THE DEVELOPMENT OF A GENETIC LINKAGE MAP FOR ALMOND BASED ON MOLECULAR AND AGRONOMIC MARKERS

A THESIS SUBMITTED BY

DAVINA GREGORY, B.SC. (HONS.)

CANDIDATE FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

THE UNIVERSITY OF ADELAIDE

SCHOOL OF AGRICULTURE AND WINE

ADELAIDE

JUNE 2004

VINCENT VAN GOGH (1853 - 1890)
ALMOND BLOSSOM
ST REMY 1890

Table of Contents

Thesis Summary

THE DEVELOPMENT OF A GENETIC LINKAGE MAP FOR ALMOND BASED ON MOLECULAR AND AGRONOMIC MARKERS

Almond, *Prunus dulcis*, is a tree nut crop that originated in central Asia and is now grown commercially worldwide. Within Australia there exists huge potential gain from optimisation of almond cultivars better suited to Australian conditions. This is the ultimate goal of the Australian Almond Breeding Program, which was established in 1997 at the University of Adelaide. As part of this breeding program a unique hybrid population was developed from a cross between the American self-incompatible cultivar 'Nonpareil' (NP) and European self-compatible cultivar 'Lauranne' (LA). The F_1 population derived from this cross is the focus of this study, the population consisted of 181 individuals, of which 93 were selected for use in the mapping study.

Investigation of a number of DNA extraction techniques was performed in order to optimise DNA extraction quality and integrity from almond leaves for future applications in molecular work.

To determine if the purported F_1 hybrids were true hybrids, derived from a cross between the cultivars NP and LA, both DNA fingerprinting with cluster analysis and *S*-

allele identification was performed, and the majority of F_1 putative hybrids clustered between the two parents when analysed using the simple matching coefficient and UPGMA. The genetic similarity between individuals comprising the mapping population ranged from 70% to 93% while the parents were 72% similar in comparison to each other. This indicated high genetic variability available for studying heritabilities and for production of a genetic map. Analysing the *S*-allele complement of all the F_1 hybrids was also performed to offer a more robust method for hybrid determination, since individuals in a breeding population with aberrant *S*-allele inheritance can be considered non-related. The inheritance of the self-fertility gene is important in breeding programs, since the majority of almond cultivars are self-incompatible, tracking the inheritance of this allele in breeding programs is therefore highly desirable.

A detailed morphological study was performed on the whole population over three growing seasons, 2001, 2002, and 2003. In 2001 tree characters such as disease prevalence, bare branches, close internodes, level of upright branches, leaf size and colour were measured. For all the seasons a number of other traits were also measured including: yield, bloom time, self-compatibility, percentage of double kernels, shell hardness, kernel weight, shape, taste, pubescence, and colour. The heritability, genetic variance, segregation and raw correlations between traits were calculated and used to establish a mode of inheritance for these traits. Rainfall and temperature maximum, minimum and monthly averages were collected and used to compare trends in the collected morphological data with these climatic data.

A preliminary investigation was undertaken to determine if the cellular structure of the kernel testa epidermis was responsible for the pubescent versus smooth mouthfeel of the F_1 hybrids. Light and scanning electron microscopy identified the presence of

cellular protuberances arising from the epidermis as a potential cause of the pubescent mouthfeel in almonds. Bulked segregant analysis using inter-simple sequence repeat (ISSR) primers identified a potential marker linked to the pubescent trait which was converted to a sequence characterised amplified region (SCAR), which was also used to screen twelve almond cultivars for this trait.

In addition to the use of BSA for the development of markers linked to traits of interest, the development of genetic linkage maps has the potential to greatly enhance current and future breeding programs by MAS. This study produced a genetic linkage map for this population, constructed using random amplified polymorphic DNA (RAPD), ISSR, and simple sequence repeats (SSR), with the mapping program Joinmap 3.0. Two parental maps were constructed, which coalesced into seven linkage groups for the female parent and eight linkage groups for the male parent, corresponding to the chromosome number of eight for almond. The marker density was 9.4 cM/marker for NP and 9.6 cM/marker for LA, covering 65% for the female and male parental maps in comparison to the highly saturated peach x almond map produced by the European *Prunus* Mapping Program (EPMP). Fourteen markers segregating in both parents were used to produce an integrated parental map for this cross, which coalesced into six linkage groups with a marker density of 11.6 cM/marker. The presence of anchor loci common to the EPMP map allowed homologous linkage groups to be established between the two populations.

This study has contributed to the understanding of key morphological traits important in almond breeding programs. The expression and influence of biotic factors on the expression of these traits was also investigated. Understanding factors

responsible for kernel taste is also an important objective and this study has contributed to this knowledge. The development of a genetic linkage map will serve as a permanent and practical resource for almond breeders in Australia, and contribute important data to the EPMP. This has significant benefit for *Prunus* breeders worldwide, and further enhances knowledge on an economically important nut crop.

Declaration

To the best of my knowledge, this thesis contains no material which has been submitted for the award of any other degree or diploma in any University; nor does it contain any previously published or written material, except where due reference is made.

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

DAVINA GREGORY

JUNE, 2004

Acknowledgements

This project was financially supported by the Australian Research Council, SPIRT project grant C00001810, 'Development of molecular tools for almond breeding' M. Sedgley and G. Collins. The scholarship was supported by an Australian Postgraduate Award (Industry) on the SPIRT grant, with industry support provided by the Almond Board of Australia (ABA). Thanks to the University of Adelaide 'Postgraduate Travel Award' for providing funds toward my research trip to Spain.

I would like to extend my gratitude to my supervisors Dr Graham Collins, Dr Michelle Wirthensohn, Dr Brent Kaiser and Prof Margaret Sedgley. I appreciate their technical advice, problem solving skills and welcome guidance. Thanks to Dr Pere Arús at IRTA, Cabrils, Spain who took over the supervisory role during my research at the IRTA research station under the linkage international travel grant LX 0242383 'Microsatellite marker development for almond breeding', I also thank to the rest of the lab for their support and patience.

Thank you to our industry partners the ABA, particularly to Chris Bennett, Andrew Lacy, and Tony Spiers for their insightful discussion and providing worldly advice and technical support. Thanks to Tony for looking after my F_1 progeny on Andrew's property at Lindsay Point.

No less importantly my greatest thanks goes to my family for their love and support, and to my fiancée Rye who has been a pillar of support, love and laughter through my 'student' years, and to my great girlfriends who even from overseas were an immense support base, I love them all.

To my lab and other Drs and Drs to be at Uni, thanks for providing me with an enormous amount of wonderful memories (and ones to come) and always providing me with an answer. Particular thanks to (in alphabetical order) Drs Cassie Collins, Kate Delaporte, and Jenny Guerin, and to Drs in training Brad Harvey, Kat Malone, Kirsty Neaylon and Meg Shelden.

This has been an invaluable learning experience that I am grateful to have been given the opportunity to undertake....thanks for reading !

List of Figures

List of Tables

List of Abbreviations

AFLP	Amplified fragment length polymorphism
BAC	Bacterial artificial chromosome
BC	Back cross
bp	basepair
BSA	Bovine serum albumin
BSA	Bulked segregant analysis
χ^2	Chi-squared
cM	Centimorgan
CPC	Cophenetic correlation
CsCl	Cesium chloride
CTAB	Cetyltrimethylammonium bromide
DETC	Diethydithiocarbamic acid
DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
DPSTC	Double pseudo-testcross
EDTA	Ethylenediaminetetraacetic acid
EMBL	European molecular biology laboratory
EPMP	European *Prunus* mapping project
EST	Expressed sequence tag
F_1	First filial generation
F_2	Second filial generation

FLB	formamide/bromophenol blue
GSI	Gametophytic self-incompatibility
HPLC	High performance liquid chromatography
IMA	Inter-microsatellite amplification
ISSR	Inter-simple sequence repeat
ITPG	isopropylthio-β-D-galactoside
kb	Kilobase
LA	'Lauranne'
LB	Luria-Bertaini
LOD	Logarithmic of the odds
LSD	Least significant difference
MAS	Marker assisted selection
MM	Molecular marker
MQ	Milli Q
NIL	Near isogenic line
NP	'Nonpareil'
PAGE	Polyacrylamide gel electrophoresis
PCA	Principal component analysis
PCR	Polymerase chain reaction
PVP	Polyvinyl-pyrolidone
QTL	Quantitative trait loci
REML	Restricted maximum likelihood
RFLP	Restriction fragment length polymorphism
RHS	Royal Horticultural Society
RIL	Recombinant inbred line
RKN	Root-knot nematode
RNA	Ribonucleic acid
RO	Reverse osmosis
RT	Room temperature
SAHN	Sequential, agglomerative, hierarchical, and nested algorithm
SC	Self-compatible
SCAR	Sequence characterised amplified region
SD	Standard deviation
SDS	Sodium dodecyl sulphate

SEM	Scanning electron microscope/microscopy
SI	Self-incompatible
SMC	Simple matching coefficient
SSR	Simple sequence repeat
STS	Sequence tagged site
Ta	Annealing temperature
TBE	Tris-borate/EDTA electrophoresis buffer
TCIF	Trichome inducing factor
TCSF	Trichome suppressing factor
TE	Tris-EDTA buffer
TEMED	N,N,N',N',-tetramethylethylenediamine
Tm	Melting temperature
TxE	'Texas' x 'Earlygold'
U	Unit
UPGMA	Unweighted pair-group method average
UV	Ultraviolet
v/v	Volume/Volume
w/v	Weight/Volume
X-gal	5-Bromo-4-chloro-3-indolyl-beta-D-galactoside

CHAPTER 1

General Introduction and Literature Review

BOTANICAL CLASSIFICATION, ORIGIN AND DOMESTICATION OF ALMOND

Class:	Dicotyledones
Subclass:	Rosidae
Order:	Rosales
Family:	Rosaceae
Subfamily:	Prunoideae
Genus:	*Prunus*
Subgenus:	Amygdalus
Section:	Euamygdalus
Species:	*dulcis*, syn. *amygdalus* (2n = 2x = 16)

The family Rosaceae, to which almond belongs, contains a number of important horticultural crops, including, apple (*Malus*), pear (*Pyrus*), strawberry (*Fragaria*), raspberry (*Rubus*), and the ornamental rose (*Rosa*). Plants within the genus *Prunus* are often referred to as stone fruit, and they are classified according to the morphology of

the fruit, which is a drupe. Of the seventy or so species belonging to the genus, almond is the only one grown as a commercial nut crop. The major diagnostic trait separating almond from the rest of the *Prunus* species, and its delineation as a nut rather than a fruit, is the dehiscence of the pericarp (Figure 1.1). In the rest of the *Prunus* species cultivated for fruit, such as peach, apricot, plum, and cherry, a fleshy mesocarp develops to form the fruit and this is not dehiscent.

Figure 1.1 Mature almond fruit showing dehiscence of the pericarp (P), exposing the endocarp (E)

The domesticated almond has been renamed on a number of occasions; it was originally designated *Amygdalus communis* by Linnaeus in 1753 (Sp PI 473). However the most commonly used and accepted names in the literature are *Prunus dulcis* (Miller, DA Webb, Feddes Rep 74: 24, 1768) literally translated as 'sweet nut' and *Prunus amygdalus* (Batsch, Betyr. Entw. Pragm. Gesch. Nat. Reiche, 1: 30, 1801) meaning 'Greek nut'.

The domestication of almond can be traced back to the third millennium B.C. (Spiegel-Roy 1986). Because almond and peach are intercompatible, Watkins (1979)

proposed that they probably originated from a primitive ancestor and evolved under different climatic conditions, which lead to their divergent characteristics. Almonds are thought to have evolved in the dry, arid, xeric conditions of central Asia, whereas the peach is more likely to have evolved in the warm, humid climes of southern Asia. Several related wild *Prunus* species are still distributed in the proposed area of origin of these two species across China and into the Middle East area (Browicz and Zohary 1996; Kester et al. 1990). Ladizinsky (1999) discussed the origin of the domestic almond from a wild progenitor and concluded that *Amygdalus fenzliana* is the most likely candidate on the basis of its morphology and natural distribution, and it is classified in the section Euamygdalus with domestic almond (Kester and Gradziel 1996). The evolution of the domesticated almond from the wild has not been well documented, although the dispersal of the domesticated form from Asia, across the Middle East and the Mediterranean basin has been discussed by a number of authors (Browicz and Zohary 1996; Kester et al. 1990; Socias i Company 1998). The general consensus is that the area where almond is likely to have undergone the majority of its conversion to a domesticated form corresponds to the location of ancient civilizations or in close proximity to historical trade routes. Due to the ease of transport of this crop in the form of seeds it may have been dispersed in a similar manner to that proposed for olives, grapes, and figs (Zohary and Spiegel-Roy 1975).

ALMOND EVOLUTION UNDER DOMESTICATION

The natural distribution of the domesticated sub-genus Amygdalus is concentrated in the Mediterranean basin, extending into Afghanistan and the Middle East, north as far as Russia, with localised pockets in China and Mongolia (Browicz and Zohary 1996; Kester and Gradziel 1996). The ideal climate for commercial almond

orchards consists of hot dry summers and cool, mild, rainy winters, typically described as Mediterranean (Quinn 1928; Woodroof 1979). Since domestication, the growing range of almond has been extended beyond its natural distribution to include regions with similar climatic conditions, such as North America, Africa, and Australia.

Since domestication, almond has undergone numerous changes to its phenotype largely due to selection pressure by humans. The most important factor that allowed almond to come into cultivation was the selection of the sweet kernel phenotype. Wild almond kernels contain a high level of cyanogenic glycosides resulting in a bitter kernel, which is toxic to many potential herbivores (Conn 1980; Jones 1998; Vetter 2000). Totally bitter almond kernels contain approximately twenty times more cyanogenic compounds than semi-bitter kernels (Dicenta et al. 2002a) indicating that there is a gradation in the level of these compounds between sweet and bitter. Humans are able to detoxify cyanide using sulphurtransferases if the protein component of the diet is adequate and the cyanogenic compound is ingested slowly over a period of time (Jones 1998). Hence almond kernels could have been consumed as a foodstuff prior to the evolution of a completely sweet form, in contrast to the hypothesis proposed by Kester and Gradziel (1996) and Ladizinsky (1999) who suggested domestication was unlikely to occur until a totally sweet phenotype evolved. The presence of bitter kernelled progeny in current breeding programs illustrates that this undesirable trait is still prominent in the genotype of modern almonds (Ledbetter and Pyntea 2000; Vargas et al. 2001; Woodroof 1979).

Domestication has also influenced the size of the kernel, resulting in an overall increase in both weight and dimensions of the nut. Other traits such as flowering time, leaf size, yield, and shell thickness have also been the target of selection pressure (Socias i Company 1998, 1999).

CULTIVATED ALMOND PRODUCTION IN AUSTRALIA

The almond is now an important agricultural crop in a number of countries, including North America, Italy, Spain, Iran, France, and Australia. Almonds were first introduced to Australia in the early 1800s, with the first documented planting occurring in 1836 at Kangaroo Island in South Australia, and later in 1842 at North Adelaide (Quinn 1928). Since this initial introduction, almonds are now cultivated commercially in South Australia, Victoria, and New South Wales from a range of cultivars, mainly from North America and Europe, although some local selections have been developed. There are still some smaller orchards in South Australia which have plantings of Australian selections taken from the initial introductions in the 1800s (Moss 1965; Quinn 1928; Quinn 1941).

Commercial almond growing in Australia is restricted to areas that have the ideal Mediterranean-style climate conducive to fruitful almond production (Baxter 1997). Between 1990 and 1999, Australian production increased by 50% (FAO, 1999) and is now approximately 1.1% of the total world production. In contrast, the major world producer, North America, produces 38% of world almond production (Figure 1.2).

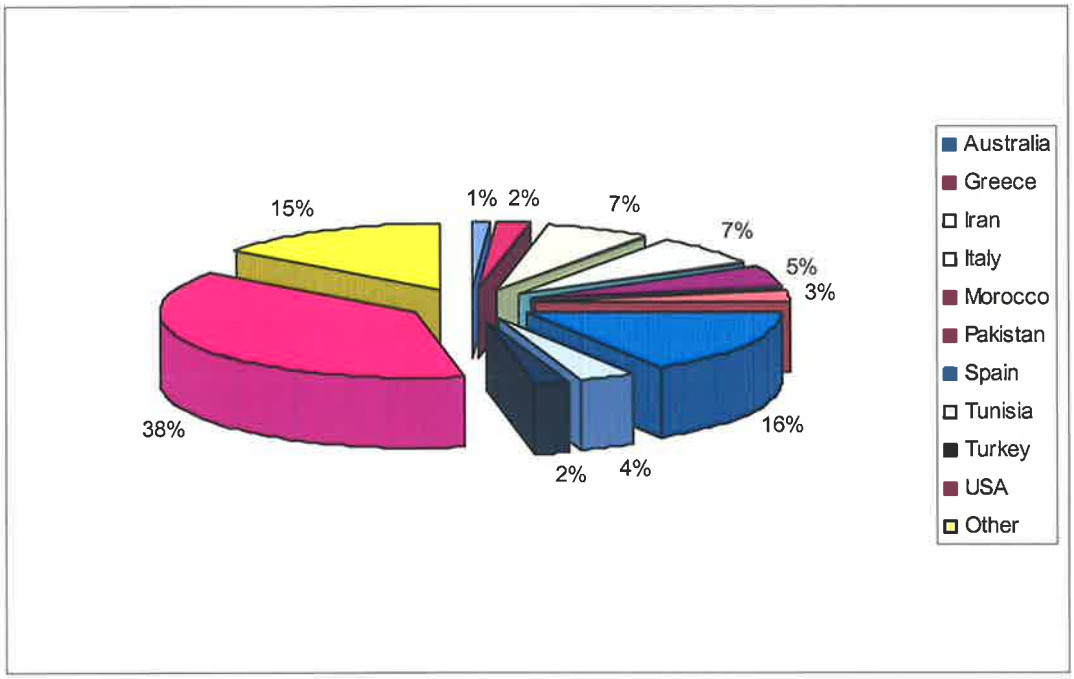

Figure 1.2 World Almond Production, on a per country basis, values are given as a percentage of the world total (FAO 2003)

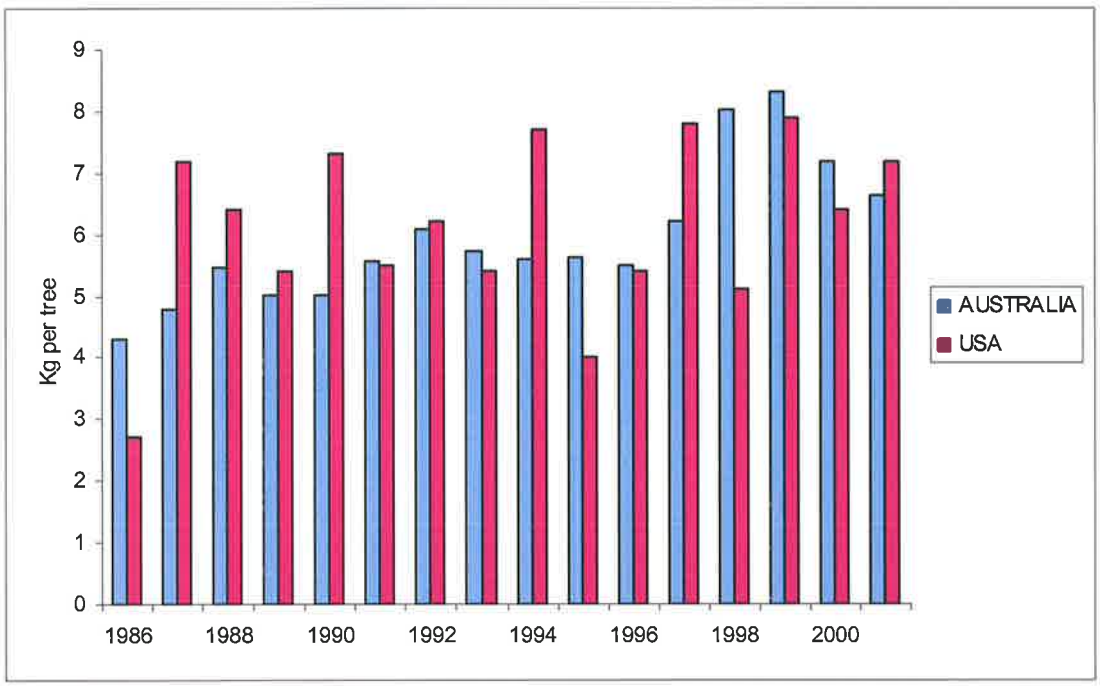

Figure 1.3 Comparative productivity of almond trees on a kg/tree basis 1986-2000 for Australia and USA (FAO 2003)

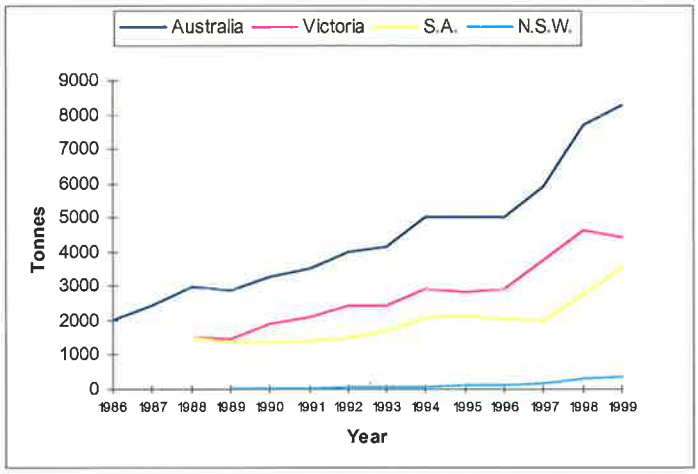

Figure 1.4 Almond production by state in comparison to total Australian value, total tonnage (ABS 2002)

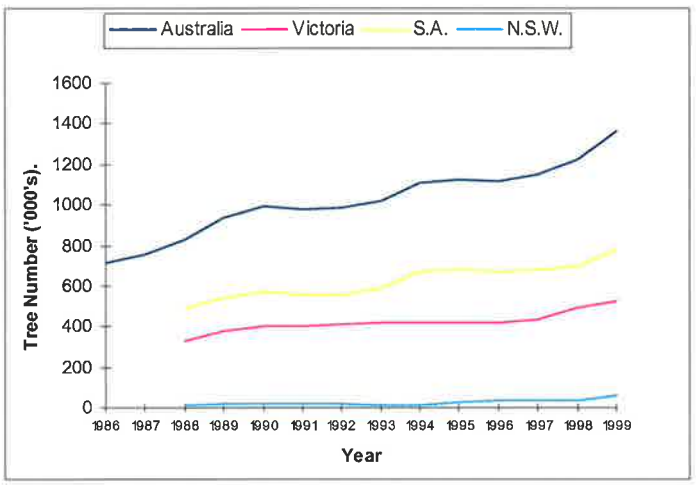

Figure 1.5 Total almond tree number (000s) per state in comparison to total Australian number (ABS 2002)

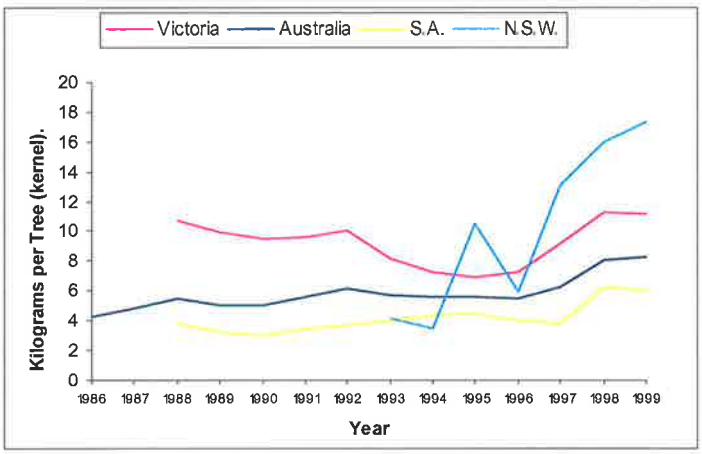

Figure 1.6 Average comparative productivity of trees (kilogram of kernel/tree) of trees aged over 6yrs in comparison to Australian average (ABS 2002)

Productivity has been comparable between Australia and North America for a number of growing seasons, averaging approximately 6.5 kg/tree (Figure 1.3). The large differences in years are due to adverse environmental conditions affecting fruit set, as illustrated in 1987, 1990, 1994, and 1998. The export value of the Australian almond crop for 1999 was estimated at $AUD10.9 m for 1255 t of kernels (Anon 2002). On a state by state comparison, Victoria is Australia's largest producer, followed by South Australia and New South Wales (Figure 1.4). South Australia has 18% more trees than Victoria (Figure 1.5), however on a per tree basis over six years of age, South Australia has the lowest production, with New South Wales the highest (Figure 1.6). The discrepancy between these two values is possibly due to the number of older plantings in South Australia, and the better climate, design and irrigation of orchards in the Riverland regions of Victoria where most of the large almond orchards are located (Bennett 1999).

AUSTRALIAN ALMOND BREEDING PROGRAM

The Australian almond improvement program, supported by the Almond Board of Australia (ABA) and Horticulture Australia Limited (HAL), commenced at the University of Adelaide in 1997 with the aim of improving both yield and quality of the local product for Australian and overseas markets. The industry has a number of horticultural limitations that the program is aiming to rectify. An important objective of this group is to develop new almond cultivars better suited to Australia's environmental conditions, by cross breeding cultivars of European and American descent and assessing the progeny under Australian conditions (Sedgley and Collins 2002; Wirthensohn and Sedgley 2002). The selection pressure is aimed at the introduction of self-compatibility, late flowering,

increased yield, and improved kernel characteristics such as appearance, organoleptic quality, and size (Wirthensohn and Sedgley 2002).

SELF-INCOMPATIBILITY OF ALMOND

Most commercial almond cultivars are self-incompatible, and therefore require a cross-compatible cultivar to be inter-planted at an approximate ratio of 1:1 within the orchard for cross-pollination and economic fruit set to occur (Kester et al. 1990; Quinn 1941; Woodroof 1979). There are several disadvantages to this system. Firstly the cross-compatible cultivar is often less commercially desirable and therefore a large proportion of the orchard is taken up with less profitable trees (Kester et al. 1990; Woodroof 1979). Secondly, pollination in almond is entomophilous, and the common honeybee (*Apis mellifera* L.) carries out pollen transfer (Jackson 1996). The introduction of hives into orchards to ensure adequate pollination by augmenting native and feral bee numbers, especially in underdeveloped areas close to rivers (Woodroof 1979), is a large cost to the grower. Honeybees are limited by unfavourable climatic conditions and the distance travelled between trees and the hive, both of which adversely affect pollination efficiency (Jackson 1996; Jackson and Clarke 1991; Lewis and Crowe 1954). Therefore the introduction of self-compatible (SC) cultivars would be highly advantageous for this industry.

Current breeding programs have introduced the SC character by using almond varieties derived from the Apulia region of Italy (Godini 2002). A high proportion of native varieties growing in this area are reported to possess the SC trait, and the commercial cultivar 'Tuono' was derived from this stock (Godini 2002). 'Tuono', and progeny derived

from it that possess the SC character, such as 'Lauranne', are popular candidates for SC introgression into breeding programs, as well as using wild and closely related species (Gradziel and Kester 1997; Kester and Gradziel 1996).

Biological Control of Self-Incompatibility

Self-incompatibility (SI) in plants is controlled by one of two mechanisms, either sporophytic or gametophytic which differ in the location of rejection of self-pollen. SI is a widespread, genetically controlled phenomenon that inhibits self-fertilisation and promotes outcrossing in plants. In natural systems the rejection of self-pollen is an important evolutionary mechanism that severely limits inbreeding depression and consequently promotes genetic variability within the breeding population. de Nettancourt (1977) described SI as "the inability of a fertile hermaphroditic seed plant to produce a viable zygote after self-pollination". The elucidation of the genetic and biochemical control of SI has been the focus of numerous research groups worldwide (Sedgley 1994; Stone and Goring 2001).

Gametophytic Self-Incompatibility

Gametophytic self incompatibility (GSI) is the most frequent form of SI and operates in more than sixty families of flowering plants (Kao and McCubbin 1996), accounting for over half of all cases of SI in the Angiospermae (Ebert et al. 1989). GSI operates in a number of important horticultural crop families including the Solanaceae, Papaervaceae, Scrophulariceae, and Rosaceae, the family to which almond belongs (Crowe 1964). The physiological basis for the GSI response involves the recognition of self-pollen by the female reproductive organs and the inhibition of pollen tube growth in the style

(Ishimizu et al. 1998). The molecular, genetic, biochemical, and morphological control of GSI has been extensively studied (Charlesworth 2000; de Nettancourt 1977; Ebert et al. 1989; Kao and McCubbin 1996; Newbigin et al. 1993; Sedgley 1994; Stone and Goring 2001).

S-allele System in *Prunus* and the Role of *S*-RNases in Self-Incompatibility

Monofactorial polymorphic GSI, such as occurs in *Prunus*, involves the recognition of the *S*-allele of the haploid pollen in the diploid tissue of the style. If the *S*-allele of the pollen matches one of the two *S*-alleles in the style, pollen tube growth is inhibited and penetration of the ovary usually does not occur (Stone and Goring 2001). Successful fertilisation in this system will be accomplished only if the *S*-allele carried by the pollen is different from the *S*-alleles present in the style.

Investigations into the molecular control of GSI indicate that pollen rejection is linked to RNA specific factors, particularly to the cytoplasmic degradation of RNA (Dickson 1994; Lee et al. 1994; Murfett et al. 1994). RNases in the styles of these plants, specifically S-glycoproteins with RNase activity, have been implicated as being involved in this response (Kao and McCubbin 1996; Tao et al. 1997; Ushijima et al. 1998).

