover is very small. For markers that are further apart from each other (> 20 cM), placement on the linkage map is possible but the assignment of their location is more tentative. Three linked markers may be placed in three different orders, and the correct order depends on which one should be placed in the middle. When multiple crossovers occur with a frequency greater than that expected by chance alone, the gene order of closely linked sites can be determined by three factor reciprocal crosses (Staub et al., 1996). If there are enough markers, the map must produce the number of linkage groups corresponding to the chromosome number of the species investigated. Because of the massive number of calculations of pairwise recombinations, the computer package usually produces a framework of markers on to which more markers are added to the framework step by step (Van Ooijen et al., 2001). If the linkage groups are considered as roads, the markers represent sign or landmarks of the roads (Paterson, 1996).

4.1.5 The *Prunus* linkage maps

Discovery of molecular based markers facilitated genome mapping studies. Several linkage maps of *Prunus* have been developed over the last a few years. As a result of the European *Prunus* Mapping project, a comprehensive map based on an interspecific almond (cv. 'Texas' syn. 'Mission') × peach (cv. 'Earlygold') F_2 population was developed using 235 RFLPs and 11 isozymes markers. This map (T × E) consists of eight linkage groups covering 491 cM (Joobeur et al., 1998). A few years later a set of 96 microsatellites were added to this map and increased the total map

distance to 522 cM with the average density of 5.4 cM / microsatellite (Aranzana et al., 2003). Finally, a further set of 220 additional markers (89 microsatellites, five STS and 126 RFLPs) was added to the (T × E) map and provided a saturated *Prunus* map with 562 markers, which defined 519 cM in total with the average density of 0.92 cM / marker (Dirlewanger et al., 2004a). The (T × E) map is widely used as an international *Prunus* reference map and creates a common terminology for linkage groups in several intraspecific and interspecific *Prunus* maps which have been developed using anchor markers (RFLPs, microsatellites, and isozymes) of this map (Dirlewanger et al., 2004a) such as: Foolad et al., 1995; Joobeur et al., 1998; Jáuregui et al., 2001; Bliss et al., 2002 or intraspecific crosses of peach (Chaparro et al., 1994; Rajapakse et al., 1995; Dirlewanger et al., 1998; Lu et al., 1998; Dettori et al., 2001; Yamamoto et al., 2001; Blenda et al., 2007), almond (Viruel et al., 1995; Joobeur et al., 2000), apricot (Hurtado et al., 2002; Lambert et al., 2004), sour cherry (Wang et al., 1998b). The synteny and shared markers between these maps allowed the integration of major genes or QTLs into a unique map (Fig. 4.1; Dirlewanger et al., 2004a). However, most of these genes are placed in low density maps so their location can be further saturated by additional markers found in homologous regions of other maps for use in MAS (Dirlewanger et al., 2004a).

4.1.5.1 The almond maps

The first genetic map of almond based on isozymes and RFLP markers was developed using an intraspecific F_1 cross between the cultivars ‘Ferragnès’ × ‘Tuono’. Eight linkage groups were constructed with 96 loci in ‘Ferragnès’ and 69 loci in ‘Tuono’ maps (Viruel et al., 1995). The second almond map integrated 54 RAPDs and

six microsatellites into this map and increased by 5% the map distance covered,'Ferragnes' (415 cM) and 'Tuono' (416 cM) (Joobeur et al., 2000). Moreover, targeted mapping in almond with the F1 population derived from cultivars 'Felisia' x 'Bertina' allowed the mapping of self-incompatibility (Ballester et al, 1998), shell hardness (Arus et al., 1999) and late blooming phenotypes (Ballester et al, 1998) on this framework map.

NOTE: This figure is included on page106 in the print copy of the thesis held in the University of Adelaide Library.

Fig. 4.1 Approximate position of 28 major genes mapped in different populations of apricot (blue background), peach (orange background), almond or almond x peach (yellow background), and Myrobalan plum (green background) on the framework of the *Prunus* reference map. Gene abbreviations correspond to: *Y*, peach flesh color; *B*, almond/peach petal color; *sharka*, plum pox virus resistance; *B*, flower color in almond x peach; *Mi*, nematode resistance from peach; *D*, almond shell hardness; *Br*, broomy plant habit; *Dl*, double flower; *Cs*, flesh color around the stone; *Ag*, anther color; *Pcp*, polycarpel; *Fc*, flower color; *Lb*, blooming date; *F*, flesh adherence to stone; *D*, non-acid fruit in peach, *Sk*, bitter kernel; *G*, fruit skin pubescence; *Nl*, leaf shape; *Dw*, dwarf plant; *Ps*, male sterility; *Sc*, fruit skin color; *Gr*, leaf color; *S**, fruit shape; *S*, self-incompatibility (almond and apricot); *Ma*, nematode resistance from Myrobalan plum; *E*, leaf gland shape; *Sf*, resistance to powdey mildew. Genes *Dl* and *Br* are located on an unknown position of G2 (Dirlewanger et al, 2004a).

Recently a low resolution map of almond based on 56 microsatellites and two major genes, kernel taste (Sk) and self-incompatibility (S), and ten QTLs has been reported based on the F_1 population obtained from a cross between cultivars 'R1000' and 'Desmayo Largueta'. The authors suggested further studies with different crosses and a more saturated map would be beneficial to MAS in breeding programs (Sánchez-Pérez et al, 2007b).

Comparison between different maps of *Prunus spp*. showed genome synteny between the markers and high levels of conservation of gene order. However, even between different cultivars of the same species variability exists in some regions of the maps (Viruel et al., 1995). The first almond map (Viruel et al., 1995) heavily relied on isozymes and RFLPs, both of which required time consuming and complex methods and showed low polymorphism. The second almond map (Joobeur et al., 2000) showed higher saturation with mostly RAPD markers which are not transferable between maps. The most recently published almond map (Sánchez-Pérez et al., 2007), which was based on microsatellites, produced a low density map with 48 microsatellites for the maternal map ('R1000') and 45 markers for the paternal map ('Desmayo Largueta').

Australian almond map

The first Australian almond maps derived from the cross NP × LA was produced based on RAPDs, ISSRs, microsatellites and self-incompatibility allele (Chapter 1). Nonpareil is the most widely cultivated almond in the world and is the main cultivar in The United States of America and Australia. 'Nonpareil' is planted in more than a third (241,086 acres) of the almond cultivation area in California (USDA, 2007). This cultivar is widely used in breeding programs but to date only a low density linkage map

of NP is available (Gregory, 2004). Three maps, for NP and LA and an Integrated map (IN) were developed. However, these maps failed to reduce the number of linkage groups to the haploid chromosome number of almond (x= 8). Moreover, many large gaps remained between the markers. The results showed that the maps are incomplete and unsaturated. For future use in MAS and similar accelerated breeding strategies, a greater marker density is needed. In this study we aimed to saturate an intraspecific NP × LA map to enable resolution of the expected number of linkage groups, and to fill the gaps in the existing map to provide maps for future QTL mapping and MAS as a short term goal and framework map for physical mapping and map-based cloning as a long term objective in almond breeding.

4.2 Materials and Methods

4.2.1 Mapping population and DNA extraction

The intraspecific F_1 population of 93 full-sibs derived from the cross between 'Nonpareil' as maternal and 'Lauranne' as paternal parent was used as the mapping population. This was the same population used by Gregory et al., (2005) for the construction of the first Australian almond linkage map. Total genomic DNA was extracted from fresh young leaves of parents and progeny using the optimised protocol of Lamboy and Alpha, (1998) by Gregory, (2004) described in Chapter 2. DNA quantity and quality was measured by Nanodrop spectrophotometer ND-1000® (Thermo Scientific, USA) at the absorption A_{260}/A_{280} ratio. Concentrated DNA then was diluted to a final stock concentration of 20 $ng.\mu l^{-1}$ with TE buffer (10mM Tris, 1 mM EDTA pH 8.0) and kept at -20°C for subsequent PCR amplification.

4.2.2 Microsatellites marker analysis

A total of 195 microsatellites isolated in different *Prunus* species were screened in the parents and selected progeny for polymorphism test. The PCR amplification followed authors' recommendations and electrophoresis was performed on conventional agarose, MetaPhor agarose, 8% (w/v) polyacrylamide, or automated capillary gels to visualise PCR products (Chapter 2). Most of the complex or multilocus markers were excluded with minor exceptions and only markers with good reproducibility and easily interpretable loci were chosen for the analyses. The genotyping data were scored and coded according to the coding system required by JoinMap 3.0 (Van Ooijen and Voorrips, 2001). All scoring was also conducted independantly by Dr Shubiao Wu, another investigator of almond mapping project, and confirmed data were used for map construction.

4.2.3 Verification of microsatellites as MAS for self-compatibility

To validate the efficiency for MAS of two markers (CPPCT021 and UDP98-412) linked to the self-compatibility alleles, the markers were tested in cultivars with different genetic backgrounds other than the mapping population and parents (Fig. 4.2). Self-compatibility was found at the end of linkage group six between microsatellites CPPCT021 and UDP98-412. The PCR amplification and electrophoresis condition on 8% polyacrylamide gel was the same as other microsatellites presented in chapter 2.

4.2.4 Linkage analysis and map construction

To evaluate the deviation from the expected Mendelian segregation ratios, chi square (χ^2) goodness-of-fit tests were performed from the '*locus genotype frequency*'

command of JoinMap program. All the suspicious markers and some of the distorted markers were then excluded from the dataset. Two datasets were organised for the parents: 'Nonpareil' and 'Lauranne'. Heterozygous markers in both parents were included in both maps and used as anchor markers between parental maps. The linkage analysis was performed using JoinMap 3.0 software (Van Ooijen and Voorrips, 2001) using the cross pollination (CP) format. For each parent the map was constructed in two phases. First, all distorted markers were excluded from the dataset and the linkage groups were established at the setting of a minimum LOD score of ≥ 6 with the '*LOD groupings'* command and recombination fraction of 0.49. Once maps were constructed, the fixed order for each linkage group was used for the next step.

The maps obtained above provide a framework map with markers placed on linkage groups accurately and confidently. In the second phase, excluded markers in the first phase were included into the dataset and maps were constructed using the fixed order from the first phase. When a marker presence caused inconsistencies in the maps, it was eliminated from the dataset before final mapping. The Kosambi mapping function (Kosambi 1944) was used to convert the recombination frequencies into genetic distances in centiMorgan (cM). An integrated parental genetic map was constructed with all loci in each parent and heterozygous loci in both parents using the map integration function of JoinMap v.3.0. The original map data were compiled for map drawing, and the map was drawn using MapChart 2.1 (Voorrips, 2002).

4.3 Results

4.3.1 Construction of the second Australian linkage maps for almond

Of the 195 microsatellites studied, 54 were clearly polymorphic and therefore selected for map construction. Three genetic linkage maps for 'Nonpareil', 'Lauranne' and an integrated map were constructed with a total of 98, 100 and 129 markers, respectively. The eight linkage groups for each map obtained correspond to the chromosome number of almond (x = 8) (Fig. 4.3 - 4.5). Microsatellites segregate co-dominantly as 1:1, 1:2:1 or 1:1:1:1 and 60 (85.7%) of them were heterozygous in both parents and showed the most informative segregation type of 1:1:1:1 which could be used as anchor loci. The distribution of mapped molecular markers across eight linkage groups with their genome length coverage, loci type, number of skewed loci, average density of markers, and anchor loci are shown in Table 4.1.

4.3.1.1 NP linkage map

The first complete NP map was developed with the total length of 541.8 cM with the longest linkage group 1 (110.2 cM) and the shortest linkage group, 2, covering 47 cM. Chi-square analysis of goodness-of-fit showed three markers did not segregate according to the Mendelian mode of inheritance (MA020, CPSCT042 linkage group 7 and RAPD marker OPB10-1033 in linkage group 8). The map comprises 73 microsatellites and three large gaps. The average density is 5.53 markers / centiMorgan.

4.3.1.2 LA linkage map

The LA map covered 534.6 cM and consist of 70 microsatellites, 22 ISSRs, five RAPDs, one SCAR and two self-compatibility alleles. This map contains five distorted

markers, three of them placed on linkage group 7. LA map covers an average map density of 5.34 cM / marker, and encompass of 42 the anchor markers from the *Prunus* reference map. Linkage group 4 had 16 markers and linkage group 8 contains the fewest markers (8). The largest gap, of 29 cM between markers EPDCU3392 and AG8YC-820, is located on linkage group 7 of the LA map.

4.3.1.3 The integrated linkage map

The integrated map, composed of all markers that segregated in the parents, produced the longest map (592.4 cM) with an average density of 4.59 cM / marker. The linkage groups consist of between 12 and 19 markers. Linkage group 1 showed the longest and the most populated linkage group, covering a map distance of 114.8 cM. Marker density varies from 3.2 in G5 to 6.5 cM / marker in linkage group 7. Three large gaps (≥ 20 cM) were found in linkage groups of 1, 3 and 7 and the remaining linkage groups showed good coverage, especially in G6 and 8.

4.3.2 Comparison among the new maps and previous Australian almond maps

The synteny between two parental maps was completely conserved, with the exception of four minor inversions in the order of loci (colinearity) in G3 (UDA-033), G6 (CPDCT006, CPPCT021) and G8 (UDP96-019). The construction of the new maps substantially increased map distance and density compared with the first Australian almond map developed by Drs Gregory and Mnejja (2004, unpublished data; Chapter 1). LA map distance increased 40% from 320 cM to 534.6, NP map increased from 482.6 cM to 541.8 (10 %) and IN map increased from 360.9 cM to 592.4 (40 %) compared with previous map. Also the number of mapped markers increased

dramatically from 50 and 59 mapped markers in previous LA and NP map to 100 and 98 markers in new LA and NP maps.

4.3.3 Marker assisted selection test for self-compatibility

Almond cultivar 'Lauranne' is self-fertile but 'Nonpareil' is self-incompatible; from 93 progeny of the LA×NP cross screened, 47 showed *Sf* and 46 self-incompatibility (Gregory, 2004), in agreement with self-incompatibility inheritance of 1:1 as a single gene locus (Gregory et al., 2005). Self-compatibility locus was located in the distal part of linkage groups 6 and closely linked to the markers UDP98-412 and CPPCT021, therefore it was assumed to be applicable for MAS but the amplification of these markers could not discriminate between self-compatible and self-incompatible cultivars and almost the same banding pattern was found in both self fertile and self-incompatible cultivars (Fig. 4.2).

4.3.4 Comparative mapping with other *Prunus* maps

The integrated map was aligned for synteny and colinearity analysis with the *Prunus* reference map (Dirlewanger et al., 2004a, Howad et al, 2005**)** and the latest published almond map (Sánchez-Pérez et al., 2007b) (Fig. 4.5). All the common markers on the *Prunus* reference map and on many other published maps of *Prunus spp*., including Myrobalan plum (Dirlewanger et al., 2004b); *Prunus* bin map in genome database for Rosaceae (GDR), and the most recent apricot map (Dondini et al., 2007) showed complete synteny. Also, the marker order almost remained conserved with some minor changes in linkage groups. 73 loci showed synteny between IN (integrated) map and the *Prunus* reference map. To the best of our knowledge, 11 microsatellites have been mapped for the first time (CPDCT006, CPDCT007, CPDCT018, EMPA004,

EPDCU2584EPDCU4658, Pchgms29, UCD-CH15, UDAp-479, UDAp-483 and UDA-014).

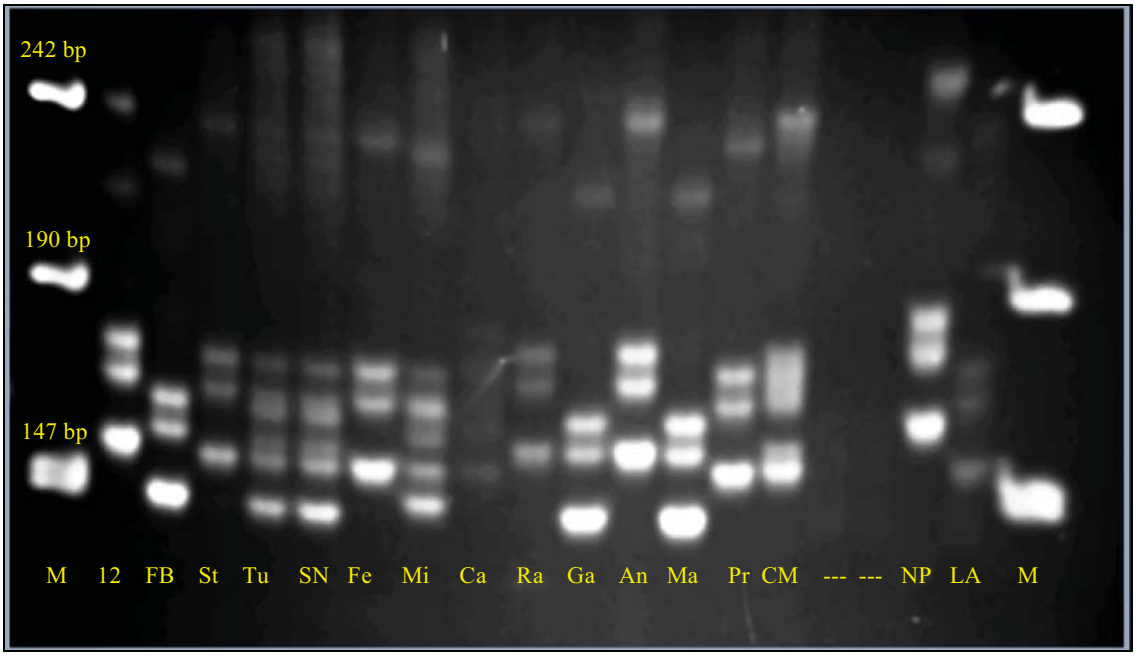


Fig. 4.2 Segregation of microsatellite UDP98-412 on Self-compatible and incompatible almond cultivars showed almost the similar banding patterns between both genotypes. Self-compatible cultivars: 12-13 (12), Falsa Barese (FB), Steliette (St), Tuono (Tu), Supernova (SN) and Lauranne (LA). Self-incompatible cultivars: Ferragnès (Fe), Mission (Mi), Carmel (Ca), Ramillete (Ra), Gabaix (Ga), Anxaneta (An), Marcona (Ma), Primorskiy (Pr), Cristomorto (CM) and Nonpareil (NP).

4.4 Discussion

Two maps for 'Nonpareil' and 'Lauranne' were constructed by the addition of selected microsatellites, 98 and 100 markers were mapped on eight linkage groups that correspond to the eight chromosomes of almond as expected (Viruel et al., 1995; Joober et al., 2000). Karyotype study by fluorescence *in situ* hybridisation (FISH) in almond demonstrated a set of eight chromosomes with a size ranging between 1.1 µm to 2.4 µm and the centromer located at median or submedian positions (Corredor et al., 2004). In karyotype studies, chromosomes have been classified as the longest chromosome receives the first number and the shortest chromosome obtains the last chromosome number. Although G1 in this study produce the longest linkage groups similar to

almond karyotypes, one-to-one association for other linkage groups is doubtful unless markers mapped in different Gs can be assigned physically on a given chromosome using the FISH technique (Corredor et al., 2004).

The maps developed in this study improved the maps of Gregory et al., (2005) by detecting the missing groups, increasing the marker density (from 10 cM/marker to 4.59 cM/marker for IN map) and extending the length of linkage groups. Homologous linkage groups to the *Prunus* reference map showed that no spurious linkage between markers had occurred. The integrated map was produced by all polymorphic markers of both parents. Three major types of DNA markers successfully employed in this study have been suggested to be useful for completing similar genetic mapping projects in other fruit trees. Well evaluated RAPDs and ISSRs marker have been used for quick establishment of a map in a short time and microsatellites can be employed for saturation of targeted regions and produce evenly distributed loci. Microsatellites covered 476.2 cM (80.4%) of the total 592.4 cM of the integrated map and provided good coverage of almond genome. Using 152 microsatellites Riaz et al., (2004) produced a genetic map of the cultivated grapevine *Vitis vinifera* that extended over 1,728 cM. Similarly, using 443 RAPD markers resulted in a map covering 1,300 cM (Lodhi et al., 1995) and with 312 AFLP and 39 microsatellites, a map with 1,639 cM of genome coverage was developed for the same species (Grando et al., 2003). It seems that compared to other classes of markers such as RAPD and RFLP, more information has been provided by the use of a smaller number of microsatellites, which can decrease the cost of map construction.