The term S-RNase has been proposed for these RNases since these S-glycoproteins have RNase activity (Newbigin et al. 1993). It is speculated that the SC genotype resulted from an alteration in the S-RNase system producing an enzyme with no ribonuclease activity (Boskŏvić et al. 1999). Studies in *Lycopersicon peruvianum* support this theory because the SC-allele encodes a non-functional protein (Kowyama et al. 1994). SC cultivars have been identified in a number of *Prunus* species, including almond

(Channuntapipat et al. 2003; Egea et al. 2000; Rovira et al. 1997), cherry (Hauck et al. 2002), apricot (Burgos et al. 1997), and Japanese apricot (Tao et al. 2000).

The SC allele, termed *Sf* for almond, is one of a number of *S*-alleles at the multi-allelic locus. Traditional methods for identifying the *S*-alleles of almond cultivars have focused on artificial hybridisation involving manual pollination followed by assessment of either pollen tube growth (Ballester 1998; Ben-Njima and Socias i Company 1995; Dicenta and García 1993a; Rovira et al. 1997), or fruit set (Ballester 1998; Kester et al. 1994; Rovira et al. 1997; Vargas et al. 1997). Isozyme or stylar ribonuclease assays have been used to identify *S*-alleles at the biochemical level, using isoelectric focusing and nonequilibrium pH gradient electrophoresis and RNase staining (Batlle et al. 1998; Boskŏvić et al. 1997, 1999; Tao et al. 1997). A combined approach is also being employed to determine *S*-alleles in many breeding programs (Ortega and Dicenta 2003). Socias i Company (1991) and Kester et al. (1994) identified *S*-allele groups by manual cross-pollination, and each of them used a different system to identify the *S*-alleles. In the case of Socias i Company (1991) an ascending series of numbers (i.e. *S1*, *S2*, *S3* etc) was used, and in the case of Kester et al., (1994) letters (S_a, S_b, S_c etc) were applied. Between them, Socias i Company (1991) and Kester et al., (1994) identified thirteen cross-incompatibility groups (CIG), which covered many of the commercially important cultivars. Further work involving biochemical (Boskŏvić et al. 1997, 1999, 2003; Tao et al. 1997) and PCR based methods (Channuntapipat et al. 2001, 2002ab, 2003; Tao et al. 1999; Ushijima et al. 1998) have increased this number to twenty three. Molecular methods for the rapid and early identification of the *S*-alleles present in breeding progeny have been applied to the Australian almond breeding program to follow introgression and to confirm hybridity (M. Wirthensohn pers. comm. 2002).

GENETIC ANALYSIS OF MORPHOLOGICAL TRAITS OF ALMOND

SI is one of only four morphological traits that have been studied extensively in almond, the others are kernel taste, bloom time, and shell hardenss. Although almond has been studied more thoroughly than the other tree nut crops, relatively little is known about inheritance of genes in comparison to other rosaceous species such as peach and apple.

Kernel Taste

The enzymatic breakdown of cyanogenic glucosides in the seed and the concomitant liberation of hydrogen cyanide and benzaldehyde molecules are responsible for the bitter taste of some kernels. However the genetics of this trait are poorly understood, especially as there exists an intermediate class between the sweet and bitter phenotypes, termed semi-bitter or slightly bitter.

Heppner (1923) studied the progeny of a number of crosses segregating for kernel taste, and found a 3:1 ratio for sweet versus bitter kernelled individuals in the progeny. A number of other studies have supported this and concluded that sweet kernel taste is dominant, and that most commercial cultivars are heterozygous (Dicenta and García 1993b; Frehner et al. 1990; Heppner 1926; Kester et al. 1977; Spiegel-Roy and Kochba 1974; Spiegel-Roy and Kochba 1981). Dicenta and García (1993b) and Dicenta et al. (2000, 2002a) grouped the intermediate individuals into the sweet group to fit with the 3:1 segregation suggested by Heppner (1923). Heppner (1926) studied 45 families for sweet versus bitter with between 1 and 64 individuals in each, showing a wide range of segregation ratios for example 3:4 to 12:0 and surmised based on the average of the data,

that sweetness segregated in a 3:1 ratio. However, the majority of families, 29%, actually showed a 2:1 ratio, with only 6.6% that showed a 3:1 ratio when examined on a per family basis. In addition, Vargas et al. (2001) reported that only 19% of families studied segregated in the 3:1 ratio. Therefore, in agreement with studies on other *Prunus* species, such as apricot (Gomez et al. 1998) and peach (Werner and Creller 1997) it is likely that the control and level of bitterness in almond kernels is more complex than that proposed in the current almond literature and is most likely polygenic.

Shell Hardness

There is a positive correlation between shell hardness and fruit weight, with the shells of hard shelled varieties having a greater proportion of fruit weight than less dense soft shells (Rugini 1986). Five categories have been classified which are directly proportional to lignin content of the shell (endocarp), these are stone, hard, semi-hard, soft and paper (Gulcan 1985). The level of shell hardness has implications for pest control in almond orchards. Hard shelled varieties that have a good shell seal are less prone to attack by insects such as navel orange worm (Soderstrom 1977) and carob moth (Gothilf 1984), and damage by birds (Baxter 1997), than paper shelled varieties such as 'Nonpareil'. Ballester (1998) proposed that a single gene, *D*, controls shell hardness. Genotype *DD* gives rise to very hard (stone), *Dd* to intermediate (hard), and *dd* to soft (paper) shelled cultivars. Modifier genes have also been implicated to account for forms that occur between these three classes (Kester and Gradziel 1996).

Bloom Time

Almond has a low winter chilling requirement and consequently has the earliest bloom time of all deciduous fruit and nut species. Both chilling temperatures and heat units regulate bloom time, and different heat requirements promote bloom time variation between cultivars (Kester 1965a; Rattigan and Hill 1986). In a study of cultivars grown in Australia, it was found that the number of chill units varied from 220 to 350, which is relatively narrow (Rattigan and Hill 1986). The major determinant of bud development after vernalisation, and thus flowering time, was the temperature above the threshold of 4 to 5°C. This stage ends at anthesis, when the cumulative exposure to warm temperature reaches the heat sum requirement for each particular cultivar. Socias i Company et al. (1999a, 1999b) proposed that there is a single dominant gene, *Lb*, controlling bloom time, where late bloom is dominant over early bloom. This is supported by Ballester et al. (2001) who found that plants with the *Lb* genotype bloomed fifteen days later than other genotypes.

USE OF MOLECULAR TECHNIQUES FOR PLANT BREEDING

Although the proposed genetics of these few morphological characters in almond have been reported, the determination of the inheritance of a particular trait in a breeding program can often only be assessed when the plant reaches maturity. This is a major problem in tree plant breeding, especially for almonds, which do not become sexually mature and thus available for assessment until approximately five years after the cross has been made. Therefore a method for assessing progeny at an earlier stage of development

would be ideal for large scale breeding programs, particularly of long-lived woody perennials, which are currently limited by the investment in cost, space, and time.

The use of molecular markers (MM) in genetic studies and for plant breeding has the capacity to greatly improve breeding programs through the development of genetic linkage maps and marker assisted selection (MAS). Markers used for genetic mapping can be divided into two basic types: those that are due to differences between expressed functional genes, such as morphological and isozyme markers, and those that reveal neutral sites of variation at the DNA sequence level (by neutral meaning not manifested in a change of phenotype).

Advantages of Molecular Techniques for Use in Genetic Mapping

The advantages that MM have over morphological and biochemical markers are numerous, as summarised below.

1. The morphological analysis of some traits, such as fruit quality, can be assessed only when the plants are sexually mature, whereas MM can be used to screen plants for the same markers at either the adult or seedling stage of development. In addition, MM can be scored using virtually any tissue type.
2. Morphological markers may require construction of a complex genetic background, such as an F_3 generation, or tedious backcrosses to establish linkage patterns, whereas MM can be used at any generational stage.
3. For most MM, allelic variation is present in non-coding regions of the genome, therefore alternative alleles at MM loci generally do not affect the phenotype of the organism. Some morphological markers are manifested after a deleterious mutation

has occurred to the genotype and are therefore not desirable in the breeding population and thus technically useless.

4. Allelic variation in MM is much higher than for morphological markers, and laboratory techniques used to identify MM are more sensitive than measurements on morphological markers. Consequently the level of variation detected by MM is far greater and more precise than data generated from morphological analysis.

5. For a linkage map to be useful for studying economically important traits it must have enough segregating markers to cover the genome at intervals of about 10 cM. This has not been possible thus far using morphological markers alone, since the genes involved represent only a small part of the genome and the number available is therefore relatively small.

6. Most morphological markers are either dominant or recessive in their phenotypic expression and frequently not all genotypes can be unequivocally established. MM techniques can be selected that produce markers that are co-dominant allowing the identification of all three morphs (homozygous dominant, heterozygous, and homozygous recessive).

7. Morphological markers frequently show epistatic effects, where one gene affects the expression of a second gene, whereas MMs are generally not affected by epistasis.

MARKERS USED FOR GENETIC MAP CONSTRUCTION

There are a variety of molecular marker types that have been developed and applied to the construction of genetic linkage maps. The selection and application of a particular marker type is based on a number of factors, including the quantity and quality of

information gained, time required for analysis, and the cost for development and application. In *Prunus*, markers have been applied for a multitude of tasks including cultivar identification, phylogenetic analysis, and genetic mapping. The following section describes each marker type, its advantages and disadvantages, and its application in studies of *Prunus*, this has also been summarised by Baird et al. (1996) and Wunsch and Hormaza (2002a).

Isozymes

Isozymes were the first non-morphological markers used in the study of almonds for cultivar identification (Cerezo et al. 1989; Hauagge et al. 1987; Vezvaei et al. 1995) as they are highly polymorphic in this species (Arulsekar et al. 1989; Byrne 1990). Isozymes are electrostatically charged protein molecules that can be separated using electrophoretic procedures and stained with histological stains. Bands visualised after staining represent species-specific variants of an enzyme system. These bands are produced when the enzymes are encoded by different alleles at the same gene locus and are differentially charged, resulting in codominant markers. This method offered breeders a way of reducing the reliance on using morphological markers as the leading system for genetic mapping. The main advantage of isozymes over morphological markers was that the identification of the marker was not based on a phenotypic change of the individual. By the mid-1990s, isozymes were being employed as a tool for mapping loci in almond (Arús et al. 1994b; Vezvaei et al. 1995; Viruel et al. 1995). However their use is limited because only a few non-random loci are represented within the genome, and they may be subject to post-translocational modifications affecting their mobility (Staub et al. 1982). The small number of isozymes also means that the parents may possess the same allele at the same locus

resulting in a monomorphic band in the progeny that has no use as a marker. Thus the search for more convenient and useful markers has continued at the molecular level.

Restriction Fragment Length Polymorphism

The first widely utilised molecular marker system for mapping was developed in 1980 and is a hybridisation-based technique, termed Restriction Fragment Length Polymorphism (RFLP) (Botstein et al. 1980; Wyman and White 1980). The development of RFLP provided the first potentially unlimited source of genetic markers in any species (Chang and Meyerowitz 1991; Tanksley et al. 1989). RFLPs are generated when DNA is digested with restriction endonucleases (enzymes that recognise and cleave DNA at specific nucleotide sequence sites) and, depending on the enzyme used, the size and number of fragments generated will vary. After digestion, the DNA fragments are loaded onto a gel and separated by electrophoresis. The DNA is then transferred to a nylon or nitrocellulose filter and probed with radiolabelled DNA fragment, in a technique referred to as Southern blotting (Southern 1975). RFLPs are codominant markers and are expressed as dominant homozygous, heterozygous, or recessive homozygous, i.e. AA, Aa, or aa, respectively. Polymorphisms result from either a mutation of a nucleotide, leading to the loss or creation of a new restriction site, or an insertion or deletion in the fragment, which will alter the size and hence change the migration pattern of the fragment.

Although RFLPs have the capacity to produce a large number of MMs, leading to the identification of homologous RFLPs across a wide evolutionary range for use in phylogenetic analysis (Ahn and Tanksley 1993) and synteny studies (Shields 1993), they are not without fault. Beckmann and Soller (1983) have discussed the costs of the method, which are relatively high due to the labour content, and the safety aspects associated with

the use of radioactivity. A further limitation of RFLPs concerns the spatial distribution of these markers on genetic maps. This has been discussed for a number of different species, including almond (Viruel et al. 1995), an almond/peach hybrid (Joobeur et al. 1998) and asparagus (Spada et al. 1998) all of which show some clustering of the markers in defined regions. However, RFLP data from other species, such as *Brassica oleracea* (Sebastian et al. 2000), faba bean (Torres et al. 1993), lettuce (Landry et al. 1987), and soybean (Lin et al. 1996b) do not confirm such clustering to be a general problem with the system, and this technique is still very popular for genetic linkage map construction, crop improvement programs and other molecular studies (Beckmann and Soller 1986; Helentjaris et al. 1985; Tanksley et al. 1989; Young 1992).

In *Prunus*, RFLP has been used for mapping in peach (Belthoff et al. 1993; Eldredge et al. 1992), almond (Viruel et al. 1995), and an almond/peach hybrid (Foolad et al. 1995; Joobeur et al. 1998).

Randomly Amplified Polymorphic DNA

In addition to RFLP there are a number of amplification-based marker systems that utilise short oligonucleotide primers to amplify one or more regions of the genome via the Polymerase Chain Reaction (PCR). Randomly Amplified Polymorphic DNA (RAPD) and DNA Amplification Fingerprinting (DAF) are two of the earliest PCR-based techniques developed (Welsh and McClelland 1990; Williams et al. 1990). Advantages that these methods have over RFLPs is that there is no need to have any prior knowledge of the genome, nor for the tedious process of blotting and hybridisation with radioactive probes. In addition most amplification techniques involve a large proportion of automation.

RAPDs have been widely utilised as marker systems, particularly for map construction mainly because of their ease of use and rapid production (Welsh and McClelland 1990; Williams et al. 1990). The RAPD protocol utilises short random oligonucleotide primers, usually 10 nucleotides long that bind to the genomic DNA and amplify short fragments of approximately 200-2000 bp. The selectivity of RAPDs is based on the fact that primers need to anneal at two sites of complementary DNA within about 2 kb of each other for amplification to occur, and therefore have the capacity to generate a large number of fragments. These fragments are then separated on either agarose or polyacrylamide gels and visualised with either ethidium bromide or by silver staining. The bands visualised after staining result from amplification of discrete DNA fragments and can be scored as presence or absence of a band, where polymorphisms are due to differences in genomic sequences between genotypes. RAPD bands act as dominant markers, which is one limitation of this technique since heterozygotes cannot be distinguished from dominant homozygotes. Problems with reproducibility of RAPDs between different laboratories have been reported by a number of authors (Fritsch et al. 1993; Jones et al. 1997; Muralidharan and Wakeland 1993; Penner et al. 1993; Williams et al. 1993).

Care must be taken to avoid artifacts when scoring RAPDs and assigning the presence and absence of bands (Penner et al. 1993). RAPDs can be cloned to generate co-dominant genetic markers, either Sequenced Characterized Amplified Regions (SCARs) (Paran and Mitchelmore 1993) or Sequence Tagged Sites (STSs) (Olson et al. 1989) that specifically amplify one sequence band, resulting in a more reproducible marker.

In the Rosaceae family genetic maps have been constructed using RAPDs alone or in combination with other marker types (Debener and Mattiesch 1999; Dirlewanger et al. 1996; Hurtado et al. 2002; Joobeur et al. 2000; Rajapakse et al. 1995; Warburton et al.

1996). RAPDs have also been used for assessing phenotypic relationships between almond (Bartolozzi et al. 1998; Martins et al. 2003; Woolley et al. 2000), and plum cultivars (Boonprakob et al. 2001), and paternity screening in cherry embryo cultures (Hormaza 1999). However, the use of RAPD data to infer relationships may lead to false conclusions unless the amplified fragments are cloned and sequenced, because an apparently single RAPD band may resolve into several similar-sized DNA fragments each with a different sequence (Brady et al. 1996; Hausner et al. 1999; Mekuria et al. 2002).

Amplified Fragment Length Polymorphism

Another increasingly popular method for generating molecular markers is a combination of RFLP and PCR techniques called Amplified Fragment Length Polymorphism (AFLP) developed by Vos et al. (1995). Polymorphisms are generated after digestion of genomic DNA with two different restriction enzymes, followed by ligation of adapters. A subset of digested fragments is then amplified by PCR using primers based on the adapters and including two to four additional bases on the 3' ends. The amplification products are separated on polyacrylamide gels by electrophoresis. Numerous polymorphisms are identified on a single gel and it has been suggested that AFLP offers a method more efficient than RFLP and RAPD for the production of molecular markers (Vos et al. 1995).

Since this procedure is PCR based it suffers from some PCR specific problems such as the high sensitivity to reaction conditions, DNA quality, PCR temperature cycles, and primer combinations (Lin et al. 1996a), which can limit the usefulness of this technique if strict experimental conditions are not applied. Furthermore, this method also requires a high level of technical expertise for pouring sequencing-sized gels and it also uses radioactive

label to detect polymorphisms on an autoradiogram. In addition, banding pattern differences in the DNA extracted from different organelles in *Triticum aestivum*, that were attributed to differences in DNA methylation affecting the initial digestion of genomic DNA, were reported by Donini et al. (1997).

The application of this technique has been reported for cultivar identification in peach (Aranzana et al. 2003a; Shimada et al. 1999), apricot (Hagen et al. 2002; Panaud et al. 2002), and plum (Goulão et al. 2001a). It has also been widely applied for genetic mapping in peach (Dirlewanger et al. 1998; Shimada et al. 2000; Wang et al. 2002) and apricot (Hurtado et al. 2002; Vilanova et al. 2003).

Microsatellites

The use of microsatellite markers has proved fruitful for the generation of genetic maps in a wide range of species (Goodfellow 1993; Rafalski and Tingey 1993). This method exploits the distribution of tandem arrangements of nucleotide repeated units, each between 2 and 10 bp in length such as $(TG)_n$ or $(AAT)_n$ where n is the number of repeated units in the cluster. The microsatellite technique was described in 1989 by Litt and Luty (1989) and a variation on this method termed Simple Sequence Repeats (SSR) was reported in the same year by Tautz (1989). Hamada et al. (1982) first reported that blocks of the microsatellite $(GT)_n$ dinucleotide repeat were present every 30-60 kb in the mammalian genome, which promoted further studies leading to the construction of a genetic linkage map for the human genome based on microsatellites (Weissenbach et al. 1992). Condit and Hubbell (1991) were the first researchers to report that microsatellites were also abundant in plant species.

Since this method provides a near infinite amount of polymorphism by varying the nucleotide arrangement and the length of the primer unit, it has facilitated the construction of genetic maps for a number of plant species including rice (Wu and Tanksley 1993), avocado (Sharon et al. 1998), and peach (Cipriani et al. 1999). The methodology for identifying microsatellite polymorphism and subsequent map construction starts with the construction of a library, which is then screened by primer hybridisation. The positive clones are sequenced and primers are designed to incorporate the unique flanking region for each locus. PCR is used to generate DNA fragments that are electrophoretically separated to identify polymorphisms between the different sized amplification products. This method is ideal for map construction as markers are codominant, PCR based, and extremely polymorphic. However, it has not been utilised as much as RFLP, RAPDs, and AFLP because the method has a long development time, is extremely expensive to establish, and requires specific primer design.

The majority of SSRs developed in *Prunus* have been derived from studies using peach (Aranzana et al. 2002; Cipriani et al. 1999; Georgi et al. 2002; Sosinski et al. 2000; Testolin et al. 2000) but SSRs have also been developed in cherry (Cantini et al. 2001) (Downey and Iezzoni 2000) and almond (P. Arús, unpublished). They have been applied for genetic mapping (Aranzana et al. 2003b; Hurtado et al. 2002; Joobeur et al. 2000) and cultivar identification in most of the *Prunus* species (Hormaza 2002; Martínez-Gómez et al. 2003ab; Serrano et al. 2002; Wunsch and Hormaza 2002b) and to assessing genetic diversity (Zhebentyayeva et al. 2003).

Inter-Simple Sequence Repeats

The technique known as inter-simple sequence repeats (ISSR) was developed based on the direct amplification of microsatellite sequences via PCR (Zietkiewicz et al. 1994). This marker system uses primers based directly on a repeated core sequence, generally of di-, tri-, or tetra-nucleotides, which amplify regions both within and between adjacent microsatellites. The primers can be anchored with one or a number of nucleotides at the 5' or 3' end, which prevents smearing by anchoring into the template DNA and preventing slipping during replication (Reddy et al. 2002; Tikunov et al. 2003). Polymorphisms are generated when there is a mutation in the nucleotide sequence between the intervening microsatellites and this is observed by the different mobilities of these fragments when gel electrophoresis is used to separate the PCR products.

This technique, unlike SSRs, requires no prior sequence knowledge, and can target any region of the genome that contains a complementary sequence. Another advantage is that it is not necessary to use either radioactivity or polyacrylamide gel electrophoresis. ISSRs are generally more reproducible than RAPDs because the longer primer sequences allow higher annealing temperatures to be used, and they have a higher multiplex ratio (Goulão et al. 2001b).

In *Prunus* the use of ISSRs has been limited and to date only six reports have been published. Goulão et al. (2001a) used ISSR and AFLP to characterise 28 plum cultivars, Martins et al. (2003) used ISSR and RAPD for phylogenetic analysis of Portuguese almond cultivars and their relation to foreign cultivars. Piagnani et al. (2002) used ISSR to differentiate between a new somaclone and wild type cherry. Cheng et al. (2001) used a slight variant on ISSR called Random Amplified Microsatellite Polymorphism (RAMP) for assessing genetic diversity in peach, which used 2 bp anchored trinucleotide repeats, which

were only 11 bp long, compared to the usual length of ISSRs of 16-25 bp (Reddy et al. 2002). Dirlewanger et al. (1998) used ISSRs, referred to in this study as Inter-Microsatellite Amplification, IMA, in combination with isozyme, RFLP, RAPD, and AFLP markers to produce a genetic map for peach. Shimada et al. (2000) produced a genetic map from the progeny of a cross between an ornamental and a rootstock peach using RFLP, RAPD, AFLP, and ISSR markers. Other studies that have used a combined approach of ISSRs with other marker types have produced maps for lentil (Rubeena et al. 2003), larch (Arcade et al. 2000), citrus (Fang and Roose 1999), soybean (Wang et al. 1998b), and wheat (Nagaoka and Ogihara 1997). However, despite the advantages of ISSRs, their main application has been for genetic diversity and phylogenetic studies (Reddy et al. 2002) rather than for genetic mapping. The use of this marker type may serve as a replacement for RAPDs since the amplified products can be converted to codominant SSR markers (van der Nest et al. 2000), and the method has the potential to be automated with double priming and fluorescent labelling (Lui and Wendel 2001).

GENETIC LINKAGE MAPS

The use of molecular markers in mapping has allowed the development of more saturated maps complementing earlier low density genetic maps based on isozymes and morphological data.

The basic concept for the construction of genetic linkage maps is to assess the comparative recombination frequencies of the markers, and then assign the relative distances between them along the linkage groups by converting the recombination frequencies to mapping units.

Development of Genetic Linkage Mapping Populations

In order to calculate recombinations and linkage, a mapping population is produced. Populations used for mapping in woody perennials are generally derived from crossing two heterozygous individuals to produce an F_1 population. This F_1 population can then be used to produce F_2, or a backcross population segregating for a number of polymorphic loci. In order to maximise the genetic information gained from the population, sufficient genetic variation in the trait of interest should exist in the two individuals that are selected as parents to ensure a high level of polymorphism and segregating markers in the progeny.

Segregation and Recombination

During meiosis, each chromosome replicates, and crossing over (chiasma formation) occurs, resulting in reshuffling of the loci. The percentage of recombined gametes for a pair of loci in the population is the recombination rate, *r*, which is related to the distance between the loci. If the loci are far apart on the chromosome, or on different chromosomes (independent assortment), then the probability of a chiasma forming between them will be close to 100%, so around 50% of the gametes will be recombinant. If two loci are closer together then the probability of a chiasma forming will be less, and so *r* will be lower than 50%. If double crossover events do not occur then marker order, or map distances based on *r*, are additive (Crow and Dove 1990).

Genetic Map Functions

If the number of loci is not saturated, then double or multi-crossover events between adjacent loci are not limited and a genetic map function must be employed (Crow and Dove

1990; Manly and Cudmore 1994). The use of genetic map functions allows conversion of *r* values into centiMorgan (cM) distances for map construction when there are large intervals between loci (Kearsey and Pooni 1996; Staub et al. 1996). The use of different genetic map functions and their calculations is discussed by de Vienne (2003) and Zhao and Speed (1996). The two most popular mapping functions used are Haldane (1919) and Kosambi (1944). The principal difference between the two is that the Kosambi mapping function takes into account the effect of interference on the calculation of genetic difference whereas the Haldane function does not (de Vienne 2003). Therefore the Kosambi function produces a shorter genetic map than the Haldane mapping function (Nillson et al. 1994)

Construction of Genetic Linkage Maps

Once all of the progeny and parents are scored for each marker, the data can be entered in an appropriate form into a mapping program for analysis. There are numerous software programs available for the construction of genetic linkage maps that vary in the algorithms used for linkage calculations. The selection of a particular mapping program will be dependant on a number of factors, including the population type (F_1, F_2, BC, RIL, outcrossing, inbreeding, etc), marker type (dominant, codominant, phase), and the mode of marker inheritance (simple or complex). The two most widely used programs are MAPMAKER (Lander et al. 1987) and Joinmap (Stam 1993), although others, such as Linkage-1 (Suiter et al. 1983), and GMENDEL (Lui and Knapp 1990) are useful in some situations.

Mapping programs calculate the recombination frequencies for all pairs of markers, and estimate the linkage between pairs of loci either by using the χ^2 test, or most commonly, by the calculation of a likelihood ratio, or LOD score (logarithmic of the odds

ratio). The LOD score calculates the ratio, to the base 10 (log), between the likelihood that linkage does not exist to the likelihood that linkage exists (Henry 1997). After the LOD is defined, the genetic distances are calculated and the linkage groups are established by linearly ordering the markers along the chromosome using the genetic map function that converts *r* into cM distances. Because of the vast number of calculations between pairwise recombinations, the software usually produces a framework of markers (usually 6 or 7) and then sequentially adds markers to this base group. The end product is a pictorial representation of the genetic markers ordered by the distance in cM between successive loci, where the number of linkage groups should correspond to the number of chromosomes if enough markers are available to saturate the map. Morphological traits are treated in the same manner after they are converted to a format for inclusion into the mapping matrix and entry into the mapping program. Like molecular markers, morphological markers will be ordered linearly onto the map and can then be targeted for marker-assisted selection.

Marker Assisted Selection

The greatest use of genetic linkage maps is for marker-assisted selection (MAS). The basic concept of associating markers with particular morphological traits is not a contemporary concept. In the early 1920s, Sax correlated seed size differences with seed coat colour and pigmentation in the common bean (Sax 1923). Also, the presence of genes for both anthocyanin pigment and trichomes in lettuce seedlings has been used by breeders as an early signal for the presence of other genes (Ryder 1971).

The use of MM for MAS is a more viable option than morphological markers for plant breeders, as discussed previously, and saturated linkage maps have been produced for a number of important crop species, including corn, *Zea mays* (Edwards et al. 1992; Stuber

1992; Stuber et al. 1987), rice, *Oryza sativa* (Yu et al. 1991) and tomato, *Lycopersicum esculentum* (Van Ooijen et al. 1994; Wing et al. 1994). The implementation of MAS in breeding programs follows four basic steps. Firstly, the selected population is screened for a number of phenotypic traits of interest and molecular markers are generated that target the polymorphisms. Secondly, a genetic linkage map is established based on these data. Thirdly, the map is screened, using computer analysis, for the detection of associations (linkage) between the morphological traits (single or quantitative (QTL)) and the molecular markers. Fourthly, the linked markers are used for screening individuals in breeding programs.

The types of traits that have been targeted for selection by MAS cover the whole spectrum of plant breeding objectives including fruit quality, disease resistance, and physiological improvement.

The Rice Example

Rice is a major food staple for a high proportion of the world's population. Therefore, along with other important cereals like wheat, barley, corn, and rye, it has been the focus for an intensive effort in crop improvement via artificial selection.

Improvements in the fragrance of rice has lead to the development of a MM used to detect the major fragrance gene, *fgr*, in common and Australian rice cultivars (Garland et al. 2000; Lorieux et al. 1996). In addition to targeting fragrance, another MM has been developed to score rice grain quality (He et al. 1999). The production of shorter-stemmed rice cultivars has also been a breeding objective and a marker linked to the semi-dwarfing gene *sd-1*, has been developed (Cho et al. 1994; Garland and Henry 2001). Disease resistance has been a major focus for molecular studies in rice and markers have been

linked to blast fungal pathogen genes (Yu et al. 1991), gall midge resistance (Nair et al. 1995), sheath blight (Che et al. 2003), bacterial leaf blight (Rao et al. 2002), and yellow stem borer (Selvi et al. 2002). Other crops where MMs have been used in breeding programs include wheat, tomato, cucumber, and sunflower (Henry 1997).

For woody perennials, such as almond, the development of MMs linked to morphological traits is not yet well-developed. This is largely due to the longevity of these plants, which results in a long inter-generational time and phenotypic assessment period. However, once these initial constraints are overcome, the use of MAS for long lived plants becomes more cost effective since the majority of costs for MAS are associated with sampling, DNA extraction, and time required for screening.