Table 4.1 Distribution of mapped markers and their coverage across linkage groups of 'Nonpareil' (NP), 'Lauranne' (LA) and Integrated (IN) maps

	G1			G2			G3			G4			G5			G6			G7			G8			Total		
	NP	LA	IN	NP	LA	IN	NP	LA	IN	NP	LA	IN	NP	LA	IN	NP	LA	IN	NP	LA	IN	NP	LA	IN	NP	LA	IN
Number of markers	17	15	19	12	10	14	13	12	18	11	16	18	13	13	17	13	14	17	9	12	14	10	8	12	98	100	129
Micro-satellites	13	11	14	11	9	12	9	8	12	7	9	9	10	8	11	10	11	11	7	10	11	6	4	6	73	70	86
ISSR	4	4	5	1	1	2	4	2	4	1	4	4	3	5	6	2	1	3	2	2	3	2	3	3	19	22	30
RAPD	--	--	--	--	--	--	--	1	1	3	3	5	--	--	--	1	--	1	--	--	--	2	1	3	6	5	10
SCAR	--	--	--	--	--	--	--	1	1	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	1	1
self incompat-ibility	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	2	2	--	--	--	--	--	--	--	2	2
Large gap ≥ 20 cM	1	1	1	--	--	--	1	--	1	--	--	--	--	--	--	--	--	--	1	1	1	--	--	--	3	2	3
Length (cM)	110.2	104.8	114.8	47	59.3	58.1	76.6	46.3	75.3	64	78.6	77.9	56.9	46.9	55.1	59.8	67.5	65.5	76.9	83.4	83.8	50.4	47.8	53.3	541.8	534.6	592.4
Ave density of markers / cM	6.5	7.0	6	3.9	5.9	4.2	5.9	3.9	4.2	5.8	4.9	4.3	4.4	3.6	3.2	4.6	4.8	3.3	8.5	6	6.5	5	6	4.4	5.53	5.34	4.59
No distorted markers	--	--	--	--	--	--	--	1	1	--	1	1	--	--	--	--	--	--	2	3	3	1	1	2	3	6	7
No anchored to T×E map	11	10	12	11	12	12	8	6	10	4	6	6	8	5	8	7	8	9	7	9	10	6	4	6	62	57	73

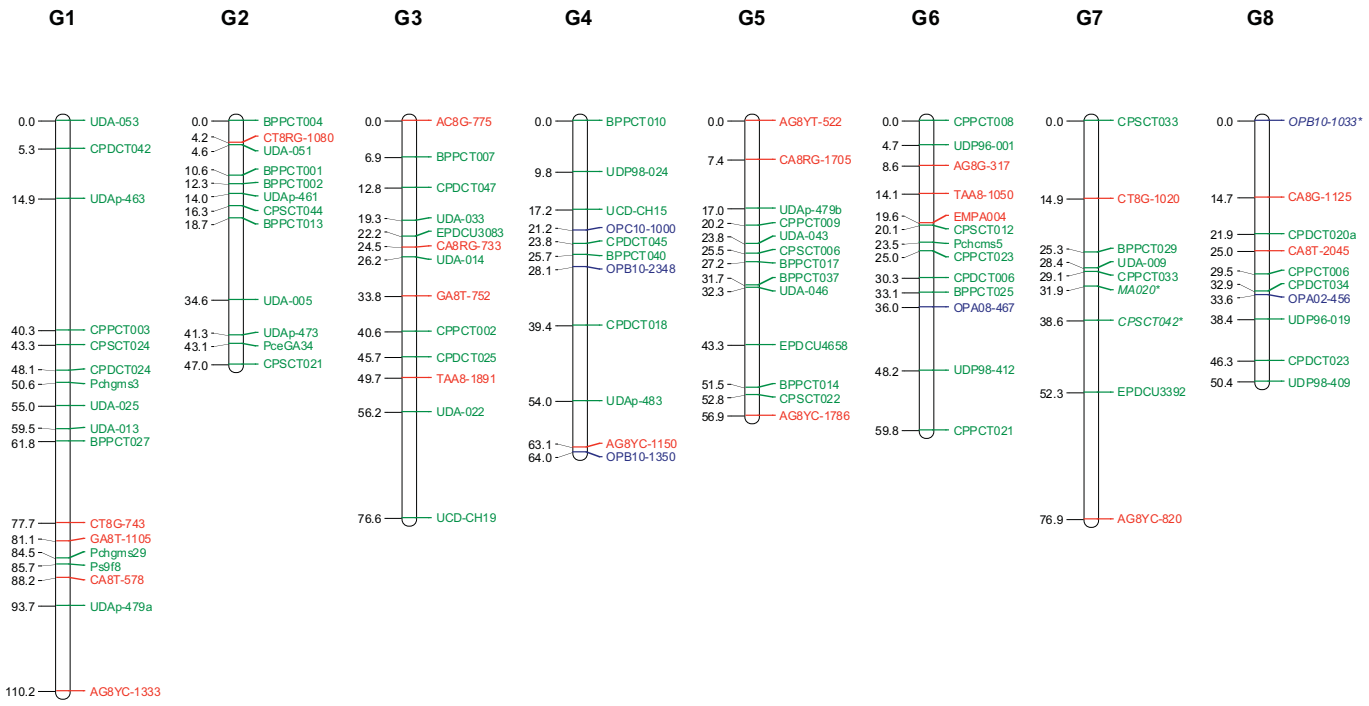

G1
0.0 UDA-053
5.3 CPDCT042
14.9 UDAp-463
40.3 CPPCT003
43.3 CPSCT024
48.1 CPDCT024
50.6 Pchgms3
55.0 UDA-025
59.5 UDA-013
61.8 BPPCT027
77.7 CT8G-743
81.1 GA8T-1105
84.5 Pchgms29
85.7 Ps9f8
88.2 CA8T-578
93.7 UDAp-479a
110.2 AG8YC-1333
G2
0.0 BPPCT004
4.2 CT8RG-1080
4.6 UDA-051
10.6 BPPCT001
12.3 BPPCT002
14.0 UDAp-461
16.3 CPSCT044
18.7 BPPCT013
34.6 UDA-005
41.3 UDAp-473
43.1 PceGA34
47.0 CPSCT021
G3
0.0 AC8G-775
6.9 BPPCT007
12.8 CPDCT047
19.3 UDA-033
22.2 EPDCU3083
24.5 CA8RG-733
26.2 UDA-014
33.8 GA8T-752
40.6 CPPCT002
45.7 CPDCT025
49.7 TAA8-1891
56.2 UDA-022
76.6 UCD-CH19
G4
0.0 BPPCT010
9.8 UDP98-024
17.2 UCD-CH15
21.2 OPC10-1000
23.8 CPDCT045
25.7 BPPCT040
28.1 OPB10-2348
39.4 CPDCT018
54.0 UDAp-483
63.1 AG8YC-1150
64.0 OPB10-1350
G5
0.0 AG8YT-522
7.4 CA8RG-1705
17.0 UDAp-479b
20.2 CPPCT009
23.8 UDA-043
25.5 CPSCT006
27.2 BPPCT017
31.7 BPPCT037
32.3 UDA-046
43.3 EPDCU4658
51.5 BPPCT014
52.8 CPSCT022
56.9 AG8YC-1786
G6
0.0 CPPCT008
4.7 UDP96-001
8.6 AG8G-317
14.1 TAA8-1050
19.6 EMPA004
20.1 CPSCT012
23.5 Pchcms5
25.0 CPPCT023
30.3 CPDCT006
33.1 BPPCT025
36.0 OPA08-467
48.2 UDP98-412
59.8 CPPCT021
G7
0.0 CPSCT033
14.9 CT8G-1020
25.3 BPPCT029
28.4 UDA-009
29.1 CPPCT033
31.9 MA020*
38.6 CPSCT042*
52.3 EPDCU3392
76.9 AG8YC-820
G8
0.0 OPB10-1033*
14.7 CA8G-1125
21.9 CPDCT020a
25.0 CA8T-2045
29.5 CPPCT006
32.9 CPDCT034
33.6 OPA02-456
38.4 UDP96-019
46.3 CPDCT023
50.4 UDP98-409

Fig. 4.3 Genetic linkage map of Nonpareil. The vertical line represents the linkage groups with marker loci on the right. Distorted loci (P<0.05) are italic and indicated by an asterisk following the locus name. Distances between markers are given in centiMorgans. Microsatellite markers are coded in green, ISSRs in red, RAPDs in blue.

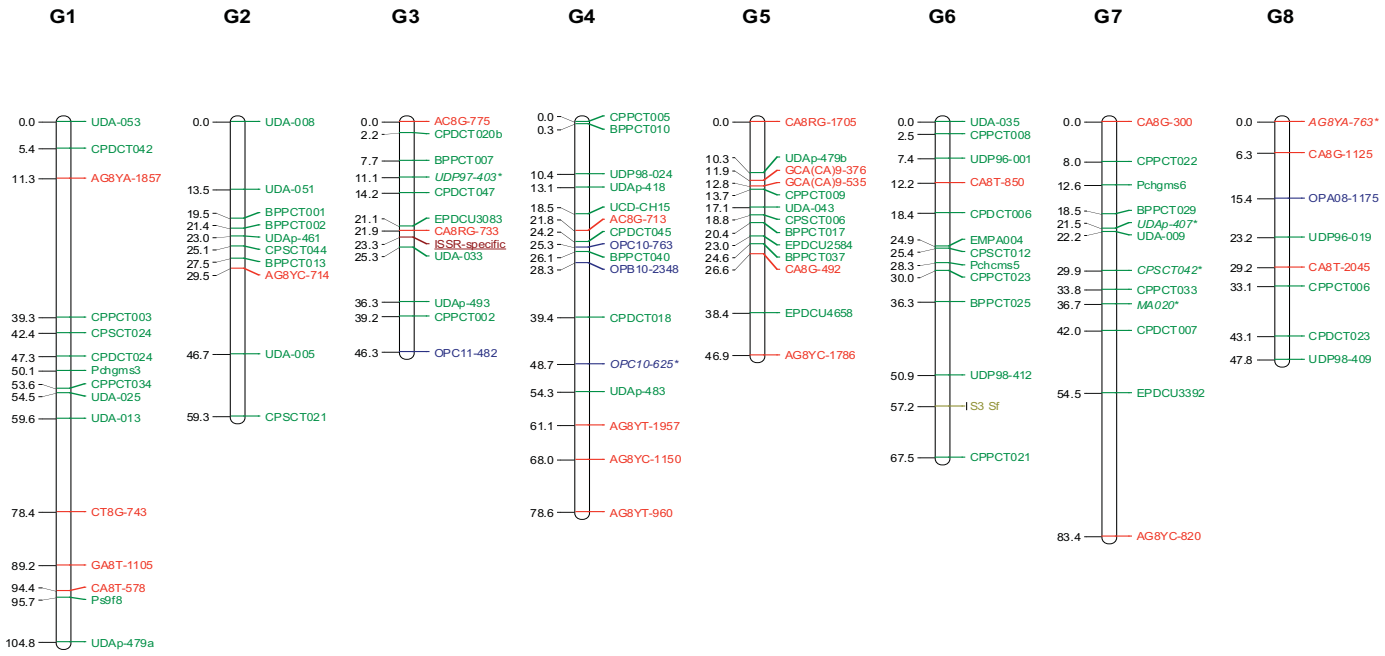
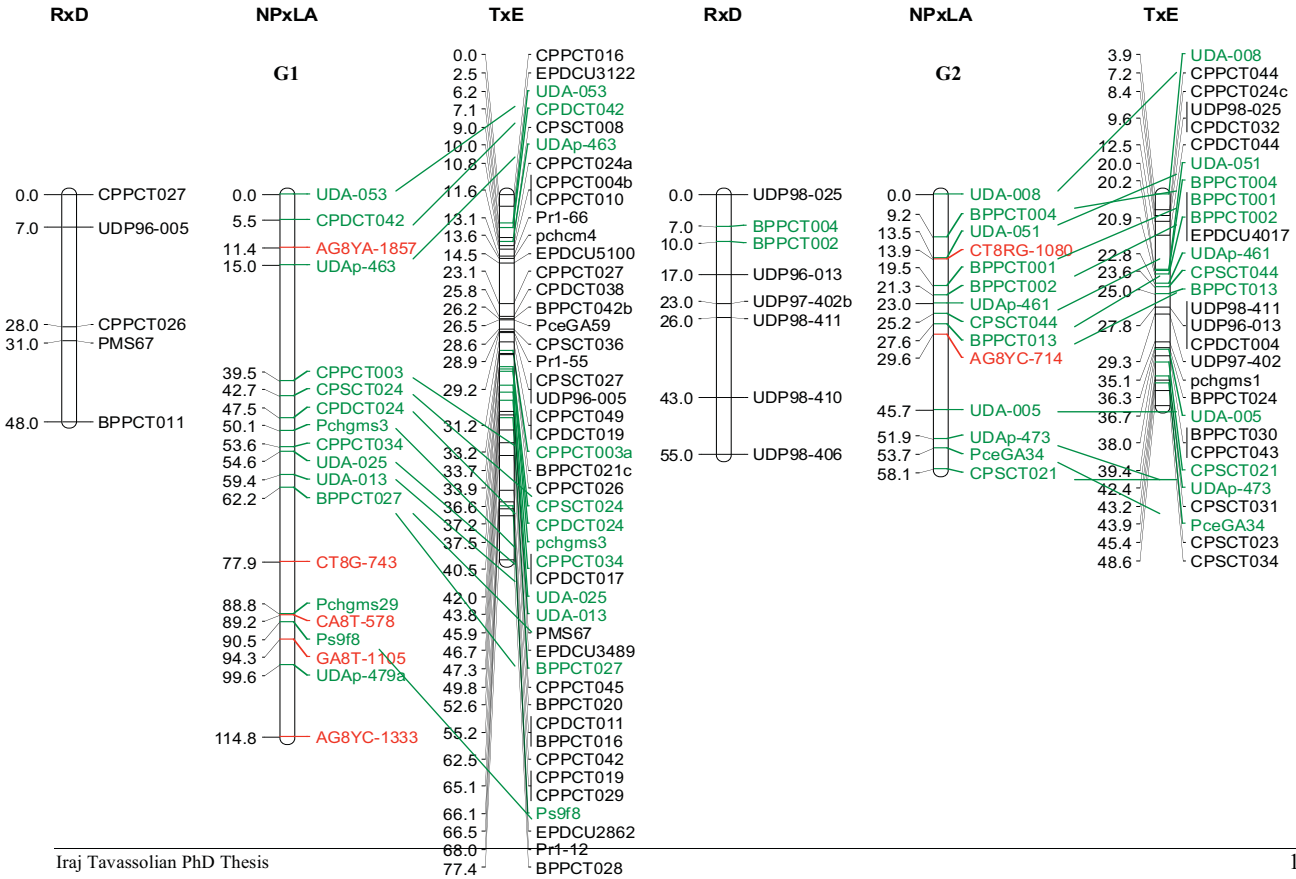
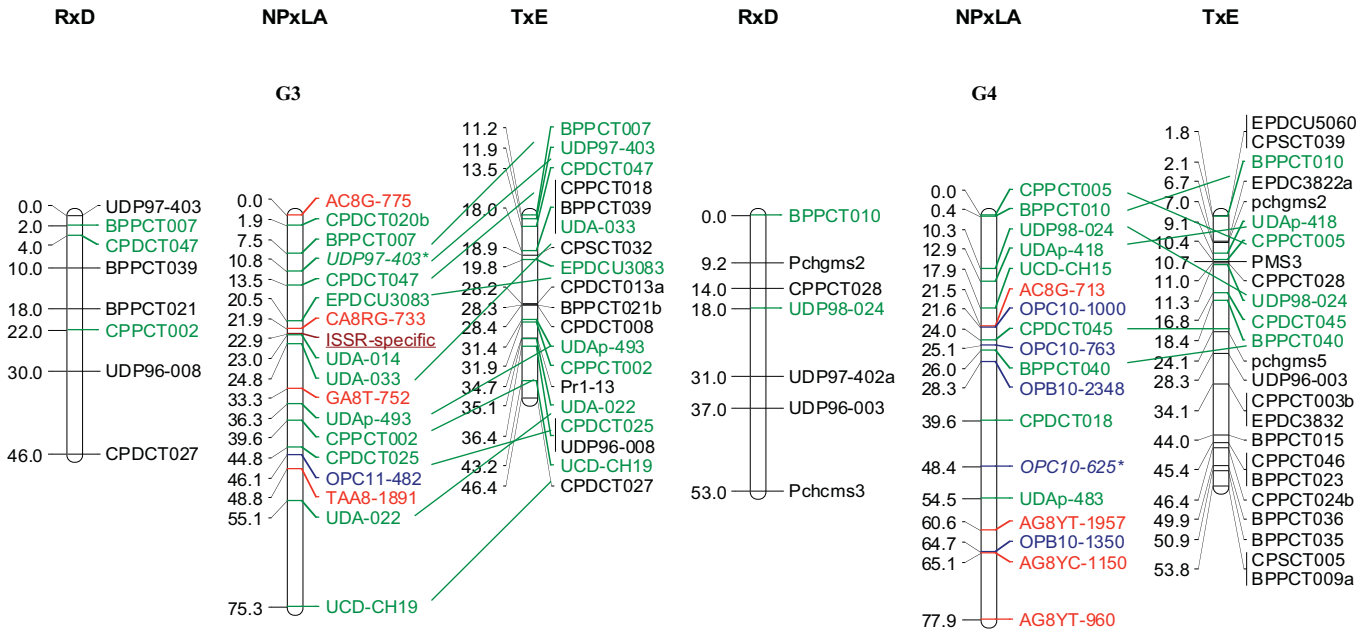
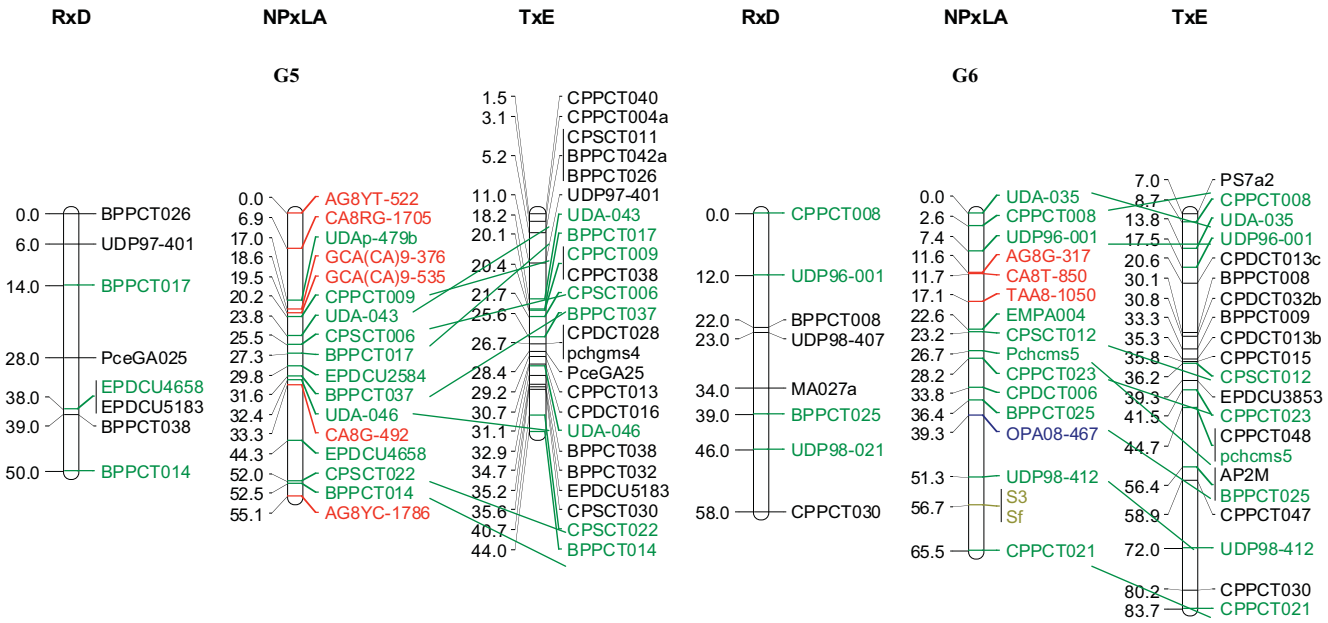


Fig. 4. 4 Genetic linkage map of Lauranne, the vertical line represents the linkage groups with marker loci on the right. Distorted loci (P<0.05) are italic and indicated by an asterisk after the name. Distances between markers are given in centiMorgans. Microsatellite markers are coded in green, ISSRs in red, RAPDs in blue, SCAR in brown and self-compatibility in olive green.