THE EUROPEAN *PRUNUS* MAPPING PROJECT

The European *Prunus* mapping project was established in 1992 to develop molecular markers for use in improvement of important *Prunus* fruit crops. It is a collaborative approach involving six research groups from France, Italy, Spain, and the U.K (Arús, 1994a). Most of the initial mapping made by the *Prunus* group was carried out on peach owing to its economic importance and wide distribution. On the latest linkage map for *Prunus* the majority of important physiological characteristics (12 out of 20 mapped) have been studied in peach, or a peach/almond hybrid (P. Arús unpublished 2001). Studies on almond have elucidated the position of only three characters specific to this species, while cherry has five characters mapped. The most important limiting factor for genome mapping via conventional methods in almond compared to peach has possibly been the lower level of cultivation of almond in Europe. In the past, the use of molecular techniques

in almond mapping has been restricted mainly to isozyme analysis (Ballester 1998; Vezvaei et al. 1995; Viruel et al. 1995) and RFLPs (Ballester 1998; Viruel et al. 1995). However, a combination of these techniques for genetic mapping has proved fruitful for genetic analysis in almond/peach hybrid populations (Foolad et al. 1995; Joobeur et al. 1998). Recent work in almond mapping has used SSRs to produce a high density map based on these markers (P. Arús, unpublished).

The first attempt to use any type of marker for MAS in almond was carried out by Asíns et al. (1994) using isozymes. Asíns et al. (1994) proposed the use of three isozyme markers for MAS linked to QTLs for late flowering and short flowering duration. Work by Graziano et al. (2000) focussed on the development of PCR markers specific for linkage to self-compatibility and shell hardness. The RFLP markers, PC9 and AG35, which are linked to shell hardness, were developed in a peach x almond hybrid cross, and these were not transportable for MAS when screened on a selection of almond cultivars with a range of shell hardness (Wirthensohn et al. 2001). This highlights the importance of developing a genetic linkage map for almond for use in the Australian almond breeding program.

BULKED SEGREGANT ANALYSIS – EXAMPLES FROM *PRUNUS*

A useful technique for developing markers linked to genes of interest is bulked segregant analysis (BSA). The technique is used for identifying markers linked to a specific trait of interest by using bulked DNA samples prepared from 8-12 individuals of alternate phenotypes. The bulked DNA samples are then screened with a range of primers to detect polymorphisms between them. This approach is useful in cases where a genetic linkage map has not been produced for the population under investigation. BSA was developed and

initially employed in an investigation to find a marker linked to disease resistance in lettuce (Michelmore et al. 1991).

In *Prunus*, BSA has been used to target markers linked to flesh colour and hardness, freestone seed, plant stature, pollen fertility, leaf colour, and nectarine character in peach (Chapparo et al. 1994; Warburton et al. 1996), male sterility and self-compatibility in apricot (Badenes et al. 2000), and late bloom in almond (Ballester et al. 2001). Therefore, although the adoption of this method in *Prunus*, particularly almond, is limited it has the potential to identify markers directly linked to traits of importance in breeding programs.

CONCLUSION

Almond is an important tree nut crop the kernel of which is used predominantly for consumption in the unprocessed state and in the baking and confectionery industry (Woodroof 1979).

The Australian almond industry produces only 1.1% of the total world production (FAO 2003), which could be increased by developing cultivars better adapted to Australian conditions. The early selection of important agronomic characters in current breeding programs by MAS would greatly improve the efficiency of the breeding program in Australia. For example, the introduction of a commercially viable self-compatible cultivar would have the capacity to alleviate problems associated with cross-pollination and the inclusion of a less commercially desirable cross-pollinator cultivar in the orchard. *S*-allele specific primers are now available (Channuntapipat et al 2001, 2002ab, 2003), and have been used to screen for self-compatible cultivars in the breeding programs (Ortega and Dicenta 2003).

The construction of genetic linkage maps is now well advanced for *Prunus*, and the development of cultivar specific maps with the inclusion of important agronomic and biological traits for MAS is now a practical tool for almond breeding. The targeting of specific traits using BSA is also a viable option and the application of both methods for MAS is now plausible for plant improvement programs. MAS can be used either for accelerated backcrossing by tracking a particular trait through a number of backcrosses, or it can be used for direct selection of a desirable trait. The ultimate application of MAS to breeding programs will allow the earlier release of commercially viable cultivars by making each round of selection more efficient.

Project Aims

The aims of this research are twofold. Firstly, to morphologically characterize a F_1 population derived from a cross between two commercially important cultivars, 'Nonpareil' and 'Lauranne'. Secondly, to develop a genetic linkage map that includes molecular markers linked to traits of agronomic interest important in current breeding programs, which can then be used for marker assisted-selection in future breeding programs.

Specifically the aims of the project are:

1. To phenotypically characterize the F_1 population for a number of phenotypically and agronomically important traits.
2. The response of phenotypic expression in relation to a number of environmental factors such as temperature and rainfall.
3. Analysis of self-compatibility by measurement of fruit set and pollen tube growth, and the use of allele specific PCR analysis.
4. The development of an optimal method for DNA extraction from almond leaves.

5. The elucidation of factors responsible for the difference in testa pubescence of the parents and F_1 hybrids using sensory panels and light and scanning electron microscopy.
6. The use of bulked segregant analysis to find an inter-simple sequence repeat marker linked to testa pubescence.
7. Development of molecular markers for use in generating a genetic linkage map.
8. Construction of a genetic linkage map for almond based on molecular and agronomic traits using a pseudo-testcross strategy.
9. Comparison of this map to other maps produced for fruit trees in the genus *Prunus*, particularly the *Prunus* reference map 'Texas' x 'Earlygold'.

CHAPTER 2

Generation of the F_1 Hybrid Population

INTRODUCTION

Commercial almond growing in Australia was first established in South Australia. Initial plantings were close to the capital Adelaide, in the south at Willunga, and to the north at Angle Vale. By the 1960s large orchards were established eastwards along the Murray River in the Riverland regions of Renmark in South Australia, and Lindsay Point in Victoria (Bennett 1999). The basis for this large shift away from the Adelaide Plains and Southern Vales was increased urban encroachment on these growing regions in parallel with water restrictions in this area, which relies heavily on bore water for irrigation.

The Riverland region is ideal for large-scale orchards as it has a hot dry Mediterranean climate. Almonds are typically cultivated in regions with poor soils and hot summers with limited water availability. They are less susceptible to drought than other fruit crops, and require less winter chilling, therefore there is a unique niche for almond tree crop growing in this region.

The majority of almond cultivars grown in Australia have been introduced from Europe and the United States (Bennett 1999; Quinn 1928; Wirthensohn and Sedgley 2002). The Australian almond improvement program was initiated in 1997 to develop almond cultivars better adapted to Australian conditions and to introduce self-fertility (Sedgley and Collins 2002; Wirthensohn and Sedgley 2002). In the first year of this project a number of families were produced by Dr Terry Bertozzi from crosses involving eleven cultivars from Australia and overseas (Wirthensohn and Sedgley 2002). A family was produced as part of this program from a cross between 'Nonpareil' (NP) and 'Lauranne' (LA). NP is an American self-incompatible, paper shelled cultivar valued for its good kernel qualities, and is the major cultivar grown for commercial almond production in both Australia and the USA. LA is a French self-compatible, hard shelled variety that is currently grown commercially in Europe. It was obtained from a cross between two European cultivars, 'Ferragnés' (French) and 'Tuono' (Italian), at INRA-Avignon, made in 1978. It is a hard-shelled cultivar which is late flowering, has early production and the self-compatible character, inherited from 'Tuono' (Godini 2002).

This family was selected for further analysis because of its potentially high segregation for important agronomic characters such as self-compatibility, flowering time, shell hardeness, testa pubescence, kernel colour, and kernel flavour.

HYBRID PRODUCTION

Classical cross hybridisations were carried out in 1997 at the Waite Agricultural Research Institute, The University of Adelaide, Adelaide, South Australia (034°53'37"S - 138°33'54"E). NP trees to be used as the female parent were kept in a bird proof cage at the Waite Claremont Orchard (Figure 2.1), under standard irrigation with nutrients applied via fertigation. Two NP clones were used for pollinations, 3-8-4-72 and 3-8-7-

72, no genetic differences were found between these clones by DNA fingerprinting using Random Amplified Polymorphic DNA (Woolley et al. 2000). The self-incompatible NP was selected as the female parent to prevent the risk of self-pollination and aberrant hybrid production in the progeny.

Branches were selected when the majority of flowers were at balloon stage, flowers that had already progressed past balloon stage were removed from the branch by hand. Branches were then enclosed in an insect proof net and left for approximately 2-5 days for flowers to open before controlled hand pollination was carried out.

'Lauranne' pollen was provided by Dr Henri Duval from the original hybrid plant and clones derived from this tree from INRA-Avignon, France in February 1997 and stored at –20°C until required. 'Lauranne' pollen was applied by hand with a camel hair paintbrush to in excess of three hundred open flowers contained in the insect proof nets. Nuts were then left to develop naturally inside the nets until fully mature (Figure 2.2).

A total of 242 nuts from this cross were collected at maturity and germinated in closed plastic bags containing wetted vermiculite and stored at 4°C for approximately 5-7 weeks. After the seeds were vernalised, imbibed and the hypocotyl had emerged from the seed, 208 of the 242 seeds that germinated, were planted in fresh soil and maintained under glasshouse conditions (25°C, 18hr daylength). After four months in the glasshouse plants were hardened off in the shade house. The 184 surviving individuals were transplanted to the field in 1998 on their own roots, in a contained plot within an established commercial almond orchard owned by Andrew Lacey at Lindsay Point, Victoria, Australia (034°15'27"S - 141°00'00"E) (Figure 2.3). Soils at this site were fertile and well drained, with a calcareous subsoil.

The planting design consisted of two adjacent rows with a staggered tree planting of 0.5 m between trees and a distance of 3 m between successive double rows, drip irrigation lines were situated between the double rows (Figure 2.4). Trees were maintained under standard orchard conditions for irrigation and pest control, and no pruning was conducted on any of the progeny.

Of the 184 hybrid trees that were transplanted to the field, three died before the first assessment in 2001. Kernel samples were collected from all hybrids that produced fruit in 2001, 2002 and 2003. A total of 115 progeny which fruited in 2001 were used for morphological assessment for this year, 181 in 2002, and 179 in 2003 (Chapter 6). Leaves for DNA extraction were collected from 171 of these 181 individuals, which were assessed for hybridity by *S*-allele analysis (Chapter 5). Molecular markers were developed for 93 of these individuals which were selected at random from the larger population (Chapter 4) and were used to construct the genetic linkage map (Chapter 8).

Figure 2.1 Waite Claremont Orchard showing trees within the bird cage and branches contained in insect proof nets

Figure 2.2 Close up photo of developing nuts inside insect proof nets after pollinations

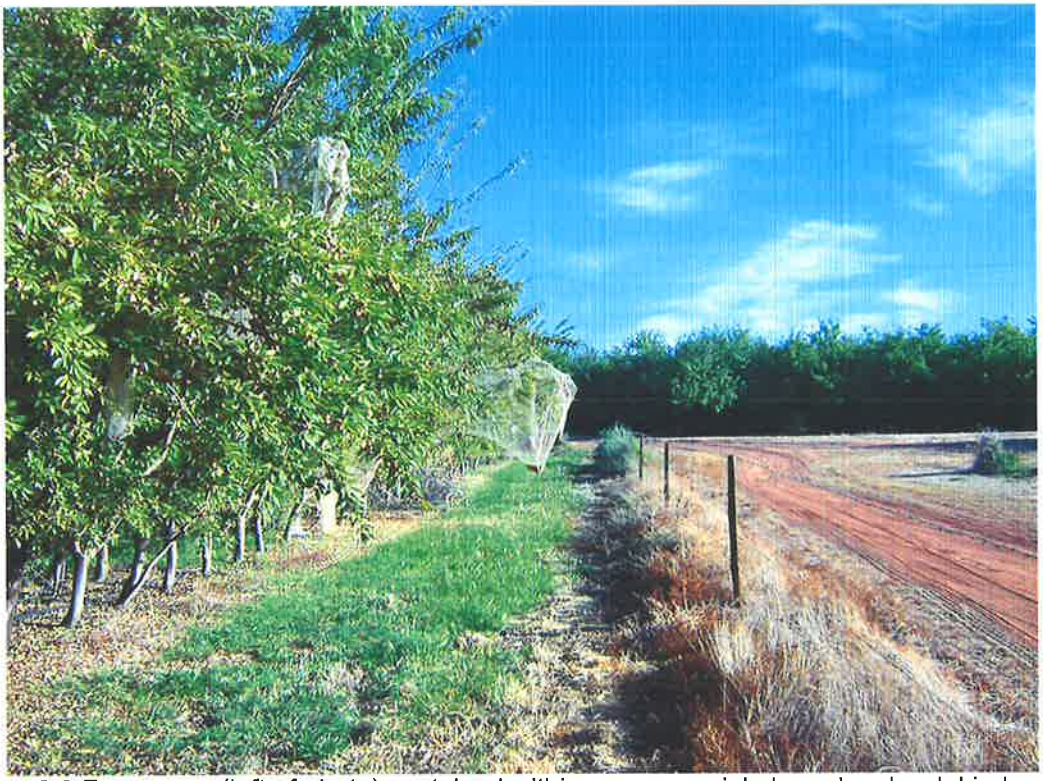

Figure 2.3 F_1 progeny (left of photo) contained within a commercial almond orchard, Lindsay Pt, Australia

Figure 2.4 Irrigation and tree spacing of F_1 progeny, staggered at a distance of 0.5 m between trees

CHAPTER 3

Comparison of Techniques for Isolation and Purification of Genomic DNA from Almond Leaves

INTRODUCTION

The success of DNA based molecular techniques for biological research is largely dependent on the preparation of pure high molecular weight DNA from the tissue. In plant species the isolation of high quality DNA is hindered by at least four classes of naturally occurring compounds present in the plant cells, including carbohydrates (mainly polysaccharides), proteins, RNA, and polyphenolic compounds. Since plants differ in the composition and levels of these compounds in their cells there is not one extraction protocol that can be universally applied to all species. Consequently, there are a number of techniques utilising different compounds for extraction of DNA from plant tissue, which can vary markedly in their effectiveness.

In general, all currently adopted extraction techniques follow the same basic scheme. Initially, the tissue is homogenised, usually in the presence of a liquid buffer, causing disruption of the cells and liberation of the DNA from the nucleus. The DNA is

then separated from the rest of the cellular components by centrifugation in conjunction with a host of appropriate chemicals and enzymes. Various compounds and procedures are also utilised to remove contaminants and maintain the integrity of the DNA throughout extraction, by protection from mechanical shearing and enzymatic degradation. The majority of popular extraction protocols are CTAB (cetyltrimethylammonium bromide) based, following the initial procedure presented by Murray and Thompson (1980) and later adapted by Doyle and Doyle (1991). CTAB is a nonionic detergent, often composing a major portion of the initial extraction buffer, which is used to bind to the DNA and displace proteins while limiting DNA degradation. Also present in many extraction buffers are chemicals that function as antioxidants which prevent phenolic oxidation of the DNA by polyphenolic compounds, including PVP (polyvinyl-pyrrolidone), β-mercaptoethanol and DETC (diethyldithiocarbamic acid). Chelating agents such as EDTA (ethylenediaminetetraacetic acid) and phenanthroline can also be included in this buffer to bind metal cations, particularly Mg^{2+}, which are essential for nuclease activity, and heavy metal cations which have the capacity to break phosphodiester bonds (covalent bonds between the phosphate of one nucleotide and the sugar of the next in the DNA helix). The second major step in DNA extraction often involves an organic extraction, coupled with centrifugation, to separate the majority of cellular debris from the suspended nucleic acids. Other solutions and treatments, such as the use of high salt levels are then used to remove contaminating secondary compounds. Finally the RNA may be precipitated or digested with RNase or lithium chloride, and the DNA precipitated using either isopropanol or ethanol, resuspended in water or TE and stored until required for analysis. In general this precis and its modifications are suitable for a wide range of plant species.

As mentioned, extraction of high quality DNA from plant cells is affected considerably by the presence of polyphenolics and polysaccharides that copurify with, or polymerise to the nucleic acid during isolation. Polyphenolics are ubiquitous compounds in the plant kingdom; they are classified by the presence of an aromatic ring bearing one or more hydroxyl groups and include anthocyanin pigments, lignin and tannins (Walker 1975). Polyphenolics are released upon cell lysis and can irreversibly bind to DNA (John 1992) or can undergo oxidation, resulting in the production of highly reactive quinones (McMurray 1994). Polyphenolic oxidases break down polyphenolic compounds to form quinonic compounds, which can sever DNA strands (Rogers 1994). Loomis and Battalie (1966) suggest that these quinonic compounds have the capacity to restrict the action of enzymes, potentially interfering with the efficiency of enzymes in later applications if not removed. The inclusion of PVP in the initial extraction buffer is particularly effective in removing polyphenolics from solution by complexing them through hydrogen bonding (Jobes et al. 1995; John 1992; Kim et al. 1997).

Carbohydrates are a large group of compounds, structurally composed of carbon, hydrogen and oxygen atoms, which includes all sugars and their polymers. Plants store carbohydrates mainly as starch, as a sugar reserve for energy. They also use other polymers predominantly cellulose as building materials for the rigid cell wall and other structural entities. The removal of carbohydrates from extraction solutions is more difficult compared to polyphenolics since they share structural similarities and have analogous physiochemical properties to DNA. The structural similarity can cause them to precipitate concomitantly with DNA effectively following the nucleic acid through each purification step. Some carbohydrate is removed in the initial chloroform: isoamyl alcohol extraction steps (Murray and Thompson 1980) but even during subsequent

ethanol precipitations the carbohydrate can persist (Segovia et al. 1965; Manning 1991). In CsCl gradients, a sensitive technique for DNA purification, the carbohydrate is seen as a hazy band at the interface with the DNA layer, suggesting that although the carbohydrate is not an intrinsic part of the nucleic acid, it still shares a propensity to behave as if it is (Edelman 1975; Segovia et al. 1965). Therefore any chemicals and conditions used to remove this contaminant must be highly selective. The most common technique used for eliminating polysaccharides is the inclusion of a precipitation step with a high molar concentration of sodium chloride (Murray and Thompson 1980; Jobes et al. 1995; Do and Adams 1991). Other researchers have reported successful removal of these contaminants using an anion exchange resin in combination with salt buffers of various concentrations (Fang et al. 1992), or a carbohydrate exchange resin (Edelman 1975). Steenkamp et al. (1994) reported a combination of CTAB and SDS (sodium dodecyl sulphate) was useful in the removal of carbohydrates.

Polysaccharides and polyphenolics are known to adversely affect DNA based molecular techniques in a variety of ways. The production of reactive oxygen species by the oxidation of polyphenolics has also been shown to break DNA strands (Hiramoto et al. 1998). Polyphenolics and polysaccharides can also bind to DNA upon cellular disruption (Manning 1990), which has consequences for enzymatic activity in molecular based analyses. Shioda and Murakami-Murofushi (1987) are widely cited for work done on the inhibition of DNA polymerase α by a polysaccharide isolated from the extracellular slime of *Physarum polycephalum*, a slime mould. The mechanism of this inhibition is thought to be via selective competition of the polysaccharide with the enzyme for the DNA. Polysaccharides can therefore potentially interfere with DNA polymerase based techniques including PCR and RAPD amplification. *Taq* polymerase activity can be affected by contaminants in a number of ways, including 1)

contaminants bound to DNA causing problems in the denaturation step of PCR (separation of the DNA strands), 2) preventing recognition of binding sites for *Taq* or 3) by obstructing primer annealing (Reiss and Rutz 1999). Studies specifically on the effects of plant polysaccharides on PCR have shown that acidic polysaccharides are the most prevalent in disrupting PCR and RAPD analysis either by changing the pH of the PCR buffer or by obscuring *Taq* polymerase activity (Demeke and Adams 1992; Pandey et al. 1996). DNA to be used for restriction digests can also be affected by polysaccharide contamination since restriction enzymes can be affected in a similar manner (Do and Adams 1991; Fang 1992).

Another lesser contaminant in DNA isolation is RNA, this nucleic acid poses a problem since it shares structural affinities with DNA and can therefore interfere with quantification of DNA by spectrophometric analysis. The use of the enzyme RNase is the general method used to enzymatically digest RNA, however if degradation is incomplete RNA fragments can potentially act as primers in PCR or show absorbance at 260 nm, giving a false reading for DNA yield. RNA contamination can also manifest itself as fast migrating bands on an agarose gel, conceivably giving false amplification patterns in RAPD analysis.

The aim of this study was to examine the efficiency of four different currently adopted extraction procedures for genomic DNA isolation from *Prunus dulcis* leaves with respect to DNA yield, DNA purity and suitability for RAPD analysis.

MATERIALS AND METHODS

Plant Material

DNA was extracted from the leaves of the F_1 population outlined in Chapter 2. Young healthy leaves, free of visible insect damage and disease, were collected from individual trees in spring 2000. They were wrapped in aluminum foil and kept under ice in the field until they were transferred to –80°C. Samples were stored for up to 12 months under these conditions until required for extraction.

Genomic DNA Extraction Protocols

Four different extraction protocols were utilised for comparison in their efficiency for extracting genomic DNA from young leaves of almond. For lypholisation, leaf material was ground in liquid nitrogen in the majority of protocols, as opposed to homogenisation in sterile buffer (Alijanabi and Martinez 1997; Messeguer et al. 1994) or the use of glass beads (Steenkamp et al. 1994), since liquid nitrogen has been reported to be more successful in producing higher DNA yields compared to grinding in buffer (Doyle and Doyle 1990). After extraction all DNA samples were stored at –20°C until required for further analysis.

A. Modified Lamboy and Alpha (1998) procedure - macrotechnique

Method A was a protocol similar to the technique presented by Lamboy and Alpha (1998) but upscaled by ten times to support extraction of DNA from up to 2 g of tissue. 1.5 to 2 g of leaf tissue, with midrib and stems removed, was placed in a

prechilled mortar and pestle (-20°C) and ground to a fine powder under liquid nitrogen. The 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% (w/v), PVP-40T, 1.5% SDS] with 50 mM DETC, 100 mM sodium ascorbate and 1.0% (v/v) β-mercaptoethanol added immediately preceding addition of ground sample. Samples were shaken by hand to form a homogeneous slurry and held on ice until all the samples were prepared. Nucleic acid extraction was initiated by incubation in a waterbath at 37°C for 30 min with gentle inversion every 10 min. After incubation, solutions were extracted with an equal volume (12 mL) of cold (4°C) chloroform : isoamyl alcohol [24 : 1 (v/v)], mixed gently for 5 min on a spinning wheel and then centrifuged at 11 000 rpm for 15 min at 4°C. The upper aqueous phase was transferred to a clean sterile tube and a 0.7 volume of 5 M NaCl was added and mixed thoroughly. The samples were then centrifuged at 11 000 rpm for 30 min at 4°C. The supernatant was decanted into a new tube, taking care not to transfer any of the viscid green pellet. A 0.66 volume of cold isopropanol (-20°C) was added, mixed gently, and placed at -20°C for 30 min to precipitate the nucleic acids. DNA was recovered by centrifuging at 11 000 rpm for 15 min at 4°C. The supernatant was gently decanted and the pellet containing the DNA washed with approximately 10 mL of cold wash buffer [76% ethanol, 10 mM NH_4Ac] overnight at 4°C, or until visibly white. The wash buffer was removed and the pellet rinsed with cold 70% ethanol (-20°C) for 5 min to remove excess salts. The pellet was air dried for 10 min at room temperature and transferred to a clean sterile 1.5 mL Eppendorf tube and dissolved in 600 µL of TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). 3 µL of DNAse-free RNase A (10 mg/mL) was added to each sample and incubated at 37°C for 30 min to digest contaminating RNA. After incubation 240 µL 7.5 M NH_4Ac was added and placed at 4°C for 30 min followed by

centrifugation at 11 000 rpm for 15 min at 4°C, to precipitate proteins, including RNase A. The supernatant was removed with a 1 mL micropipette, taking care not to dislodge the diffuse pellet of protein, and transferred equally to two sterile Eppendorf tubes. DNA was precipitated by adding 2 volumes of cold absolute ethanol (-20°C) to each tube and allowed to stand for 30 min at -20°C, after which the solution was gently inverted to concentrate the DNA. DNA was recovered by centrifuging at 11 500 rpm for 15 min at 4°C, the ethanol decanted, and the white pellet washed with 400 µL cold wash buffer (-20°C) for 15 min, followed by 200 µL of cold 70 % ethanol (-20°C) for 5 min. The pellet was left to air dry at room temperature, and the DNA resuspended in 200 µL of TE buffer and the corresponding samples recombined to give 400 µL total.

B. Modified Lamboy and Alpha (1998) procedure – microtechnique

According to the original protocol of Lamboy and Alpha (1998) approximately 100 mg of leaf tissue was used for extraction. A slight modification included placing the samples in a prechilled 2 mL Eppendorf tube in a plastic rack containing liquid nitrogen (to maintain the small surface area at -180°C) and ground to a fine powder with a prechilled (-20°C) blunt-ended stainless steel rod. Once ground, samples were capped and kept under liquid nitrogen until all samples were prepared. Samples were thawed at room temperature for approximately 30 sec and 600 µL of extraction buffer [250 mM Tris-HCl (pH 8.0), 250 mM NaCl, 50 mM EDTA, 4.0% (w/v), PVP-40T and 1.5% SDS] with 50 mM DETC, 100 mM sodium ascorbate and 1.0% (v/v) β-mercaptoethanol added immediately prior to use, was added to the ground samples, and mixed thoroughly by shaking. The samples were then incubated in a waterbath at 37°C for 30 min and inverted every 10 min, after which, 1 volume of chloroform : isoamyl alcohol

[24 : 1 (v/v)] was added to each tube and mixed gently on a spinning wheel for 5 min. The phases were separated by centrifuging at 11 500 rpm for 15 min at 4°C, and the upper aqueous phase transferred to a fresh 1.5 mL Eppendorf tube. A 0.7 volume of 5 M NaCl was added, mixed gently and centrifuged at 11 500 rpm for 30 min at 4°C. The supernatant was removed to a fresh tube using a 1 mL micropipette with approximately 1 mm of the tip removed to avoid shearing the DNA contained in solution. The DNA was precipitated with a 0.66 volume of cold isopropanol (-20°C), mixed and left at -20°C for 30 min to increase precipitation efficiency. After centrifuging at 11 500 rpm for 15 min at 4°C the supernatant was carefully decanted without disturbing the DNA pellet, and subsequently washed with 400 µL of cold wash buffer [-20°C, 76% ethanol, 10 mM NH_4Ac] for 5 min to 1 hr depending on the level of discolouration. After washing, the buffer was decanted and the pellet rinsed with 200 µL cold 70% ethanol (-20°C) for 5 min, and then left to air dry at room temperature until no liquid appeared visible in the tubes. The air dried DNA pellet was redissolved in 300 µL of TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) containing 15 µg of DNAse-free RNase A and incubated in a 37°C waterbath for 30 min. After incubation a 0.4 volume of 7.5 M NH_4Ac was added and placed at 4°C for 30 min and then centrifuged at 11 500 rpm for 15 min at 4°C to remove precipitated proteins. The supernatant was carefully transferred to a clean tube with a 1 mL micropipette without disturbing the pellet and the DNA precipitated with 1.5 volume of cold absolute ethanol (-20°C) and placed at -20°C for 30 min. DNA was coalesced by inverting the tubes gently until the DNA appeared as a viscous blob in solution, and concentrated further by microfuging at 11 500 for 15 min at 4°C. The alcohol was decanted and the pellet washed for 5 min with 100 µL wash buffer, rinsed briefly with 100 µL cold 70% ethanol (-20°C), air dried, and finally redissolved in 200 µL TE buffer.

C. Alijanabi and Martinez (1997) procedure

50 – 100 mg of young leaf tissue was ground to a fine powder in liquid nitrogen using a chilled blunt ended stainless steel rod as described in procedure B, and stored capped in liquid nitrogen until all samples were triturated. The samples were then removed from the liquid nitrogen, uncapped to relieve internal pressure and thawed at room temperature for approximately 30 sec. 400 µL of sterile salt buffer [10mM Tris-HCl (pH 8.0), 0.4 M NaCl, and 2mM EDTA (pH 8.0)] was added to each tube followed by 40 µL 20% SDS and 6.4 µL of 25 mg/mL proteinase K (Sigma Chemicals, Australia), and mixed thoroughly. The samples were incubated at 65°C for 90 min, removed, cooled slightly, and 300 µL 6M NaCl added. The mixture was agitated and the aqueous phase recovered by centrifugation at 7000 rpm for 25 min at 4°C and transferred to a fresh tube. An equal volume of cold isopropanol (–20°C) was added, mixed gently and placed at –20°C for 1 hr to allow the DNA to precipitate. The DNA was then collected by centrifugation at 7000 rpm for 20 min at 4°C, the supernatant was disposed of and the pellet washed with 200 µL of cold 70% ethanol (–20°C). After washing the alcohol was removed, the pellet left to air dry, and resuspended in 100 µL TE buffer.

D. Messeguer et al. (1994) procedure

Leaves, either fresh or after storage at –80°C, were used for DNA extraction. Using a Waring blender, 6-12 g of leaf tissue was liquified in 80 mL of extraction buffer [0.5 M sorbitol, 0.1 M Tris-HCl (pH 8.0) and 25 mM EDTA] with 3.8 g/L sodium

bisulphite added immediately prior to use. The sample was then transferred to a sterile 250 mL centrifuge tube and centrifuged at 2500 rpm for 15 min at 4°C. After centrifugation, the supernatant was decanted and the pellet resuspended in 20 mL of the above extraction buffer, to which 25 mL of nuclei lysis buffer [200 mM Tris-HCl (pH 7.5), 50 mM EDTA, 2% CTAB (w/v), 2 M NaCl] and 5% Sarcosil (w/v) were added. Each sample was mixed gently and incubated in a shaking waterbath for 20 min at 65°C. The samples were divided into three equal volumes to which 1 volume of chloroform : isoamyl alcohol [24 : 1 (v/v)] was added and shaken for 20 min. The mixture was centrifuged at 2000 rpm for 15 min at room temperature and the aqueous phase transferred to fresh 50 mL falcon tubes. Nucleic acids were precipitated by adding a 0.66 volume of ice cold isopropanol (-20°C) and the tubes inverted until the DNA was visible as a precipitate. Using a heat sealed pasteur pipette as a hook, the DNA was spooled and transferred to a 14 mL polypropylene tube and washed with 70% ethanol. Before redissolving the DNA it was dried on filter paper and resuspended in HPLC grade water or TE buffer (pH 8.0) at 65°C for 1 hr.