RxD
NPxLA
TxE
G1
0.0 CPPCT027
7.0 UDP96-005
28.0 CPPCT026
31.0 PMS67
48.0 BPPCT011
0.0 UDA-053
5.5 CPDCT042
11.4 AG8YA-1857
15.0 UDAp-463
39.5 CPPCT003
42.7 CPSCT024
47.5 CPDCT024
50.1 Pchgms3
53.6 CPPCT034
54.6 UDA-025
59.4 UDA-013
62.2 BPPCT027
77.9 CT8G-743
88.8 Pchgms29
89.2 CA8T-578
90.5 Ps9f8
94.3 GA8T-1105
99.6 UDAp-479a
114.8 AG8YC-1333
0.0 CPPCT016
2.5 EPDCU3122
6.2 UDA-053
7.1 CPDCT042
9.0 CPSCT008
10.0 UDAp-463
10.8 CPPCT024a
11.6 CPPCT004b
CPPCT010
13.1 Pr1-66
13.6 pchcm4
14.5 EPDCU5100
23.1 CPPCT027
25.8 CPDCT038
26.2 BPPCT042b
26.5 PceGA59
28.6 CPSCT036
28.9 Pr1-55
29.2 CPSCT027
UDP96-005
CPPCT049
CPDCT019
31.2 CPPCT003a
33.2 BPPCT021c
33.7 CPPCT026
33.9 CPSCT024
36.6 CPDCT024
37.2 pchgms3
37.5 CPPCT034
40.5 CPDCT017
42.0 UDA-025
43.8 UDA-013
45.9 PMS67
46.7 EPDCU3489
47.3 BPPCT027
49.8 CPPCT045
52.6 BPPCT020
55.2 CPDCT011
BPPCT016
62.5 CPPCT042
CPPCT019
65.1 CPPCT029
66.1 Ps9f8
66.5 EPDCU2862
68.0 Pr1-12
77.4 BPPCT028
RxD
NPxLA
TxE
G2
0.0 UDP98-025
7.0 BPPCT004
10.0 BPPCT002
17.0 UDP96-013
23.0 UDP97-402b
26.0 UDP98-411
43.0 UDP98-410
55.0 UDP98-406
0.0 UDA-008
9.2 BPPCT004
13.5 UDA-051
13.9 CT8RG-1080
19.5 BPPCT001
21.3 BPPCT002
23.0 UDAp-461
25.2 CPSCT044
27.6 BPPCT013
29.6 AG8YC-714
45.7 UDA-005
51.9 UDAp-473
53.7 PceGA34
58.1 CPSCT021
3.9 UDA-008
7.2 CPPCT044
8.4 CPPCT024c
9.6 UDP98-025
CPDCT032
12.5 CPDCT044
20.0 UDA-051
20.2 BPPCT004
BPPCT001
20.9 BPPCT002
EPDCU4017
22.8 UDAp-461
23.6 CPSCT044
25.0 BPPCT013
UDP98-411
27.8 UDP96-013
CPDCT004
29.3 UDP97-402
35.1 pchgms1
36.3 BPPCT024
36.7 UDA-005
38.0 BPPCT030
CPPCT043
39.4 CPSCT021
42.4 UDAp-473
43.2 CPSCT031
43.9 PceGA34
45.4 CPSCT023
48.6 CPSCT034


RxD
NPxLA
TxE
G3
0.0 UDP97-403
2.0 BPPCT007
4.0 CPDCT047
10.0 BPPCT039
18.0 BPPCT021
22.0 CPPCT002
30.0 UDP96-008
46.0 CPDCT027
0.0 AC8G-775
1.9 CPDCT020b
7.5 BPPCT007
10.8 UDP97-403*
13.5 CPDCT047
20.5 EPDCU3083
21.9 CA8RG-733
22.9 ISSR-specific
23.0 UDA-014
24.8 UDA-033
33.3 GA8T-752
36.3 UDAp-493
39.6 CPPCT002
44.8 CPDCT025
46.1 OPC11-482
48.8 TAA8-1891
55.1 UDA-022
75.3 UCD-CH19
11.2 BPPCT007
11.9 UDP97-403
13.5 CPDCT047
18.0 CPPCT018
BPPCT039
UDA-033
18.9 CPSCT032
19.8 EPDCU3083
28.2 CPDCT013a
28.3 BPPCT021b
28.4 CPDCT008
31.4 UDAp-493
31.9 CPPCT002
34.7 Pr1-13
35.1 UDA-022
36.4 CPDCT025
UDP96-008
43.2 UCD-CH19
46.4 CPDCT027
RxD
NPxLA
TxE
G4
0.0 BPPCT010
9.2 Pchgms2
14.0 CPPCT028
18.0 UDP98-024
31.0 UDP97-402a
37.0 UDP96-003
53.0 Pchcms3
0.0 CPPCT005
0.4 BPPCT010
10.3 UDP98-024
12.9 UDAp-418
17.9 UCD-CH15
21.5 AC8G-713
21.6 OPC10-1000
24.0 CPDCT045
25.1 OPC10-763
26.0 BPPCT040
28.3 OPB10-2348
39.6 CPDCT018
48.4 OPC10-625*
54.5 UDAp-483
60.6 AG8YT-1957
64.7 OPB10-1350
65.1 AG8YC-1150
77.9 AG8YT-960
1.8 EPDCU5060
CPSCT039
2.1 BPPCT010
6.7 EPDC3822a
7.0 pchgms2
9.1 UDAp-418
10.4 CPPCT005
10.7 PMS3
11.0 CPPCT028
11.3 UDP98-024
16.8 CPDCT045
18.4 BPPCT040
24.1 pchgms5
28.3 UDP96-003
34.1 CPPCT003b
EPDC3832
44.0 BPPCT015
45.4 CPPCT046
BPPCT023
46.4 CPPCT024b
49.9 BPPCT036
50.9 BPPCT035
53.8 CPSCT005
BPPCT009a


RxD
NPxLA
TxE
G5
0.0 BPPCT026
6.0 UDP97-401
14.0 BPPCT017
28.0 PceGA025
38.0 EPDCU4658
EPDCU5183
39.0 BPPCT038
50.0 BPPCT014
0.0 AG8YT-522
6.9 CA8RG-1705
17.0 UDAp-479b
18.6 GCA(CA)9-376
19.5 GCA(CA)9-535
20.2 CPPCT009
23.8 UDA-043
25.5 CPSCT006
27.3 BPPCT017
29.8 EPDCU2584
31.6 BPPCT037
32.4 UDA-046
33.3 CA8G-492
44.3 EPDCU4658
52.0 CPSCT022
52.5 BPPCT014
55.1 AG8YC-1786
1.5 CPPCT040
3.1 CPPCT004a
CPSCT011
5.2 BPPCT042a
BPPCT026
11.0 UDP97-401
18.2 UDA-043
20.1 BPPCT017
20.4 CPPCT009
CPPCT038
21.7 CPSCT006
25.6 BPPCT037
26.7 CPDCT028
pchgms4
28.4 PceGA25
29.2 CPPCT013
30.7 CPDCT016
31.1 UDA-046
32.9 BPPCT038
34.7 BPPCT032
35.2 EPDCU5183
35.6 CPSCT030
40.7 CPSCT022
44.0 BPPCT014
RxD
NPxLA
TxE
G6
0.0 CPPCT008
12.0 UDP96-001
22.0 BPPCT008
23.0 UDP98-407
34.0 MA027a
39.0 BPPCT025
46.0 UDP98-021
58.0 CPPCT030
0.0 UDA-035
2.6 CPPCT008
7.4 UDP96-001
11.6 AG8G-317
11.7 CA8T-850
17.1 TAA8-1050
22.6 EMPA004
23.2 CPSCT012
26.7 Pchcms5
28.2 CPPCT023
33.8 CPDCT006
36.4 BPPCT025
39.3 OPA08-467
51.3 UDP98-412
56.7 S3
Sf
65.5 CPPCT021
7.0 PS7a2
8.7 CPPCT008
13.8 UDA-035
17.5 UDP96-001
20.6 CPDCT013c
30.1 BPPCT008
30.8 CPDCT032b
33.3 BPPCT009
35.3 CPDCT013b
35.8 CPPCT015
36.2 CPSCT012
39.3 EPDCU3853
41.5 CPPCT023
CPPCT048
44.7 pchcms5
AP2M
56.4 BPPCT025
58.9 CPPCT047
72.0 UDP98-412
80.2 CPPCT030
83.7 CPPCT021

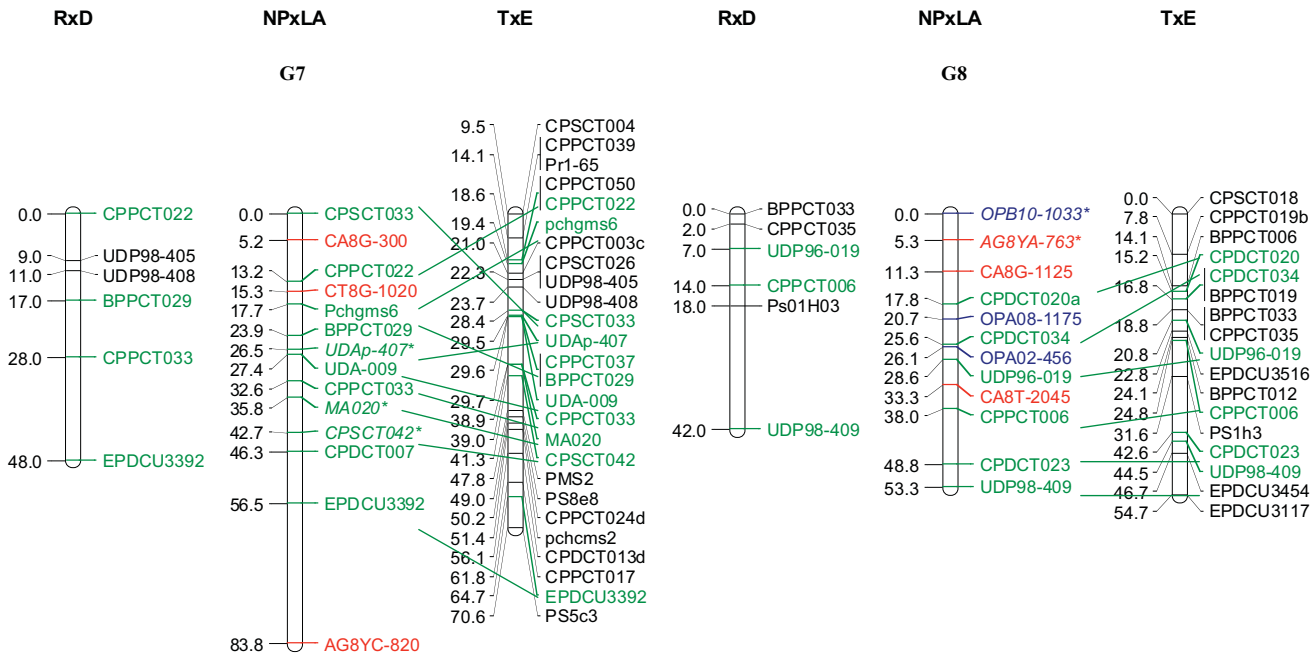

RxD
NPxLA
TxE
G7
0.0 CPPCT022
9.0 UDP98-405
11.0 UDP98-408
17.0 BPPCT029
28.0 CPPCT033
48.0 EPDCU3392
0.0 CPSCT033
5.2 CA8G-300
13.2 CPPCT022
15.3 CT8G-1020
17.7 Pchgms6
23.9 BPPCT029
26.5 UDAp-407*
27.4 UDA-009
32.6 CPPCT033
35.8 MA020*
42.7 CPSCT042*
46.3 CPDCT007
56.5 EPDCU3392
83.8 AG8YC-820
9.5 CPSCT004
14.1 CPPCT039
Pr1-65
18.6 CPPCT050
19.4 CPPCT022
21.0 pchgms6
CPPCT003c
22.3 CPSCT026
UDP98-405
23.7 UDP98-408
28.4 CPSCT033
29.5 UDAp-407
CPPCT037
29.6 BPPCT029
UDA-009
29.7
38.9 CPPCT033
39.0 MA020
41.3 CPSCT042
47.8 PMS2
49.0 PS8e8
50.2 CPPCT024d
51.4 pchcms2
56.1 CPDCT013d
61.8 CPPCT017
64.7 EPDCU3392
70.6 PS5c3
RxD
NPxLA
TxE
G8
0.0 BPPCT033
2.0 CPPCT035
7.0 UDP96-019
14.0 CPPCT006
18.0 Ps01H03
42.0 UDP98-409
0.0 OPB10-1033*
5.3 AG8YA-763*
11.3 CA8G-1125
17.8 CPDCT020a
20.7 OPA08-1175
25.6 CPDCT034
26.1 OPA02-456
28.6 UDP96-019
33.3 CA8T-2045
38.0 CPPCT006
48.8 CPDCT023
53.3 UDP98-409
0.0 CPSCT018
7.8 CPPCT019b
14.1 BPPCT006
15.2 CPDCT020
16.8 CPDCT034
BPPCT019
18.8 BPPCT033
CPPCT035
20.8 UDP96-019
22.8 EPDCU3516
24.1 BPPCT012
24.8 CPPCT006
31.6 PS1h3
42.6 CPDCT023
44.5 UDP98-409
46.7 EPDCU3454
54.7 EPDCU3117

Fig. 4.5 Comparison of the IN map with almond RxD Sánchez-Pérez et al, (2007) and *Prunus* reference map (Dirlewanger et al., 2004a**;** Howad et al, 2005). Distorted loci ($P<0.05$) are italic and indicated by an asterisk after the name. Distances between markers are given in centi Morgans. Microsatellite markers are coded in green, ISSRs in red, RAPDs in blue and self-compatibility in olive green. Homologous loci are connected by green lines.

The first complete linkage maps of 'Nonpareil' as the world's most popular and cultivated almond and self-compatible 'Lauranne' including more than 70 co-dominant microsatellites and self-compatibility gene were developed. The minimum LOD score 6 used to produce a framework map without any distorted markers for subsequent use as the fixed order for final map construction increased the reliability of map generation. Moreover, all microsatellite data was scored independently by two investigators to ensure accuracy of the maps. Both maps covered similar lengths, 541.8 cM for NP and 534.6 for LA map. A similar result was obtained in the myrobalan plum map (Dirlewanger et al., 2004b). However, the length of parental maps in the population derived from a cross between the almond cultivars was shorter, 415 cM for 'Ferragnès' and 416 cM for 'Tuono' map (Joobeur et al., 2000). More than 60 of the markers used in the present study were in common between both parents and allowed study of the comparative maps between parents. Marker order was generally identical on homologous groups with four exceptions in linkage groups 2, 6 and 8. Few distorted markers were mapped on the linkage groups, three for NP and five on the LA map. The order-changes between NP and LA maps were not due to the inclusion of distorted markers, because these changes occurred even when all the distorted markers were removed. Different map orders between the parents of mapping population were also found in other *Prunus* maps. Five of the 68 microsatellites mapped in progeny of the cross between myrobalan plum (P.2175) and almond-peach hybrid (GN22) were not located in the homologous linkage groups of another parent and two markers were inverted (Dirlewanger et al., 2004b). In the almond map derived from the cross 'R1000' and 'Desmayo Largueta', three markers were placed in different parental map order (Sánchez-Pérez et al., 2007b). Donidi et al., (2007) observed in an apricot map

reordering of a closely linked marker in G4. It is still unclear if the discrepancy observed in marker order was caused by minor chromosomal rearrangement, error due to the small numbers of progeny in the mapping population, or both. It has been speculated that most order differences have a low impact in genome studies because they occur most often in closely linked areas, possibly due to the low number of informative meioses. If this is the case, then larger populations should solve the problem (Dondini et al., 2007).

The length of the NP (541.8 cM) and LA maps (534.6 cM) is comparable with the other *Prunus* maps. A map of the cross between myrobalan plum 'P.2175' with 94 markers and 'GN22' map includes 1 major gene and 166 markers defining 462.8 and 458.9 cM, respectively (Dirlewanger et al., 2004b). The apricot map of 'Polonias' was 538 cM with 110 markers and for 'Stark Early orange', 699 cM with 141 markers (Lambert et al., 2004). Another apricot map produced parental maps of 'Lito' covering 504 cM genome and 'Bo81604311' covering 620 cM (Donidi et al., 2007) and almond maps of 'R 1000' covering 388.8 cM and of Spanish cultivar 'Desmayo Largueta' covering 345 cM (Sánchez-Pérez et al., 2007b). The comparison between these maps indicates that the parental maps in this study cover most of *Prunus* genome. Most of the *Prunus* maps are constructed with MapMaker program and it is well-known that those maps are usually longer than the maps developed by JoinMap such as the maps presented in this study (Personal communication with Dr Pere Arús).

The construction of IN map allows the study of different markers in combinations, and both shared genes and QTLs in both parents can be analysed. The integrated map was compared with the *Prunus* reference map (T × E; *Prunus* bin map ((Dirlewanger et al., 2004a; Howad et al, 2005), and the latest intraspecific almond map

(R × D), and in both cases showed considerable homology (Fig. 4.5). A few closely linked markers were found with different map orders compared to T × E map. For instance, PceGA34 and CPSCT021 on G2 were reordered in IN map. Since none of these markers were distorted and the results of the first framework map without any distorted markers showed the identical order for these markers, the order change may be due to the inversion of almond genome. The different marker order in the integrated map compared with parental maps was also observed in grape map 'Riesling' × 'Cabernet Sauvignon'. This divergence could be attributed to the different recombination frequency in some segments of chromosome (Riaz et al., 2004). The small number of individuals (75) in the *Prunus* reference map and the high proportion of distorted markers (46%) may cause the incorrect genetic distance and consequently inaccurate marker order (Aranzana et al., 2003; Lambert et al., 2004).

The IN map comprises 30 more microsatellites than the R×D map but compared to the T × E map it has 99 less microsatellites. The *Prunus* reference map was developed over 10 years as a result of a European project with six research groups in different countries, which enabled to produce the map with higher density (Arús et al., 1994; Joober et al., 1998; Aranzana et al., 2003; Dirlewanger et al., 2004a; Howad et al, 2005). Distribution of markers in the linkage groups was quite irregular, ranging from 19 in G1 to 12 markers on G8, a similar pattern was observed in other *Prunus* maps (Joobeur et al., 2000; Bliss et al., 2002; Aranzana et al., 2003, Dirlewanger et al., 2004a; Dirlewanger et al., 2004b; Dondini et al., 2007). The few markers placed on G8 of IN map is in agreement with several published *Prunus* maps (Dettori et al., 2001; Bliss et al., 2002; Dirlewanger et al., 2004b; Decroocq et al., 2005, Dondini et al., 2007). Dirlewanger et al., (2006) screened 397 SSRs which 33 were already reported to be

located in G8 but none of them was polymorphic on the F2 population of peach and G8 could not be generated.

Uneven allocation of markers could be the result of different recombination on the chromosome, non-random sampling of the genome due to uneven selection of markers (Riaz et al., 2004). Although no obvious cluster of loci similar to the maps with AFLPs (Verde et al., 2005) observed in this study, regions with high density or co-segregating markers were found in all linkage groups. These areas of low recombination frequency often cluster around the centromic regions (Crouzillat et al., 1996). Heterogeneity in recombination up to tenfold less than other parts of the genome was reported in tomato (Tanksley et al., 1992). Furthermore, areas with lower marker density or gaps appeared mostly in the distal part of linkage groups, which could be related to a high recombination rate, or 'hot spot', which the segments of a chromosome in which many recombination events tend to occur and often are followed by 'cold spots', regions of the chromosome with lower than average recombination rates (Hey, 2004). Construction of high resolution maps is often easier in regions with high recombination, but for suppressed recombination areas, larger populations are required allowing rare crossover events to occur (Tanksley et al., 1992).

Segregation distortion is referred to as a departure from the expected Mendelian ratio for a given genotype within a segregating population (Dufour et al., 2001). Distorted segregation is a common feature that has been reported in many species. In this study, seven (5.4%) of the 129 mapped markers (IN map) showed significant segregation distortion. Distorted loci distributed among linkage groups 3, 4, 7 and 8 with the highest (3) clustered on linkage group 7. Deviation from the expected Mendelian inheritance has been reported in other species such as: peach (Verde et al.,

2005), tomato (Chetelat, and Meglic, 2000; Paterson et al., 1988), apple (Yamamoto et al., 2002), almond (Joober et al., 2000), and *Eucalyptus* sp. (Verhaegen and Plomion, 1996). In the *Prunus* reference map, 41.5% of microsatellites and 43% of all markers deviated from the expected ratio (Aranzana et al., 2003; Dirlewanger et al., 2004a), which was suggested to be attributed to the interspecific nature and divergence of the parental cross. It is speculated that compared with dominant markers, segregation distortion has less impact on the recombination frequencies of co-dominant markers (Lorieux et al., 1995). The percentage of distorted markers in our population (5.3%) is less than most of the other distorted markers reported in *Prunus* species such as: peach, 15% in Lu et al., (1998); 18.5% in Dettori et al., (2001); 23.1% (Verde et al., 2005); apricot, 11% in Hurtado et al., (2002), 11.3% in Lambert et al., (2004); almond, 11.1% (Joobeur et al., 2000); and peach × almond, 45% in Bliss et al., (2002). Several mechanisms have been hypothesised for the occurrence of segregation distortion including the presence of sub-lethal or deleterious genes segregating in the mapping population, heterogeneity within the parents, i.e. greater genetic distance between parents causing more marker distortion (Grandillo and Tanksley, 1996), chromosomal rearrangements (Tanksley, 1984), reproductive differences between the two parents (Foolad et al., 1995), and pre or post zygotic selection of genes flanking the distorted loci (Joober et al., 2000).