E. Modified Steenkamp et al. (1994) protocol

Approximately 2 g of leaf tissue was ground to a fine powder under liquid nitrogen as described in protocol A, and added to 7.5 mL of cold buffer [3% w/v CTAB, 1.4 M NaCl, 20 mM EDTA, 1.0 M Tris-HCl (ph 8.0)] with 15 µL of 2-mercaptoethanol and 15 mg PVP-40T added prior to the addition of the sample. This mixture was gently inverted until all the leaf powder was submerged in buffer and kept on ice until all samples were prepared. The samples were incubated at 60°C for 30 min (with gentle inversion every 10 min), 7.5 mL of chloroform : isoamyl alcohol [24:1 (v/v)] was added

and mixed thoroughly for 10 min. To separate the phases, the sample was centrifuged at 3000 rpm for 20 min at RT. If there was not a distinct biphasic mix separated by a layer of plant debris then the sample was centrifuged for a further 20 min. The upper aqueous phase was transferred to a clean sterile tube, to which a 0.66 volume of ice cold isopropanol (-20°C) was added, and gently mixed until the DNA precipitated. Using a heat sealed pasteur pipette shaped into a hook, the DNA was spooled out of solution, partially dried by squeezing the globule against the side of the tube and transferred to a tube containing approximately 20 mL of wash buffer [76% ethanol, 10 mM NH_4Ac]. The DNA was left washing in this solution until white (in some cases left overnight). After washing, the DNA was gently squeezed against the tube to remove excess wash buffer, and dissolved in 1 mL of TE buffer. 2 μL of DNAse-free RNase A (10 mg/mL) was added and incubated for 30 min at 37°C. Proteins were precipitated by adding a further 2 mL of TE and 1 mL of 7.5 M NH_4Ac, placing on ice for 20 min, and centrifuging at 10 000 rpm for 20 min at 4°C. The supernatant was mixed with two volumes of cold ethanol (-20°C), gently inverted, and centrifuged at 8000 rpm for 10 min at 4°C. The DNA pellet, was air dried, dissolved in 1 mL TE buffer and transferred to a 1.5 mL Eppendorf tube.

Spectrophometric Assessment of Nucleic Acid Concentration and Purity

The DNA content and quality of the samples was determined by ultraviolet irradiation spectrometry, where the levels of ultraviolet absorbance is an indication of both yield and purity according to the Beer-Lambert Law:

$$A = \varepsilon cl$$

Where A = absorbance at a specific wavelength

ε = extinction coefficient ($\varepsilon = 0.020\ (\mu g/mL)^{-1}cm^{-1}$ for ds DNA)

NOTE: page 55 is missing from the print and digital copies of the thesis

min at 95°C, 10s at 50°C, 15s at 45°C, 20s at 40°C, 1 min at 35°C, 30s at 45°C and 1 min 45s at 72°C, with a final extension step of 5min at 72°C.

Gel Electrophoresis

The reaction products were subjected to electrophoresis on 1.5% agarose gel in 1 x Tris-borate electrophoresis buffer (89 mM Tris HCl, pH 8.0, 89 mM boric acid, 5 mM EDTA) for 35 min to 60 min at a constant current of 65 mA (small gels) or 80 mA for 90 min (large gels). To aid in fragment size identification, 1.5 μL of 100 bp ladder (GeneWorks, Adelaide, Australia) was also loaded on the gel. Bands were visualised after staining with ethidium bromide (0.5 μg/mL) for 10 min, followed by washing with RO water for 10 min and exposure to UV light. Images were captured using both Polaroid® film (667) and using the Tekcap computer program (Version 1.1, Tekram Corporation, 1998) converted to bitmap files using Paintshop Pro software® (Version 5.0, Jasc Software Inc. 1998).

All fragment sizes were calculated using GelPro Analyser (version 3.1, Media Cybernetics, Maryland, USA) by comparison to the 100 bp ladder loaded with each gel.

RESULTS

DNA Purity

The level of contamination of DNA extractions was determined by the spectrophometric absorbance values expressed as ratios of A_{260}/A_{230} and A_{260}/A_{280}, presented in Table 3.1.

DNA extracted using excess of 1 g of leaf tissue had absorbance values ranging from 0.73 to 1.77 (A_{260}/A_{230}) and 1.25 to 1.85 (A_{260}/A_{280}). Only method A showed an

acceptable level of purity for both ratios (above 1.7), with method D showing the lowest ratios, indicating a high level of contamination present in the extraction.

Isolation of DNA without the inclusion of a method to remove RNA resulted in RNA contamination. This was observed using agarose gel electrophoresis, where RNA contamination is visualised as fast migrating bands or a smear at low molecular weight (Figure 3.1). This contamination is effectively removed by RNAse treatment and ammonium acetate precipitation (Figure 3.2).

DNA Yield and Integrity

Overall, variations on the method presented by Lamboy and Alpha (1998) (methods A and B) produced the highest yields on a ng/g fw basis, which was greatly improved when using young, actively growing leaves free from physical damage (data not shown).

Table 3.1 Comparison of absorbance values expressed as A_{260}/A_{230} and A_{260}/A_{280} as an indication of purity of DNA extracted from leaf tissue of *Prunus dulcis*, using four different protocols as described in the materials and methods

			Averages			
	Technique	n	260/230	SD	260/280	SD
A	Lamboy and Alpha	100	1.77	0.29	1.85	0.09
B	Micro Lamboy and Alpha	4	1.50	0.34	1.84	0.14
C	Alijanabi and Martinez	3	1.10	0.33	1.58	0.23
D	Messeguer et al	16	0.73	0.25	1.25	0.17
E a	Steenkamp	15	0.84	0.20	1.65	0.27
E b	Micro Steenkamp	4	1.23	0.49	1.44	0.32

Method D produced the lowest yield of all the macrotechniques (using in excess of 1 g of leaf tissue), while the micro version of method E produced the lowest yield of all the microtechniques (using less than 100 ng of leaf tissue).

DNA integrity was assessed using agarose gel electrophoresis, in all cases there was no sign of degraded DNA which indicates that all methods are capable of producing intact high molecular weight DNA, visualised as discrete bands (Figure 3.1).

Method A, although the most time consuming, produced the best results with respect to yield and absence of contamination, while method C was the least time consuming but resulted in RNA contamination.

The Effect of Purification Method on RAPD Amplification

The DNA from two extraction methods, A and D were compared to evaluate the action of impurities on RAPD amplification patterns. In general, the RAPD amplification pattern was not affected greatly by the presence of observed impurities between the techniques or replicates. On average there were between one or two more bands present on the samples extracted by method A than method D. However the greatest variation was observed between replicates of the same sample indicating that although the PCR program and position within the heating block were held constant, the variation observed is likely to be an artifact of PCR cycling.

DISCUSSION

The preparation of nucleic acids from plant tissues is often complicated by the presence of secondary compounds and cell walls (Hattori et al. 1987). The occurrence of contaminants such as carbohydrates, proteins, polyphenolics and unwanted nucleic acids can reduce the efficiency of downstream applications such as PCR, RAPDs or restriction digests. The choice of DNA extraction method should therefore be evaluated not only on purification efficiency but also with respect to the tolerance of the DNA-

based genetic method to be used. The time, effort and monetary costs must also be taken into account when assessing overall extraction efficiency.

There are numerous established protocols for the extraction of DNA from *Prunus dulcis* leaves (Aljanabi and Martinez 1997; Bartolozzi et al. 1998; Foolad et al. 1995; Messeguer et al. 1994; Woolley et al. 2000). However the adaptation of the technique presented by Lamboy and Alpha (1998) for the extraction of DNA from leaves of *Vitis vinifera* has proved the most fruitful approach with respect to DNA quality, quantity and suitability for RAPD analysis in this study. However, comparison with additional extractions for methods B, C, and Eb would give a more significant level of confidence in these results. The basic outline of methods A and E (Lamboy and Alpha 1998; Steenkamp et al. 1994, respectively) are very similar, although it is likely that the success of the former is mainly attributable to the inclusion of a high salt precipitation step. This step precipitated a large proportion of visible cellular debris from the sample, which may include pieces of cellular membrane, organelles, and free lipids. The use of high molecular concentrations of NaCl in plant DNA extractions has also been reported to increase the solubility of polysaccharides in solution. When the DNA is then precipitated with either isopropanol or ethanol the polysaccharides remain in solution and are discarded with the supernatant (Fang et al. 1992; Jobes 1995; Murray and Thompson 1980). The removal of these compounds is important since it is widely recognised that they interfere with a number of enzymatic systems. Acidic polysaccharides can affect the digestion of DNA by restriction enzymes (Boiteux et al. 1999; Do and Adams 1991; Fang et al. 1992). Polysaccharides can also inhibit PCR and RAPD analysis, in particular the action of *Taq* DNA Polymerase. PCR is based on the efficient amplification of target DNA by *Taq*. If contaminating substances prevent *Taq* from recognising binding sites or interfere with primer annealing, then PCR can fail

Figure 3.1 Agarose (1.5%) gel electrophoresis of genomic DNA extracted from leaf tissue of *Prunus dulcis*. Comparison of DNA integrity using five different extraction protocols as described in the materials and methods. (L-R) Lanes 1 and 2, (A) Lamboy and Alpha; Lanes 3 and 4, (E) Steenkamp et al; Lane 5, Blank; Lane 6 (B) Lamboy and Alpha microtechnique; Lane 7 (E) Steenkamp et al. microtechnique; Lane 8 (C) Alijanabi and Martinez

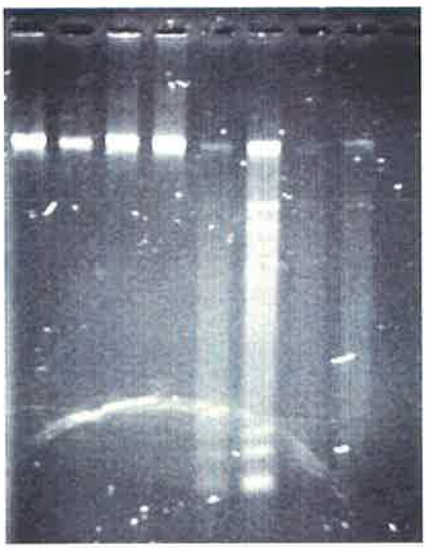

Figure 3.2 Agarose (1.5%) gel electrophoresis of genomic DNA extracted from leaf tissue of *Prunus dulcis*. Comparison of DNA extracted by Lamboy and Alpha protocol (A) treated with RNase A (Lanes 1-4) and not treated with RNase A (Lanes 5-8).

(Reiss and Rutz 1999). It has been proposed that acidic polysaccharides can change the pH of the PCR buffer or inhibit *Taq* directly (Demeke and Adams 1992; Pandey et al. 1996). The exact nature of this inhibition is not known but Shioda and Murakami-Murofushi (1987) have proposed that because DNA polymerase α activity is inhibited exponentially with increasing carbohydrate levels it is likely that the carbohydrate complexes with either the template DNA or the enzyme itself and thus restricts amplification. Since the optimisation of PCR is paramount to the success of RAPD analysis, DNA purity is important.

In the present study comparing the RAPD amplification patterns between two extraction procedures using three RAPD primers assessed the effect of DNA contaminants. These results suggest that the polysaccharides (or other contaminants) present did not greatly affect the amplification pattern between samples of poor quality and relatively high quality. Pandey et al. (1996) have shown that RAPD profiles of spinach DNA inoculated with neutral polysaccharides showed no inhibition of amplification. It may be likely that the majority of the polysaccharide contamination identified by the A_{260}/A_{230} ratio is composed of neutral polysaccharides. The slight variation observed between samples could be an artifact of RNA contamination since Yoon and Glawe (1993) demonstrated that poor RAPD amplification resulted when DNA solutions not treated with RNase were used for PCR. When RNase was used to treat the same samples prior to PCR, strong bands were visualised, indicating that RNA can inhibit RAPD amplification. Since method D did not include a step specific for the removal of RNA it is plausible that the slight variation observed in the RAPD profiles could be attributable to RNA contamination. Some variation was observed between replicates, which may be an artifact of the PCR cycling temperatures since all samples between replicates were subject to the same conditions and identical position within the

heating block to limit machine variation, which can influence reproducibility (Penner et al. 1993).

In summary, optimising the isolation of DNA from plant tissue can be achieved by using young plant tissue (free from stems and major veins) which is lower in polyphenolics and polysaccharides, using liquid nitrogen to lyse cells, since this deactivates DNA degrading nucleases and limits protein coprecipitation (Doyle and Doyle 1990) and the use of appropriate chemicals, solvents and enzymes. Failure to adhere to these protocols results in low DNA yields and substandard DNA purity. Limiting the variables for RAPD analysis is important for reproducibility. Using a DNA extraction technique that produces high quality intact DNA is a good starting point. The technique of Lamboy and Alpha (1998) upscaled to extract 1-2 g of tissue, has also been successful for extracting DNA of high quality from other plant species including *Vitis vinifera* (Lamboy and Alpha 1998), *Eucalyptus sp.* (Neaylon et al. 2001) and *Olea europaea* (Mekuria et al. 2001b). This indicates it is a robust technique suitable for extracting DNA from plant species containing high levels of secondary metabolites.

CHAPTER 4

DNA Based Molecular Markers for Genetic Mapping

INTRODUCTION

DNA-based molecular markers (MM) are frequently employed for genetic mapping, and several PCR-based methods have been developed and applied in plant research. A MM is a variation at the DNA sequence level that is polymorphic and inherited in a Mendelian fashion. The advantages for DNA markers over morphological and protein-based markers for mapping are numerous. They are not influenced by the environment, are selectively neutral, that is these variations are not reported to affect the phenotype of the individual, their number is essentially unlimited and they are detectable in all tissues. The types of MM that are available to molecular biologists are numerous with each marker type having its own unique benefits for each application.

In *Prunus,* initial studies used isozymes for map production (Vezvaei et al. 1995; Viruel et al. 1995), but the use of this type of marker has been superseded by molecular based techniques because of the limited number of biochemical markers, the

small number of polymorphisms these markers are able to reveal, and the possibility of environmental effects (Staub et al. 1982). MMs for genetic mapping studies in *Prunus* have generally been produced by using Random Amplified Polymorphic DNA (RAPD), Restriction Fragment Length Polymorphism (RFLP) and Simple Sequence Repeats (SSRs) (Aranzana 2003ab; Arús et al. 1994a; Arús 1996; Baird et al. 1996; Ballester et al. 2001; Dirlewanger and Bodo 1994; Dirlewanger et al. 1998; Joobeur et al. 1998; Joobeur et al. 2000; Rajapakse et al. 1995; Tao et al. 2000; Viruel et al. 1995; Wang et al. 1998a; Wang et al. 2002; Warburton et al. 1996).

RAPDs are a class of polymorphic genetic markers that utilise short oligonucleotide primers to amplify one or more regions of the genome via PCR. RAPDs have been widely utilised as a marker system for genetic map construction because of their ease of use since they do not require any prior sequence information. In the reaction mix there is only one primer, usually a decamer, which has the capacity to bind randomly at complementary regions along the genome potentially producing a high number of polymorphic bands per reaction.

The application of this technique for genetic map construction in plants covers a large number of species including woody perennials including *Citrus* (Cai et al. 1994), *Eucalyptus* (Byrne et al. 1995; Grattapaglia et al. 1995, 1996; Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996), *Vaccinium* (Qu and Hancock 1997) and *Malus* (Conner et al. 1997), and other important crops including *Lycopersicon* (Foolad et al. 1993) and *Citrullus* (Levi et al. 2001). In the Rosaceae family a number of genetic maps have been constructed using RAPDs alone or in combination with other marker types (Debener and Mattiesch 1999; Dirlewanger and Bodo 1994; Joobeur et al. 2000; Rajapakse et al. 1995; Warburton et al. 1996), illustrating their usefulness for this purpose.

The identification of repetitive DNA sequences, often called microsatellites, opened a new avenue for MM generation. Two popular techniques that exploit repetitive DNA sequences as a source of polymorphisms are Inter-Simple Sequence Repeats (ISSR) and SSRs. The ISSR technique uses primers that contain short repetitive sequences and therefore amplify regions of genomic DNA that occur between microsatellite sequences (Zietkiewicz et al. 1994). ISSRs are reported to be random in their distribution in the genome and are advantageous because, like RAPDs, no sequence information is required, and they are highly polymorphic. However, they are reported to be more reproducible than RAPDs because the primers are longer and therefore have higher annealing temperatures, resulting in more stringent PCR conditions (Bornet and Branchard 2001; Godwin et al. 1997; Goulão et al. 2001; Nagaoka and Ogihara 1997; Tikunov et al. 2003).

ISSRs are useful for genetic linkage mapping since the primers target different regions of the genome compared to RAPDs (Sankar and Moore 2001), and hence the marker types of the two techniques complement each other for map saturation, together with RFLPs, SSRs and isozymes. Substantial genetic linkage maps have been produced using ISSRs in combination with other marker types for *Citrus* (Sankar and Moore 2001), *Larix* (Arcade et al. 2000), *Castanea* (Casasoli et al. 2001), *Fragaria* (Cekic et al. 2001) and *Triticum* (Nagaoka and Ogihara 1997). To date only one study has utilised ISSRs for genetic mapping in *Prunus*, for peach (Dirlewanger et al. 1998), but these markers have not been used in almond studies prior to this.

Although RAPDs and ISSRs are extremely versatile for genetic mapping and other molecular based studies, their usefulness is limited by their dominant nature. Therefore it is advisable in many genetic linkage studies to include marker types that are codominant, such as SSRs and/or RFLPs, which can distinguish heterozygotes from

homozygotes.

SSR primers amplify repetitive DNA because they are designed to target the flanking regions of microsatellites. SSRs were originally utilised for genetic mapping, but are now used for a wide range of applications such as assessing genetic diversity, gene flow, genetic fingerprinting, establishing pedigree and phylogenetic relationships and analysis of genotypes in many fruit crops (Gupta et al. 1996; Wunsch and Hormaza 2002a) including almond and peach, cherry, apricot, pear and apple (Cantini et al. 2001; Hokanson et al. 1998; Hormaza 2002; Martínez-Gómez et al. 2003; Serrano et al. 2002; Wunsch and Hormaza 2002b; Yamamoto et al. 2001, 2002a; Zhebentyayeva et al. 2003). However, the use of SSRs in many plant species is limited because of the high cost of developing the primers, which often yield very few useable SSR loci (Rafalski and Tingey 1993). In some cases, primers that have been developed for a particular species have been used successfully in a closely related species (Rossetto 2001). The transportability of such markers within the Rosaceae family is widely documented; especially those that initially were developed for peach (Aranzana et al. 2002; Cipriani et al. 1999; Joobeur et al. 2000; Martínez-Gómez et al. 2003ab; Sosinski et al. 2000). Peach is the most closely related species to almond in the genus *Prunus* (Watkins 1979), and therefore it shows the highest rate of cross-species amplification since the conserved primer region can be transcribed across interspecies boundaries. This cross-species transportability has also been observed for other genera including *Brassica* (Plieske and Struss 2001) and *Vitis* (Di Gaspero et al. 2000). However Devos and Gale (1997) report that cross-species transferability of SSRs is not possible for grasses.

Another approach for developing SSRs for a particular species is to screen public sequence data bases such as EMBL and GenBank with repeat motifs (Scott 2001), which has been used for SSR development in almond (Ma et al. 2003; Mnejja et

al. 2003)

In the present study, 181 primers were screened for MM generation, and from this, 6 RAPD, 13 ISSR, and 14 SSR markers were selected to generate polymorphic MM in the progeny and parents of an F_1 almond hybrid population. These markers will be used to produce a molecular map for this population.

MATERIALS AND METHODS

Primer Selection and PCR Conditions

RAPD

DNA from the parents of the F_1 hybrid population; 'Nonpareil' and 'Lauranne' was initially screened with a total of 60 RAPD primers from Operon Kits A, B and C (Operon Technologies, USA, Appendix A). Primers were assessed for the clarity, reproducibility, and number of segregating amplified products. The selected primers were then tested on both parents again, and on 10 progeny to check for segregation. Finally the reproducible polymorphic primers were screened on the parents and the 93 F_1 progeny. Duplicate and if necessary triplicate repeats were performed to ensure reproducibility of the amplified products. Primers that showed high amplification efficiency in the first PCR, but did not show good amplification in the second PCR were excluded from analysis based on the lack of reproducibility.

The PCR reaction mix contained 40 ng of genomic DNA, 3 mM $MgCl_2$, 0.25 μM primer, 0.2 mM of each dNTP, 1 x cresol red dye (2% w/v sucrose, 0.1 mM cresol red, sodium salt, Aldrich Chemicals, USA), 1 x PCR buffer (Invitrogen, Australia) and 1.1 U of Taq Polymerase (Invitrogen, Australia), made to 20 μL with sterile MQ water.

PCR amplification was performed in a programmable thermocycler (MJ Research Inc., USA) fitted with a hot bonnet, using the PCR program of 2 min at 95 °C,

followed by 45 cycles of 30 sec at 95 °C, 45 sec at 48 °C, and 1 min at 72 °C, with a final extension step of 10 min at 72 °C. Negative controls containing no DNA were used for all PCR reactions to check for DNA contamination.

ISSR

Twenty three ISSR primers were selected from those reported previously (Casasoli et al. 2001; Sankar and Moore 2001; van der Nest et al. 2000), and purchased from Genset Pacific Pty Ltd, Australia, listed in Table 4.1 in shorthand notation, for example $(CT)_8A$ corresponds to primer sequence CTCTCTCTCTCTCTCTA. Twenty were dinucleotide anchored (at either the 3' or 5' end), two were trinucleotide non-anchored, and one trinucleotide was anchored at the 5' end. All primers were initially screened to see if they could successfully amplify using the parents and 10 progeny, the PCR conditions listed in Table 4.1, and PCR cycling conditions listed in Table 4.2. To optimise PCR amplification for a number of primers several different PCR cycling conditions were trialled on this subset of individuals (Table 4.1). The addition of ammonium sulphate at a final concentration of 10 mM was trialled on a number of primers to aid in amplification efficiency (Table 4.1). Once an optimised set of PCR conditions were established all successful primers were screened on the parents and 93 progeny.

The PCR reaction mix contained 50 ng of genomic DNA, 1.5 mM $MgCl_2$, 0.5 mM primer, 0.2 mM of each dNTP, 1 x cresol red dye (Aldrich Chemicals), 1 x PCR buffer (Invitrogen, Australia) and 1.0 U of Taq Polymerase (Invitrogen, Australia) made up to 20 μL with sterile MQ water.

Table 4.1 Optimal PCR conditions for ISSR primer amplification

Primer Type	Primer	bp length	Annealing Temp Tm	Annealing Temp Ta	Worked Yes	Worked Polymorphic	Final PCR protocol used (Table 4.2)	Other PCR protocols trialled
3' anchored dinucleotide	$(CT)_8A$	17	50	-	×	-	-	MED-PCR, $ISSR_{55}$, $ISSR_{55}$ + NH_4
	$(CT)_8G$	17	52	48	✓	✓	MED-PCR	
	$(GA)_8T$	17	50	48	✓	✓	MED-PCR	
	$(CA)_8G$	17	52	48	✓	✓	MED-PCR	
	$(CA)_8T$	17	50	48	✓	✓	MED-PCR	
	$(AG)_8G$	17	52	55	✓	✓	$ISSR_{55}$ + NH_4	MED-PCR
	$(AC)_8G$	17	52	48	✓	✓	MED-PCR	
	$(TG)_8G$	17	52	-	×	-	-	MED-PCR, $ISSR_{55}$ + NH_4
3' anchored dinucleotide with degenerate base	$(AG)_8YC$	18	55	48	✓	✓	MED-PCR	
	$(AG)_8YT$	18	53	55	✓	✓	$ISSR_{55}$	
	$(AG)_8YA$	18	53	55	✓	✓	$ISSR_{55}$	$ISSR_{55}$ + NH_4
	$(GA)_8YC$	18	55	55	✓	×	$ISSR_{55}$	$ISSR_{55}$ + NH_4
	$(CT)_8RC$	18	55	48	✓	✓	MED-PCR	
	$(CA)_8RG$	18	55	48	✓	✓	MED-PCR	
5' anchored dinucleotide	$GCA(CA)_9$	21	51	60	✓	✓	$ISSR_{60}$	MED-PCR, $ISSR_{65}$
	$CAG(CA)_9$	21	51	-	✓	×	-	$ISSR_{60}$, $ISSR_{65}$
	$GTC(CA)_9$	21	51	-	✓	×	-	$ISSR_{60}$, $ISSR_{65}$
	$GAC(CA)_9$	21	51	-	✓	×	-	$ISSR_{60}$, $ISSR_{65}$
	$CTG(CA)_9$	21	51	-	✓	×	-	$ISSR_{60}$, $ISSR_{65}$
	$CGA(CA)_9$	21	51	-	✓	×	-	$ISSR_{60}$, $ISSR_{65}$
Trinucleotide unanchored	$(ACC)_6$	18	60	55	✓	×	$ISSR_{55}$	
	$(TAA)_8$	16	33	42	✓	✓	$ISSR_{42}$	
5' anchored trinucleotide with degenerate bases	$BDB(ACA)_5$	18	49	-	×	-	-	BDB, $ISSR_{55}$

B = C or G • D = A or G or T • R = A or G • Y = C or T

$ISSR_{(XX)}$ denotes the annealing temperature used for ISSR PCR cycling profile (Table 4.2)

+ NH_4 indicates if ammonium sulphate was used in PCR reaction mix

Table 4.2 PCR amplification profiles for selected ISSR and SSR primers

STEP	PCR Program - ISSR						PCR Program - SSR					
	MED-PCR		*$ISSR_{(AT)}$*		*BDB*		*Method A*		*Method B*		*Method C*	
	(°C)	(min)	(°C)	(min)	(°C)	(min)	(°C)	(min)	(°C)	(min)	(°C)	(min)
1.Initial denaturation	95	2.00	94	7.00	92	1.00	94	3.00	94	5.00	94	1.00
2. Denaturation	95	0.30	93	1.00	92	1.00	94	1.00	94	0.45	94	0.30
3. Annealing	48	0.45	(Table 4.1)	1.00	58	1.00	55	1.00	57	0.45	(Table 4.3)	0.30
4. Extension/elongation	72	1.00	72	2.00	72	1.00	72	1.00	72	0.45	72	1.00
5. Cycling	34 times to 2		41 times to 2		34 times to 2		24 times to 2		34 times to 2		35 times to 2	
6. Final Extension	72	10.00	72	8.00	72	10.00	55 72	2.00 8.00	72	8.00	72	5.00

* $ISSR_{(AT)}$ refers to annealing temperature used for each specific primer listed in Table 4.1

PCR amplification was performed in a PTC-100 programmable thermocycler (MJ Research Inc., USA) fitted with a hot bonnet holding a 96 well plate, using one of three cycling profiles (Table 4.2). Negative controls containing no DNA were used for all experiments to check for DNA contamination in the PCR.

SSR

The 36 SSR primer pairs used are shown in Table 4.3. Twenty of these primer pairs were developed from an almond-enriched genomic library prepared from 'Texas', twelve from peach-enriched genomic libraries from cultivars 'O'Henry', 'Bicentennial', and 'Redhaven', and four from a cherry-enriched genomic library from 'Napoleon'.

SSRs were first tested on the two parents, and those that showed unclear banding patterns, were monomorphic, or produced more than one locus were eliminated and not screened on the progeny.

The PCR protocol used for each individual is shown in Table 4.3 together with the primer sequence data. PCR cycling conditions are listed in text and in table format (Table 4.2).

SSR - Method A

For each PCR, a total of 10 ng of genomic DNA was amplified with 1 mM $MgCl_2$, 0.4 μM each of the forward and reverse primers, 0.4 mM of each dNTP, 1x Taq Buffer (75 mM Tris-HCl, pH 8.8, 20 mM $(NH_4)_2SO_4$, 0.01% Tween-20) (Institut de Recerca i Tecnologia Agroalimentàries (IRTA) Cabrils, Spain) and 1.0 U of Taq DNA Polymerase (IRTA, Spain). 0.16 µM α-^{33}P dCTP (2000 Ci/mmol, Amersham) was added and the volume made up to 15 µL with HPLC grade water. The PCR protocol was 94°C for 1 min,

Table 4.3 *Prunus* derived SSR markers tested in this study, primer sequence, repeat motif, and PCR cycling conditions

Origin	Primer	Primer Sequence (5' → 3')	Repeat Motif	Reported bp length	PCR Program^	AT (°C)	Result		Reference
							Amplified	Polymorphic	
Almond enriched library of 'Texas'	CPACT117	F: TCC AAG GAG AAG GCC TGA AA R: ATT GTG GGT TCC AAC CAATG	$(CT)_{14}$	114	Method A	62	✓	×	(P. Arús, personal communication 2003)
	CPACT119	F: GTC CTC CTC CCA GCT TCT CT R: GGT TTA GCG CAA AAG CTT CA	$(CT)_{12}$-$(CT)_5$	190	"	62	✓	✓	
	CPACT120	F: TGC AGG TTG AAT GTG GCA AT R: CTT TGG GTA GTG CAG GGA TG	$(GA)_{19}$	164	"	62	✓	✓	
	CPACT122	F: GAA GCA GCC ATT CCT AGT GC R: TGT TTA TGG ACC TTA GTA GTC TGG	$(GA)_{18}$	191	"	62	✓	×	
	CPACT132	F: ATG GTC TAA AAA CCG CGA AG R: GGA GAT CAA GAC CGC CTG T	$(TC)_6$-$(CT)_8$	176	"	62	✓	*	
	CPACT137	F: TGC AAA GAA AAA CGG AGA GG R: GAA ACT CAG TGG CAC AAT CG	$(GA)_{25}$	154	"	62	✓	×	
	CPACT161	F: GGA AAC CTG ATT AGG GCA CTT R: GGT CTG CTA TAC TGA CCT AGG ATT	$(GA)_{19}$	196	"	62	✓	*	
	CPACT165	F: AAT TTC TTT TGT TAG GAT AAT ACA R: TTT GCA TAT TGA AAA TTT GTG G	$(CT)_{15}$	200	"	47	✓	✓	
	CPACT166	F: AAA ACT CCT CTC CTT TTC CCT TT R: TCT TCC TCA CCA CCT CAA GC	$(CT)_{24}$	153	"	62	✓	✓	
	CPACT173	F: TGA TCG GCG TCT CCT TTA TC R: AAA GCA AGC AGG CAA ATG AA	$(CT)_{14}$	152	"	52	✓	✓	
	CPACT175	F: GTG GCA AAT GTT GGC AAA G R: AAC ACA AAG CAG CAC CAA GA	$(CT)_{17}$	172	"	62	✓	✓	
	CPACT177	F: TGA AAT CTT TAA ATC ACC CGA CT R: CTT GCT TGC TTG CTTCAC CT	$(CT)_{19}$-$(GT)_3$	188	"	62	✓	✓	
	CPACT179	F: GAC CTC ATC AGC ATC ACC AA R: TTC CCT AAC GTC CCT GAC AC	$(CT)_{10}$	172	"	62	✓	*	
	CPACT185	F: TGA GGA GAG CAC TGG AGG AG R: CAA CCG ATCCCT CTA GAC CA	$(CT)_{19}$	174	"	62	✓	✓	
	CPACT191	F: TGA ACG TTG CAC TCC TTC AC R: ACC ACC ACC ATA ACC ACC AT	$(GA)_{19}$	171	"	62	✓	✓	
	CPACT314	F: AAT TCA TAA ATC AAC AAA TCA ACA R: GCA GAG CTT TTG GGT CAA CT	$(GA)_{22}$	179	"	62	✓	*	

Continued

Origin	Primer	Primer Sequence (5' → 3')	Repeat Motif	Reported bp length	PCR Program^	AT (°C)	Result Amplified	Result Polymorphic	Reference
Almond enriched library of 'Texas'	CPACT319	F: GAG AAC CTT TTG TTT GGC CTT A R: CGT CGT ATT TAG TGC CGT TG	$(GT)_4$-$(CT)_7$-$(CT)_6$	165	"	60	✓	✓	(P. Arús unpublished)
	CPACT326	F: TCG AAG GAG GAT GAA GTT GC R: ATA TCA CGA GGG GCA AAA TG	$(GA)_{17}$	146	"	62	✓	✓	
	CPACT387	F: TGG ACA TCG ATT CAG AGA AAA A R: CGC AAG GTC AAA CTT TCT CA	$(GA)_{21}$	166	"	60	✓	✓	
	CPACT391	F: TCA AAA ACA CCC ATT ATT GAA R: AAA CAT TTA GGG CTT GTT TGG	$(CT)_{10}$	174	"	57	✓	✓	
Peach (AG/CT) enriched library of 'O'Henry'	CPPCT005	F: CAT GAA CTC TAC TCT CCA R: TGG TAT GGA CTC ACC AAC	$(CT)_{25}$	160	Method B	52	✓	✓	(Aranzana et al. 2002)
	CPPCT030	F: TGA ATA TTG TTC CTC AAT TC R: CTC TAG GCA AGA GAT GAG A	$(CT)_{30}$	198	Method A	55	✓	×	
Peach genomic library of 'Bicentennial'	pchgms1	F: GGG TAA ATA TGC CCA TTG TGC AAT C R: GGA TCA TTG AAC TAC GTC AAT CCT C	$(AC)_{12}$-$(AT)_6$	194	"	57	✓	*	(Sosinski et al. 2000)
	pchgms3	F: ACG GTA TGT CCG TAC ACT CTC CAT G R: CAA CCT GTG ATT GCT CCT ATT AAA C	$(CT)_{14}$	179	"		✓	×	
	pchgms4	F: ATC TTC ACA ACC CTA ATG TC R: GTT GAG GCA AAA GAC TTC AAT	$(CT)_{21}$	174	"	52	✓	×	
Peach (AG/CT) or (AC/GT) enriched library of 'Redhaven'	UPD96-001	F: AGT TTG ATT TTC TGA TGC ATC C R: TGC CAT AAG GAC CGG TAT GT	$(CA)_{17}$	120	Method C	57	✓	✓	(Cipriani et al. 1999)
	UDP96-003	F: TTG CTC AAA AGT GTC GTT GC R: ACA CGT AGT GCA ACA CTG GC	$(CT)_{11}$-$(CA)_{28}$	143	"	57	✓	×	
	UDP96-005	F: GTA ACG CTC GCT ACC ACA AA R: CAC CCA GCT CAT ACA CCT CA	$(AC)_{16}TG$ $(CT)_2CA$ $(CT)_{11}$	155	"	57	✓	✓	
	UDP97-401	F: TAA GAG GAT CAT TTT TGC CTT G R: CCC TGG AGG ACT GAG GGT	$(GA)_{19}$	130	"	57	✓	✓	
	UDP98-405	F: ACG TGA TGA ACT GAC ACC CA R: GAG TCT TTG CTC TGC CAT CC	$(AG)_9$	104	"	57	✓	✓	
	UDP98-409	F: GCT GAT GGG TTT TAT GGT TTT C R: CGG ACT CTT ATC CTC TAT CAA CA	$(AG)_{19}$	129	"	57	✓	✓	
	UDP98-412	F: AGG GAA AGT TTC TGC TGC AC R: GCT GAA GAC GAC GAT GAT GA	$(AG)_{28}$	129	"	57	✓	*	

Continued

Table 4.3 Continued

Origin	Primer	Primer Sequence (5' → 3')	Repeat Motif	Reported bp length	PCR Program^	AT (°C)	Result		Reference
							Amplified	Polymorphic	
Cherry enriched genomic library of 'Napoleon'	PS7a2/ PS07A02	F: CAG GGA AAT AGA TAA GAT G R: TCT AAT GGT GGT GTT CAT T	?	164	-	52	×	-	(Sosinski et al. 2000)
	PS8e8/ PS08E08	F: CCC AAT GAA CAA CTG CAT R: CAT ATC CAA TCA CTG GGA TG	?	183	Method B	55	✓	×	(Cantini et al. 2001)
	PS9f8	F: GGT TCT TTG GTT ATT ATG A R: ACA TTT CTA TGC AGA AGT A	?	156	-	-	×	-	(Joobeur et al. 2000)
	PS12e2/ PS12A02	F: GCC ACC AAT GGT TCT TCC R: AGC ACC AGA TGC ACC TGA	?	164	Method C	55	✓	✓	(Downey and Iezzoni 2000)

? = not listed in literature
^ = PCR program listed in Table 4.2
* = polymorphism not clear or multiallelic

followed by 35 cycles of 94°C for 30 sec, annealing temperature for 30 sec (Table 4.3) and 72°C for 1 min, followed by a final extension step of 72°C for 5 min. The PCR reactions were carried out in a GeneAmp PCR system 9700 (Applied Biosystems, USA) in a 96 well plastic disposable plate.

SSR – Method B Adapted from Joobeur et al. (2000)

For each PCR a total of 40 ng of genomic DNA was amplified with 2 mM $MgCl_2$, 0.2 mM each of the forward and reverse primers, 0.36 mM of each dNTP, 0.9 x cresol red dye, 1 x Taq Buffer (Invitrogen, Australia), and 1.0 U of Taq DNA Polymerase (Invitrogen, Australia) made up to 20 µL with sterile MQ water. The PCR protocol was 94°C for 3 min, followed by 24 cycles of 55°C for 1 min, 72°C for 1 min, and 94°C for 1 min, followed by 2 min at 55°C, and a final extension step of 72°C for 8 min. The PCR reactions were carried out in a Programmable Thermal Controller, (MJ Research Inc., USA) fitted with a hot bonnet in a 96 well plastic disposable plate.

SSR – Method C Adapted from Cipriani et al. (1999)

For each PCR, a total of 100 ng of genomic DNA was amplified with 1.5 mM $MgCl_2$, 0.2 µM each of the forward and reverse primers, 0.2 mM of each dNTP, 0.8 x cresol red dye, 1.5 x Taq Buffer (Invitrogen, Australia) and 1.1 U Taq Polymerase (Invitrogen, Australia) made up to 25 µL with sterile MQ water. The PCR protocol was 95°C for 5 min, followed by 34 cycles of 94°C for 45 sec, 57°C for 45 sec and 72°C for 45 sec, followed by a final extension step of 72°C for 8 min. The PCR reactions were carried

out in a in a 96-well plastic disposable plate in a Programmable Thermal Controller (MJ Research Inc., USA) fitted with a hot bonnet.

Gel Electrophoresis

RAPD

PCR products were separated on a 1.5% w/v agarose gel (LE, Promega, Australia) for band visualisation and band size was assessed by comparison to a 100 bp ladder using the same conditions described in Chapter 3.

ISSR

PCR products were resolved on 1.75% w/v agarose gels (LE, Promega, Australia), using the same conditions described in Chapter 3.

SSR

PCR products that were to be separated by polyacrylamide gel electrophoresis (PAGE), were mixed with an equal volume of 2 x formamide/bromophenol blue (FLB) loading dye (95% v/v formamide, 10 mM EDTA pH 8.0, 0.02% w/v bromophenol blue and 0.02% w/v xylene cyanol FF). Immediately prior to loading on the gel, the mixture was heated at 95°C for 5 min and then cooled on ice (Collins and Symons 1993).

SSR - Method 1 – PAGE with autoradiogram detection

Electrophoresis of the denatured PCR products was performed in a Biorad polyacrylamide gel electrophoresis (PAGE) tank. Prior to gel preparation, one plate of the gel system was coated with Sigmacote® (Sigma Chemicals) to aid in release of the gel from

the plate after electrophoresis. The polyacrylamide gel consisted of 6% polyacrylamide (19:1, w/w acrylamide:bis-acrylamide), 8 M urea, 80 µL of TEMED (N,N,N',N',-tetramethylethylenediamene, Sigma Chemicals), 400 µL of 10% w/v freshly prepared ammonium persulphate, and 1 x TBE buffer (90 mM Tris-borate, 2 mM EDTA, pH 8.0) in a volume of 100 mL. Gels were preheated to 50°C in 1x TBE buffer at 120 W before the samples were loaded, and then run at a constant 120 W for 1½ to 2 hours, until the loading dye reached the bottom of the gel. For fragment size identification a 30-330 bp ladder (AFLP Ladder, Invitrogen) radiolabelled with α-^{33}P dATP was loaded into the middle and end lanes of the gel. Direct comparison of the PCR product size with this ladder was used to estimate the sizes of the SSR loci.

Gels were transferred to blotting paper (GB 002, Schleicher and Schuell) and dried under vacuum at 80°C for 2 hours. The gel was then exposed to AGFA autoradiography film for 24-48 hours in a light proof cassette (Sigma Chemicals) to expose amplification products. Bands were visualised by developing the exposed film in standard photographic solutions for this film type.

SSR - Method 2 – PAGE with ethidium bromide detection

Electrophoresis of the denatured PCR products was carried out using vertical 16 x 13 cm gel plates and the associated tank. Prior to gel preparation one plate was coated with Repel-Silane (2% w/v dimethyldichlorosilane solution in 1,1,1-trichloroethane, LKB Bromma, Sweden) to aid in gel separation from the plate. The polyacrylamide gel (SequaGel® 8, National Diagnostics, USA) consisted of 8% polyacrylamide, 7 M urea, 50 µL of TEMED (N,N,N',N',- tetramethylethylenediamene), and 400 µL of freshly prepared

10% w/v ammonium persulphate (Sigma Chemicals) added prior to casting in a total volume of 50 mL. Gels were left to set for approximately 1 hr and then preheated at 25 mA in 1 x TBE buffer for 20 to 40 min before loading. After the samples were loaded they were run at a constant current of 25 mA for approximately 3 hr, or until the line of bromophenol blue reached the bottom of the gel. For fragment size identification, a ladder consisting of *p*UC19 DNA restricted with *Hpa* III (Geneworks, Australia) was loaded into the middle and end lanes of the gel. Direct comparison of the PCR product size with this ladder was used to estimate the sizes of the SSR loci.

Gels were stained with ethidium bromide solution (0.5 μg/mL) for 10 min and destained for a further 10 min in RO water. Bands were visualised by exposure to UV light and the image captured by polaroid photography or a video camera print out with a Sony digital thermal printer.

SSR – Method 3 – Fluorescent detection

Seven forward primers were labelled at the 5'-end with a fluorescent dye. FAM (FAM-UDP97-401, FAM-UDP98-409) and HEX (HEX-UDP96-001, HEX-CPPCT005, HEX-UDP98-405) labelled primers were purchased from Geneworks, Australia NED (NED-PS12e2, NED-UDP96-005) labelled primers were purchased from Applied Biosystems, USA. The PCR conditions followed those presented in Method B, with an annealing temperature of 57°C for all except HEX-CPPCT005, which was 52°C, all PCRs had a final extension step after the initial PCR of 45 min at 72°C to aid in incorporation of the fluorescent dyes.

Prior to sample electrophoresis, 10 μL of water was added to the PCR reaction. 2.5

μL of this diluted reaction was mixed with 2.5 μL of loading buffer [200 μL Hi-Di-Formamide (Applied Biosystems), 40μL Blue Dextran (Applied Biosystems), 30 μL Genescan-500 ROX® or 50μL Genescan-500 TAMRA®, dependant on the filter set] and denatured at 95°C for 3 min, then placed on ice for 5 min. 1 μL of the denatured solution was resolved on a 377 ABI Prism DNA sequencer (36 cm long, 0.2 mm thick, 96 lane gel, 4.5% polyacrylamide (19:1 acrylamide-bis) with 6 M urea, using 1 x TBE running buffer) at 3 000 V for 2.8 hrs at 51°C (AGRF, Melbourne, Australia). Data were analysed using the Genotyper software (PE Applied Biosystems, USA).

Calculation of Multiplex Ratio

The multiplex ratio was calculated by dividing the total number of scoreable segregating bands identified in the population, by the total number of clear and reproducible bands amplified in the population.

RESULTS

Reproducible polymorphic banding patterns were obtained from 10% of the 60 RAPD primers screened, 56% of the 23 ISSR primers, and 14% of the 98 SSR primers. From the 6 RAPD primers scored on the progeny 15 polymorphic bands were produced, and for the 13 ISSR primers, 36 polymorphic bands were produced (Table 4.4). The multiplex ratio (calculated as a ratio of polymorphic products in relation to total number of bands produced) for RAPDs and ISSRs showed a difference in the efficiency between these two techniques. RAPDs showed a higher average number of bands (8.3) compared to ISSRs (6.6) but the number of polymorphic bands were 2.5 and 2.8 respectively. Therefore,

for ISSRs the multiplex ratio was 0.42 compared to 0.30 for RAPDs (Table 4.5). Because of the nature of SSRs only one locus can be scored for each primer, and therefore RAPDs and ISSR produced approximately two and a half to three times more polymorphisms per primer.

Several different methods were trialled for the optimisation of ISSR primers. Table 4.2 lists the sequences of primers tested, the presence or absence of anchor nucleotides, whether the primers produced a PCR product and if this was polymorphic, and the PCR protocols used for optimisation. The PCR cycling conditions were selected based either on previous publications (Cai et al. 1994; Casasoli et al. 2001; Sankar and Moore 2001), or the calculated theoretical annealing temperatures. Although ammonium sulphate was used for a number of primers to optimise PCR conditions there appeared to be no obvious trends for the requirement of this reagent.

Three ISSR primers did not produce any amplification products using any PCR conditions trialled, and seven primers produced an amplified product that was not polymorphic. Primers based on 5'-anchored dinucleotide sequences produced very similar banding patterns. Table 4.6 shows the average multiplex ratio for the various primers used. Both the polypurine primers, and the 3'-anchored dinucleotides with a 2 bp anchor and a degenerate base, showed a higher than average multiplex ratio (0.56 for both). Conversely the polypyrimidine and 1 bp 3'-anchored dinucleotides showed the lowest average multiplex ratios (0.26 and 0.38, respectively).

Of the 43 almond-derived SSR primer pairs sampled on the parents, 81.4% produced amplification fragments in the total population. 51% of these were polymorphic and the remainder were either multiallelic, or either not polymorphic or not clear. For the

primers generated from other *Prunus* species (peach and cherry), 82% produced an amplifiable product (Table 4.3). On a per species basis, 100% of the peach-derived SSR primers were able to produce a product, and 50% of these produced a polymorphic banding pattern. For cherry only 50% of primers tested were able to amplify in almond, and only 25% of these were polymorphic (Table 4.3). Primers that were not optimised did not produce a PCR product or were unclear in their banding profile.

Although 45 additional peach primers were selected and screened on the parents, a significantly high number did not produce a product when amplified at their suggested optimal annealing temperatures. Due to resource and time restrictions these primers were not fully optimised and are not included in the discussion.

PCR products produced by both RAPD and ISSR primers ranged from 300 to 2220 bp in the scored area, whereas those produced by SSR primers were much smaller, ranging from 104 to 198 bp.

DISCUSSION

The production of MMs by both the RAPD and ISSR techniques is based on the binding of the primers to random complementary sequences in the genome. The percentage of RAPD primers initially screened for this study that were able to generate polymorphisms (10%) was lower than values reported for other *Prunus* species, for example 12% for almond (Joobeur et al. 2000), 17% for peach (Dirlewanger et al. 1998) and 27.5% for apricot (Hurtado et al. 2002). In comparison to values obtained for other fruit crops it was significantly lower when compared to apple, 82% (Conner et al. 1997), 71% for olive (la Rosa et al. 2003) and chestnut, 58% (Casasoli et al. 2001). ISSRs were more efficient in

Table 4.4 Summary of number of primers screened on the mapping population

Primer Type	screened on parents[a]	screened on progeny[b]	screened and scored[c]
RAPD	60	11	6
ISSR	23	13	13
SSR	98	30	14
TOTAL	**181**	**54**	**33**

[a] number of primers screened on both parents
[b] number of primers screened on both parents and the F_1 progeny
[c] number of primers from [b] that were able to be scored clearly and were reproducible

Table 4.5 Multiplex ratio of RAPD and ISSR primers used in this study, based on total number of polymorphic bands as a ratio of the total number of bands produced per primer. RAPD primer sequences liseted in Appendix A.

Primer	Total bands/primer	Polymorphic bands/primer	Multiplex ratio
OPA-02	5	1	0.20
OPA-08	6	3	0.50
OPB-10	11	3	0.27
OPC-07	4	1	0.25
OPC-10	14	3	0.21
OPC-11	10	4	0.40
AVERAGE RAPD	**8.3**	**2.5**	**0.30**
$(CT)_8G$	8	3	0.37
$(GA)_8T$	9	3	0.33
$(CA)_8G$	13	3	0.23
$(CA)_8T$	7	3	0.43
$(AG)_8G$	2	1	0.50
$(AC)_8G$	9	4	0.44
$(AG)_8YC$	7	6	0.86
$(AG)_8YT$	5	3	0.60
$(AG)_8YA$	6	3	0.50
$(CT)_8RC$	6	1	0.16
$(CA)_8RG$	3	2	0.66
$GCA(CA)_9$	7	2	0.28
$(TAA)_8$	4	2	0.50
AVERAGE ISSR	**6.6**	**2.76**	**0.42**

Table 4.6 Comparison of multiplex ratio based on ISSR nucleotide primer composition, and presence of anchor bases

Primer Composition	Number	Average multiplex ratio
Polypurine (A/G containing)	5	0.56
Polypyrimidine (C/T containing)	2	0.26
Purine/Pyrimidine (A/C containing)	5	0.41
3' anchored dinucleotide (1 bp anchor)	6	0.38
3' anchored dinucleotide (2 bp anchor with degenerate base)	5	0.56

this study (56%) than those reported in the literature for chestnut, 50% (Casasoli et al. 2001), citrus, 50% (Sankar and Moore 2001), larch, 36% (Arcade et al. 2000), and wheat, 33% (Nagaoka and Ogihara 1997). SSR efficacy was considerably lower in this study (15%) compared with values reported for other *Prunus* species such as peach, 80% (Aranzana et al. 2003), 75% for almond (Joobeur et al. 2000) and 31% for apricot (Hurtado et al. 2002). This indicates that the efficiency for each primer type is species and cultivar specific. For this study ISSRs were the most proficient marker type for polymorphism generation on a per primer basis. These markers do however suffer from the disadvantage that they are dominant. However some papers have reported the conversion of these marker types to co-dominant SSRs by cloning and sequencing the PCR product and designing primers to the flanking regions of the amplified ISSR (Lian et al. 2001; van der Nest et al. 2000).

Although RAPDs produced on average, a higher number of bands per primer than ISSRs they had a lower multiplex ratio because the bands were less polymorphic. This finding is supported by work on chestnuts where the number of polymorphic bands per

primer was found to be 2.6 for RAPDs and 3.0 for ISSRs (Casasoli et al. 2001). For pear the multiplex ratio for ISSRs was calculated to be 0.79 (Monte-Corvo et al. 2001) which is significantly higher than observed in this study. Significantly more bands were scored in pear, on average 42 bands per primer, of which 80% were polymorphic, compared to 6.6 bands/primer in almond (Table 4.5). Therefore the conversion of primer efficiency into a multiplex ratio provides a comparable number for use between different studies where the number of PCR products is significantly different.

No poly (TA) or poly (GC) ISSR primers were used, although poly (AT) dinucleotide repeats are the most abundant in plants (Morgante and Olivieri 1993; Wang et al. 1994). This is because both these poly-repeated nucleotides are self-complementary and can potentially form dimers and hairpin structures during amplification (Wang et al. 1994). Vijayan and Chatterjee (2003) showed that the amplification of genomic DNA in mulberry using AT rich primers was difficult due to the self-complemental nature of these primer types. Primers with one to three anchor bp were used for all ISSR PCRs to prevent smearing, as proposed by Tikunov et al. (2003). In contrast, Arcade et al. (2000) observed that 5' anchored ISSR primers did not show a clear amplification product but rather inconclusive smearing. This was not observed in this study which used 3 bp 5' anchored dinucleotide ISSRs compared to the 4 bp 5' anchored dinucleotide primers used by Arcade et al. (2000). Bornet and Branchard (2001) found that using an annealing temperature (T_a) above the calculated primer melting temperature (T_m) reduced smearing for tri- and tetra-nucleotide repeats, but no smearing was ever observed for any ISSRs trialled in this study, even when the T_a was less than T_m (Table 4.1). Although Dirlewanger et al. (1998) used eight ISSR primers for mapping in peach, predominantly CA/AG containing dinucleotides

with various 5' and 3' anchor bases, they do not state the primer efficiency based on composition, therefore no comparison can be made with other *Prunus* species. For the closely related pear, primers $(CA)_8G$ and $(AG)_8YT$ did not produce any amplification (Monte-Corvo et al. 2001), however they did produce scoreable polymorphic amplification products in almond. Therefore the optimisation and amplification efficiency of ISSR primers appears to be somewhat specific to the species under investigation.

For the almond-derived SSRs, 18.6% did not produce any amplification product, but since optimisation was not employed this number may be inflated. Of the 17 cross species derived SSR primers, 82% showed amplification (100% of peach and 50% of cherry), supporting the conclusion that primer transportability is proportional to genetic relatedness (Rossetto 2001; Wunsch and Hormaza 2002a). Cherry is less closely related to almond than peach, with both almond and peach being classified in the subgenus *Amygdalus* (Watkins 1979) therefore transferability appears to be directly proportional to taxonomic distance as proposed by Dayanandan et al. (1997) for the Leguminosae. However, in contrast, Joobeur et al. (2000) found that cherry primers were highly transportable for the amplification of almond DNA obtained from an F_1 population derived from a cross between 'Ferragnes' and 'Tuono'. Because these two cultivars are the parents of 'Lauranne', one of the parents used in the present study, it was expected that amplification with primers developed from cherry would be successful since the populations are quite genetically similar. One possible explanation for this anomaly may be that a cross-over event occurred in 'Lauranne' that either disrupted some of the primer-binding sites or resulted in the presence of a null allele. Therefore no amplification would be expected in this individual or any of the F_1 progeny that were derived from it. Part of the

variation in primer efficiency for SSRs may be related to the GC content of the primer sequence. The cherry primers have a lower average percentage GC content (38%) compared to the peach derived primers (46%). Dayanandan et al. (1997) observed a correlation between low GC content and negative transferability of SSRs in tropical tree species of the Leguminosae family, which is in the same subclass, Rosidae as Rosaceae. However cross species developed SSRs are still valuable, since the initial development of SSRs is prohibitive for small scale mapping populations, and thus the cross species transfer of SSRs is a procedure which allows a more cost effective and time enhanced method for SSR amplification.

This study shows that there is a high transportability between SSR primers derived from almond cultivars and closely related species. The use of nonspecific primer types such as RAPDs and ISSRs has the capacity to generate a large number of polymorphic markers cheaply and rapidly. Therefore the combined application of these primer types to mapping in almond has the capacity to enhance new and developed maps in the *Prunus* genus.

CHAPTER 5

Confirmation of 'Nonpareil' x 'Lauranne' F_1 Population Hybridity by S-allele Identification and Molecular Fingerprinting

INTRODUCTION

Historically, the progeny arising from crosses between almond cultivars have been identified by analysis of morphological traits such as tree architecture, fruit and other agronomic characters (Quinn 1928). There are a number of problems associated with the determination of hybridity by ampelographic methods, since environmental conditions and the growth stage of the plant often affect the phenotypic expression of morphological characters under investigation (Asíns et al. 1994; Barbera et al. 1988; Nanos et al. 2002). An additional problem is that although hybrids may have a phenotype intermediate to both parents, genetic dominance or epistasis may affect some characters, posing difficulty in adequately distinguishing the progeny. In an effort to refine almond cultivar identification further, several groups have concentrated on biochemical analysis. In particular, the use of chemometric analysis has proved fruitful

for varietal identification. Cordeiro et al. (2001) used the nutritive components of almond kernels, including the levels of proteins, lipids, fibre, ash, and sugars to separate a number of Portuguese almond cultivars. Martín-Carratalá et al. (1999) analysed levels of triglycerides in almond kernel oil for cultivar classification, and using a series of statistical tests, was able to discriminate individual cultivars, as well as groupings for regional areas of origin, i.e. USA, Spain, Italy. However the major disadvantage of biochemical tests is that the levels of the compounds being examined may vary depending on the tissue, the developmental stage of the plant and environmental conditions (Nanos et al. 2002; Schirra 1997). Therefore the identification of a procedure to correctly classify individual cultivars or individuals *per se* has been a priority.

The development of DNA based techniques has allowed researchers to open up a new avenue for cultivar identification and genetic fingerprinting at the molecular level in the breeding of fruit trees and other crops (Wunsch and Hormaza 2002). Two techniques based on molecular methods are now available to breeders to check hybridity of individuals, based on the segregation of known PCR products, such as *S*-alleles, or unknown DNA sequences such as molecular markers in tandem with cluster analysis.

In almond, self-incompatibility (SI) is controlled by a single polymorphic locus (the *S*-locus) with multiple alleles. Diploid tissue of the plant contains two *S*-alleles, while haploid tissue such as pollen contains only one of these *S*-alleles. Based on controlled pollinations, pollen tube growth tests, fruit set, detection of stylar *S*-RNases and molecular analysis it is now established that *S*-alleles are inherited independently (Ortega and Dicenta 2003). Using a new method based on Polymerase Chain Reaction (PCR) developed by Channuntapipat et al. (2001), some *S*-alleles can now be identified at the DNA level using molecular techniques. The variation in both length and

nucleotide composition between the individual *S*-alleles can be used to discriminate between them, and presently there are 13 *S*-alleles that can be identified by this technique (Channuntapipat et al. 2001; 2002ab; 2003; Tamura et al 2000). The genetic variation of these alleles, particularly in the intron regions can be targeted to track inheritance and hence hybridity in a breeding population using allele specific PCR primers. In a fully compatible cross (no *S*-alleles in common between the maternal and paternal tissues) segregation of *S*-alleles will approximate to a 1:1:1:1 ratio, with one *S*-allele inherited from the female parent and one from the male (Figure 5.1a). In a semi-compatible cross (one *S*-allele in common) the ratio will be close to a 1:1 segregation (Figure 5.1b).

1a) *S7S8* x *S1S3* = *S1S7* : *S3S7* : *S1S8* : *S3S8*

1 : 1 : 1 : 1

1b) *S7S8* x *S3S7* = *S3S7* : *S3S8*

1 : 1

Figures 5.1a and 5.1b Inheritance and segregation of theoretical *S*-allele types in an independently assorting population between a fully compatible cross (1a), and a semi-compatible cross (1b).

Randomly amplified polymorphic DNA (RAPD) markers have become a routine tool for genetic fingerprinting in a number of plants including almond (Bartolozzi et al. 1998; Woolley et al. 2000), plum (Boonprakob et al. 2001), *Alstromeria* (De Benedetti et al. 2000), *Eucalyptus* (Keil and Griffin 1994; Neaylon et al. 2001), walnut (Nicese et al. 1998), and olive (Mekuria et al. 1999; Wu et al 2004). Inter-simple sequence repeats (ISSR) are a relatively new technique to be adopted for this type of analysis, although

there are reports for genetic identification of rice (Blair et al. 1999), chestnut (Goulão et al. 2001), pear (Monte-Corvo et al. 2001), and walnut (Potter et al. 2002).