The selection of desirable traits with linked markers has been routinely used in breeding for superior cultivars of tree fruits (Martínez-Gómez et al., 2003b). It was suggested that tight linkage (< 10 cM) between the trait of interest and the closest marker can be useful for MAS (Staub et al., 1996). Microsatellites UDP98-412 (4.6 cM) and CPPCT021 (7.1 cM) were found in the distal part of linkage group 6, which were

closely linked to self-incompatibility allele in other *Prunus* maps (Ballester et al., 1998; Joobeur et al., 1998; Bliss et al., 2002; Vilanova et al., 2003; Dirlewanger et al., 2004a and Sánchez-Pérez et al., 2007b). The similarity of results provides further evidence for homology of self-incompatibility allele in different *Prunus* species. However, when the markers linked to the self-fertility allele were amplified in self-incompatible and self-compatible cultivars they could not distinguish the differences. This trait allele has been reported to be single gene controlled (Gregory et al., 2005). Moreover the flanking markers are in complete synteny and colinearity with other *Prunus* maps; therefore, it could be concluded that the distance between the linked markers and the gene itself is larger than that able to be used for MAS. However, the unique banding pattern was observed in some self-compatible cultivars such as Tuono (Tu) and Supernova (SN). Considering the small number of self-compatible cultivars (6) which was used in this test it is possible with an increase in the sample number of self-compatible cultivars, a more exclusive banding pattern could be achieved. Furthermore, using more markers or a larger population may provide closer markers to be used for MAS. A CAP marker linked (9.5 cM) to the single dominant gene for root-knot nematode resistance (*mi*) mapped in an F_2 population of peach allowed susceptible homozygotes to be identified and culled from the population, but the utility of this marker was evaluated in commercial peach rootstocks results showed the same pattern for resistant and susceptible cultivars (Gillen and Bliss., 2005). The result suggests that the efficiency of MAS largely depends partly on evaluation in different cultivars. Probably the distance of linked marker in each species has a key role in the reliability of MAS.

As proposed by Joobeur et al., (2000) microsatellites are the best markers for comparative mapping. Microsatellites from different *Prunus* maps showed good

transferability to the almond map and revealed substantial level of synteny between *Prunus* genome. The homology observed in this study using comparative mapping supports the conservation of genomes among the *Prunus* species. These comparisons also suggest the possibility of chromosome rearrangement in almond. Although the maps obtained in this study have the highest number of microsatellites among almond maps and one of the most populated maps within the *Prunus*, some gaps still appeared which need further saturation. From 86 microsatellites, 73 act as anchor loci between this map (IN) and *Prunus* reference map, and 11 of them (BPPCT007, CPPCT002, UDP98-024, BPPCT017, BPPCT014, UDP96-001, BPPCT025, UDP98-412, CPPCT022, CPPCT033, UDP98-409) are part of the "genotyping set" which was proposed by Aranzana et al., (2003) for universal coverage and genotyping of the *Prunus* genome. Markers in this set originally come from single loci, are highly polymorphic and well distributed throughout *Prunus* genome. The successful amplification of around half of the microsatellite markers (11 of 24) suggested in the *Prunus* genotyping set in almond has been reported in present study and suggests that they can be can be used for fast comparison and data integration from different almond cultivars. The considerable synteny of homologous loci allows the integration and comparative genome studies among the *Prunus* genus. The almond map constructed in the present study eventually will be used for incorporation and identification of the desirable genes or QTLs which subsequently will increase the precision and the speed of breeding programs.

Chapter 5

General Discussion

Almond (*Prunus dulcis* (Mill.) D.A. Webb) is a species of the family Rosaceae, which has been domesticated from around 2000 BC (Moldenke and Modenke, 1952). Almond is the most important nut crop in terms of world production, the harvest exceeding 1.76 million tonnes in 2006 (FAO, 2007). Although the consumption of fresh fruit such as apple and other stone fruit is steady or decreasing, almond consumption has almost doubled during the last decades (Almond Board of California, 2005). Part of the reason for this is undoubtedly due to the recognition by consumers of health benefits and the high nutritional values of almond. Due to the similarity of continental climates with the native location of this species and almost pest-free growing areas, Australian almond cultivation has increased significantly over the past years. By world standards, the almond industry in Australia is still small, producing around 1% of the total world crop (FAO, 2007). For Australia to become more competitive in the international markets and in the domestic market against imports, crop production and nut quality must be maximised. The achievement of these goals is possible via two ways: first, the area of production and the number of trees should be increased; second, better cultivars

must be developed. Although almond has been generally considered as drought tolerant species, each cultivar responds differently during periods of water stress (Herralde et al.., 2003). Application of genomic approaches such as genetic mapping and marker assisted selection can contribute to accelerate the identification and the release of superior almond cultivars with higher water use efficiency. However, this issue is beyond the scope of the present study. As part of the Australian almond breeding program (Wirthensohn et al., 2002), the first generation of the intraspecific almond mapping population was produced, as described more fully in Chapter 1. Since the advent of the first genetic map in *Drosophila* 95 years ago by Alfred Sturtevant, this technique has become an essential tool for genetic analysis in both fundamental research and practical applications by the discovery of new genes and their function, marker assisted selection and comparative mapping (Liu, 1998). Genetic improvement in fruit trees like almonds is hampered by large plant size, self-incompatibility and long reproduction cycles that make the evaluation of nut characters and selection of superior seedlings expensive and time consuming. A map with markers tightly linked to the traits of interest allows identification and discarding of a significant proportion of the inferior progeny in early stages. High heterozygosity in almond provides diverse, superior genotypes, which can be used as parents in mapping studies that can assist in the breeding of cultivars with outstanding kernel quality ('Nonpareil' or Marcona'), a late blooming trait ('Tardy Nonpareil') or self compatibility ('Lauranne' and 'Tuono'). The first Australian almond map was developed by Gregory et al., (2005) using the F_1 progeny derived from a cross between the cultivars 'Nonpareil' and 'Lauranne'. Three genetic linkage maps of almond produced by Gregory et al., (2005) from this mapping population were not saturated and the information in the maps was limited. The linkage

groups obtained did not match the chromosome number of almond (x = 8), genome coverage was poor and many large gaps were observed. To be useful in breeding programs, there was a requirement that these maps be more saturated.

5.1 Polymorphism detected by microsatellites

Microsatellite markers were chosen over the other DNA based markers to complete the first almond maps reported by Gregory et al., (2005), because microsatellites are abundant, co-dominant, robust, very polymorphic and informative, PCR-based and easily transferable across related taxa, which makes them the best choice for comparison and mapping studies (Testolin, 2003). Before the commencement of this project, the quality and quantity of the concentrated DNA of the mapping population extracted by Gregory (2004) was measured by Nanodrop spectrophotometer ND-1000® prior to any microsatellite amplification and results demonstrated the very good quality and stability of DNA after three years in which DNA was kept in TE buffer at -20ºC.

5.1.1 Comparison of microsatellite detection systems

Comparison of the three methods used for microsatellite detection showed that generally polyacrylamide gel electrophoresis (PAGE) compared with MetaPhor® agarose and automated sequencer capillary electrophoresis, despite its tedious and time consuming nature, is the most efficient and economical method for analysing microsatellites in small-scale mapping programs. MetaPhor® agarose is cheaper but it cannot resolve the differences less than five bp. Capillary fluorescent tagged electrophoresis resolves to the same or better accuracy than PAGE, but the cost of

fluorescent primers and equipment for analysis is much higher (around 10 times). Similar findings were also reported by Sánchez-Pèrez et al., (2006b).

5.1.2 Genotyping of mapping population by microsatellites

In total, 195 microsatellites isolated from different *Prunus* species were tested in this study. 54 markers detected easily scorable and reliable loci. Importantly, to provide added confidence in the experimental data, another member of the almond breeding program (Dr Shubiao Wu) independently confirmed all the scoring. The highest number of amplified and polymorphic markers was detected in microsatellites isolated from almond, and the sweet cherry microsatellites produced the lowest number of amplified and polymorphic markers. The results suggest that although microsatellites are transferable across closely related species, the efficiency of this process decreases as the phylogenetic distance increases. Substantial reproducibility of microsatellites markers was observed during screening of the mapping population, which showed that the usefulness of microsatellite markers could be transferred across different research groups. Microsatellite markers are widely used in paternity analysis for tracking the origin of alleles (Lian et al., 2001; Mookerjee et al., 2005; Slavov et al., 2005). Each locus detected by a microsatellite is produced from two alleles, one from each parent, which can be used to determine the hybridity of progeny. All the polymorphic microsatellites in this study showed the correct amplification of parents' alleles in progeny, and no aberrant bands were observed. From this result it can be concluded that all the progeny in mapping population are true hybrids. Seven of eight microsatellites originally isolated from a cDNA library of almond were amplified but none of them showed polymorphism in the population screened. These microsatellites come from the

coding region of chromosomes, where it is likely that fewer mutations occur than in non-coding regions, and that they are therefore less polymorphic. Our finding is in agreement with Dendini et al., (2007) who reported a low frequency of polymorphic cDNA microsatellites in apricot. As a result, it can be suggested that this type of microsatellite is less suitable for mapping studies. However, more than a quarter of microsatellite markers (27.7%) were transferred successfully from other *Prunus* species to the almond. Successful transfer of microsatellites from species with the high number of microsatellites to less studied species will greatly facilitate the wide variety application of this marker particularly in mapping program.

5.2 Pomological analysis of the mapping population

To plan an efficient breeding program it is required to understand the genetic and environmental factors that influence the traits, coupled with analysis of the inheritance data (Kester et al., 1977). Kernel quality characters such as size, taste and attractiveness improve the market price of kernels, and at the same time bring about a consistent reduction in the price of low quality kernels (Socias i Company et al., 2008). To develop a comprehensive picture of the interactions between genotype and environment operating on the progeny of the cross used for map development, twelve quality and pomological traits were measured and their inheritance and correlations analysed during three consecutive years from 2005 to 2007.

5.2.1 Phenotypic and inheritance studies of the shell hardness, kernel size and weight

Shell hardness is assumed to be controlled by a major dominant gene *D* (Grasselly, 1972; Socias i Company, 1998). The result of this study showed that around half of the progeny had a semi-hard shell phenoype, which cannot be explained by a monogenic mode of inheritance, suggesting that a more complex and possibly additive gene effect could be responsible for the expression of this trait. The more complex mode of inheritance for shell hardness is in agreement with the observations of Dicenta et al., (1993) and Sánchez-Pèrez et al., (2007a). Heritability of 0.752 was estimated for shell hardness in accordance with previous studies in almond $H^2 = 0.82$ (Spiegel-Roy and Kochba, 1981) and $H^2 = 0.55$ (Kester, 1977) and $H^2 = 0.91$ in walnut families (Hansche et al., 1972).

Kernel shape, an important commercial parameter, could be classified into a number of different classes. The high heritability of traits estimated for kernel shape ($H^2 = 0.92$) and its components (kernel length $H^2 = 0.85$, width $H^2 = 0.83$, Kernel Thickness $H^2 = 0.72$) was observed in this study. In fact, kernel shape is characteristic of cultivar and has been used in traditional cultivar identification. Variability in this trait among the years studied remained very low, suggesting the high potential for manipulation of this trait by appropriate selective strategies. High heritability of these traits indicating that they are under strong genetic control. Therefore, the selection of parents for these traits based on their phenotypic performances was recommended and the gain attained in this way expected to be high and quick (Yao and Mahlenbacher, 2000).

Significant correlation was observed between the in-shell and kernel weight ($r = 0.74$), among kernel size components such as kernel length / kernel width ($r = 0.67$), kernel weight / kernel length ($r = 0.78$) and kernel width ($r = 0.80$). The high correlation

coefficient estimated in the present study suggests that almond breeding programs may benefit from the apparent linkage of these traits. Identification of markers segregating for any one trait may reasonably be expected to be good predictors of the inheritance of the additional linked traits. In addition, even with conventional selection methods, selection of one trait culminated in the selection of other trait that reduces the cost and effort of additional measurement. In some cases, kernels in the progeny were larger than found in either parent (hybrid vigor), suggesting the opportunity exists for improvement of this trait even in populations derived from crosses between parents with relatively small kernels. Continuous phenotypic variation of the progeny indicates that many loci or genes are responsible for the variation, each with a small effect that along with environmental effects express the phenotype.

5. 2.2 Analysis of kernel flavour inheritance and amygdalin contents

Kernel taste is a major factor in the Australian almond breeding program. Heppner (1923) and other authors (Kester et al., 1977; Dicenta and García, 1993) indicated that kernel bitterness in almond is a monogenic trait in which bitter kernel is a recessive allele *ss* and sweet kernels *SS* or *Ss*. Confusingly, the inheritance of the slightly bitter kernel phenotype cannot be explained by this mode of inheritance. The result of χ^2 analysis of data collected in all three years of the present study showed that mode of inheritance does not correspond to a 3:1 segregation ratio. Consequently, a more complex mechanism for this trait can be considered and it is speculated that more than one gene is controlling the kernel taste phenotype. The bitter taste of the almond kernel is suggested to be due to the high level of the cyanogenic diglucoside amygdalin (Kester and Gradziel, 1996). Very few studies have been performed on the association

of amygalin with semi bitter kernel phenotype. To elucidate the role of amygdalin in slightly bitter kernels, the amount of amygdalin in sweet and slightly bitter kernels in the mapping population was measured. The mean of amygdalin content in sweet kernels (3.67 mg kg^{-1} FW) was lower than slightly bitter kernels (20.34 mg kg^{-1} FW) but some sweet kernels had higher levels of amygdalin than slightly bitter kernel. However, no sweet kernel with amygdalin more than 14.5 mg kg^{-1} FW was found. This threshold may be useful in assessment by chemical analysis. Dicenta et al., (2002) hypothesised that other compounds than amygdalin must be involved in the production of slightly bitter taste, which is in agreement with our finding.

5.2.3 Phenotypic variation of double kernels

Both ovules in almond can be viable and produce double kernels that are deformed and so decrease the commercial value of the crop. The highest variation recorded in this study was observed for this trait. It has been reported that environmental conditions during ovule production have a strong influence on the production of double kernels. In this study the lowest pre-blossom temperature occurred in June and July 2006 and resulted in the highest proportion of double kernels in the subsequent harvest in 2007. This finding is in agreement with a previously reported high correlation between low temperatures and the production of double kernels (Egea and Burgos, 1995). Both of the parents of the mapping population produced a very low proportion of double kernels. Considering the strong environmental effect on the production of undesirable double kernel, which was found both by Gregory, (2004) and the present study, it can be speculated that the trait is under low genetic control and many genes with small effects may control its expression. Interaction of environment ×

modifier genes ultimately determines the phenotypic variation of double kernel. As mentioned in Chapter 3 even cultivars with a very low percentage of double kernels, in conditions of low pre- blossom temperatures, produce a high proportion of double kernels. Therefore, the use of parents with zero double kernels may further reduce the risk of inheritance in this trait, but the best way to evade this problem is to avoid establishment of almond orchards in unfavourable regions.

5.2.3.1 Benefit of double kernels in genetic studies

In other studies, Martínez-Gómez and Gradziel, (2003), some of the almond double kernels were found to be haploids or aneuploids and thus valuable for genetic studies such as locating, isolating and transferring genes. The data presented in this study, and that reported earlier, suggests that artificially low pre-blossom temperatures can be used to promote the production of double kernels. In this way, a higher proportion of aneuploid seedlings will be produced, which can be used to find the association between linkage groups and chromosome number via FISH (Fluorescence in-situ hybridisation).

Estimations of heritability in this study demonstrated common results compared to other nut crops and therefore, it can be suggest that this information is applicable in other less-studied nut species. The inheritability and correlation estimated in the present study can predict the phenotypic variation of progeny in segregating population although not as accurate as the QTL mapping. The pomological traits showed the quantitative and normally distribution, two pre requirements for QTL (Quantitative trait loci) study. With this data and saturated map for QTL mapping, the location of complex traits on chromosomes and phenotypic variation on each region can be identified, which

allows a better understanding of gene function and phenotype. When this trait is mapped, traits can be used for MAS and map-based cloning to improve the cultivars.

5.3 Almond saturated map

Map-based strategies such as linkage analysis, QTL mapping and positional cloning are very useful for the understanding of genome structure and mechanism of the trait inheritance (Oraguzie and Wilcox, 2007). Before the start of this project several interspecific almond and peach linkage maps had been generated (Foolad et al., 1995; Jáurgui et al., 2001; Bliss et al., 2002; Aranzana et al., 2003). An earlier genetic map of almond, developed by Joobeur et al., (2000), produced the expected eight linkage groups but was mainly based on RFLP and RAPD markers, with only six microsatellite markers. This made the comparison of this map with those that were subsequently produced very difficult. In the present work, moderately dense genetic maps, with the majority of markers composed of highly informative microsatellites, were constructed for the cultivars 'Nonpareil' and 'Lauranne' with 98 and100 markers, and an integrated map with a total of 129 markers. Each map produced eight linkage groups equal to the basic chromosome number of almond, with a map length of 541.8 cM for NP, 534.6 cM for LA and 592.4 cM for the integrated map as shown in Chapter 4. Almost even and random distribution of microsatellite markers was observed across the linkage groups and confirmed the usefulness of these markers for map saturation and as an excellent complement for ISSR and RAPD markers. The addition of new markers increased the density of maps to 5.53, 5.34 and 4.59 cM/marker for NP, LA and integrated map, respectively.

5.3.1 Comparative mapping of the integrated map with *Prunus* reference map

Between 57 to 73 anchor microsatellites in the LA and integrated maps allowed the alignment of the eight linkage groups with other *Prunus* maps, particularly the most extensively studied *Prunus* reference map (Dirlewanger et al., 2004a; Howad et al., 2005). A high degree of synteny and colinearity was observed between our maps and other *Prunus* maps. All the common microsatellites in the presented map showed complete synteny with the *Prunus* reference map (100%), although 13 of the 73 shared markers (17.8%) did not show complete collinearity. On the other hand, comprehensive collinearity and synteny (100%) between a previous almond map (Sánchez-Pèrez et al., 2007b) and our map was observed. The discrepancy between *Prunus* reference map and the integrated map can be due to the different parental genotypes. The *Prunus* reference map is derived from an interspecific cross between peach × almond, but the integrated map derives from the almond × almond cross. It can be considered that some rearrangement in the almond genome has occurred, and has resulted in small changes in a few chromosome regions. Dirlewanger et al., (2004a) suggest that the *Prunus* genus can be considered as single genetic system. The result of our study is apparently support this idea and the comparative mapping between the almond maps with other *Prunus* species confirm this claim. This homology can be used to speed up the transfer of information from other high density *Prunus* maps to the almond maps.

5.3.2 Marker assisted selection for self-compatibility

One of the most practical and potentially powerful uses of linkage map is marker assisted selection, where the degree of success depends on the minimisation of the frequency of recombination between marker and the trait (Luby and Shaw, 2001). Self-

compatibility is an important trait in almond, adding the new microsatellite decrease the distance between the flanking markers and the self-compatibility. Self fertility allele is located at the distal part of linkage group six between the microsatellites CPPCT021 (8.9 cM) and UDP98-412 (5.4 cM). When these markers were screened across different self-compatible and self-incompatible cultivars, among six self fertile cultivars just two specific banding patterns was observed to distinguish the gene. Many studies have mentioned the use of linked markers for MAS, but only a few have confirmed the validity of their proposed marker in different populations. One example of this is the RAPD and SCAR marker linked to the root-knot nematode resistance (Lecouls et al., 1999); also, genetic markers linked to the olive leaf spot resistance and susceptibility (Mekuria et al., 2001). The reason for failing to identify all the self-fertile alleles can be the distances between the closest mapped markers and the *Sf* trait, which are not tightly linked. Therefore, recombination may occur between markers and trait-specific genes, suggesting that these markers cannot be used for selection of self-fertility in this stage.

Most almond cultivars are highly outcrossing due to the obliged self-incompatibility. This trait in almond is gametophytic, in which a pistil rejects pollen from genetically related cultivars (de Nattancourt, 2001). Selection on self-compatibility is a typical example of negative frequency-dependent selection, i.e. the rare *S-alleles* have advantage over common alleles because of successful fertilisation. Consequently, the rare SI alleles spread quickly in the population. This system over the time produces stable linkage equilibrium with many alleles at equal frequency (Silvertown, and Charlesworth, 2001). As a result, the linkage disquilibrium decrease in each generation.

Channutapipat et al., (2003) reported the use of PCR primers linked to the self-compatibility locus. If these primers show the polymorphism in mapping population then they can be mapped in the identical place of self-compatibility locus in the maps.

Ten new microsatellite loci located on our map which have not been previously mapped in any *Prunus* map could be useful to improve the density of other maps, especially the *Prunus* reference map. However, a few markers (seven) still remain unlinked, implying that the maps need more saturation.

5.4 Future Prospect of the almond map

All the available microsatellites from the *Prunus* reference map were used in this study, therefore, for further saturation recently released microsatellites from other *Prunus* species or maps should be used. In addition, specific microsatellites from almond can be isolated, along with other markers such as those generated by AFLP (Amplified Fragment Length Polymorphism), DArT (Diversity Array Technology) and SNP (Single Nucleotide Polymorphism) be used for further map saturation (Edwards and Mogg, 2008). DArT markers can produce over one hundred genomic loci without any prior sequence information (Wenzl et al., 2004). The costs of DArT per data point have been reported to be 10-fold lower than the cost of microsatellites (Xia et al., 2005).

As mentioned earlier, many of almond characters studied in the present study showed a complex mode of inheritance when describing the quantitative traits. It is suggested that a certain region of chromosome contains the genes that affect significantly the expression of complex traits. QTL mapping involves the identification of these regions with a set of genetic markers (Gardnier et al., 2007). An essential requirement for QTL identification is a saturated genetic map with reliable markers.