RAPD and ISSR primers are versatile tools for DNA fingerprinting analysis since they are non-specific in their binding sites and display a random binding pattern (Williams et al. 1990; Zietkiewicz et al. 1994). Amplification of DNA sequences using these primers is extremely sensitive since even small variations in the primer binding sequence, such as a point mutation or deletion, can change the primer-target site and hence the amplification products. When a marker is detected in one parent and not the other, it is termed polymorphic, or monomorphic if a band is present in all individuals, both these types of bands can be used for genetic fingerprinting analysis. The reasoning behind using these molecular techniques for this type of analysis concentrates on the fact that closely related individuals have accumulated few genetic changes, whereas distantly related individuals are more genetically diverse with respect to polymorphisms (Nadeau and Sankoff 1998). DNA markers can be used to measure the level of genetic similarity between individuals within artificial or natural breeding populations by cluster analysis. RAPDs and ISSRs are reported to be more suitable for large populations of closely related individuals compared to RFLPs and SSRs (Cipriani et al. 1994; Williams et al. 1993), because the latter techniques show a lower level of polymorphism (Nicese et al. 1998), have a relatively long and intensive development period (Beckmann and Soller 1983) and require specific primer design (Lin et al. 1996; Rafalski and Tingey 1993).

The main objective of this study was to determine the genetic similarities of the 'Nonpareil' x 'Lauranne' F_1 hybrid progeny in comparison to the two parents as a method for verifying hybridity of the mapping population using *S*-allele identification and RAPD and ISSR based cluster analysis.

MATERIALS AND METHODS

Plant Material and DNA Extraction

DNA was extracted using a technique modified from Lamboy and Alpha (1998) (outlined in Chapter 3), from both parents and the total hybrid population of 171 putative hybrid progeny.

S-Allele Identification

PCR Primers and Conditions

Allele specific primers were designed from the published intron sequences of the *S3 (*Ma and Oliveira 1999) and *S7*, *S8*, *Sf* alleles (Channuntapipat et al. 2001) (Table 5.1). The intron sequences were obtained using primers designed from the conserved DNA sequences published by Ushijima (1998) (*S7* and *S8*) and Ma and Olivera (1999) (*Sf*). Each primer pair listed was used for allele specific PCR amplification of genomic DNA extracted from both parents and the whole 171 F_1 progeny in 2001. The annealing temperatures for each primer pair for *S*-allele analysis were individually optimised by modifying the temperature profile to limit non-specific binding of the primers (Table 5.1).

PCR reactions were performed in either 10 or 20 µL volumes containing 40 ng of genomic DNA, 1.25 µM $MgCl_2$, 0.25 µM each of the forward and reverse primers, 200 µM each of dNTPs, 1 x cresol red dye [2 % sucrose w/v, 0.1 mM cresol red, sodium salt (Aldrich Chemicals, USA)], 1 x Taq polymerase buffer (20 mM Tris-HCl, pH 8.4, 50 mM KCl, Invitrogen, Australia) and 1.1 U of Taq DNA polymerase (Invitrogen, Australia) made up to the appropriate volume with sterile MQ water.

Table 5.1 Primer sequences, PCR product sizes and annealing temperatures of the four *S*-allele specific primer pairs used for *S*-allele identification of the F_1 almond progeny, 'Nonpareil' x 'Lauranne'.

S-allele	Primer	Nucleotide Sequence (5' → 3')	PCR product (bp)	Annealing Temperature (°C)
S3	S3F	CAC TTA CTC TTT AGC ACT AGC C	529	60
	S3R	CGA ATA ACA TTG TTT CTT TCC		
S7	S7F	ACC ATA TAA CAT CGT GTT GC	438	60
	S7R	GAG GAT AAT ATG GTA CAT TC		
S8	S8F	CAA ATG GTC CTT CAG GTT TTC	650	65
	S8R	CCC AAA TCG CAG ACT CAC TCT		
Sf	SfF	GTG CCC TAT CTA ATT TGT TGA C	459	60
	SfR	GAC ATT TTT TTA GAA AGA GTG		

Amplification was performed in a PTC-100 programmable thermocycler (Programmable Thermal Controller, MJ Research Inc., USA) fitted with a hot bonnet holding a 96 well plate. Cycling parameters consisted of an initial denaturation step of 95°C for 30 sec, then 34 cycles of 95°C for 30 sec, 45 sec at the optimum annealing temperature (Table 5.1), 72°C for 1 min, and a final extension step of 72°C for 10 min (Channuntapipat et al. 2001). PCR products were resolved on 1.5 % agarose gels, in 1x TBE buffer [89 mmol/L Tris-HCl, 89 mmol/L boric acid, 2 mmol/L EDTA, pH 8.0] at a constant voltage of 180 V and current of 160 mA for 90 minutes. Band identification was determined by comparison to a 100 bp ladder (Geneworks, Australia) loaded concomitantly in lanes on the gels.

Reproducibility of the *S*-alleles was confirmed by triplicate PCR reactions, using different reagents on 10 individuals selected randomly from the population, using the conditions described above. The identification of *S*-alleles was repeated in 2002 to check reproducibility from re-extracted tissue using DNA extracted from a subset of eight individuals selected at random from the progeny. The extraction of DNA, PCR amplification and electrophoresis were as described above.

RAPD and ISSR Cluster Analysis

Polymorphism Generation

Six RAPD primers (OPA-02, OPA-08, OPB-10, OPC-07, OPC-10 and OPC-11, Operon Technologies, USA, Appendix A) and ten ISSR primers ($(AC)_8G$, $(GA)_8T$, $(AG)_8G$, $(CA)_8T$, $(CA)_8G$, $(CT)_8G$, $(TAA)_8$, $(AG)_8YA$, $(AG)_8YT$, $GCA(CA)_9$ Proligo, Australia) were selected based on amplification efficiency from the primer survey described in Chapter 5 and screened across the parents and 93 of the F_1 progeny. PCR and the separation of amplification products were carried out according to the

procedures described in Chapter 4. Gels were stained with ethidium bromide and visualised under UV illumination. Band size was determined by comparison to the 100 base pair ladder (Geneworks, Australia) loaded into lanes on either side of the samples for each gel.

Duplicate and if necessary triplicate PCR reactions were performed to ensure reproducibility in the results, only bright reproducible bands were used for analysis.

Data Analysis

The PCR products were visually scored as 1 for presence or 0 for absence, and the resulting binary matrix consisted of a total of 75 markers (43 polymorphic and 32 monomorphic) from a total of 6 RAPD and 10 ISSR primers. A similarity matrix was produced using the NTSys-pc (Numerical Taxonomy and Multivariate Analysis System, Version 2.2j, Exeter Software, Setauket, N.Y.) developed by Rohlf (1998), based on the number of shared amplification products using the simple matching coefficient (SMC). The Sequential, Agglomerative, Hierarchical and Nested (SAHN) algorithm which clusters individuals with the pair-group method, and can assign individuals only to non-overlapping bifurcations, was used in tandem with the Unweighted Pair-Group Method Average (UPGMA) to compute the average genetic similarity or dissimilarity of a candidate individual, with equal weighting between individuals in the cluster (Sneath and Sokal 1973). These calculations were then used to display a tree matrix in the form of a dendrogram. Principal component analysis (PCA) was also performed on the data to identify the most important variables from the set. In order to test the efficiency of the graphical representation the cophenetic correlation (CPC) was calculated. The CPC is a measure of the goodness of fit of the similarity implied by the dendrogram with the

original similarity matrix, a high CPC indicates that the pictorial depiction is a close representation of the original similarity matrix.

RESULTS

S-allele Analysis

For each primer combination, the genotypes of the two parents were used as either a positive control, where a band was expected based on the known phenotype, or a negative control, where no amplification was expected. The PCR amplification products for the *S*-alleles of *S3, S7, S8*, and *Sf* produced a discrete band of the expected size (based on the known intron sequence) of 529 bp, 438 bp, 650 bp, and 459 bp, respectively, as shown in Figure 5.2.

All of the progeny produced a single PCR amplification product, using two of the four primer pairs tested, that corresponded to the *S*-allele inherited from each parent (Figure 5.2). Thus, four different recombinant phenotypes segregated in the progeny, *S3S7, S3S8, S7Sf,* and *S8Sf* consistent with a hybrid population of the parentage 'Nonpareil' (*S7S8*) and 'Lauranne' (*S3Sf*). The segregation corresponded to a 1:1:1:1 ratio ($\chi^2 = 2.17$, 3 df), with the numbers in each class shown in Table 5.2.

In three different PCR reactions the *S*-alleles were the same, illustrating the reproducibility of this technique for this type of analysis. All eight of the individuals that were selected randomly in 2002 for confirmation gave the same PCR products as for 2001, which corresponded to the alleles, identified for each individual in the initial *S*-allele survey.

RAPD and ISSR Analysis

Six RAPD and ten ISSR primers were used to amplify DNA from the parents and 93 of the F_1 hybrid progeny, producing a total of 75 bands for cluster analysis. The two parents showed a genetic similarity of 72%, genetic similarity values for the F_1 progeny ranged from 70% to 93% (Figure 5.3). The dendrogram showed three clusters at the 70% level of genetic similarity. Cluster A included those individuals most similar to the male parent 'Lauranne' (LA), cluster B contained individuals most similar to the female parent 'Nonpareil' (NP). Cluster C included eight individuals that did not cluster with either of the parents. 91% of the F_1 individuals clustered between the rankings of the two parents. A further test using PCA did not provide a clearer representation of the data with respect to grouping of the outliers.

DISCUSSION

Verification of hybridity in populations used for applications such as mapping is important to limit errors associated with the inclusion of non-hybrid data, which has the potential to skew segregation data and recombination values.

Both fruit set and pollen tube growth in the style after individuals have been pollinated with self or cross pollen have been used as classical measures of SI in woody perennial species. The long juvenile stage of these plants limits the usefulness of these measures with respect to time constraints and influence by environmental factors. A molecular method based on PCR was used to test for the SI groups in the F_1 hybrid population using primers designed from the unique intron sequence of each *S*-allele. Both parents of known *S*-allele type were subject to PCR analysis with these primers, and all four *S*-alleles could be successfully distinguished from each other. These results were in

Table 5.2 Number of F_1 almond seedlings segregating for four recombinant *S*-allele genotypes from 'Nonpareil' (*S7S8*) x 'Lauranne' (*S3Sf*) cross.

S-allele	*S7S3*	*S8S3*	*S7Sf*	*S8Sf*
N	40	41	51	39
%	23.4	23.9	29.8	22.8
	mean = 42.8		Total n = 171	
(χ^2 = 2.17, 3 df)				

Table 5.3 *S*-allele complement of the eight outlying individuals from Figure 5.3 based on cluster analysis

Individual	*S*-allele	
	Female derived	Male derived
1bT31	*S8*	*S3*
1bT33	*S7*	*Sf*
3aT10	*S7*	*Sf*
3aT22	*S7*	*S3*
3aT37	*S7*	*S3*
3bT20	*S8*	*S3*
3bT25	*S7*	*S3*
3bT44	*S8*	*S3*

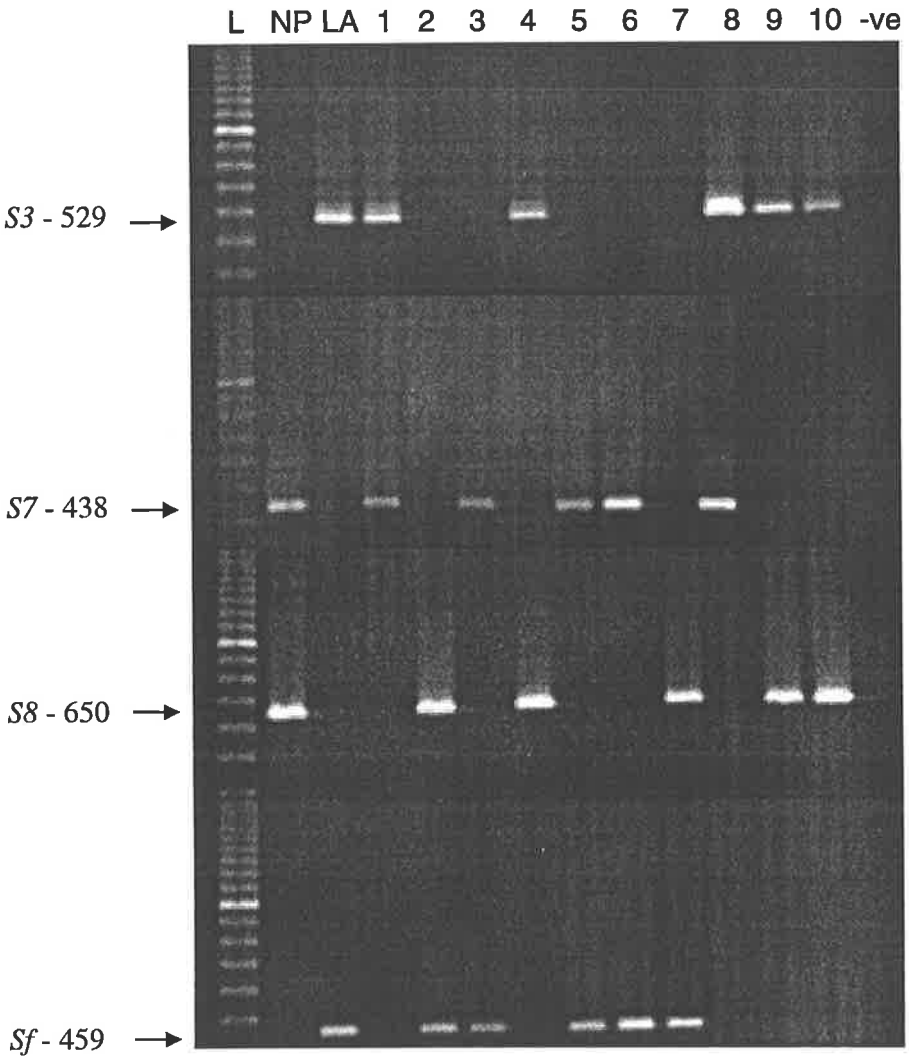

Figure 5.2 *S*-allele inheritance of parents 'Nonpareil' (NP) and 'Lauranne' (LA) and 10 F_1 hybrids, lanes 1 – 10, L = 100bp ladder, -ve = negative control. Arrows indicate *S*-allele and bp size

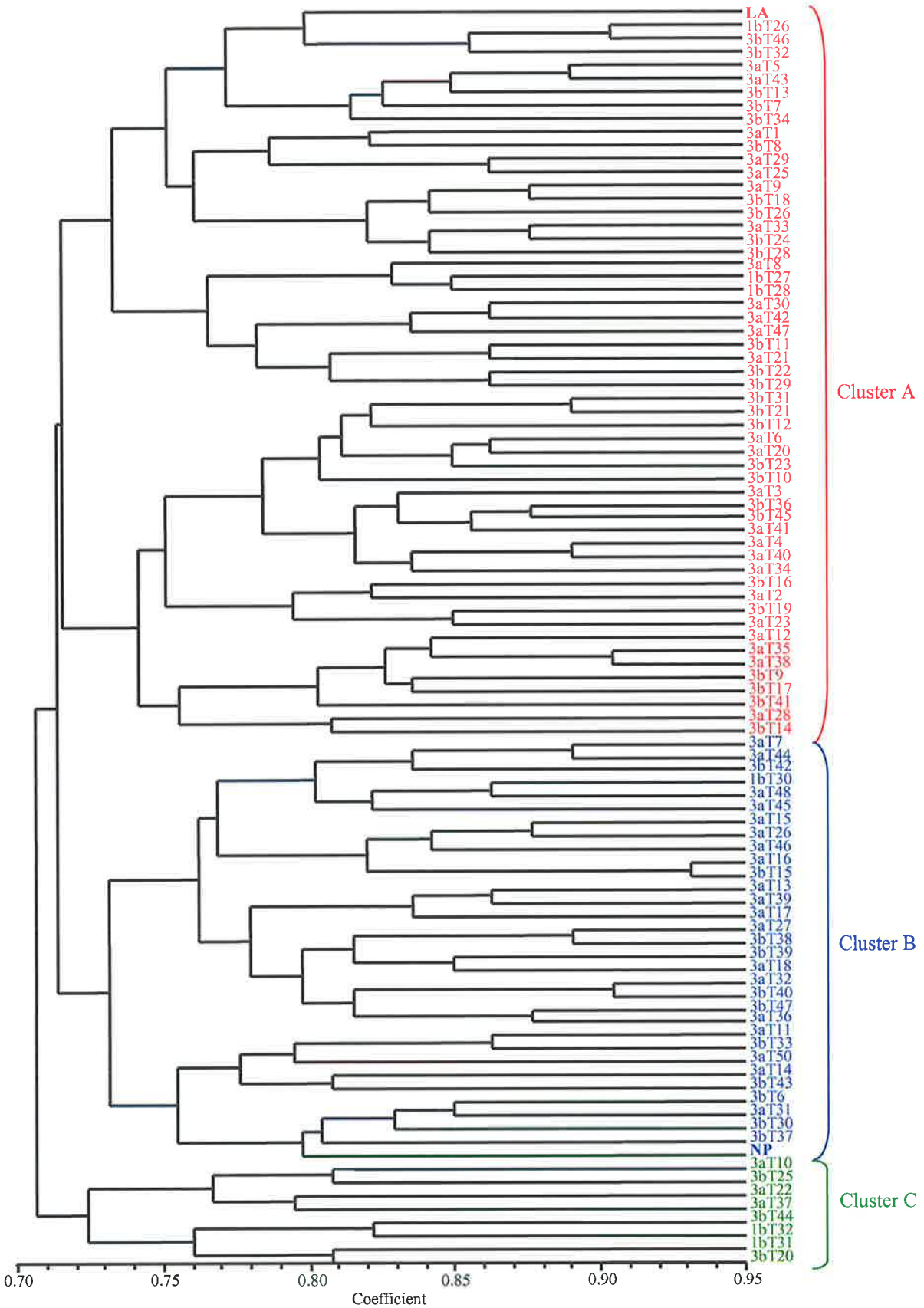

Figure 5.3 Dendrogram of 93 F_1 hybrid progeny and the two parents 'Nonpareil' (NP) and 'Lauranne' (LA), generated by UPMGA cluster analysis using the simple matching coefficient. Cluster A (red) includes hybrids most similar to the male parent LA, cluster B (blue) those with similarity to the female parent NP, and cluster C (green) those hybrids grouped as outliers.

accordance with the published data on the *S*-alleles of these cultivars (Bŏsković et al. 1997; Channuntapipat 2002)

The method described here for the identification of *S*-alleles provides a more direct and selective technique than traditional techniques, and requires only a small amount of DNA from vegetative tissue for analysis. This circumvents the need for the trees in the breeding program to become sexually mature before testing for *S*-allele groups can be carried out. A similar approach using PCR in combination with restriction enzymatic digestion to identify *S*-allele types has also been used in almond (Ushijima et al. 1998) and other Rosaceous species including apple (Janssens et al. 1995), cherry (Tao et al. 1999; Wiersma et al. 2001), sour cherry (Yamane et al. 2001), Japanese apricot (Tao et al. 2000) and Japanese pear (Ishimizu et al. 1999). The procedure described here does not rely on digestion by restriction endonucleases, since the use of allele specific primers amplifying a fragment of a unique and known size does not require this additional step for precise identification, further enhancing its usefulness as a more efficient technique for *S*-allele identification in almond. The accurate reproducibility of this technique also supports its application as a veritable method for *S*-allele identification in this species.

In addition to identifying the *S*-allele genotypes of the population, this technique also provides a novel approach for the testing of true hybridity amongst the progeny since a particular *S*-allele can be inherited only from one or the other parent. If the parental allele is absent in the progeny it can be assumed that it is not a true offspring derived from the original parental cross.

The work presented here for the *S*-allele identification of the total F_1 progeny of 171 individuals derived from a cross between two horticulturally and economically important almond cultivars 'Nonpareil' and 'Lauranne' indicates independent

segregation of the recombinant genotypes in a 1:1:1:1 ratio, suggesting monofactoral control. The segregation of the *Sf* allele in the F_1 population was consistent with previous studies since it is inherited independently with a frequency of 50% in the progeny (Dicenta and García 1993; Ortega and Dicenta 2003).

On the basis of cluster analysis using 2-D graphical representation eight individuals did not align themselves between the two parents and hence with the majority of the progeny. The existence of these eight individuals as apparent outliers from the main grouping could be accounted for by a number of factors. Firstly the individuals may not be true hybrids based on the segregation patterns of the selected RAPD and ISSR DNA markers. Secondly the number of markers used may not have been sufficient to group these few outlying individuals with the main group, and the type of markers chosen for analysis may have affected the results. RAPDs have been reported to produce non-parental derived bands by formation of heteroduplex molecules (Ayliffe et al. 1994), which may have skewed the scoring, and assignment of some individuals. Hackett et al. (2000) showed that bands present in the progeny but absent in the parents was linked to the introduction of foreign pollen based on cluster analysis. However the *S*-allele complement of the eight outliers in this dendrogram all showed the correct hypothesised parental inheritance and segregation (Table 5.3). Gu et al. (1999) have illustrated RAPD bands often contain multiple DNA sequences and therefore the presence of two bands at an identical position on an agarose gel may be misinterpreted as a single product resulting in scoring error. Scoring error is estimated to be responsible for approximately 2% of variation in genetic distance calculations (Skroch and Nienhuis 1995). Although this may account for some variation, the number of outliers from the main group is higher than 2%. Therefore thirdly and the most plausible reason for this distortion is that the use of any clustering method imposes

certain constraints on the relationship between pair-wise groupings, and choice of clustering algorithm directly affects the result. In this case the SAHN algorithm is limited to producing only bifurcations, and one individual may not become a member of another partition, which may result in some prejudice between groupings. It has also been noted that the order of the individuals in the original data matrix does have some affect on an individual's position within the dendrogram after similarity calculations (data not shown). The CPC was calculated as 0.5, which corresponds to a very poor fit to the similarity matrix, the poor fit of this data may be due to error in recording character states from the original gel pictures and the limitations of the clustering methods selected. Since this value is extremely low it means individual clusters cannot be distinguished from random clustering events, this was expected since all the individuals are full-sibs, and therefore it is not expected that one group be more closely related than another.

In conclusion, all of the 171 F_1 progeny DNA was extracted from, including the 93 individuals used as the mapping population (Chapter 8) showed an *S*-allele complement which is believed to be derived from hybridisation of the two parents 'Nonpareil' and 'Lauranne'. The construction of a similarity matrix based on RAPD and ISSR marker data analysed with cluster analysis methods supports this since the majority of putative F_1 progeny were placed between the two parents when depicted in a 2-D form. This strongly suggests that these individuals are true hybrids of the genetic cross between the female and male parent, 'Nonpareil' and 'Lauranne' respectively.

CHAPTER 6

Morphological Analysis of the Full-Sib F_1 Hybrid Population

INTRODUCTION

Almond is an economically important nut crop and is ranked second in world production behind walnut (FAO 2004). The dehiscence of the pericarp is the major diagnostic trait separating almond from the rest of the *Prunus* species and its commercial delineation as a nut rather than a fruit. It is the only member of the stone-fruit where the kernel is the most valued part of the crop, in the rest of the *Prunus* species, a fleshy mesocarp develops to form the fruit and is not dehiscent. Therefore, unlike other *Prunus* species such as peach, apricot, plum, and cherry where improvement of the mesocarp through breeding programs is of prime importance, the target of almond breeding is the improvement of the kernel. Numerous almond cultivars are grown worldwide which differ widely in a number of important vegetative and reproductive traits which are often affected by genetic and environmental conditions (Kester and Gradziel 1996).

In numerous groups worldwide, research has concentrated on understanding the genetic basis and inheritance of important morphological traits in almond. Perhaps the most extensively studied and economically important characteristic that breeding programs have focused on has been the control of flowering time and other associated phenological traits (Asíns et al. 1994; Kester 1965a; Vargas and Romero 2001). The reason for this intense focus is because frosts, especially at bloom time, have the capacity to greatly reduce productivity and yield by arresting pollen tube growth and destroying pistils, consequently reducing fertilisation and subsequent fruit set (Socias i Company 2001). Inclement weather conditions also restrict the activity and foraging of honeybees, the chief pollinating vector in almond orchards (Jackson 1996). Bloom date varies from season to season, in response to various climatic cues (Dicenta et al. 1993) and is considered a qualitative trait (Socias i Company 1999; Vargas and Romero 2001). The qualitative variation may be due to modifier genes and the differences in both chilling and heat requirements for each cultivar (Rattigan and Hill 1986; Samish 1954). Studies on the genetic control of bloom time in almond have proposed a single gene *Lb*, regulates this trait (Ballester et al. 2001; Socias i Company et al. 1999ab). The incorporation of late flowering into breeding programs is especially important for reducing crop losses in areas where frost is a risk, such as inland regions of the Mediterranean (Socias i Company et al. 1999b).

Improvement of yield on an individual tree basis and orchard level is another important trait, particularly for growers of this crop. Analysis of the yield of a particular cultivar or individual tree is highly subjective based on the methodology adopted for its analysis, i.e. total weight of nuts or kernels per tree; number of fruits per tree; or kg/tree or kg/ha of saleable kernels (Kester and Gradziel 1996; Quinn 1928). Although yield itself is

skewed by the method used for its calculation the parameters affecting the final value are numerous. Initially the age of the tree has a significant influence on the productivity of the tree, since almond trees do not reach bearing age until about three years of age and will not reach full production until approximately five years (Woodroof 1979). Productivity declines generally after twenty years of commercial production under standard orchard conditions, where most trees of this age are removed and replanted with juvenile trees (T. Spiers, pers. comm. 2001).

After age effects, environmental factors are the most important affecting yield. Several physical factors have been suggested as being involved either negatively or positively in influencing yield. At a gross level, the total number of flowers determines the number of nuts per tree, which is concomitantly determined by flower initiation the previous season. Flower initiation and development are heavily influenced by heat and chilling requirements, water and nutrient availability, tree architecture or bearing habit, and also orchard planting design, particularly tree spacing. These factors are also important for controlling nut development. Pollination is essentially the most important factor after flowering for fruit development, since almonds are unable to bear fruit parthenocarpically and therefore each flower needs to be cross-pollinated entomophilously (in the case of self-incompatible cultivars) which is usually carried out by honeybees (*Apis mellifera*). Weather conditions, and distance between trees and the hive influence honeybee pollination efficiency (Jackson 1996). Alternate bearing also influences fluctuations in yield of individual trees between growing seasons (Hill et al. 1987). Alternate bearing is positively influenced by irrigation depravation in previous years (Klein et al. 2001).

At the biological level, each of the parameters affecting yield, such as heat and chilling requirement, self-incompatibility, flower initiation and development, tree architecture and alternate bearing have a genetic component. The interaction between these genetic components with abiotic conditions predisposes this trait to a complex mode of inheritance heavily influenced by temporal and seasonal conditions.

Some studies have centered on the control and development of fruit and kernel characteristics. Almond cultivars vary widely in a number of traits relating to kernel and shell attributes, including, size, shape, percentage doubles, kernel taste, pubescence, and colour, shell hardness, markings and suture line as well as disease and pest resistance.

The occurrence of two kernels within the one endocarp, each possessing a fully intact seed coat is referred to as double kernels, or doubles (Figure 6.1). Double kernels are found in all stone fruits (Asensio and Socias i Company 1996) and are produced when both ovules present within the flower are fertilised (Spiegel-Roy and Kochba 1974), rather than one of the ovules degenerating (Egea and Burgos 2000; Pimienta and Polito 1982). Doubles are not ideal for the industry since they are often grossly misshapen, and hence not commercially desirable. Therefore factors affecting the development and fertilisation of ovules and subsequent fruit development are of economic significance to the industry.

The predisposition to double production does appear to be cultivar dependant since some cultivars such as Nonpareil show no doubling (Spiegel-Roy and Kochba 1974) whereas cultivars such as Romana can show up to 48% doubles dependant on the season (Barbera et al. 1988). It is however this seasonal and locational variation that has led to the assumption of an environmental influence on the expression of this character. Studies on the heritability of this trait have all shown a high degree of variability from season to

season (Asíns et al. 1994; Dicenta et al. 1993; Kester et al. 1977; Spiegel-Roy and Kochba 1974, 1981).

Figure 6.1 Double kernel production in almond, double kernels on right are deformed and misshapen compared to normal nut on left. Scale bar = 1 cm

Shell hardness has been identified as being controlled by a single gene with two alleles, *D/d* (Grasselly 1972). The lignification of the endocarp and the concomitant ratio of shell weight to kernel weight is used to determine shell hardness (Rugini 1986). Most European cultivars have a shelling percentage of over 35%, which is classified as a hard shell (e.g. 'Ferragnès'), while American varieties tend towards values of below 35% classified as a soft or paper shell (e.g. 'Nonpareil') (Kester et al. 1990). The Australian almond industry has been influenced by both growing regions and hence has cultivars with a range of shelling percentages (e.g. 'Chellaston', soft shell; 'Pethick Wonder', hard shell) (Quinn 1941). Paper and soft shelled varieties have a tendency to have a weak shell seal, resulting in a predisposition to infestation by insect pests such as Navel Orangeworm, *Paramyelois transitella* (Soderstrom 1977) Leaffooted Bug, *Leptoglossus clypealis* (Gradziel and Martínez-Gómez 2002) and Carob Moth, *Ectomyelois ceratoniae* (Pound and Collins 2003). A mix of shell hardness types within the orchard poses problems for

mechanical crackout since the machines need to be set to optimise crackout percentage while maintaining the integrity of the kernel.

The saleable part of almond production is the kernel, and hence interest in factors affecting the appearance of the kernel is of great economic importance. Almonds vary greatly between cultivars in a number of nut characters such as size, shape, weight, colour, rugosity, pubescence and general appearance. All of these traits have been used in combination for cultivar identification at the morphological level and as targets for improvement by breeding programs. The appearance of the kernel and its subsequent consumer appeal is influenced by a number of factors, including the shape, colour, rugosity, pubescence and damage by external forces, mechanical, environmental or biological.

Individual kernel weight is determined by both total yield of the tree and nutrient assimilation in the last phase of kernel growth until maturation (Brooks 1939; Hawker and Buttrose 1980). It has been noticed in tree nut crops, such as almond (Hill et al. 1987), and pecan (Dodge 1946) that with an increase in yield (measured as the total number of fruit) there is a concomitant decrease in individual fruit size and weight. In almond, although weight does vary between individuals on the same tree and between the same cultivars in different locations, this trait shows a high heritability suggesting a strong genetic component (Chandrababu and Sharma 1999; Dicenta et al. 1993; Kester 1965b; Kester et al. 1977). However Kumar and Das (1996) have shown a xenic effect on nut size, dependant on the nut size of the pollen parent, in a positive or negative direction, also supporting a genetically transmissible factor.

Kernel shape is a function of the linear measurements of length, width, and thickness and can be categorised based on the length/width ratio or a visual observation of round to narrow (Gulcan 1985). Kester (1965b) has shown that even over a number of years the ratio of length to width of the kernel in a number of cultivars is constant even between different locations, supporting a genetic influence on the control of this trait.

Kernel colour, expressed as the colour of the testa or seed coat, is also important for the consumer and varies from very pale to dark. Although some cultivars tend towards one of the two extremes, colour can also be altered in part by the stage of maturity and also harvest time, length of storage and exposure to light (Kester et al. 1977).

The study of testa pubescence has been a slightly more neglected trait in the study of almond. It appears the level of pubescence may in part be influenced by the cellular structure of the outer epidermal layer of the testa. Some of these epidermic cells are projected from the surface (Pascual-Albero et al. 1998) and may possibly be a factor in this trait. Hawker and Buttrose (1980) has proposed the enlargement and ensuing collapse of these epidermal cells, is responsible for the formation of a pubescent or furry covering surrounding the embryo. Variation of pubescence is graded from low (e.g. 'Nonpareil') to extremely high (e.g. 'Chellaston').

The rugosity or wrinkling of the kernel and seed coat can be influenced biologically by the endosperm failing to develop (Kester et al. 1990). Environmental conditions, most importantly irrigation restrictions at either initial fruit development or in late summer when the kernels are undergoing the final development of the embryo also influence the level and severity of kernel rugosity (Baxter 1997). Failure of the kernels to develop normally can also be influenced by infection by phytoplasma, which results in the associated almond

shriveled kernel disease (Uyemoto et al. 1999). The presence of a discrete crease in one side of the kernel is distinct from the rugose character. This character shows a high heritability (Kester 1977) and is more prevalent in some cultivars (pers. obs.) suggesting a strong genetic predisposition to this abnormality. Kester (1977) proposed a single gene may be responsible for this trait based on a high heritability value of 0.79.

Taste is perhaps the second most important characteristic for the consumer behind appearance. The flavour of the kernel is influenced by the presence of cyanogenic glycosides, which are secondary metabolites derived from α-amino acids (Conn 1980; Vetter 2000). Overall the inheritance of kernel flavour in almond is proposed to be under monogenetic control with bitterness being due to the recessive allele (Dicenta and García 1993; Heppner 1923; Heppner 1926). Heppner (1923) first proposed the segregation of this trait as a 3:1 ratio of sweet to bitter. However after further investigation Dicenta and García (1993) proposed a more complex mode of inheritance that accounted for the presence of slightly bitter kernel types, as potentially heterozygotes with an altered expression of the dominance system, a conclusion subsequently supported by (Ledbetter and Pyntea 2000).

This study focuses on the study of the inheritance and pomological character variation with respect to genetic and environmental control of a hybrid F_1 population for three phenological, eight morphological and one molecular trait. Understanding the inheritance and genotype-by-environment interactions of these characters is important both for the industry and breeding programs, and also the application of morphological data for integration into molecular maps and subsequent marker-assisted selection and quantitative trait loci (QTL) analysis.

MATERIALS AND METHODS

In this study a number of morphological traits were measured for three growing seasons (2000 – 2003) on a total of 181 progeny of a F_1 hybrid cross. Weather data for the nearest meterological station (Renmark station #24048, South Australia) was collected as a monthly average for 2000 – 2003. Both the mean average rainfall and mean maximum and minimum air temperatures are presented in Appendix B.

Plant Material

The F_1 population was derived from a cross between two almond cultivars 'Nonpareil' x 'Lauranne', further details of this cross are presented in Chapter 2. The population were maintained under standard orchard conditions for irrigation and pest control and no pruning was performed.

Nut Harvest

Nuts were harvested when approximately 90% of the nuts on the tree were mature, indicated by drying and dehiscence of the husk (mesocarp). Moisture content of the kernels is approximately 5% or less at this stage.

In the first fruiting year 2001, the crop was relatively light so the total number of nuts from all fruiting trees were harvested by hand. In 2002 and 2003 the crop was sizeable and so only 30 – 40 fruit were randomly harvested from each tree.

After the nuts were harvested they were dehulled and stored in paper bags at room temperature (20 – 25 °C) until required for analysis. Nuts were cracked for kernel examination and shell hardness measurements using a manual nut cracker.

Morphological Traits

A random ten nut sample was selected for analysis for kernel measurements for each individual, and only well filled nuts were used. A digital caliper was used for all nut measurements (Mitutoyo, Japan) and a pan balance used for weight measurements.

Vegetative Traits

For the 2001 season, vegetative characters were scored after the tree had been harvested of fruit. Tree habit was scored on a scale from 1 (extremely upright) to 9 (weeping) based on the almond descriptors list (Resources 1985) presented in Table 6.1a. Tree height was scored from ground level (base of trunk) to the tallest branch emerging from the canopy. Bare branches and closely spaced internodes were scored subjectively as yes or no. Leaf size and colour were scored as small, medium or large and light, medium or dark green, respectively. Pest and disease presence was scored for the condition (generally rust, *Tranzschelia discolor* or shothole, *Wilsonomyces carpophilus*) where the disease affected more than 50% of the tree canopy area.

Bloom Time and Precocity

Bloom time was taken as the day 50% of the flowers on the tree were open, for data analysis the number of trees flowering on a given day were divided by the total number of flowering trees and expressed as a percentage of the total.

Precocity was determined by examining the first year of flowering for individual trees. The majority of trees flowered in 2001, trees that flowered in 1999 were scored as high. Those trees that didn't flower in 2001 were scored as low, all other trees were scored as average if they flowered in 2000 or 2001 for the first time.

Table 6.1a Descriptor list for nine vegetative traits scored over the three years on the F_1 progeny, * in AD indicates trait is listed in the almond descriptors list (Gulcan 1985)

Code	Character	Description	AD
BLO	Bloom time	1. Early	
		2. Early-Mid	
YLD	Yield/Cropping Efficiency	1. Low	*
		2. Intermediate	
		3. High	
PRC	Precocity	1. High	
		2. Low	
THA	Tree Habit	1. Extremely Upright	*
		2. Upright	
		3. Spreading	
		4. Drooping	
		5. Weeping	
BBR	Bare Branches	1. Yes	
		2. No	
THT	Tree Height		
LSZ	Leaf Size	1. Small	
		2. Medium	
		3. Large	
LCO	Leaf Colour	1. Light Green	
		2. Medium Green	
		3. Dark Green	
DIS	Disease Prevalence	1. Low Susceptibility	
		2. High Susceptibility	

Yield and Fecundity

In the 2001 season all the nuts from each individual tree were harvested and yield was calculated as the total inshell weight of all kernels from the tree. In subsequent years, a visual observation based on total fruit number and fruiting intensity was made and scored as high, medium or low.

Fecundity was determined by visual assessment of individual trees for all three years. Trees that showed a very high fruiting intensity at nut maturity compared to the overall population average were scored as high, those with very poor fruit set were scored as low. This was a similar measure to yield, but for this assessment a score of high or low for at least two of the three years was the individual final score for this trait.

Self-Compatibility

Self-compatibility was determined by allele-specific molecular analysis for the *Sf* allele as outlined in Chapter 5.

To verify this technique, both pollen tube growth and fruit set were observed for ten individuals for one year (2003) for pollen tube data and two years for fruit set (2002 and 2003).

Pollen Tube Analysis

Two branches with approximately 50 unopened flowers were excised from each of the ten individuals. These branches were placed in a 5% sucrose solution in the laboratory at 22°C, and 20-30 flowers at balloon stage were selected and emasculated, all other flowers were then removed from the branch. Emasculated flowers were hand pollinated when receptive (24 hrs after emasculation) with either self or fully compatible 'Somerton'

pollen applied with a small paint brush. The pollinated flowers were harvested from the branch 7 days after pollination, and placed in fixing solution. Flowers were fixed in Carnoys fluid (absolute ethanol: chloroform: acetic acid, at a ratio of 6:3:1) for 24 hrs. The flowers were then transferred through a 70 – 95% ethanol gradient and stored at 4°C until analysis.

Fixed samples were transferred through a rehydrating gradient of 50% ethanol to 30% ethanol to distilled water for 30 min in each solution. To soften the styles for stain uptake the samples were soaked in 0.8 M NaOH for six hours and then washed overnight in running water. The pistils were dissected from the flower, and the ovary removed. Both styles and ovary were stained with 0.1% aniline blue in 0.1 M potassium phosphate for 1 hour. Samples were mounted in 90% glycerol for observation. The total number of pollen grains both ungerminated and germinated and tubes in the stigma, upper style, lower style and at the ovary were determined by means of fluorescence microscopy using a Ziess Axiophot Photomicroscope (exciter filter 395-400, interface beam splitter FT 460, and barrier filter LP 470). Digital images were captured from an attached colour video camera (JVC, Japan) using the computer program AcQuis (Syncroscropy, USA). For each pistil, pollen tubes at each section of the style were expressed as a percentage of pollen grains on the stigma surface.

Fruit Set

Two individual branches with approximately 100 flowers were selected at bud-burst stage for each of the ten individuals, each branch was then enclosed in an insect proof net. One to two weeks later, between 50 – 150 flowers at balloon stage on each branch were

hand pollinated using a small paint brush with either self-pollen (collected from open flowers on the day, or from flowers opened the week before) or fully compatible pollen collected from 'Somerton'. All other flowers that were not pollinated were removed from the branches, and the nets were left on the branches until the nuts were fully mature.

The number of developing nuts was evaluated at 4 and 12 weeks for 2002, and 12 weeks and maturity for 2003.

Percentage Doubles

The percentage of double kernels for each individual was calculated as the total number of nuts containing two kernels, divided by the total number of nuts collected and expressed as a percentage.

Kernel Weight and Shell Hardness

The average inshell weight and kernel weight was obtained for ten randomly selected nuts (not including obvious doubles), the shell (endocarp) was then removed and the average weight from the ten kernels calculated. Shell hardness is calculated as the percentage of the inshell kernel proportion of the total nut weight by the following equation (Rugini 1986), and sorted into classes based on the percentage values (Table 6.1b)

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

Kernel Size and Shape

Kernel length was measured longitudinally from the basal end to the kernel apex, kernel width was measured as the widest point across the kernel face for ten kernels. The lateral dimension between the thickest point of the kernel for ten kernels was taken as

kernel thickness. Kernel shape was calculated by the kernel width/length ratio of the mean of ten nuts, classes are represented in Table 6.1b.

Testa Characteristics

Both testa colour and testa pubescence were scored subjectively, using both parents as comparisons. Testa colour was scored as light 'Nonpareil' or dark 'Lauranne' for years 2002 and 2003, in 2001 testa colour was scored using RHS colour charts (RHS 1966) and then grouped into one of these two categories. Testa pubescence or roughness of the seed coat was scored as pubescent or smooth, according to mouth-feel. Scoring consistency was determined by a sensory evaluation panel, this panel consisted of three trained individuals and is outlined in Chapter 5.

Crease

Crease was scored as the presence or absence of a distinct indentation on one face of the kernel in more than 10% of the sample.

Rugosity

Rugosity or shrivelling of the testa and kernel was assessed subjectively based on the severity and percentage of nuts in the sample affected, given as an approximate severity level, high or low.

Kernel Taste

Kernel flavour was determined by tasting an average of two kernels from each individual for each growing season. Taste was scored subjectively as sweet or semi-bitter (marzipan flavour). For all years a subset of hybrids, representing approximately 10%

Table 6.1b 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 (Gulcan 1985). KSH calculated according to the width/length ratio.

Code	Character	Description	AD
DBL	Percentage Doubles	1. Low (0 - 5%)	*
		2. High (> 5%)	
KWE	Kernel Weight/Size	1. Extremely Small (< 0.9g)	*
		2. Small (1.0 - 1.1g)	
		3. Medium (1.2 - 1.4g)	
		4. Large (1.5 - 1.8g)	
		5. Very Large (> 1.9g)	
SHH	Shell Hardness	1. Soft (> 55%)	*
		2. Paper (56 - 45%)	
		3. Semihard (44 - 35%)	
		4. Hard (34 - 25%)	
		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. Semi-bitter	
TPU	Testa Pubescence	1. Smooth	*
		2. Pubescent	
CRE	Kernel Crease	1. Yes	
		2. No	
TRU	Testa Rugosity	1. High	*
		2. Low	

of the population, were submitted to a tasting panel as described for kernel pubescence to determine consistency for the scoring of this trait.

Statistical Analysis

To determine the differences between averages for fruit set χ^2 analysis was performed. To test for significance between treatments and individuals for the pollen tube data, mean values were analysed using the t-test (two sample assuming unequal variance).

For vegetative, bloom time, and nut data, significant segregation ratios were determined by χ^2 analysis.

Broad sense heritability was calculated by variance according to the following equation:

$$h^2 = \sigma^2 g/(\sigma^2 g + \sigma^2 e)$$

Where $\sigma^2 g$ = genetic variance and $\sigma^2 e$ = environmental variance. Genetic and environmental variances were estimated using REML estimation of a Mixed Model fitted in S-Plus (Insightful Corp 2002), and SAMM (Butler et al. 2000). For the calculation of correlations between traits, the following formula was used (Anderson et al. 1994):

$$r_{xy} = \frac{S_{xy}}{S_x S_y}$$

Where $S_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}$, $S_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$ and $S_y = \sqrt{\frac{\sum (y_i - \bar{y})^2}{n-1}}$

RESULTS

Vegetative Traits

Tree height ranged from 2 m to 5 m, with the average 3.81 m ± 0.45. Trees showed a range of tree habits from 1 to 5. The majority of trees were spreading, 42.6%, followed by drooping, 31.1%, upright 20.2%, weeping, 3.3% and extremely upright 2.7%. Leaf size was predominantly medium 70.8% with small and large leaves present in 14.6% of trees. The bulk of trees had medium green leaf colour, 92.3%, while only 7.2% had light green leaves and 0.5% dark green leaves. Only 2.2% of trees showed close internodes, and approximately one third of trees had bare branches. In 2001 disease was prevalent in the orchard, 83.1% of trees showed symptoms of rust, and 16.9% of trees showed symptoms of both rust and shothole.

Bloom Time

The flowering period spanned 20 days in 2000 and 19 days in 2002 (Figure 6.2). In 2000, a bimodal distribution of flowering was observed with over half the trees flowering on two days, 8 days apart. In 2002 a similar pattern was observed, but the interval between the two groups was only 3 days. In 2002 half the trees reached full bloom a week before the majority of trees in 2000, resulting in an overall earlier bloom time for this season.

Precocity

Only 13 trees flowered for the first time in 1999, representing only 7% of the total population, and in 2000 a total of 131 trees flowered (72.3%). In 2001, the year taken as first fruiting year for the population, 44 trees failed to flower (24.3%) and these were scored low for precocity. Of the trees that first flowered in 1999, 61.5% were self-fertile,

for the other two years there was no significant correlation with self-fertility and high or low precocity.

Fecundity

To measure fecundity, yield measurements were taken over three years and correlated to the presence of self-fertility. For trees that scored high yields for two out of three years, 46% had the *Sf* allele, while for trees with a low overall yield, 46% also had the *Sf* allele indicating no correlation between self-fertility and high yield

Fruit Set

Fruit set was scored at 4 and 12 weeks after pollination and at maturity. Only a small difference was observed in the percentage of fruit set between the first and final observation in both years (Table 6.2). Comparing the first and final values for fruit set for all treatments, a greater decrease was observed in the percentage fruit set levels between week 4 to week 12 in 2002 than week 12 to maturity in 2003, where only one of the four average values changed by 1%. It is likely based on these data that the majority of fruit abortion occurs before week 12. A significant difference was observed between the percentage fruit set of self-compatible (SC) individuals and self-incompatible (SI) individuals pollinated with both self and fully compatible pollen. The lowest fruit set was observed in SI individuals pollinated with self-pollen which was 4.1% and 2.4% at 12 weeks for 2002 and 2003, respectively. At 12 weeks, close to 50% of SC flowers pollinated with self-pollen produced fruit, this value was slightly lower for 2002 but was still the highest fruit set observed for this time period. In 2002 the percentages for all fruit set were

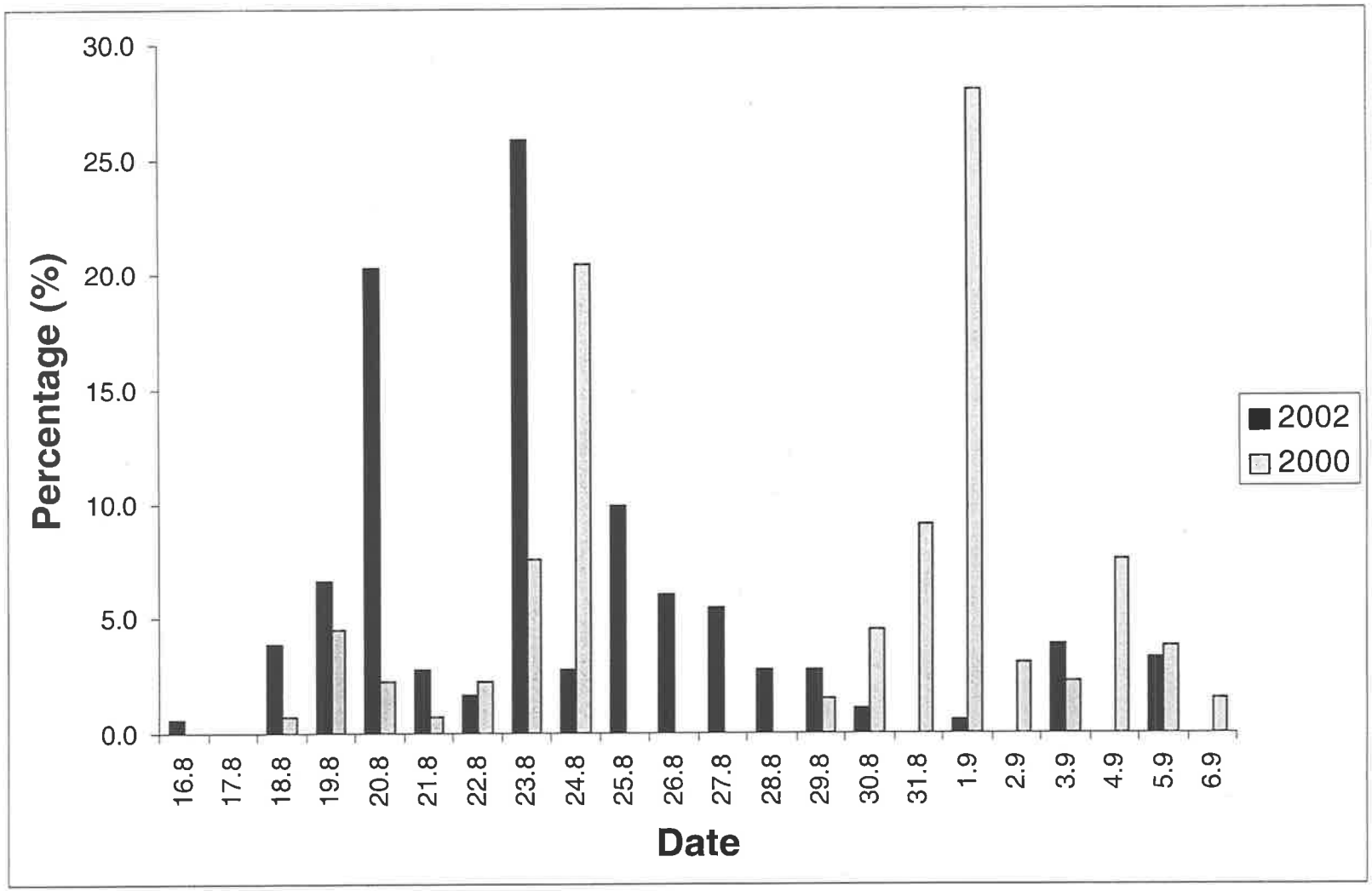

Figure 6.2 Percentage of trees in full bloom between 16th August and 6th September for 2000 and 2002 season

lower than the comparative groups in 2003 at week 12, except for selfed SI individuals. SC individuals had a higher fruit set when pollinated with fully-compatible pollen than SI individuals, but this difference was greatly magnified when SC individuals were pollinated with self-pollen, compared to SI individuals, which showed an approximate 10 to 20 fold decrease in fruit set in comparison.

Pollen Tube Analysis

Figure 6.3 shows the average percentage of pollen tubes reaching each section of the pistil, for each of the ten individuals analysed. No obvious trends were evident except for an overall decrease in the percentage of pollen tubes reaching the upper and lower portions of the style, with the lowest levels of pollen tubes observed at the ovule. In most cases no significant difference was observed between self- and cross-pollination, or between the different pollen applied to SC or SI individuals. The only significant difference observed was in the percentage of pollen tubes reaching the ovule in cross-pollination was three times higher than in self-pollination. A significant difference was also observed between the percentage of pollen tubes reaching the ovule in SI and SC individuals pollinated with self-pollen, which was six times higher in the SC individuals.

Percentage Doubles

The percentage of doubles broken into percentage categories decreased exponentially for all the three years (Figure 6.4). For all years approximately 60% of individuals showed 0% doubles, with the next largest group containing approximately a quarter of all individuals in the 0.1 – 10% doubles range.

Table 6.2 F_1 hybrid fruit set data collected from 2002 and 2003 at 4, 12 weeks and maturity, for self-pollination and fully compatible cross (*cv.* 'Somerton'). For self-compatible hybrid individuals (1aT26, 1aT30, 1bT32, 1bT51 and 1bT47) and self-incompatible hybrid individuals (1aT4, 1aT9, 1aT40, 1bT31, and 1bT42).

Fruit set experiment on 'Nonpareil' x 'Lauranne' trees 2002 - 4 weeks

Tree	*S*-alleles	Self cross @ 4 weeks: No. flowers for fruit set	No. fruit set	% fruit set (S.E.)	Compatible cross (Somerton S_1S_{21}) @ 4 weeks: No. flowers for fruit set	No fruit set	% fruit set (S.E.)
1aT26	S_7S_t	28	21	75	60	45	75
1aT30	S_8S_t	154	74	48	64	48	75
1bT32	S_8S_t	66	57	86	62	57	92
1bT51	S_7S_t	49	15	31	68	12	18
1bT47	S_8S_t	53	27	51	35	17	49
	Total	**350**	**194**	**58 (9.9)**	**289**	**179**	**62 (13.0)**
		Self cross @ 4 weeks			Compatible cross (Somerton S_1S_{21}) @ 4 weeks		
1aT4	S_7S_3	50	3	6	48	21	44
1aT9	S_7S_3	28	3	11	28	11	39
1aT40	S_8S_3	14	1	7	24	4	17
1bT31	S_8S_3	55	2	4	60	39	65
1bT42	S_8S_3	112	4	4	85	45	53
	Total	**259**	**13**	**6.2 (1.3)**	**245**	**120**	**44 (8.0)**

Fruit set experiment on 'Nonpareil' x 'Lauranne' trees 2002 - 12 weeks

Tree	*S*-alleles	Self cross @ 12 weeks: No. flowers for fruit set	No. fruit set	% fruit set (S.E.)	Compatible cross (Somerton S_1S_{21}) @ 12 weeks: No. flowers for fruit set	No fruit set	% fruit set (S.E.)
1aT26	S_7S_t	28	21	75	60	39	65
1aT30	S_8S_t	154	39	25	64	22	34
1bT32	S_8S_t	66	34	52	62	15	24
1bT51	S_7S_t	49	6	12	68	6	9
1bT47	S_8S_t	53	24	45	35	17	49
	Total	**350**	**124**	41.9 (10.8)	289	99	36.2 (9.7)
		Self cross @ 12 weeks			Compatible cross (Somerton S_1S_{21}) @ 12 weeks		
1aT4	S_7S_3	50	3	6	48	15	31
1aT9	S_7S_3	28	0	0	28	6	21
1aT40	S_8S_3	14	1	7	24	4	17
1bT31	S_8S_3	55	2	4	60	33	55
1bT42	S_8S_3	112	4	4	85	33	39
	Total	**259**	**10**	**4.1 (1.2)**	**245**	**91**	**32.6 (6.8)**

Fruit set experiment on 'Nonpareil' x 'Lauranne' trees 2003 - 12 weeks

Tree	*S*-alleles	Self cross @ 12 weeks: No. flowers for fruit set	No. fruit set	% fruit set (S.E.)	Compatible cross (Somerton S_1S_{21}) @ 12 weeks: No. flowers for fruit set	No fruit set	% fruit set (S.E.)
1aT26	S_7S_t	50	26	52	50	29	58
1aT30	S_8S_t	50	37	74	50	32	64
1bT32	S_8S_t	50	18	36	50	23	46
1bT51	S_7S_t	50	13	26	50	3	6
1bT47	S_8S_t	28	17	61	41	15	37
	Total	**228**	**111**	**49.7(8.6)**	**241**	**102**	**42.1(10.2)**
		Self cross @ 12 weeks			Compatible cross (Somerton S_1S_{21}) @ 12 weeks		
1aT4	S_7S_3	50	1	2	50	23	46
1aT9	S_7S_3	7	0	0	18	0	0
1aT40	S_8S_3	50	0	0	50	27	54
1bT31	S_8S_3	50	3	6	50	25	50
1bT42	S_8S_3	50	2	4	50	18	36
	Total	**207**	6	**2.4(1.2)**	**218**	**93**	**37.2(9.8)**

Fruit set experiment on 'Nonpareil' x 'Lauranne' trees 2003 - kernel maturity

Tree	*S*-alleles	Self cross @ maturity: No. flowers for fruit set	No. fruit set	% fruit set (S.E.)	Compatible cross (Somerton S_1S_{21}) @ maturity: No. flowers for fruit set	No fruit set	% fruit set (S.E.)
1aT26	S_7S_t	50	26	52	50	29	58
1aT30	S_8S_t	50	37	74	50	32	64
1bT32	S_8S_t	50	18	36	50	23	46
1bT51	S_7S_t	50	13	26	50	3	6
1bT47	S_8S_t	28	17	61	41	13	32
	Total	**228**	**111**	**49.7(8.6)**	**241**	**100**	**41.1(10.4)**
		Self cross @ maturity			Compatible cross (Somerton S_1S_{21}) @ maturity		
1aT4	S_7S_3	50	1	2	50	23	46
1aT9	S_7S_3	7	0	0	18	0	0
1aT40	S_8S_3	50	0	0	50	27	54
1bT31	S_8S_3	50	3	6	50	25	50
1bT42	S_8S_3	50	2	4	50	18	36
	Total	**207**	**6**	**2.4(1.2)**	**218**	**93**	**37.2(9.8)**

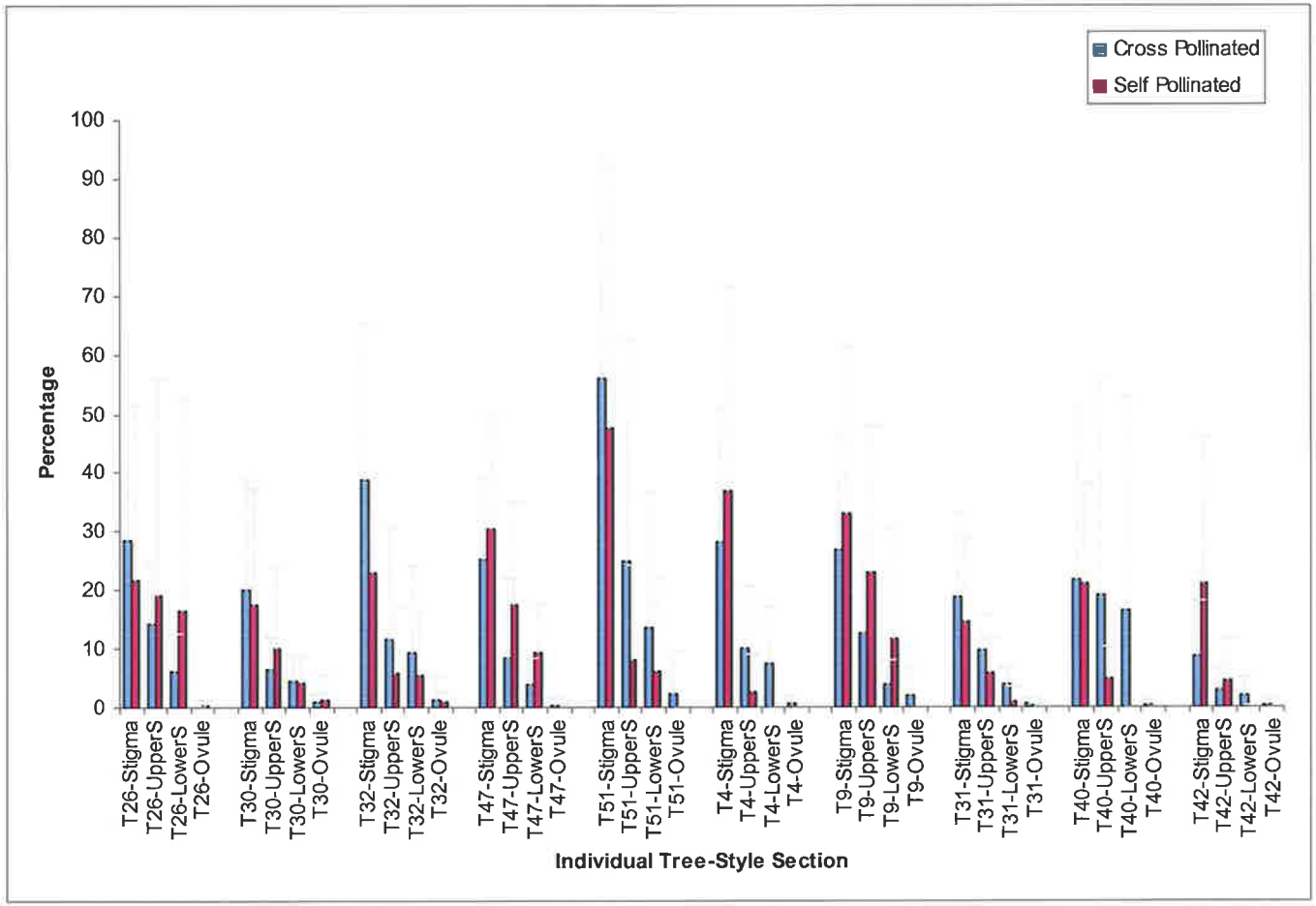

Figure 6.3 Graph of percentage of pollen tubes reaching each section of the style, T26, T30, T32, T47 and T51 are self-compatible individuals, T4, T9, T31, T40 and T42 are self-incompatible individuals. Codes for style section are *Stigma* – pollen tubes in the stigma, *Upper S* – pollen grains penetrating the upper section of the style, *Lower S* – pollen tubes penetrating the lower section of the style, *Ovule* – pollen grains present at the ovule. Bars correspond to standard deviation.