Selection of progeny with superior nut trait in almond would greatly facilitate breeding programs. 12 quantitative traits in almond had been mapped in a low-resolution genetic linkage map of almond (Sánchez-Pèrez et al., 2007b). QTL analysis of a population can detect only the difference between parental alleles in the population, to completely estimate the phenotypic effect of identified QTL further studies with different genetic background should be established. However, a QTL is not a true gene; it only putatively indicates the genomic region that with high probability contains genes of desired traits (Liebhard et al., 2003). Considering the very small nuclear genome of almond (0.66 ± 0.06 pg/2C), approximately twice the size of the *Arabidopsis* genome (Baird et al., 1994), and the availability of a physical map in the closely related peach (Zhebentyayeva et al., 2008), BAC (Bacterial Artificial Chromosome) and YAC (Yeast Artificial Chromosome) based systems offer the possibility of physical mapping and map-based cloning for major and quantitative traits in almond. The maps constructed in this study can be a starting point in this direction and act as model species for other nut crops. As stated by Goodfellow, (1993) "It is obvious that microsatellite maps should be produced for every species studied by geneticists". At the present time, the maps presented in this study are the most saturated maps based on microsatellites in almond species and we report the first completed 'Nonpareil' map. These maps provide a tool for achieving the goal to produce improved cultivars in almond and ultimately in *Prunus* and other nut crops.

References

Ahmad, R., Potter, D. and Southwick, S.M. (2004). Genotyping of peach and nectarine cultivars with SSR and SRAP molecule markers. Journal of the American Society for Horticultural Science, **129:** 204-210

Alijanabi, S. and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research, **25:** 4692-4693

Almond Board of California. (2005). Almond Almanac. Almond Board of California, Modesto, CA, USA.

Anon (2007). Almond investors report. www.almondinvestors.com.au

Aranzana, M.J., Garcia-Mas, J., Carbó, J., Arús, P. (2002). Development and variability analysis of microsatellite markers in peach. Plant Breeding, **121:** 87-92

Aranzana, M.J., Pineda, A., Cosson, P., Dirlewanger, E., Ascasibar, J., Cipriani, G., Ryder, C. D., Testolin, R., Abbott, A., King, G.J., Iezzoni, A.F. and Arús, P. (2003). A set of simple sequence repeat (SSR) markers covering the Prunus genome. Theoretical and Applied Genetics, **106:** 819-825

Arteaga, N. and Socias i Company, R. (2001). Heritability of fruit and kernel traits in almond. Acta Horticulturae, **591:** 269-274

Arús, P., Ballester, J. Jáuregui, B., Joobeur, T., Truco, M.J. and de Vicente, M.C. (1999). The European *Prunus* mapping project: update of marker development in almond. Acta Horticulturae, **484:** 331-336

Asíns, M.J., Mestre, P., García, J.E., Dicenta, F. and Carbonell, E.A. (1994). Genotype x environment interaction in QTL analysis of an intervarietal almond cross by means of genetic markers. Theoretical and Applied Genetics, **89:** 358-364

Australian Bureau of Statistics (2007). Agricultural Commodities., Canberra, Publication No. 7.303, 34 pp

Baird, W.V., Esteager, A.S. and Wells, J. (1994). Estimation of nuclear DNA content in peach and related diploid species using flow cytometry and DNA hybridization. Journal of the American Society for Horticultural Science, **119**: 1312-1316

Baird, W.V., Ballard, R.E., Rajapakse, S., and Abbott. A.G. (1998). Progress in *Prunus* mapping and application of molecular markers to germplasm improvement. HortScience, **31:** 1099-1106

Ballester, J., Bošković, R., Batlle, I., Arús, P., Vargas, F. and De Vicente, M.C. (1998). Location of the self-incompatibility gene on the almond linkage map. Plant Breeding, **117:** 69-72

Ballester, J., Socias i Company, R. Arús, P. De Vicente, M.C. (2001). Genetic mapping of a major gene delaying blooming time in almond. Plant Breeding, **120:** 268–270

Becker, J. and Heun, M. (1995). Mapping of digested and undigested random amplified microsatellite polymorphisms in barley. Genome, **38:** 991-998

Blenda, A.V., Verde, I., Georgi, L.L., Reighard, G.L., Forrest, S.D., Muñoz-Torres, M., Baird, W.V. and Abbott, A.G. (2007). Construction of a genetic linkage map and identification of molecular markers in peach rootstocks for response to peach tree short life syndrome. Tree Genetics and Genomes, **3:** 341-350

Bliss, F.A., Arulsekar, S., Foolad, M.R., Becerra, V., Gillen, A.M., Warburton, M.L., Dandekar, A.M., Kocsisne, G.M. and Mydin, K.K. (2002). An expanded genetic linkage map of *Prunus* based on an interspecific cross between almond and peach. Genome, **45:** 520-529

Bruford, M.W., and Wayne, R.K. (1993). Microsatellites and their application to population genetic studies. Current Opinion in Genetics and Development, **3:** 939-943

Cantini, C., Iezzoni, A.F., Lamboy, W.F., Boritzki, M. and Struss, D. (2001). DNA fingerprinting of tetraploid cherry germplasm using simple sequence repeats. Journal of the American Society for Horticultural Science, **126:** 205-209

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

Castiglioni, P., Ajmone-Marsan, P., van Wijk, R. and Motto, M. (1999). AFLP markers in a molecular linkage map of maize: codominant scoring and linkage group distribution. Theoretical and Applied Genetics, **99:** 425-431

Chandrababu, R.J. and Sharma, R.K. (1999). Heritability estimates in almond [*Prunus dulcis* (Miller) D. A. Webb]. Scientia Horticulturae, **79:** 237-243

Channuntapipat, C., Wirthensohn, M., Ramesh, S.R., Batlle, I., Arús, P., Sedgley, M. and Collins, G. (2003). Identification of incompatibility genotypes in almond (*Prunus dulcis* Mill.) using specific primer based on the introns of the S-alleles. Plant Breeding, **122:** 164-168

Chaparro, J.X., Werner, D.J, O'Malley, D. and Sederof, R.R. (1994). Targeted mapping and linkage analysis of morphological, isozyme, and RAPD markers in peach. Theoretical and Applied Genetics, **87:** 805-815

Chetelat, R.T. and Meglic, V. (2000). Molecular mapping of chromosome segments introgressed from Solanum lycopersicoidesinto cultivated tomato (*Lycopersicon esculentum*). Theoretical and Applied Genetics, **100:** 232–241

Ciofi, C., Funk, S.M., Coote, T., Cheesman, D.J., Hammond, R.L. Saccheri, I.J. and Bruford, M.W. (1998). Genotyping with microsatellite markers. pp. 195-211. In: Karp, A., Issac, P.G. and Ingram, D.S. (eds). Molecular tools for screening biodiversity. Chapman & Hall, London.

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

Clarke, J.B. and Tobutt, K.R. (2003). Development and characterization of polymorphic microsatellites from *Prunus avium* 'Napoleon'. Molecular Ecology Notes, **3:** 578-580

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

Conn, E.E. (1980). Cyanogenic compounds. Annual Review of Plant Physiology, **31:** 433-451

Conner, P. Brown, S. and Weeden, N. (1997). Randomly amplified polymorphic DNA based genetic maps of three apple cultivars. Theoretical and Applied Genetics, **95:** 350-359

Corredor, E., Román, M., García, E, Perera, E. Arús, P. and Naranjo, T. (2004). Physical mapping of rDNA genes establishes the karyotype of almond. Annals of Applied Biology, **144:** 219-222

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

Dakin, E.E. and Avise, J.C. (2004). Microsatellite null alleles in parentage analysis. Heredity, **93:** 504-509

Dayanandan, S., Rajora O.P. and Bawa, K.S. (1998). Conservation of microsatellites among tropical trees (Legominosae). American Journal of Botany, **84:** 1658-1663

Debener, T. and Mattiesch, L. (1999). Construction of a genetic linkage map for roses using transferability of apricot and grape EST microsatellite markers across taxa. Theoretical and Applied Genetics, **99:** 350-359

Decroocq, V., Favè, M.G., Hagen, L., Bordenave, L. and Decroocq, S. (2003). Development and transferability of apricot and grape EST microsatellite markers across taxa. Theoretical and Applied Genetics, **106:**912–922

De Herralde, F., Biel, C., Batlle, I. and Savè, R. (2003). Gas Exchange Ander water stress conditions in three almond cultivars. III GRAMPA meeting on almonds and pistacios, p. 45

de Nettancourt, D. (2001). Incompatibility and Incongruity in Wild and Cultivated Plants. Springer-Verlag, Berlin, Germany.

Dendini, L., Lain, O., Geuna, F., Banfi, R., Gaiotti, F., Tartarini, S., Bassi, D., and Testolin, T. (2007). Deveolopment of a new SSR-based linkage map in apricot and analysis of synteny with existing *Prunus* map. Tree Genetics and Genomes, **3:** 239-249

Denisov, V.P. (1988). Almond genetic resources in the USSR and their use in production and breeding. Acta Horticulturae, **224:** 299-306

Dettori, M.T., Quarta, R. and Verde, I. (2001). A peach linkage map integrating RFLPs, SSRs, RAPDs and morphological markers. Genome, **44:** 783-790

Dicenta, F. and García, J.E. (1992). Phenotypical correlations among some traits in almond. Journal of Genetics and Breeding, **46:** 241–248

Dicenta, F. and García, J.E. (1993). Inheritance of the kernel flavour in almond. Heredity, **70:** 308-312

Dicenta, F., Martínez-Gómez, P., Grané, N., Martín M.L., León A., Cánovas, J.A. and Berenguer, V. (2002). Relationship between cyanogenic compounds in kernels, leaves and roots of sweet and bitter kernelled almonds. Journal of Agriculture, Food and Chemistry, **50:** 2149-2152

Dirlewanger, E. and Bodo, C. (1994) Molecular genetic mapping of peach. pp 309-311. In: Schmidt, H, Kellerhals, M. (eds). Progress in temperate fruit breeding. Kluwer, The Netherlands.

Dirlewanger, E., Pronier, V., Parvery, C., Rothan, C. Guye, A. and Monet, R. (1998). Genetic linkage map of peach [*Prunus persica* (L.) Batsch] using morphological and molecular markers. Theoretical and Applied Genetics, **97:** 888-895

Dirlewanger, E., Cosson, P., Tavaud, M., Aranzana, M.J., Poizat, C., Zanetto, A., Arús, P. and Laigret, F. (2002). Development of microsatellite markers in peach

[*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theoretical and Applied Genetics, **105:** 127-138

Dirlewanger, E., Graziano, E,E., Joobeur, T,T., Garriga-Caldere, F., Cosson, P., Howard, W., and Arús, P. (2004a). Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proceedings of the National Academy of Sciences, **101:** 9891-9896

Dirlewanger, E., Cosson, P., Howad, W., Capdeville, G., Bosselut, N., Claverie, M., Voisin, R., Poizat, C., Lafargue, B., Baron, O., Laigret, F., Kleinhentz, M., Arús, P. and Esmenjaud, D. (2004b). Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid location of root-knot nematode resistance genes. Theoretical and Applied Genetics, **109:** 827–838

Dirlewanger, E., Cosson, P., Boudehri, K., Renaud, C., Capdeville, G., Tauzin, Y., Laigret, F. and Moing, A. (2006). Development of a second-generation genetic linkage map for peach [*Prunus persica* (L.) Batsch] and characterization of morphological traits affecting flower and fruit. Tree Genetics and Genomes, **3:** 1-13

Dondini, L., Lain, O., Geuna, F., Banfi, R., Gaiotti, F., Tartarini, S., Bassi, D., and Testolin, T. (2007). Deveolopment of a new SSR-based linkage map in apricot and analysis of synteny with existing *Prunus* map. Tree Genetics and Genomes, **3:** 239-249

Doyle, J. and Doyle, J. (1991). Isolation of plant DNA from fresh tissue. Focus. **12:** 13-15

Downey, S.L. and Iezzoni, A.F. (2000). Polymorphic DNA markers in black cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach and sour cherry. Journal of the American Society for Horticultural Science, **125:** 76-80

Dufour, P., Johnsson, C., Antoine-Michard, S., Cheng, R., Murigneux, A. and Beckert, M. (2001). Segregation distortion at marker loci: variation during microspore emberyogenesis in maize. Theoretical and Applied Genetics, **102:**993-1001

Echt, C., Knapp, S. and Liu, B.H. (1992). Genome mapping with non-inbred crosses using Gmendel 2.0. Maize Genetic Cooperation Newsletter, **66:** 27-29

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

Edwards, K.J. and Mogg, R. (2008). Plant genotyping by analysis of single nucleotide polymorphisms. pp. 1-15. In: Henry, R.J. (Ed). Plant genotyping II. CABI Publications.

Egea, J. and Burgos, L. (1995). Double kernelled fruits in almond *(Prunus dulcis,* Mill.) as related to pre-blossom temperatures. Annals of Applied Biology, **126:** 163-168

Egea, J. and Burgos, L. (2000). Ovule differences between single-kernelled and double-kernelled fruits in almond *(Prunus dulcis)*. Annals of Applied biology, **136:**291-295

Falconer, D.S., and Mackay, T.F.C. (1996). Introduction to Quantitative Genetics, 4th edn. Addison-Wesley Longman Ltd, London

FAO (2007). FAO Statistical Databases: Agriculture. http://www.fao.org

Foolad, M.R., Arulsekar, S., Becerra, V. and Bliss, F.A. (1995). A genetic map of *Prunus* based on an interspecific cross between peach and almond. Theoretical and Applied Genetics, **91:** 262-269

Franks, T.K., Yadollahi, A., Wirthensohn, M.G., Guerin, J.R., Kaiser, B.N., Sedgley M. and Ford, C.M. (2008). A seed-coat cyanohydrin glucosyltransferase is

associated with bitterness in almond (*Prunus dulcis* Mill) kernels. Functional Plant Biology, **35:** 236-246

Gardiner, S.E., Zhu, J.M., Whitehead, H.C.M. and Madie, C. (1994). The New Zealand apple genome mapping project. A progress report. Euphytica, **77:** 77-81

Gardiner, S.E., Bus, V.G. M. Rusholme, R.L., Chagné and Rikkerink, E.H.A. (2007). Apple. In: Kole, C. (Ed.) Genome mapping and molecular breeding in plants, Volume 4, Fruit and Nuts. Springer-Verlag Berlin Heidelberg

García, J.E., Dicenta, F. and Carbonell, E.A. (1994). Combining ability in almond. Plant breeding, **112:** 141-150

Gathercole, F.J. (2002). An introductory to commercial almond growing in Australia. SARDI publication. Loxton, SA. Australia,

Gillbert, A. (2005). Almonds. pp. 5-14. In: Moore, B. (ed). Just nuts. Hyland House, Australia.

Gillen, A.M. and Bliss, F.A. (2005). Identification and mapping of markers linked to the *Mi* Gene for root-knot nematode resistance in peach. Journal of American Society for Horticultural Science, **130:** 24-33

Godini, A. (2002). Almond fruitfulness and role of self-fertility. Acta Horticulturae, **591:** 191-199

Goodfellow, P.N. (1993). Microsatellites and the new genetic maps. Current Biology, **3:** 149-151

Grandillo, S. and Tanksley, S.D. (1996). Genetic analysis of RFLPs, GATA microsatellites and RAPDs in a cross between *L. esculentum* and *L. pimpinellifolium.* Theoretical and Applied Genetics, **92:** 957–965

Grando, M.S., Bellin, D., Edwards, K.J., Pozzi, C., Stefanini, M. and Velasco, R. (2003). Molecular linkage maps of *Vitis vinifera* L. and *Vitis riparia* Mchx. Theoretical and Applied Genetics, **106:** 1213-1224

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

Grasselly, C., (1972). L' Amendier; caracteres morphologiques et physiologiques des varietes, modalite de leurs transmissions chez les hybrids de premeriere generation. PhD Thesis. University of Bordeaux.

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

Gregory, D. (2004). The development of a genetic linkage map for almond based on molecular and agronomic markers. PhD thesis. The University of Adelaide, Australia.

Gregory, D., Sedgley, M., Wirthensohn, M.G., Arús, P., Kaiser, B. and Collins, G. G. (**2005**). An integrated genetic linkage map for almond based on RAPD, ISSR, SSR and morphological markers. Acta Horticulturae, **694:** 67-72

Gülcan, R., (1985). Descriptor List for Almond (*Prunus amygdalus*), revised edition. International Board of Plant Genetic Resources, Rome, Italy.

Hamwieh, A., Udupa, S.M., Choumane, W., Sarker, A., Dreyer, F., Jung, C. and Baum, M. (2005). A genetic linkage map of *Lens* sp. based on microsatellite and

AFLP markers and the localization of fusarium vascular wilt resistance, Theoretical and Applied Genetics, **110:** 669-677

Han, Y., Gasic, K., Marron, B., Beever, J. E., and Korban, S.S. (2007). A BAC-based physical map of the apple genome. Genomics, **89:** 630-637

Hancock, J.M. (2001). Microsatellites and other simple sequences: genomic context and mutational mechanisms. pp. 1-9. In: Microsatellites evolution and applications. Goldstein, D. B and Schlötterer, C. (Eds). Oxford University Press.

Hansche, P.E., Beres, V. and Brooks, R.M. (1966). Heritability and genetic correlation in the sweet cherry. Journal of the American Society for Horticultural Science, **88:** 173-183

Hansche, P.E., Bringhurst, R.S. and Voth, V. (1968). Estimate of genetic environmental parameters in the strawberry. Journal of American Society for Horticultural Science, **92:** 338-345

Hansche, P.E., Beres, V. and Forde, H.I. (1972). Estimates of quantitative genetic properties of walnut and their implications for cultivar improvement. Journal of the American Society for Horticultural Science, **97:** 279-285

Hardner, C., Winks, C., Stephenson, R. and Gallagher, E. (2001). Genetic parameters for nut and kernel traits in macademia. Euphytica, **117:** 151-161

Hauge, B.M., Hanley, S.M., Cartinhour, S., Cherry, J.M., Goodman, H.M. Koorneef, M., Stam, P., Change, C., Kempin, S., Medrano, L. and Meyerowitz, E. M. (1993). An integrated genetic/RFLP map of the *Arabidopsis thaliana* genome. Plant Journal, **3:** 745-754

Hawker, J. S. and Buttrose, M.S. (1980). Development of the Almond Nut *{Prunus dulcis* (Mill.) D. A. Webb). Anatomy and chemical composition of fruit parts from anthesis to maturity. Annals of Botany, **46:** 313-321

Hemmat, M., Weeden, N.F. and Brown, S.K. (2003). Mapping and evaluation of *Malus* × *domestica* microsatellites in apple and pear. Journal of the American Society for Horticultural Science, **128:** 515-520

Henry, R.J. (2001). Plant Genotyping: Henry, R. J. (eds) The DNA fingerprinting of Plants, CABI Publishing, New York. p. 325

Heppner, J. (1923). The factor for bitterness in the sweet almond. Genetics, **8:** 390-392

Heppner, J. (1926). Further evidence on the factor for bitterness in the sweet almond. Genetics, **11:** 605-606

Hey, J. (2004).What's so hot about recombination hotspots? PLoS Biology **2:** 730-733

Hormaza, J.I. (2002). Molecular characterization and similarity relationships among apricot (*Prunus armeniaca* L.) genotypes using simple sequence repeats. Theoretical and Applied Genetics, **104:** 321–328

Howad, W., Yamamoto, T., Dirlewanger, E., Testolin, R., Cosson, P., Cipriani, G., Monforte, A.J., Georgi, L., Abbott, A.G. and Arús, P. (2005). Mapping with a few plants: Using selective mapping for microsatellite saturation of the *Prunus* reference map. Genetics, **171:** 1305–1309

Hurtado, M., Romero, C., Vilanova, S., Abbott, A., Llácer, G., and Badenes, M. (2002). Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.), and mapping of PPV (sharka) resistance. Theoretical and Applied Genetics, **105:**182–191

Jacobs, J.M.E., Van Eek, H.J., Arens, P., Verkerk-Bakker, B., Lintel Hekkert, B. te., Bastiaanssen, H.J.M., EI-Kharbotly, A., Pereira, A., Jacobsen, E. and Stiekema, W.J. (1995). A genetic map of potato *(Solanum tuberosum)* integrating molecular markers, including transposons, and classical markers. Theoretical and Applied Genetics, **91:** 289-300

Jackson, J. (1996). Gene flow in commercial almond orchard. Sexual Plant Reproduction, **9:** 367-369

Jáuregui, B., de Vincnte, M.C., Messeguer, R., Felipe, A., Bonnet, A., Salesses, G. and Arús, P. (2001) A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. Theoretical and Applied Genetics, **102:** 1169-1176

Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., van deWiel, 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

Joobuer, T., Viruel, M.A., de Vicente M.C., Jauregui, B., Ballester, J., Dettori, M. T., Verde, I., Truco, M.J., Messeguer, R., Batlle, I., Quarta, R. Dirlewanger, E. and Arús, P. (1998). Construction of a saturated linkage map for *Prunus* using an almond x peach F_2 progeny. Theoretical and Applied Genetics, **97:**1034-1041

Joobuer, T., Periam, N., de Vicente M.C., King, G.J. and Arús, P. (2000). Development of second generation linkage map for almond using RAPD and SSR markers. Genome, **43:** 649-655

Jung, S., Jesudurai, C., Staton, M., Du, Z., Ficklin, S., Cho, I., Abbott, A., Tomkins, J. and Main, D. (2004). GDR (Genome Database for Rosaceae): integrated

web resources for Rosaceae genomics and genetics research. BMC Bioinformatics, **5:**130

Kader, A.A. (1996). Almond production manual. pp. 274-277. In: Micke, C.W. (ed). In: plant storage. Publication No.3364. University of California, Division of Agriculture and Natural Resources

Karp, A., Seberg, O. and Buiatti, M. (1996). Molecular techniques in the assessment of botanical diversity. Annals of Botany, **73:** 143-149

Kester, D. E. (1965). Size, shape and weight relationships in almond kernels, Journal of the American Society for Horticultural Science, **87:** 204-213

Kester, D.E. and Asay, R. A. (1975). Almond breeding. In: Janick, J. and Moore, J. N. (eds). Advances in fruit breeding. Purdue University Press, West Lafayette, Indiana, pp.382-419

Kester, D.E., Hansche, P.E., Beres, V. and Asay, R.N. (1977). Variance components and heritability of nut and kernel traits in almond. Journal of the American Society for Horticultural Science, **102:** 264-266

Kester, D.E., Gradziel, T.M. and Grasselly, C. (1990). Almonds. Acta Horticulturae, **290:** 699-758

Kester, D.E., Gradziel, T.M. and Grasselly, C. (1991). Almonds (*Prunus*). pp. 701-758. In: Moore, J.N., Ballington, H.J. (eds). Genetic resources of temperate fruit and nut crops. International Society of Horticultural Science, The Netherlands

Kester, D.E. and Gradziel, T.M. (1996). Almonds. pp.1-97. In: Janick, J. and Moore, J.N. (Eds). Fruit Breeding, Vol. III, Nuts. John Wiley & Sons, New York

Kosambi, D.D. (1944). The estimation of map distances from recombination values. Annals of Eugenics, **12:**172–175

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

Kwok, S., Kellog, D.E., McKinney, N, Spasic, D., Goda, L. and Levenson, C. (1990). Effect of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus1 model studies. Nucleic Acids Research, **18:** 999-1005

Lambert, P., Hagen, L.S., Arús, P. and Audergon, J.M. (2004). Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond 'Texas' × peach 'Earlygold' reference map for *Prunus*. Theoretical and Applied. Genetics, **108:** 1120-130

Lamboy, W.F. and Alpha, C.G, (1998). Using simple sequence repeats (SSRs) for DNA fingerprinting germplasm accessions of grape (*Vitis* L.) species. Journal of the American Society for Horticultural Science, **123:** 182-188

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

Lecouls, A.C., Bosselut, N., Bergougnoux, V., Rubio-Cabetas, M.J., Bosselut, N., Voisin, R., Poessel, J.L., Faurobert, M., Bonnet, A.,G., Salesses, Dirlewanger, E. and Esmenjaud, D. (2004). Marker-assisted selection for the wide-spectrum resistance to root-knot nematodes conferred by the *Ma* gene from Myrobalan plum (*Prunus cerasifera*) in interspecific *Prunus* material. Molecular Breeding, **13:** 113-124

Ledbetter, C.A. and Palmquist, D.E. (2006). Comparing physical measures and mechanical cracking products of 'Nonpareil' almond (*Prunus dulcis* [Mill.] D. A.

Webb) with two advanced breeding selections, Journal of food engineering, **76:** 232-237

Lian, C., Miwa, M. and Hogetsu, T. (2001). Outcrossing and paternity analysis of *Pinus densiflora* (Japanese red pine) by microsatellite polymorphism. Heredity, **81:** 88-98

Li, C.D., Rossnagel, B.G. and Scoles, G.J. (2000). The development of oat microsatellite markers and their use in identifying relationships among avena species and oat cultivars. Theoretical and Applied Genetics, **101:** 1259-1268

Liebhard, R. Gianfranceschi, L., Koller, B., Ryder, C.D. Tarchini, R, Van de Weg, E. and Gessler, C. (2003). Mapping quantitative physiological traits in apple (*Malus* × *domestica* Borkh.). Plant Molecular Biology, **52:** 511-526

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

Lodhi, M.A., Ye, G.N., Weeden, N.F., Reisch, B.I. (1995). A molecular marker-based linkage map of *Vitis*. Genome, **38:** 786–794

Lorieux, M., Terrier, X., Goffinet, B., Lanaud, C., Gonzalez de Leon, D. (1995). Maximum-likelihood models for mapping genetic markers showing segregation distortion. 2. F2 populations. Theoretical and Applied Genetics, **90:** 81-89

Liu, B.H. (1998). Statistical genomics: linkage, mapping, and QTL analysis. CRC Press LLC, New York.

Lu, Z.X., Sosinski, B., Reighard, G.L., Baird, W.V., and Abbott, A.G. (1998). Construction of a gen**et**ic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstock. Genome, **41:** 199-207

Luby, J.J. and Shaw, D.V. (2001) Does marker-assisted selection make dollars and sense in a fruit breeding program? HortScience, **36:** 872-879

Maliepaard, C., **Alston, F.H., van Arkel, G., Brown, L.M., Chevreau, E., Dunemann, F., Evans, K.M., Gardiner, S.**, **Guilford, P., van Heusden, A.W., Janse, J., Laurens, F., Lynn, J.R., Manganaris, A.G., den Nijs, A.P.M., Periam, N., Rikkerink, E., Roche, P., Ryder, C., Sansavini, S., Schmidt, H., Tartarini, S., Verhaegh, J.J., Vrielink-van, M., Ginkel, G.** and **King, J.** (1998). Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theoretical and Applied Genetics, **97:** 60-73

Manly, K.F., Cudmore, Jr, R.H. and Meer, J.M. (2001). Map Manager QTX cross-platform, software for genetic mapping. Mammalian Genome, **12:** 930-932

Martínez-Gómez, P., Arulsekar, S. and Gradziel, T.M. (2002). Characterization of twin embryos in almond. Acta Horticulturae, **591:** 257-262

Martínez-Gómez, P., Arulsekar, S., Potter, D. and Gradziel, T.M. (2003a). Relationships among peach and almond and related species as detected by SSR markers. Journal of the American Society for Horticultural Science, **128:** 667-671

Martínez-Gómez, P., Sozzi, G. O., Sánchez-Pérez, R., Rubio, M. and Gradziel, T. M. (2003b). New approaches to *Prunus* tree crop breeding. Food, Agriculture and Environment, **1:** 52-63

Martínez-Gómez, P., Sánchez-Pérez, R., Dicenta, F. Howad, W., Arús, P. and Gradziel, T. M. (2007). Almond. pp. 230-242. In: Kole, C. (Ed). Genome mapping and molecular breeding in plants, Volume 4. Fruits and Nuts. Springer-Verlag Berlin Heidelerg.

Mekuria, G.T., Collins, G.G., Sedgley, M. and Lavee, S. (2001)**.** Identification of genetic markers in olive linked to olive leaf spot resistance and susceptibility. Journal of the American Society for Horticultural Science, **126:** 305-308

Messeguer, R., Viruel, M., de Vincent, M., Dettori, M. and Quarta, R. (1994). RFLPs. In: Messeguer R (ed). Methods of molecular marker analysis in *Prunus*. IRTA Technical publication, pp 26-38

Messina, R., Lain, O., Marrazzo, M.T., Cipriani, G. and Testolin, R. (2004). New set of microsatellite loci isolated in apricot. Molecular Ecology Notes, **4:** 432-434

Micke, W. and Kester, D. (1978). Almond orchard management. Agricultural Scientific Publications, University of California, Berkely, USA

Miller, R. (1997). Linkage mapping of plant and animal genomes. pp. 27-48. In: Dear P. H. (Ed). Genome mapping: A practical approach. Oxford University Press

Mnejja, M., Garcia-Mas, J., Howad, W., Badenes, M.L. and Arús, P. (2004). Simple sequence repeat (SSR) markers of Japanese plum (*Prunus saliciana* Lindl.) are highly polymorphic and transferable to peach and almond. Molecular Ecology Notes, **4:** 163-166

Mnejja, M., Garcia-Mas, J., Howad, W., Badenes, M.L. and Arús, P. (2005). Development and transportability across *Prunus* species of 42 polymorphic almond microsatellites. Molecular Ecology Notes, **5:** 531–535

Mohan, A., Goyal, A., Singh, R., Balyan, H.S., Pushpendra K. and Gupta, P.K. (2007). Physical mapping of wheat and rye expressed sequence tag–simple sequence repeats on wheat chromosomes. Crop Science, **47:** 3-13

Moldenke, H.H. and Modenke, A.L. (1952). Plants of the bible. Chronica Botanica, Co., Waltham, MA.

Monselise, S.P., and Goldschmidt, E.E. (1982). Alternate bearing in fruit trees. Horticultural Reviews, **4:** 128-73

Mookerjee, S., Guerin, J., Collins, G., Ford, C. and Sedgley, M. (2005). Paternity analysis using microsatellite markers to identify pollen donors in an olive grove. Theoretical and Applied Genetics, **111:** 1174-1182

Morgante, M. and Oliviery, A.M. (1993). PCR amplified microsatellites as markers in plant genetics, Plant Journal, **3:** 175-182

Murua, J. R. (1994). Yield model for Spanish almonds. Acta Horticulturae, **373:** 325-332

Murray, M. and Thompson, W. (1980). Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research, **8:** 4321-4325

Myburg, A.A., Griffin, 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

Nanos, G.D., Kazantzis, I., Kefalas, P., Patrakis, C., and Stavroulakis, G.G. (2002). Irrigation and harvest time affect almond kernel quality and composition. Scientia Horticulturae, **96:** 249-256

Nier, S., Simpson, D.W., Tobutt, K.R. and Sargent, D.J. (2006). A genetic linkage map of an inter-specific diploid *Fragaria* BC1 mapping population and its comparison with the *Fragaria* reference map (FV × FN). Journal of Horticultural Science and Biotechnology, **81:** 645–650

Oraguzie, N.C. and Wilcox, P.L. (2007). An overview of association mapping. Eds Oraguzie, N. C. In: Association mapping in plants. Springer Science + Business Media, New York, USA

Oliveira, E.J., Vierira, C.L.M., Garcia, A.A.F., Munhoz, G.F., Margarido, G.R. A. Consoli, L., Matta, F. P., Moraes, M.C., Zucchi, M.I., and Fungaro, M.H.P. (2008). An integrated molecular map of yellow passion fruit based on simultaneous maximum-likelihood estimation of linkage and linkage phases, Journal of American Society for Horticultural Science, **133:** 35-41

Pablo, J., Oritiz, A., Pessino, S.C., Bhat, V., Hayward, M.D. and Quarin, C.L. (2001). A genetic linkage map of diploid *Paspalum notatum*. Crop Science, **41:** 823-830

Paterson, A.H., Lander, E.S., Hewitt, J.D., Peterson, S., Lincoln, S.E. and Tanksley, S.D. (1988). Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature, **335:** 721-726

Paterson, A.H. (1996). Making genetic maps. pp. 41-54. In: Paterson A.H. (ed.) Genome mapping in plants. Landes Company, San Diego, California; Academic Press, Austin, Texas.

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

Pederse, B.H. (2006). DNA fingerprints of 51 sweet and sour *Prunus* accessions using Simple Sequence Repeats. Journal of Horticultural Science and Biotechnology, **81:** 118–124

Phippen, C. and Ockendon, D.J. (1990). Genotype, plant, bud size and media factors affecting anther culture of cauliflowers (*Brassica oleracea* var. *botrytis*). Theoretical and Applied Genetics, **79:** 33-38

Pimienta, E. and Polito, V.S. (1982). Ovule abortion in 'Nonpareil' almond (*Prunus dulcis* [Mill.] D.A. Webb). American Journal of Botany, **69:** 913-920

Pound, L.M. and Collins, G.C. (2003). The carob moth. A literature review with particular reference to its effect on almond. Nutgrower, **17:** 30-33

Powell W., Morgante, M., Andre, C., Hanafey, M. ,Vogel, J., Tingey, S. and Rafleski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR markers for germplasm analysis. Molecular Breeding, **2:** 225-238

Rabinson, J.P. and Harris S.A. (1999). Amplified fragment length polymorphism and microsatellites: A phylogenic perspective. In: Gillet, E.M. (ed.).Which DNA Marker for which purpose? Final Compendium of the Research Project Development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union

Rafleski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR markers for germplasm analysis. Molecular Breeding, **2:** 225-238

Rajapakse, S., Belthoff, L.E., He, G., Estager, A. E., Scorza, R., Verde, I., Ballard, R. E. Baird, W.V., Callhahan, A., Monet, R., and Abbott, A.G. (1995). Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. Theoretical and Applied Genetics, **90:** 503-510

Ramsay, L., Macaulay, M., Ivanissevich, S., MacLean, K., Cardle, L. Fuller, J., Edwards, K.J. , Tuvesson, S. , Morgante, M., Massari, A. Maestri, E., Marmiroli, N., Sjakste, T., Ganal, M., Powell, W. and Waugh, R. (2000). A Simple sequence repeat-based linkage map of barley. Genetics, **156:** 1997-2005

Riaz, L.S., Dangl, G.S., Edwards, K.J. and Meredith, C.P. (2004). A microsatellite marker based framework linkage map of *Vitis vinifera* L. Theoretical and Applied Genetics, **108:** 864-872

Riley, R. and Law, C.N. (1984). Chromosome manipulation in plant breeding: progress and prospects. pp 301-322. In: Gustafson JP (Ed). Gene manipulation in plant improvement. Plenum Press, New York.

Ritter, E., Gebhardt, G. and Salamini, F. (1990). Estimation of recombination frequencies and construction of RFLP linkage map in plant from crosses between heterozygous parents. Genetics, **12:** 645-654

Rossetto, M. (2001). Sourcing of SSR markers from related plant species. pp 211–224. In: Henry, R. (Ed). Plant genotyping the DNA fingerprinting of plants. CAB International, Wallingford, UK.

Roy, S.P. and Kochba, J. (1974).The inheritance of bitter and double kernel characters in the almond *(Prunus amigdalus,* Batsch). Plant Breeding, **71:** 19-329

Rugini, N. (1986). Almond. pp. 507-611. In: Evans, A., Sharp, D., Williams, R. and Phillip, A. (Eds). Handbook of plant cell culture. MacMillan Publishing, NY.

Sabate, J., Bell, H.E.T. and Fraser, E.G. (1996). Nut composition and coronary heart disease risk. pp. 145-152. In: Spiller, G.A.(ed). Lipids in human nutrition. CRC Press, Boca Raton, FL

Sambrock, J. and Russell, D.W. (2001). Molecular cloning: A laboratory manual. Vol 3. Cold Spring Harbour Laboratory Press, USA

Sánchez-Pérez, R., Howad, W., Arús, P., Martínez-Gómez, P. and Dicenta, F. (2006a). New findings in the genetic control and expression of kernel bitterness in almond. NUCIS, **13:** 20-22

Sánchez-Pèrez, R., Ballester, J., Dicenta, F., Arús, P. and Martínez-Gómez, P. (2006b). Comparison of SSR polymorphisms using automated capillary sequencers, and polyacrylamide and agarose gel electrophoresis: Implications for the assessment of genetic diversity and relatedness in almond. Scientia Horticulturae, **108:** 310–316

Sánchez-Pérez, R., Ortega, E. Duval, H., Martínez-Gómez, P. and Dicenta, F. (2007a). Inheritance and relationships of important agronimc traits in almond. Euphytica, **155:** 381-391

Sánchez-Pèrez, R., Howad, W., Dicenta, F., Arús, P. and Martínez-Gómez, P. (2007b). Mapping major genes and quantitative trait loci controlling agronomic traits in almond. Plant Breeding, **126:** 310-318

Sang, S., Lapsley, K., Jeong, W., Lachance, P., Ho, C. and Rosen, R. (2002). Antioxidative phenolic compounds isolated from almond skins (*Prunus amygdalus* Batsch). Journal of Agriculture, Food and Chemistry, **50:** 2459–2463

Schneider, K. (2005). Mapping populations and principles of genetic mapping. pp. 1-21. In: Meksem, K. and Kahl, G. (eds). The handbook of plant genome mapping, genetic and physical mapping. Willey-VCH, Weinheim,

Schubert, R. and Mjller-Starck, G. (2002). Development and application of molecular markers in conifers. In: Molecular techniques in crop improvement. pp. 85-137. (Eds), Jain, S.M. Brar, D.S. and Ahloowalia, B.S. Kluwer Academic Publishers

Scorza, R. (1996). Genome mapping of horticultural crops: introduction to the colloquim. HortScience: **31:** 1098

Scorza, R. (2001). Progress in tree fruit improvement through molecular genetics. HortScience, **36:** 855-857

Sedgley, M. and Collins, G. (2002). Almond improvement in Australia. Fruits, **57:** 129-134

Semagn, K., Bjørnstad, Å. and Ndjiondjop, M.N. (2006). Principles, requirements and prospects of genetic mapping in plants. African Journal of Biotechnology, **5:** 2569-2587

Serrano, B., Gomez-Aparisi, J. and Hormaza, J.I. (2002). Molecular fingerprinting of Prunus rootstocks using SSRs. Journal of Horticultural Science & Biotechnology, **77:** 368-372

Silva, C., Garcia-Mas, J., Sánchez, A.M., Arús, P. and Oliveira, M.M. (2005). Looking into flowering time in almond (Prunus dulcis (Mill) D. A. Webb): the candidate gene approach. Theoretical and Applied Genetics, **110:** 959–968

Silvertown, J. and Charlesworth, D. (2001). Introduction to plant biology. 4th edition. Blackwell Publishing. 347 pages.

Slavov, G.T., Howe, G.T., Gyaourova, A.V., Birkes, D.S. and Adams, W.T. (2005). Estimating pollen flow using SSR markers and paternity exclusion: accounting for mistyping. Molecular Ecology, **14:** 3109-3121

Socias i Company, R. (1997). Almond genetics: Past, present and future, Acta Horticulturae, **470:** 57-65

Socias, i Company, R. (1998). Fruit tree genetics at a turning point: the almond example. Theoretical and Applied Genetics, **96:** 588-601

Socias i Company, R., Kodad, O. and Alonso, J.M. (2008). Almond quality: A breeding perspective. (Ed) Janick, J. Horticultural Reviews, **34:** 197-238

Soderstrom, E. (1977). Seal of almond shells and resistance to navel orangeworm. Journal of Economic Entomology, **70:** 467-468

Sosinski, B., Gannavarapu, M., Hager, L.D., Beck, L.E., King, G.J., Ryder, C.D., Rajapakse, S. Baird, W.V., Ballard, R.E. and Abbott, A.G. (2000). Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. Theoretical and Applied Genetics, **101:** 421–428

Spiegel Roy, P. and Kochba, J. (1974). The inheritance of bitter and double kernel characters in the almond (*Prunus amygdalus* Bstsch). Journal of Plant Breeding, **71:** 319-329

Spiegel-Roy, P. and Kochba, J. (1981). Inheritance of nut and kernel traits in almond (*Prunus amygdalus* Bstsch). Euphytica, **30:** 167-174

Spiegel-Roy, P. (1986). Domestication of fruit trees. pp. 201-211. In: Barigozzi, C. (Ed). Origin and Domestication of cultivated Plants. Elsevier, Amsterdam.

Srinivasan, C., Padilla, I.M.G. and Scorza, R. (2005). *Prunus* spp. Almond, Apricot, Cherry, Nectarine, Peach and Plum. pp. 512-542. In: Litz, R.E. (Ed), Biotechnology of Fruit and Nut Crops. CABI Publishing, Wallingford, UK.

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

Staub, J.E., Serquen, F.C. and Gupta, M. (1996). Genetic markers, map construction and their application in plant breeding. Hortscience, **31:** 729-741

Steenkamp, J. Wild, I., Lourens, A. and Van Heldon, P. (1994). Improved method for DNA extraction from *Vitis vinifera*. American Journal of Enology and viticulture **45:** 102-106

Struss, D., Ahmad, R. and Southwick, S.M. (2003). Analysis of sweet cherry cultivars (*Prunus avium*) using SSR and AFLP markers. Journal of the American Society for Horticultural Science, **128:** 904-909

Suiter, K.A., Wendel, J.F. and Case, J.S. (1983). Linkage: A pascal computer program for the detection and analysis of genetic linkage. Journal of Heredity, **74:** 203-204

Tang, H.E., Bowers, J.E., Wang, X., Ming, R., Alam, M. and Paterson, A.H. (2008). Synteny and collinearity in plant genomes. Science, **320:** 486 - 488

Tanksley, S.D. (1984). Linkage relationships and chromosomal locations of enzyme-coding genes in pepper, *Capsicum annuum.*. Chromosoma, **89:** 352-360

Tanksley, S.D., Young, N.D., Paterson, A.H. and Bonierbale, M.W. (1989). RFLP mapping in plant breeding: new tools for an old science. Biotechnology, **7:** 247-264

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., Roeder, 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

Tartarini, S. and Sansavini, S. (2003). The use of molecular markers in pome fruit breeding. Acta Horticulturae, **622:** 129-140

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acid Research, **17:** 6463-6471

Testolin, R., Marrazzo, T., Cipriani, G., Quarta, R., Verde, I., Dettori, M.T., Pancaldi, M. and Sansavini, S. (2000). Microsatellite DNA in peach *Prunus Persica* L. Batsch) and its use in fingerprinting and testing the genetic origin of cultivars. Genome, **43:** 512-520

Testolin, R. (2003). Marker assisted selection (MAS) in stone fruits. Acta Horticulturae, **633:** 163–176

Thompson, M.M. (1977). Inheritance of nut traits in filbert (*Corylus avellana* L.). Euphytica, **26:** 465-474

Traynor, J. (1966). Increasing the pollinating efficiency of honeybees. Bee World, **47:** 101-110

United State Department of Agriculture (USDA). (2007). National Agricultural Statistics Service, California Field Office

Valdes, A.M., Slatkin, M. and Freimer, N.B. (1993). Allele frequencies at microsatellite loci: The stepwise mutation model revisited. Genetics, **133:** 737-749

Van Ooijen, J.W., and Voorrips, R.E. (2001). JoinMap® 3.0 Software for the calculation of genetic linkage maps. Wageningen, the Netherlands.

Vavilov, N.I. (1930). Wild progenitors of the fruit trees of Turkistan and Cacausus and the problem of the origin of fruit trees. pp. 271-286. IX International Horticultural Congress reports and proceeding. London.

Verde, I., Lauria, M., Dettori, M.T., Vendramin, E., Balconi, C., Micali, S., Wang, Y. Marrazzo, M.T., Cipriani, G., Hartings, H., Testolin, R., Abbott, A.G., Motto, M. and Quarta, R. (2005). Microsatellite and AFLP markers in the *Prunus persica* [L.

(Batsch)] × *P. ferganensis* BC_1 linkage map: saturation and coverage improvement. Theoretical and Applied Genetics, **111:** 1013-1021

Verhaegen, D. and Plomion, C. (1996). Genetic mapping in Eucalyptus urophylla and Eucalyptus grandis using RAPD markers. Genome, **39:** 1051-1061

Vezvaei, A., Hancock, T.W., Giles, L.C., Clarke, G.R. and Jackson, J.F. (1995). Inheritance and linkage of isozyme loci in almond. Theoretical and Applied Genetics, **91:** 432-438

Vilanova, S., Soriano, J.M., Lalli, D.A., Romero, C., Abbott, A.G. , Llacer, G. and Badenes, M.L. (2006). Development of SSR markers located in the G1 linkage group of apricot (*Prunus armeniaca* L.) using a bacterial artificial chromosome library. Molecular Ecology Notes, **6:** 789–791

Viruel, M.A., Messeguer, R., de Vicente, M.C., Garcia-Mas, J., Puigdomènech, P., Vargas, F. and Arús, P. (1995). A linkage map with RFLP and isozyme markers for almond. Theoretical and Applied Genetics, **91:** 964-971

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., Raijans, M., Lee, T.V., 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

Wang, G., Mahalingan, R. and Knap, H.T. (1998a). (C-A) and (GA) anchored simple sequence repeats (ASSRs) generated polymorphism in soybean, *Glycine max* (L.) Merr. Theoretical and Applied Genetics, **96:** 1086-1096

Wang, D., Karle, R., Brettin, T.S. and Iezzoni, A. F. (1998b). Genetic linkage map in sour cherry using RFLP markers. Theoretical and Applied Genetics, **97:** 1217-1224

Wang, D., Karle, R. and Iezzoni, A.F. (2000). QTL analysis of flower and fruit traits in sour cherry. Theoretical and Applied Genetics, **100:** 535-544

Wang, Q., Zhang, K., Qu, X., Jia, J., Shi, J., Jin, D., Wang, B. (2001). Construction and characterization of a bacterial artificial chromosome library of peach. Theoretical and Applied Genetics, **103:**1174-1179

Wang, Y., Georgi, L., Zhebentyayeva, T., Reighard, G.L., Scorza, R. and Abbott, A.G. (2002a). High-throughput targeted SSR marker development in peach (*Prunus persica*). Genome, **45:** 319-328

Wang, Y., Garay, L., Reighard, G.L., Georgi, L.L., Abbott, A.G. and Scorza, R. (2002b). Development of bacterial artificial chromosome contigs in the evergrowing gene region in peach. Acta Horticulturae, **592:** 183-189

Warner, D.J. and Creller, M.A. (1997). Genetic studies in peach: inheritance of sweet kernel and male sterility. Journal of American Society for Horticultural Science, **122:** 215-217

Watkins, R. (1995). Cherry, plum, peach, apricot and almond. pp. 423-428. In: Smartt, J. and Simmonds, N.W. (Eds), Evolution of crop plants. Longman Scientific and Technical, England.

Wattier, R., Engel, C.R., Saumitou-Laprade, P. and Valero, M. (1998). Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the Dinucleotide locus *Gv1CT* in *Gracilaria gracilis* (Rhodophyta). Molecular Ecology, **7:** 1569-1573

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

Weeden, N.F., Timmerman, G.M. and Lu, J. (1994). Identifying and mapping genes of economic significance. Euphytica, **73:** 191-198

Weising, K., Nybom, H., Wolff, K. and Kahl, G. (2005). Linkage analysis and genetic maps. pp. 277-291. In: Weising (ed). DNA Fingerprinting in Plants: Principles, Methods, and Applications, CRC Press

Wenzl, P., Carling, J., Kudrna, D., Jaccoud, D., Huttner, E., Kleinhofs, A. and Kilian, A. (2004). Diversity Arrays Technology (DArT) for whole-genome profiling of barley. Proceedings of the National Academy of Science of the United States of America, **101:** 9915-9920

Wirthensohn, M. and Sedgley, M. (2002). Almond breeding in Australia. Acta Horticulturae, **591:** 245-248

Wirthensohn, M., Collins, G. and Sedgley, M. (2002). International collaboration in almond genome mapping. Australian Nutgrower, **15:** 26-27

Wirthensohn, M.G., Chin, W.L., Franks, T., Baldock, G. and Ford, C.M. (2008). Characterising the flavour phenotypes of almond (*Prunus dulcis* Mill.) kernels. Journal of Horticultural Science and Biotechnology, **83:** 462-468

Woolley, F., Collines, G., Sedgley, M. (2000). Application of DNA fingerprinting for the classification of selected almond [Prunus dulcis (Miller). D. A Webb] cultivars. Australian Journal of Experimental Agriculture, **40:** 995-1001

Wu, S., Collins, G. and Sedgley, M. (2004). A molecular linkage map of olive (*Olea europea* L.) based on RAPD, microsatellite and SCAR markers. Genome, **125:** 59-67

Wunsch, A. and Hormaza, J.I. (2002). Cultivars identification and genetic fingerprinting of temperate fruit tree species using DNA markers. Euphytica **125:** 59-67

Xia, L., Peng, K., Yang, S., Wenzl, P., de Vicente, M.C., Fregene, M. and Kilian, A. (2005). DArT for high-throughput genotyping of cassava (*Manihot esculenta*) and its wild relatives. Theoretical and Applied Genetics, **110:** 1092-1098

Xie, H., Sui, Y., Chang, F.Q. Xu, Y. and Ma, R.C. (2006). SSR allelic variation in almond (*Prunus dulcis* Mill.). Theoretical and Applied Genetics, **112:** 366–372

Yamamoto, T., Shimada, T., Imai, T., Yaegaki, H., Haji, T., Matsuta, N., Yamaguchi, M. and Hayashi, T. (2001) Characterization of morphological traits based on a genetic linkage map in peach. Breeding Science, **51:** 271–278

Yamamoto, T., Kimura, T., Shoda, M., Imai, T., Saito, T Sawamura, Y. Kotobuki, K., Hayashi, K. and Matsuta, N. (2002). Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. Theoretical and Applied Genetics, **106:** 9-18

Yao, Q. and Mahlenbacher, S. A. (2000). Heritability, variance components and correlation of morphological and phenological traits in hazelnut. Plant Breeding, **119:** 369-381

Zhebentyayeva, T.N., Swire-Clark, G. Georgi, L.L., Garay, L., Jung, S., Forrest, S., Blenda, A.V., Blackmon, B., Mook, J., Horn, R., Howad, W., Arús, P., Main, D., Tomkins, J.P., Sosinski, B., Baird, W.V., Reighard, G.L. and Abbott, A.G. (2008). A framework physical map for peach, a model Rosaceae species. Tree Genetic and Genomes, (In press).

APPENDIX

Appendix Microsatellite markers tested in present study, primer name, sequences, repeat motif and annealing temperature (T_a)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
AP2M	GGG AGT GGA GGC TGA CAT AAA T CCT TGC CAT TGC ACT TGA TTG C	?	288	√	--	60	Personal comm. Dr Howad, 2006
BPPCT001	AAT TCC CAA AGG ATG TGT ATG AG CAG GTG AAT GAG CCA AAG C	$(GA)_{27}$	159	√	√	57	Dirlewanger et al., 2002
BPPCT002	TCG ACA GCT TGA TCT TGA CC CAA TGC CTA CGG AGA TAA AAG AC	$(AG)_{25}$	229	√	√	57	"
BPPCT004	CTG AGT GAT CCA TTT GCA GG AGG GCA TCT AGA CCT CAT TGT T	$(CT)_{22}$	200	√	√	57	"
BPPCT008	ATG GTG TGT ATG GAC ATG ATG A CCT CAA CCT AAG ACA CCT TCA CT	$(GA)_{36}$	148	√	√	57	"
BPPCT009	ATT CGG GTC GAA CTC CCT ACG AGC ACT AGA GTA ACC CTC TC	$(CT)_{14}$	171	√	√	57	"
BPPCT010	AAA GCA CAG CCC ATA ATG C GTA CTG TTA CTG CTG GGA ATG C	$(AG)_4GG(AG)_{10}$	131	√	√	57	"
BPPCT012	ACT TCC ATT GTC AGG CATC A GGA GCA ACG ATG GAG TGC	$(CT)_{13}CC(CT)_7$	164	--	--	57	"
BPPCT013	ACC CAC AAA TCA AGC ATA TCC AGC TTC AGC CAC CAA GC	$(AG)_{28}$	183	√	√	57	"
BPPCT015	ATG GAA GGG AAG AGA AAT CG GTC ATC TCA GTC AAC TTT TCC G	$(AG)_{13}$	150	--	--	57	"
BPPCT016	GAT TGA GAG ATT GGG CTG C GAG GAT TCT CAT GAT TTG TGC	$(AG)_{14}$	96	√	--	57	"
BPPCT019	TGA TAC CAC CAT CCA ATC TAG C TTG CTG GGA CAT GGT CAG	$(CT)_{28}$	194	--	--	57	"
BPPCT020	CGT GGA TGG TCA AGA TGC ATT GAC GTG GAC TTA CAG GTG	$(AG)_{14}GG(AG)_7$ $AT(AG)_8$	200	√	√	57	"
BPPCT021	TGC ATG AGA AAC TTG TGG C CCA AGA GCC TGA CAA AGC	$(GA)_{24}$	240	√	√	57	"
BPPCT023	TGC AGC TCA TTA CCT TTT GC AGA TGT GCT CGT AGT TCG GAC	$(CT)_{21}$	224	√	--	57	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
BPPCT024	GAG GAA TGT GCC TCT TCT GG CTC CCG TAC GCG TTT ACC	$(AG)_{15}$	96	--	--	57	Dirlewanger et al., 2002
BPPCT026	ATA CCT TTG CCA CTT GCG TGA GTT GGA AGA AAA CGT AAC A	$(AG)_8GG(AG)_6$	134	√	--	57	"
BPPCT027	CTC TCA AGC ATC ATG GGC TGT TGC CCG GTT GTA ATA TC	$(GA)_{11}$	249	√	√	57	"
BPPCT029	GGA CGG ACA GAA ATG AAG GT CCT TAA CCC ACG CAA CTC C	$(GA)_{12}(CAGA)_4$	159	√	√	57	"
BPPCT030	AAT TGT ACT TGC CAA TGC TAT GA CTG CCT TCT GCT CAC AC C	$(AG)_{25}$	175	√	--	57	"
BPPCT032	TTA AGC CAC AAC ATC CAT GAT AAT GGT CTA AGG AGC ACA CG	$(AG)_{10}CG(AG)_{13}$	203	√	√	57	"
BPPCT033	GTA GCC GGA GCC GTG TAT CTA GAA CCC TAT AAA CAC ATG GC	$(AG)_{32}$	180	√	√	57	"
BPPCT035	TGA AGG ATG GCT CTG ATA CC AAT TCA TCT ACT TCT TCC TCA AGC	$(GA)_{33}$	113	--	--	57	"
BPPCT036	AAG CAA AGT CCA TAA AAA CGC GGA CGA AGA CGC TCC ATT	$(AG)_{11}$	253	√	--	57	"
BPPCT037	CAT GGA AGA GGA TCA AGT GC CTT GAA GGT AGT GCC AAA GC	$(GA)_{25}$	155	√	√	57	"
BPPCT038	TAT ATT GTT GGC TTC TTG CAT G TGA AAG TGA AAC AAT GGA AGC	$(GA)_{25}$	135	--	--	57	"
BPPCT039	ATT ACG TAC CCT AAA GCT TCT GC GAT GTC ATG AAG ATT GGA GAG G	$(GA)_{20}$	154	√	√	57	"
BPPCT040	ATG AGG ACG TGT CTG AAT GG AGC CAA ACC CCT CTT ATA CG	$(GA)_{14}$	135	√	√	57	"
BPPCT042	AAC CCT ACT GGT TCC TCA GC GAC CAG TCC TTT AGT TGG AGC	$(CT)_{25}$	243	--	--	57	"
CPDCT004	TCT CAG GTT CGT ATC CCC TCT GCC CAT TTT GTG TGT GTC AA	$(CT)_{19}$	151	√	--	62	Mnejja et al., 2005

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
CPDCT011	ATG GTC TAA AAA CCG CGA AG GGA GAT CAA GAC CGC CTG T	$(TC)_6$-$(CT)_8$	176	√	√	62	Mnejja et al., 2005
CPDCT013	GTT TTA GAA ACC TCA TTC CAA CTT AAT TCT AAC ACT GGG GTA TTG T	$(CT)_{14}$	100	--	--	62	"
CPDCT016	GGA AAC CTG ATT AGG GCA CTT GGT CTG CTA TAC TGA CCT AGG ATT	$(GA)_{19}$	196	√	√	62	"
CPDCT017	CGT GCC ACG AGA ATG AGA AT CCA GGA CTT AGG AGG TGT CG	$(GA)_{13}$	179	√	--	62	"
CPDCT032	TCA GCT CTC TTT CTC CTC ACG GGA AAT CGG CTA GCC TTG AT	$(CT)_5$-$(CT)_6$- $(CT)_6$-$(CT)_6$	185	√	√	62	"
CPDCT038	ATC ACA GGT GAA GGC TGT GG CAG ATT CAT TGG CCC ATC TT	$(GA)_{25}$	181	√	√	62	"
CPDCT044	ACA TGC CGG GTA ATT AGC AA AAA ATG CAC GTT TCG TCT CC	$(GA)_{21}$	175	√	√	62	"
CPPCT003	GTA ACG AAG AAG TTA CGG G AAC TGT CGC TGC TGG GTT	$(TC)_{17}$	160	√	√	52	Aranzana et al., 2002
CPPCT004	TCA TTC GAA GAC GAC CGTGTC TAG GCA CGT TGC TAG	$(CT)_{22}$	250	--	--	52	"
CPPCT006	AAT TAA CTC CAA CAG CTC CA ATG GTT GCT TAA TTC AAT GG	$(CT)_{16}$	190	√	√	59	"
CPPCT007	GTAGTCCTTATGGTGTACAGG CATAATGGTGCAAATTCATA	$(CT)_8$	147	--	--	59	"
CPPCT008	GAG CTC TCA CGC ATT AGT TT TTT GAC TGC ATA ACA AAA CG	$(CT)_6$ ... $(CT)_7$	161	√	√	59	"
CPPCT009	CAATGC AAT TAT GCA AGT GT TGG ACA TTG TTT AAG CAT CA	$(CT)_{12}$	171	√	√	59	"
CPPCT010	GAA TAT TTG GAT TGC AAA GG GGA ATA TAA GCT CTG CTG CT	$(CT)_{18}$	142	√	√	59	"
CPPCT013	GCA TTT CGA GAG CTG TAT TT GTC TTA CGT GCA GCT TCA TT	$(CT)_{14}$	150	√	√	59	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
CPPCT015	TGG AGT GCC AAT ACT ATT TA CAT ATG CAT GGT TAT GGT	$(CT)_{31}$	200	--	--	50	Aranzana et al., 2002
CPPCT016	AAT TCC CTA TGG AAA TTA GA CGC ATA TTA TAG GTA GGA AA	$(CT)_{24}$	191	√	--	50	"
CPPCT017	TGA CAT GCA TGC ACT AAA CAA TGC AAA TGC AAT TTC ATA AAG G	$(CT)_{18}$	177	--	--	60	"
CPPCT018	TAC GTG CAC CCT ACT GCT TG TC CAA AGT TAG TCA ATT TCT TTC	$(CT)_{18}$	142	--	--	60	"
CPPCT019	AAT TCA ATG TCA AGA CAC A TCA TCA AAA TAA ATA TCC AGT	$(CT)_{22}$	184	√	--	50	"
CPPCT021	CGG ATC CCA GTT GTA TTA AAT G GAG GAA CTG GTT ATC ACC TTG G	$(CT)_{10}$	142	√	--	60	"
CPPCT023	CAT GGT TTG CAA CTG TCT TCA GAC ACA GGT GTG TAG ATC ATT GG	$(CT)_{9}$	166	√	√	55	"
CPPCT024	TTC TCC CAA AAA CCA AAA CC TCA TTG GCT GCT AAG TGT CCT	$(CT)_{35}$	180	√	√	50	"
CPPCT026	AGA CGC AGC ACC CAA ACT AC CAT TAC ATC ACC GCC AAC AA	$(CT)_{22}$	180	√	√	55	"
CPPCT027	GAG CAG TTC ATA AGT TGG AAC AA CGA TAA AGA TTT TGA CTG CAT GA	$(CT)_{30}$	114	--	--	55	"
CPPCT028	ACA TAT GCC TTA TCA GCT T ATT GAA GAG AAA GCA GTG T	$(CT)_{16}$	130	--	--	50	"
CPPCT029	CCA AAT TCC AAA TCT CCT AAC A GA TCA ACT TTG AGA TTT GTT GAA	$(CT)_{24}$	190	√	√	55	"
CPPCT033	TCA GCA AAC TAG AAA CAA ACC TTG CAA TCT GGT TGATGT T	$(CT)_{16}$	151	√	√	50	"
CPPCT034	TCG GTT TTT AAA ATT CCA AAA GTT ACC CTT ATT TGC ACC CAA CA	$(CT)_{14}$	193	√	√	60	"
CPPCT035	CTA CCC ATT AGC CAC CAA GC TCC CAA TTC GTT GCA ATC TT	$(CT)_{21}$	176	--	--	50	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
CPSCT004	GCT CTG AAG CTC TGC ATT GA TTT GAA ATG GCT ATG GAG TAC G	$(GA)_8$	130	√	√	62	Mnejja et al., 2004
CPSCT005	CTG CAA GCA CTG CGG ATC TC CCC ATA TTC CCA ACC CAT TA	$(CT)_{15}$	175	√	√	62	"
CPSCT008	TGG ATC CAA TCC AAG AGT CTG GCA AAA GCA GAG ATG GAG AGA	$(GA)_{17}$	198	--	--	62	"
CPSCT011	ATT TGG GTT TGC GAC TCA AG ACT CAT CCC TTG CCC TTT CT	$(CT)_{21}$	186	√	--	62	"
CPSCT023	CGG GTT GAC TCA GTT CCT TC ATT TGA GCA GAA GCC CAG AA	$(CT)_8$	135	√	√	46	"
CPSCT026	TCT CAC ACG CTT TCG TCA AC AAA AAG CCA AAA GGG GTT GT	$(CT)_{16}$	195	--	--	46	"
CPSCT027	CCC ATG CTC CTG TGG TAA GT TTT AGA ATC CCA ACC CCA CA	$(GA)_{23}$	166	√	--	62	"
CPSCT030	CAA CAG CGA GTG TCA CGT TT AGG CAA CGG ACA AAA ATC TG	$(CT)_{12}$	191	--	--	46	"
CPSCT031	TTC AGA TGA AAA AGA AAA AGA AAG T AAA GAA ACG CTT GTC TTG CAC	$(GA)_{18}$	191	√	√	46	"
CPSCT032	CAT CAT CAT CCT CAC CCA AA TGC TGA TCC GTG AGA TCT TG	$(GA)_{10}$	197	√	√	62	"
CPSCT033	TCC TCA TTT GAG TGT TGT GGA TGC CCA ATT TGA AAA CTT TGT	$(GT)_{11}(GA)_{13}$	158	√	√	62	"
CPSCT034	AGG TGG ACA ATA GCC GTG AT TTT CCA GAC CCT GAG AAA GC	$(CT)_5GT(CT)_6 \ldots (CA)_5$	182	√	--	62	"
CPSCT036	TCG AAG ACA GAC CAG ACA AAG A TTG CCT TAT GCC GGT AAT TT	$(GA)_5 \ldots (GA)3CGG(GA)_7$	134	--	--	62	"
CPSCT037	TCA AAT AAA AGA GAG AAA GAG AGA G GCT TTG CTT AAT TTT CTA TGA TCG	$(GA)_{17}$	118	--	--	62	"
CPSCT039	GCC GCA ACT CGT AAG GAA TA TCC ACC GTT GAT TAC CCT TC	$(GA)_{18}$	122	--	--	62	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results Amplified	Results Polymorphic	T_a °C	References
CPSCT042	TGG CTC AAA AGC TCG TAG TG CCA ACC TTT CGT TTC GTC TC	$(GA_{)10}$	164	√	√	62	Mnejja et al., 2004
CPSCT044	CCA GCA CAG AGA AAA CGA TG GAG CTC CTA CTC TGA GTC TGT AAA A	$(GT)_8(AT)_4$ … $(GA)_7CA(GA)_4$	200	√	√	62	"
EMPA004	TAC GGT AGG CTT CTG CAA GG TTG GCA GGT TCT GTT CAC AT	$(GA)_4AA(GA)_4$ $AA(GA)_{15}$	189	√	√	55	Clark and Tobutt, 2003
EMPA009	CAT GAT CTG CAC TGG GAA TC AGA GGG CAA GAA AAA GAC GC	$(AG)_{12}$	205	√	--	55	"
EMPA011	TGT GCT CAC TCT CTG CTG CT TGT GTG GGT TCA CAG TCT CC	$(AG)_{16}$	245	√	--	55	"
EPDC3853	TGT CGC CAC CAT GTC TGT TT ACA TTG TGT TTG CGG TGG T	?	197	√	--	57	Pers. Comm. with Dr Arus
EPDCU2862	GTG GAA AAA CCT GCT CCA GA TCA TTC TCT TCC CCA GAT GC	?	146	√	--	57	"
EPDCU3117	CAG AGG GAA CAG TGT GAG CA TGT TGT TGT CGA CCC TGA AA	?	164	√	--	57	"
EPDCU3122	AGC GGA GTG TAC AGC AAG GT TAT GTT GTT TCC GGC ATT GA	?	177	√	--	57	"
EPDCU3489	AAA TCA GCT CCC ATC ACT CC AGC TGA GTG GAA CCA GAG GA	?	169	--	--	57	"
EPDCU3516	ACC GTT AAC GAG GCT CAG TC ACC TCC ACT GCC ATA TCC AC	?	183	√	--	57	"
EPDCU4017	CTG AGC CCA CCT CTG TTC TC CAA GCA GGT GGT TGT GTC TG	?	152	√	--	57	"
EPDCU5100	CTC TTC TCG CCT CCC AAT TT TGC TTA GCC CTG GGT ACA AG	?	174	√	--	57	"
MA020	CTT GCC CAT TTA TGT ACT GA TAT ATC GCA TAA TCA CGG TC	$(AG)_{23}$	180	√	√	55	Yamamoto et al., 2002
PacC3	TGA CTT GAT CAG ACT CGA CA TTG CAT TTG CAT TTA CAA TAG A	$(GA)_{16}$	?	√	--	55	Decroocq et al., 2003

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
PacA33	TCAGTCTCATCCTGCATACG CATGTGGCTCAAGGATCAAA	$(GA)_{16}$	188	√	--	56	Decroocq et al., 2003
PceGA25	GCA ATT CGA GCT GTA TTT CAG ATG CAG TTG GCG GCT ATC ATG TCT TAC	?	?	--	--	58	Downey and Iezzoni, 2000
PceGA34	GAA CAT GTG GTG TGC TGG TT TCCACT AGG AGG TGC AAA TG	?	?	√	√	50	"
PceGA59	AGA ACC AAA AGA ACG CTA AAA TC CCT AAA ATG AAC CCC TCT ACA AAT	?	?	√	√	60	"
Pchcms2	AGG GTC GTC TCT TTG AC CTT CGT TTC AAG GCC TG	$(CA)_8$	180	--	--	52	Sosinski et al., 2000
Pchcms3	CGC CCA TGA CAA ACT TA GTC AAG AGG TAC ACC AG	$(CA)_{9}(TA)_8$	246	√	√	52	"
Pchcms5	CGC CCA TGA CAA ACT TA GTC AAG AGG TAC ACC AG	$(CA)_{9}(TA)_8$	246	√	√	52	"
Pchgms2	GTC AAT GAG TTC AGT GTC TAC ACT C AAT CAT AAC ATC ATT CAG CCA CTG C	$(CT)_{24}$	163	--	--	60	"
Pchgms5	CCA GTA GAT TTC AAC GTC ATC TAC A GGT TCA CTC TCA CAT ACA CTC GGA G	$(CA)_9$	160	√	√	55	"
Pchgms6	CAT TGT TCA TGG GAG GAA TT AGA ACA TTC CTA AAG GAG CA	?	?	√	√	58	Pers. Comm. with Dr Arus, 2006
Pchgms27	GGC TTT GTG TGG TTG AGG TT GCC CAA GTC AAC TCG TAA GG	$(TTA)_7$	204	√	--	55	Wang et al., 2002b
Pchgms28	GCG CCA TTG TCA CAA AAT C CGA GCC ATC TGT CAG GTA CA	$(GA)_{24}$	194	--	--	58	"
Pchgms29	CCT GAA GAA GGT GGA CCA GA CCT CCC AAT TCA AAT TCC CT	$(GA)_{21}$	128	√	√	58	"
Pchgms41	GGA AAT TCC CTG TCC TTC CT CCT CGA ACT AGT TGC CTT TGA	$(AG)_9$	217	√	√	57	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
PMS2	CAC TGT CTC CCA GGT TAA ACT CCT GAG CTT TTG ACA CAT GC	?	?	--	--	55	Cantini et al., 2001
PMS3	TGG ACT TCA CTC ATT TCA GAG A ACT GCA GAG AAT TTC ACA ACC A	?	?	√	√	57	"
PMS67	AGT CTC TCA CAG TCA GTT TCT TTA ACT TAA CCC CTC TCC CTC C	?	?	√	√	57	"
PS5C3	AGA TCT CAA AGA AGC TGA AGC TTA TGC ATA TAC CTG	?	200	--	--	50	Sosinski et al., 2000
Ps08e8	CCC AAT GAA CAA CTG CAT CAT ATC AAT CAC TGG GAT G	?	?	√	√	50	Personal comm Dr Howad
UCD-CH12	AGA CAA AGG GAT TGT GGG C TTT CTG CCA CAA ACC TAA TGG	$(CA)_{14}$	173-200	√	--	55	Struss et al., 2003
UCD-CH13	ACC CGC TTA CTC AGC TGA AC TTA GCA CTA AGC CTT TGC TGC	$(CA)_{10}$	133-139	--	--	55	"
UCD-CH14	GTA CAC GGA CCC AAT CCT G TCT AAC ATC ATG TTA AAC ATC G	$(CT)_{18}$	139-147	--	--	55	"
UCD-CH15	TCA CTT TCG TCC ATT TTC CC TCA TTT TGG TCT TTG AGC TCG	$(CT)_{15}$	94-112	√	√	55	"
UCD-CH17	TGG ACT TCA CTC ATT TCA GAG A ACT GCA GAG AAT TTC CAC AAC CA	$(CT)_{11}$	186-190	√	--	55	"
UCD-CH19	GTA CAA CCG TGT TAA CAG CCT G ACC TGC ACT ACA TAA GCA TTG G	$(CA)_{12}$	122	√	√	55	"
UCD-CH21	TTG TTG ACC ATC GAA TAT GAA G GAA GGT ACA TGG CGT GCC	$(CA)_{18}$	114-122	--	--	55	"
UCD-CH31	TCC GCT TCT CTG TGA GTG TG CGA TAG TTT CCT TCC CAG ACC	$(CT)_{26}$	111-148	√	--	55	"
UDA-001	CAT ATA GGG TCA AGG GAG TG AAA TAA ATA TAT ACA CAC ACA CACAC	$(GT)_{15}$	175	--	--	50	Testoline et al., 2004
UDA-002	AAA CGT GAG GTC TCA CTC TCT C GCC ATT TAA GGG TCT GGT CA	$(TC)_{17}(AC)_{18}$	146	√	√	56	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
UDA-003	CCT GCA GAA CGG TTT CTT TC CCA GAT GGA CCA ACT CAA GC	$(GT)_{15}$	191	√	√	50	Testoline et al., 2004
UDA-004	TGG TAC ATT ATC CCC CGG TA GAA GAA GCT CCA TTC TTG TGA	$(GT)_{8}$	156	√	--	50	"
UDA-005	CAT CAC ACA CAA ACA CAA ATG C GCA TTG TGC TCT TCA TGG AC	$(AC)_{14}(TC)_{21}$	172	√	√	56	"
UDA-006	ATT CTC CAA GGC GAT AAG CA TTA GGC ACC TGT CCC CTA CA	$(CA)_{11}$	156	√	--	50	"
UDA-008	AGA CGC TTT GCA TAC ATA CAA GT TGC AGG AAC TGG GAT TAG AGA	$(AC)_{18}$	149	√	√	56	"
UDA-009	AAA ACA TCT CTC TCC TCC ATG C AGT TCT CTG GCA GCA CAA GC	$(AC)_{14}$ ATA $(CT)_{21}$	220	√	√	56	"
UDA-010	GAC TCA CAT ACA CGT GGG TTT C GGT GTG ATT TGT GTG TGT GC	$(ACCCC)_4$CCA $(AC)_9$	157	√	--	56	"
UDA-012	CCT CCG GGG CTC TTA TAA AT ATG TGT GAT GGC CAG AGC TT	$(CA)_{13}A(CA)_{12}$ $(CT)_{11}$	199	√	√	50	"
UDA-013	CTC CCT GTC ACA AGA AA CCC CAT TTT CCA GTA GTC CA	$(CA)_5(N)_7(CA)_9$	162	√	√	50	"
UDA-014	TAA AAT ACA CAC GCG CAC AC ACC AAG CAT CGT CAC TAG CC	$(AC)_{14}$	165	√	√	56	"
UDA-015	ACT CCA TCG CTT GCA TTT TC GCT CCG TGT GTG TTT GTG TG	$(CA)_{22}$	140	√	√	50	"
UDA-018	TTT AAC TAT AAA ATA CAC ACA CACACA CTG CAT CAT CGG CTT TAT TAG	$(AC)_{15}$	193	√	√	56	"
UDA-020	TGT GCA CCA AAC ACA ACT GA GCA GTG TTG CCA ATG TTG AT	$(AT)_5(AC)_{15}$ $(AT)_5$	172	√	--	50	"
UDA-022	GCC GTC TCA TTT TCC CAT TA GTG CGA TGG AGG AGC ACT	$(TC)_6(N)_6(TC)_9$ $(N)_{14}(TC)_{22}$	174	√	√	50	"
UDA-023	TTG CCG TGA TAC ACT AAC AAC T ACC TGC CAA GTA AGT GCC TA	$(CT)_{11(}CA)_{21}$	175	√	--	50	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
UDA-025	TCG AGA AAG CTG CAC TGG TA AAA GCT GCT TAT TCG TGT GTG	$(CA)_{18}$	138	√	√	50	Pers.Comm. Dr Testolin, 2005
UDA-026	AAA AAC CTG AAA ACA CAC ACA C GCT CGG CTG TTC AAA AAT AG	?	?	√	√	56	"
UDA-027	GGA GAC AGA CGG AGC AAC AT CTT CAA TCT CGC TCC CAA AG	?	?	√	√	56	"
UDA-029	GAA TCT CAT ATT CTG CAC CAC A TTG GCA TTC CGT AAG GTA CA	?	?	√	--	56	"
UDA-031	AAC AAC ATC CTA CAG GTC TCT CTC GTC CAT CTC TGC ACA CCA GA	?	?	√	--	56	"
UDA-033	AAT TCA CTT CAT CTC CTC TCT CTC TTG TTC AGA GCT GAA ATC CAG A	?	?	√	√	56	"
UDA-035	GGT GGA TGA GGG TTT CAC AC GCC ATC TCA AAG CCC ATA AC	?	?	√	√	56	"
UDA-036	AAT TCA CAT ATA TAC CCG TAC ACAC TGT TGG ATT GTT TCC TCT GG	?	?	√	√	56	"
UDA-037	ATA TGC ACA ACT CTC CAT CA AAA AAA TGA CAT ACA TAG AGA GAG	?	?	√	√	56	"
UDA-038	CCA TCC ATG TAT ATC TTA TGT CTAAGT TCT TGA CAA CCC AAA GTG GA	?	?	√	--	56	"
UDA-039	CCC GGT TGA AGA GTG TGT GT ATT TCA AAT GTT CAT ACC TGT GC	?	?	√	--	56	"
UDA-040	CTT AAC TCT CAC CAA CCC ACA CCC AAA ATG GGT CTC CAA TA	?	?	√	√	56	"
UDA-041	TTC CGAAGG CAA CAT AAT CC CCA CAA TCT CTT GGA CTC CTG	?	?	--	--	56	"
UDA-042	CCA GAG CTC GTC CAC TTA AC AGA GCT AGA GAT GTA AAT ACA CAC AC	?	?	√	--	56	"
UDA-043	GAA TAC ATA AAT GGG ATA CCA AGG A TTT GGA CTC ATA CCA TTT TGT G	?	?	√	√	56	"
UDA-044	TTT CAA AGA GCT CAG GCA CA ATC GTA CAA TGC GTC CCT GT	?	?	--	--	56	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
UDA-045	CCA TCC CAG GCC TTA GTA CA GGA GGA TGC TAT TGG GGT CT	?	?	√	--	56	Pers.Comm. Dr Testolin, 2005
UDA-046	ATT TGA ATC CAGGCA AAT G TGA GTA AGA TAA AGA TAT CAA ACA CAC	?	?	√	√	56	"
UDA-048	AGA CGC AGG ATA GCA ACA GG CCA AAC CAC TGT ATA TAC TTT AAC TGC	?	?	√	√	56	"
UDA-050	TTC ATC AGC TTC ATC AGT G TAA CAA TAA AAT ACA CAC ACA CAC	?	?	√	--	56	"
UDA-051	AGT GCA AGA TGT TCG GCA TA TGC CCC AAG AGC TTA ACG TA	?	?	√	√	56	"
UDA-053	TGA CAT GAT AAC ATA TGA CCA AGA GAA TGT GCG TGC ATA CTC AG	?	?	√	√	56	"
UDA-056	AAT TTA ACA ATA AAA TAC ACA CAC ACA AGG CGA CTG CTT CAT CTG	?	?	√	√	56	"
UDA-059	AGT CCC CAT GTG GTA CAT CT GTG CAA TCT TTC ACA CAC ACA	?	?	√	--	56	"
UDAp-401	AAA CCC TAG CCG CCA TAA CT GCT AAA GGC CTT CCG ATA CC	$(TC)_{23}$	201	√	√	56	Messina et al., 2004
UDAp-405	CAT TAC CCA ACC ACC TCC AC TTA GTT TTG GAG TTT GAT GAG AGA G	$(CT)_{26}$	149	--	--	56	"
UDAp-407	TTC TGC TAC TTA CAA TCG TGT TCT C AGA GCA CCA GGT CTT TCT GG	$(TC)_8TT(TC)_9$	188	√	√	56	"
UDAp-409	TGG CCA CAC AAA GAT GAA GA GGT TTT GGA CTG GTT GAG CA	$(AG)_{21}AC$ $(AG)_{13}$	167	√	--	56	"
UDAp-410	TTG TTG ACA AGA AGA AAA CAA AGC CAA CGG GTT GGT TTC AGA AG	$(AG)_{24}$	155	√	--	56	"
UDAp-415	AAC TGA TGA GAA GGG GCT TG ACT CCC GAC ATT TGT GCT TC	$(GA)_{21}$	156	√	--	56	"
UDAp-418	CAG AAA TAG CCC CAG CAC AT TTC TTG CGC CAA AAA CAA CT	$(AG)_{25}$	169	√	√	56	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
UDAp-454	ACA AAC CCT AGC CGC CAT A AGA AAC AAT AGC TAA AAAGGGTAC G	?	?	√	√	56	Pers.Comm. Dr Testolin, 2005
UDAp-456	CCA AGC TCC TCT GGA AAA TG CCA CGT GTC CGT CTC TCT TT	?	?	√	√	56	"
UDAp-457	TCT TGC AGC CTT TCC TAG ACA CGC TGG CAA TGT AAC AAA AA	?	?	√	--	56	"
UDAp-460	TCA TCA GTC AGG TGG TGC TC TGA CAG CCT AAT CAG CCA TTT	?	?	√	--	56	"
UDAp-461	ACG GTT TCA AGA GGT TGG TG AGT GCA CGT GTG ACC CTG TA	?	?	√	√	56	"
UDAp-462	CTG TGT GTA ATA CAT GGG GAG AG CTT CAA GCC CTT CTT CGT TG	?	?	√	--	56	"
UDAp-463	ATA TTG CTC AAG GGG TGC TT ACC AAT CCA TCT TGC AGC TC	?	?	√	√	56	"
UDAp-465	GGA GGG GTT CGT GAT TCA G GCC CCC TTT TCT CTC TGT CT	?	?	--	--	56	"
UDAp-467	TGC ACG AAC TAC TAT GCA ATC A TTA CCA CGA TTT CGC TGA CC	?	?	--	--	56	"
UDAp-469	AAGGCCATTTTGAAACTCCA TGCTCAGGCCCTTTCTTAAA	?	?	√	--	56	".
UDAp-471	CGA TGA ACA GAC ACC GTT GA AGT CCG TCT TTG CTC AGC TC	?	?	√	--	56	"
UDAp-473	AAA GGC CAT CAT CCA TCA AG AAG AAG TCC CTT TGC TGC TG	?	?	√	√	56	"
UDAp-476	AAT TCA ACA AGC CCT GAA CA AAA CTC CAA GGG TGT GTT GC	?	?	√	--	56	"
UDAp-479	GGC CCC TTT CTA AAC CCT AA GGT GCT GGT GCC TAG AAG AA	?	?	√	√	56	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
UDAp-483	CCC AAG AAG GAC GCA AAT AC TCC TCC CAA CTC TAC ATA CGG	?	?	√	√	56	Pers.Comm. Dr Testolin, 2005
UDAp-485	TCA TGA CGA CAT TTT CTT CTC C GGT CTC TTT AGC CCT GAT GG	?	?	√	√	56	"
UDAp-489	CTT CAA CCC CAA GAA AGC AA CAC ATC GAG TGG CCA AAA C	?	?	√	√	56	"
UDAp-493	TCC ACA GCT AGC ATA ATA GGA CA TGA GCC AAA GAT CTC TCA AGC	?	?	√	√	56	"
UDAp-496	AAT TTC TCT TAA AAC GTG AGG TCT CAC TGG CCA TTT AAG GGT CT	?	?	√	--	56	"
UDAp-499	GAG CTG AGC ATC TCC ACT CA TTC GGT CTC GGA CTT GAA TC	?	?	--	--	56	"
UDAp-503	CAA AAT GGC CAA ATG GGT AG GCC CAA AAC AAG CAA AGT TC	?	?	--	--	56	"
UDP96-008	TTG TAC ACA CCC TCA GCC TG TGC TGA GGT TCA GGT GAG TG	$(CA)_{23}$	165	√	--	57	Cipriani et al., 1999
UDP96-013	ATT CTT CAC TAC ACG TGC ACG CCC CAG ACA TAC TGT GGC TT	$(AG)_{22}(TG)_8TT(TG)_{10}$	198	√	--	57	"
UDP96-019	TTG GTC ATG AGC TAA GAA AAC A TAG TGG CAC AGA GCA ACA CC	$(TG)_{18}(AG)_7$	216	√	--	57	"
UDP97-402	TCC CAT AAC CAA AAA AAA CAC C TGG AGA AGG GTG GGT ACT TG	$(AG)_{17}$	136	√	√	57	"
UDP97-403	CTG GCT TAC AAC TCG CAA GC CGT CGA CCA ACT GAG ACT CA	$(AG)_{22}$	150	√	√	57	"
UDP98-024	CCT TGA TGC ATA ATC AAA CAG C GGA CAC ACT GGC ATG TGA AG	$(GT)_{19}TC(TG)_7$	105	√	√	57	"
UDP98-025	GGG AGG TTA CTA TGC CAT GAA G CGC AGA CAT GTA GTA GGA CCT C	$(CA)_{19}$	135	--	--	57	"

Appendix (Continued)

Microsatellite locus	Forward and reverse primer (5'-3')	Repeat motif	Reported length bp	Results		T_a °C	References
				Amplified	Polymorphic		
UDP98-409	GCT GAT GGG TTT TAT GGT TTTC CGG ACT CTT ATC CTC TAT CAA CA	$(AG)_{19}$	129	√	√	57	Cipriani et al., 1999
UDP98-411	AAG CCA TCC ACT CAG CAC TC CCA AAA ACC AAA ACC AAA GG	$(TC)_{16}$	150	√	√	57	"

? = not reported in literature -- = no amplification or no polymorphism