The mean percentage doubles decreased by approximately half in 2002 and 2003 compared to the value of 10.31% in 2001 (Table 6.3), which also showed the largest variation in the mean (Figure 6.5a). The highest individual value observed was 100% in 2001, and the lowest was 0% observed in all years. When grouped into two classes, the averages for all three years was consistently 60% to 40%, for total absence of doubles versus doubles present above 1%, respectively (Figure 6.6a). Chi-squared analysis was used to determine if the trait was inherited in a 1:1 or 3:1 ratio and the inheritance of this trait did not correspond to a simple mendelian inheritance pattern.

This trait showed a high genetic variance of 52.45, and a heritability of 0.686 (Table 6.4), no correlation was observed with any of the other traits tested (Table 6.5).

Shell Hardness

Shell hardness varied from 37.4% in 2002 to 32.5% in 2003, which is classified as semi-hard to hard respectively (Figure 6.5b). When classified on a category basis the majority of individuals in 2001 and 2003 were classified as hard shelled, while in 2002 the largest group was those with semi-hard shells (Figure 6.7a). Approximately 90% of the individuals for all years showed a shell hardness value of one of these two levels. In 2002, 9% of individuals had a soft shell which was approximately three times more than levels observed in the other two years. Less than 1% had a paper shell, which is the shell type of 'NP' (Figure 6.7a). Shell hardness showed a genetic variance of 19.6 and a heritability of 0.348 (Table 6.4), with no correlation to any other trait (Table 6.5).

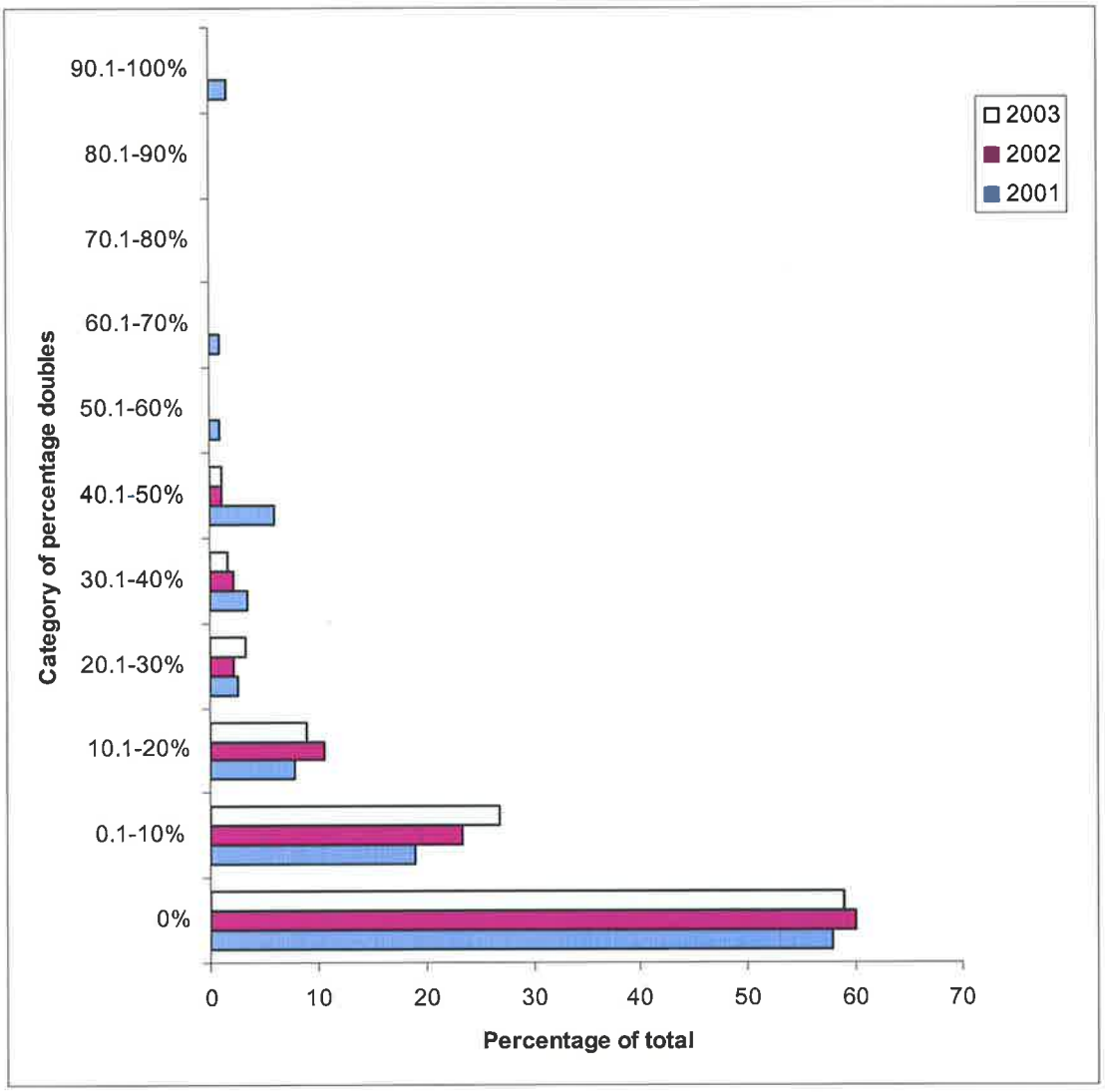

Figure 6.4 Percentage of total individuals showing 0 – 100% doubles, in 10% increment categories, distribution over three years (2001-2003)

Inshell Weight

The average inshell weight was highest in 2002 at 3.19 g (Figure 6.5c), with the highest individual average in the same year at 5.27 g, and the lowest in 2001 at 0.77 g, illustrating a wide variation across the F_1 progeny which is shown in Figure 6.8. The genetic variance and heritability for this trait was calculated at 0.368 and 0.626 respectively (Table 6.4). A positive correlation of 0.60 was observed between inshell weight and kernel weight (Table 6.5).

Table 6.3 Morphological characters measured on F_1 hybrid population of cross between 'Nonpareil' and 'Lauranne', average values for data collected for 3 years (2001-2003)

2001

Character	n	mean	SD	range		conf. interval
				lowest value	highest value	95%
% Doubles	115	10.31	20.88	0	100	3.78
% Shell hard [a]	115	33.43	6.56	17	62	1.21
Kernel wt (g) [b]	115	1.08	0.26	0.52	2.09	0.05
Shell wt (g)	115	2.21	0.74	0.77	4.79	0.17
Kernel shape [c]	115	0.57	0.05	0.46	0.68	0.14
Kernel thickness [d]	115	7.19	0.78	4.7	9.43	0.15

2002

Character	n	mean	SD	range		conf. interval
				lowest value	highest value	95%
% Doubles	181	5.00	9.27	0	48	1.35
% Shell hard [a]	181	37.40	6.58	22	72	0.96
Kernel wt (g) [b]	181	1.16	0.20	0.67	2.11	0.03
Shell wt (g)	181	3.19	0.68	1.67	5.27	0.10
Kernel shape [c]	181	0.59	0.05	0.4	0.73	0.01
Kernel thickness [d]	181	8.46	0.74	6.33	13.33	0.11

2003

Character	n	mean	SD	range		conf. interval
				lowest value	highest value	95%
% Doubles	179	4.20	8.29	0	43.8	1.21
% Shell hard [a]	179	32.50	6.24	14.2	72.4	0.91
Kernel wt (g) [b]	179	1.00	0.21	0.53	1.57	0.03
Shell wt (g)	179	3.10	0.73	1.4	4.83	0.11
Kernel shape [c]	179	0.61	0.08	0.47	0.78	0.01
Kernel thickness [d]	179	7.70	0.64	6.01	9.89	0.09

[a] Shelling percentage (proportion of inshell almond weight that is the kernel). Papershell: 55-65%+, Soft shell: 45-55%, Semihard: 35-45%, Hard-shell: 25-35%, and Stone-shell: 20-25% (Rugini, 1986)

[b] grams, avg.weight of 10 kernels :Very small=0.9 or less, small=0.9-1.1, medium=1.1-1.4, large=1.4-1.8, very large>1.8

[c] width/length ratio of 10 kernels: very narrow (<.40), narrow (.40-.48), medium (.49-.55), broad (.56-.65), very broad (>.65)

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

Kernel Weight

The average individual kernel weight ranged from 0.52 g in 2001 to 2.11 g in 2002 (Table 6.3). The average value for all individuals was highest in 2002 at 1.16 g (Figure 6.5d). Kernel weight had a low genetic variance of 0.031 and a heritability of 0.384 (Table 6.4). This trait showed a positive correlation with both inshell weight (0.60) and kernel length (0.68) (Table 6.5).

Kernel Shape

Kernel shape is a function of the kernel width/length ratio and is classified into five classes. For all years the majority of individuals were classified as broad, followed by the medium class. In 2003 the number of very broad individuals approximately doubled compared to previous years, with a corresponding drop in the numbers in the medium category (Figure 6.7b). In 2001 there was a nearly even split between the medium and broad categories with only a small percentage of very broad individuals. There was an overall trend toward broader individuals from 2001 to 2003. The lowest individual value of 0.40 (very narrow) was observed in 2002, with the highest of 0.78 (very broad) in 2003 (Table 6.3).

Kernel shape showed an extremely low genetic variance of 0.003 and a heritability of 0.551 (Table 6.4), and a positive correlation of 0.60 was observed with kernel width (Table 6.5).

Kernel width was highest in 2003 with an average of 13.5 mm, this trait has a genetic variance of 0.864 and a heritability of 0.434 (Table 6.4), and is positively correlated with kernel shape (0.60) (Table 6.5).

Average kernel length was highest in 2001 at 23.24 mm, it showed a high heritability of 0.797 and a genetic variance of 4.894. A positive correlation was observed with kernel weight (0.68).

Kernel thickness was highest in 2002, and the thickest individual average of 13.33 mm was also observed in this year, while in 2001 the thinnest individual average was 4.7 mm (Table 6.3). This trait is divided into five classes, in 2001 and 2003 the majority of individuals were in the medium class, while in 2002 the majority were thick (Figure 6.7c and Figure 6.9). Generally in 2001, individuals were grouped in thinner classes than in 2002 and 2003, and in 2002 a much higher proportion of very thick individuals was observed. Kernel thickness showed a low genetic variance and heritability of 0.253 and 0.256, respectively (Table 6.4). This trait did not show any correlation with any other traits studied (Table 6.5).

Testa Appearance

This trait was scored as a two-class trait; pale or dark, although a gradation between nine different shades was detected when this trait was measured with the RHS colour charts. There was a dramatic shift from a majority of dark kernels in 2001 to nearly equal levels of pale and dark in 2003. Therefore testa colour showed a high amount of variance over the three years with respect to the percentage in each class (Figure 6.6b). In 2001 and 2003 the values did not correspond to a 3:1 ratio, in 2002 the values corresponded to a 3:1 ratio as tested by $\chi 2$ analysis. This correspondence is likely to be an artifact given this trait did not segregate as expected in the other two years.

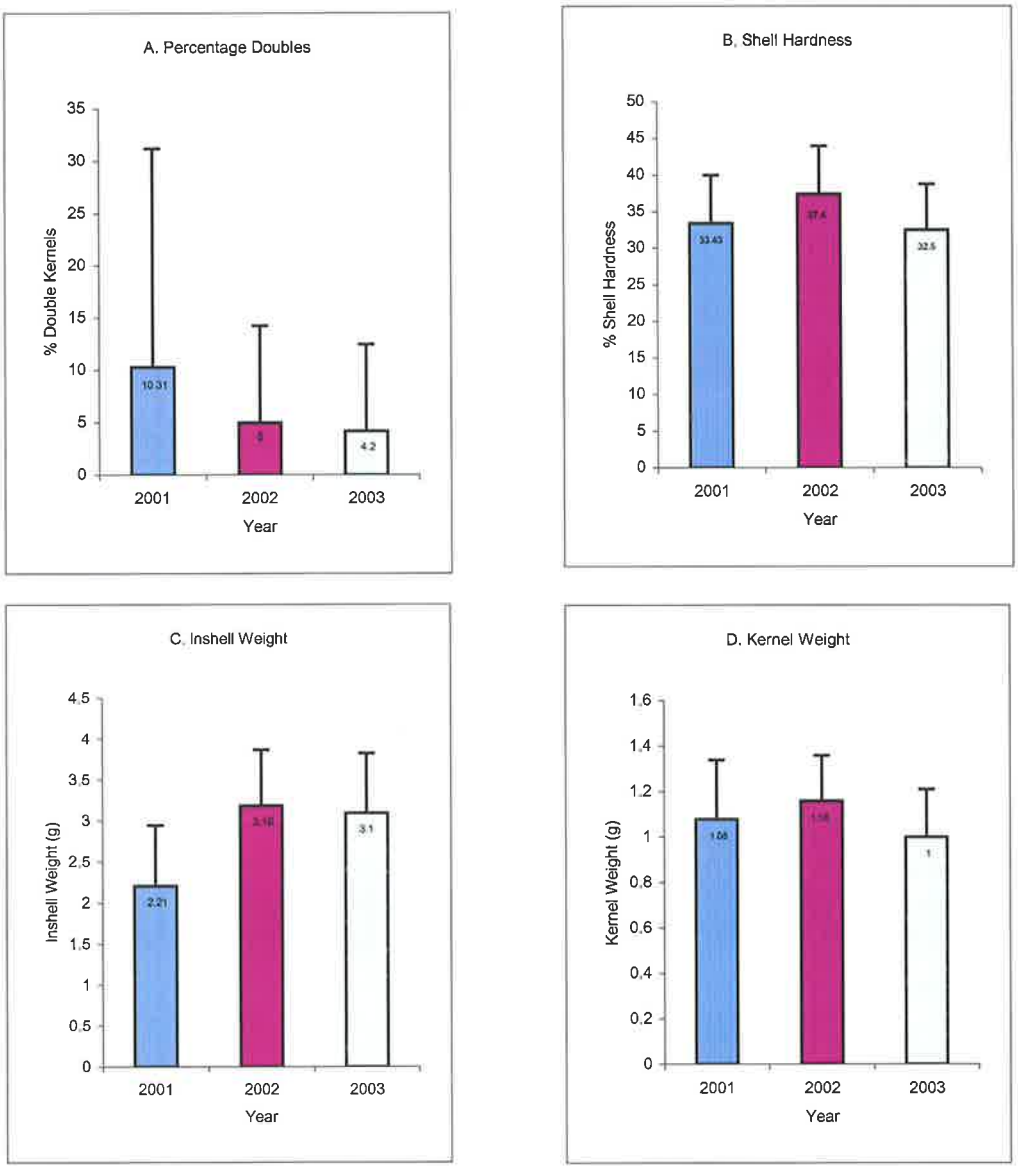

Figure 6.5 A,B,C,D Averages of characters over three years. A. Percentage Doubles, B. Shell Hardness, C. Inshell Weight, D. Kernel weight. Bar corresponds to standard deviation.

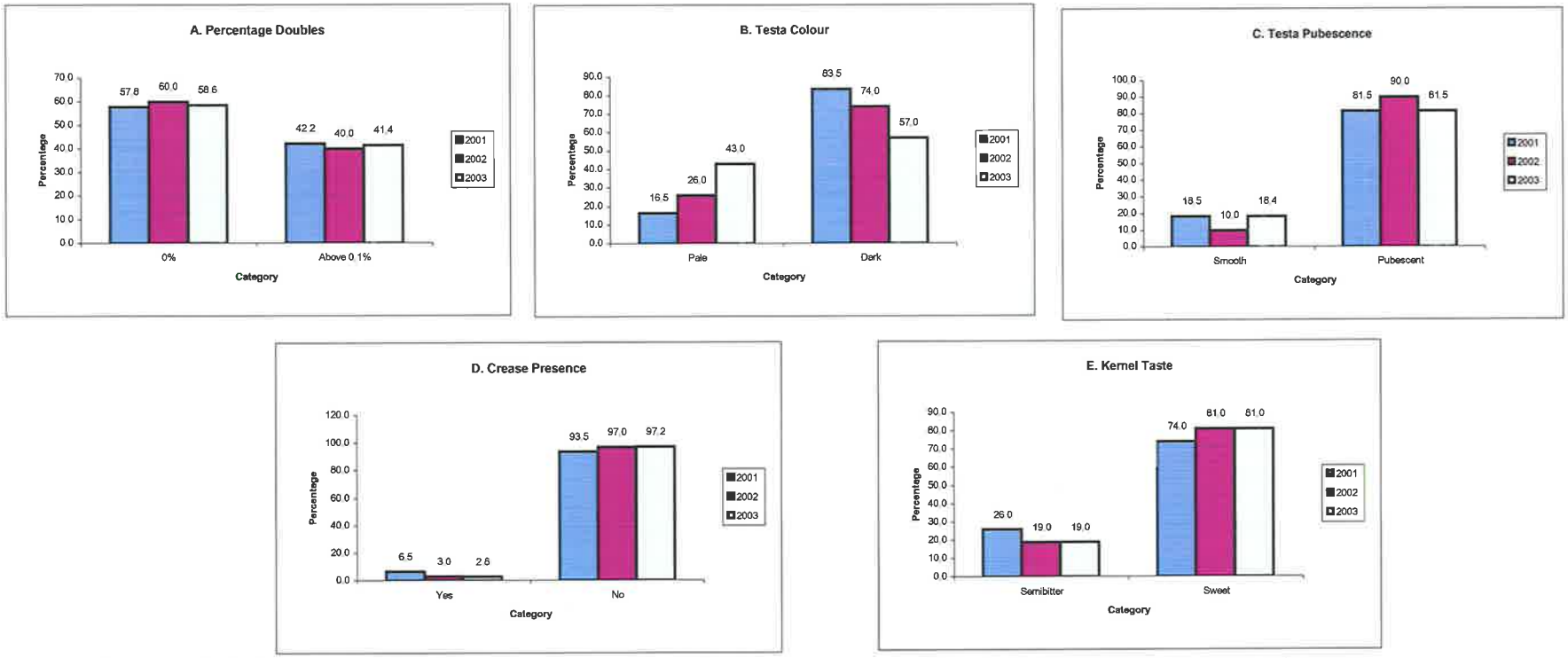

Figure 6.6 A,B,C,D,E Percentage of hybrids grouped into two classes for A. percentage doubels, B. Testa Colour, C. Testa pubescence, D. Crease presence, E. Kernel taste

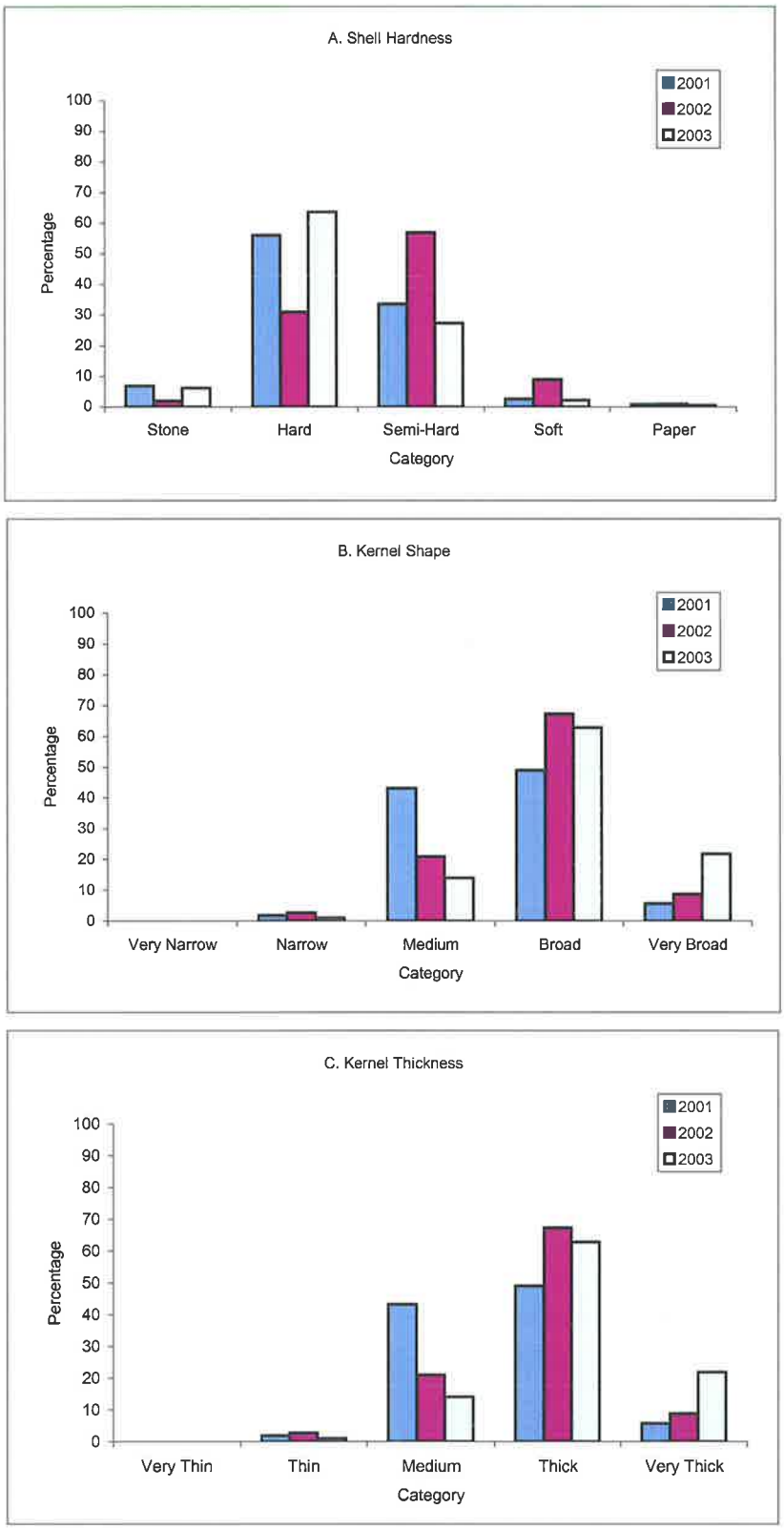

Figure 6.7 A,B,C Percentage of hybrids grouped into categories for A. Shell hardness, B. Kernel shape, C. Kernel thickness

Figure 6.8 Variation in inshell size and kernel weight in F_1 progeny of 'Nonpareil' x 'Lauranne'. Scale Bar = 5cm

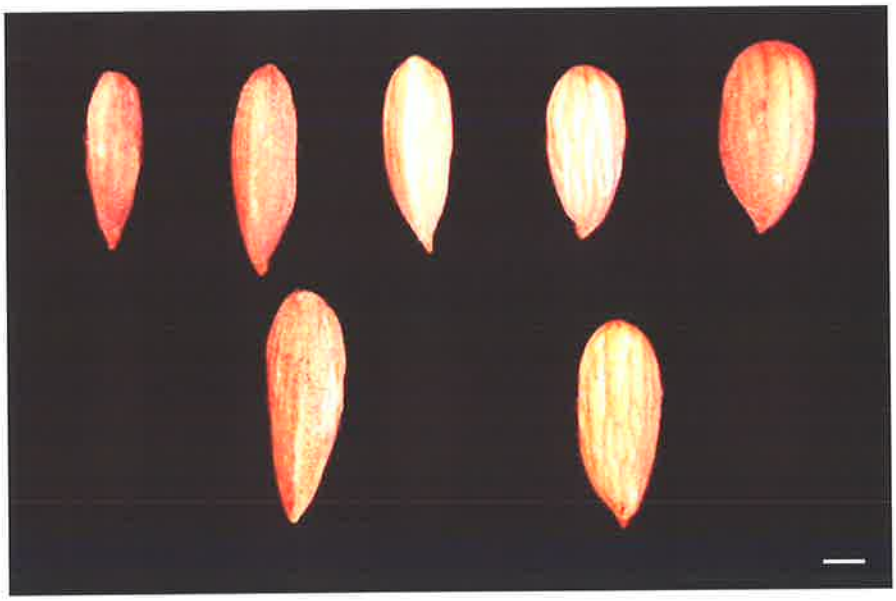

Figure 6.9 Variation in kernel thickness, five categories Top Row: L-R very thin, thin, medium, thick, very thick. Bottom Row: L-R 'Nonpareil' 'Lauranne'. Scale Bar = 1cm

Table 6.4 Broad sense heritabilities for characters scored in the F_1 hybrid 'Nonpareil' x 'Lauranne' cross for three years (2001-2003)

Trait	h^2
Doubles	0.686
Shell Hardness	0.348
Kernel Weight	0.031
Shell Weight	0.368
Kernel Shape	0.551
Kernel Width	0.864
Kernel Thickness	0.253
Kernel Length	0.797

Table 6.5 Phenotypic raw correlations among traits calculated from F_1 hybrid population of 'Nonpareil' x 'Lauranne' cross over three years

Trait	Correlation (r)
Inshell Weight/Kernel Weight	0.6
Inshell Weight/Shell Hardness	0.34
Inshell Weight/% Doubles	0.13
Inshell Weight/Kernel Length	0.49
Inshell Weight/Kernel Thickness	0.11
Inshell Weight/Kernel Width	0.55
Kernel Length/Kernel Width	0.35
Kernel Length/Kernel Thickness	-0.09
Kernel Width/Kernel Thickness	0.01
% Doubles/Shell Hardness	-0.07
% Doubles/Kernel Weight	0.06

Testa pubescence did not segregate to a 3:1 ratio, but for 2001 and 2003 the percentage of individuals in each class were identical, in 2002 the numbers of individuals with the pubescent character increased by 10% (Figure 6.6c).

Crease presence was seen only in a very small percentage of individuals, and decreased by approximately 50% in 2002 and 2003 compared to the 2001 value (Figure 6.6d).

Kernel Taste

The number of sweet individuals was approximately four times the number of semi-bitter individuals in 2002 and 2003. In 2001 the number of sweet individuals was 7% lower than the other the other two years (Figure 6.6e). Consequently only 2001 showed a $\chi 2$ value corresponding to a 3:1 segregation ratio.

DISCUSSION

Vegetative traits were assessed only in 2001 since the majority of traits were highly subjective, difficult to score and showed little useful information. Tree architecture has the potential to affect yield and disease resistance (C. Bennett pers. com. 2002), the majority of trees were spread across three tree habit classes. However this trait was difficult to assess accurately because trees were so closely planted it is unlikely that they are assuming their natural stature under such close conditions. This was compounded in 2002 and 2003 as the trees grew larger. Leaf colour has little economic value, and leaf size varies with the leaf position on the plant (Kester and Gradziel 1996). Tree height is also likely to have been

affected by the artificial close planting arrangement. These traits would be better assessed once the trees were grafted and transplanted to an orchard style growing system.

Bloom time was later in 2000 compared to 2002, however both years showed two major peaks in bloom time accounting for approximately 50% of all hybrids. Bloom date is expected to change according to season since it is triggered by environmental cues (Vargas 2001), and in this study the number of days between the two major flowering peaks was reduced in 2002 but the overall pattern remained the same. DiGrandi-Hoffman et al. (1994) have shown different incremental changes in temperature requirements for single blossom opening for various cultivars, which ranges from 1.1°C for 'Ne Plus Ultra' to 8.9°C for 'Mission'. The mean average air temperature in 2002 was 0.7°C warmer for August (appendix B) which may have promoted earlier petal opening in this year. In addition the variation in the years may have been due to cooler temperatures which slow the rate of petal opening, conversely delaying petal opening in 2000, as shown in studies by Bernad and Socias i Company (1995)

The correlation between precocity and self-fertility was examined as it has been suggested that precocity is correlated with the inheritance of self-fertility (R. Socias i Company pers. com. 2003). This was observed for plants that flowered for the first time in 1999, however self-fertility was only observed in 61.5% of the flowering plants suggesting a significant but weak correlation. More studies will be required to determine if this is a casual observation or a true phenomenon, since this is only a relatively small sample size.

The inheritance of self-fertility (Sf) or self-compatibility (SC) in the whole population was assessed by allele specific PCR. Pollen tube analysis and fruit set was also carried out on a selection of 10 individuals, 5 identified as self-incompatible and 5

identified as SC to provide a comparison for the molecular technique. Inheritance of the *Sf*-allele followed a 1:1 inheritance ratio and as expected 50% of the progeny inherited this allele from the male parent. This is consistent with other studies on *S*-allele inheritance in almond (Dicenta and García 1993a; Ortega and Dicenta 2003; Rovira et al. 1997). Fruit set was significantly higher in individuals with the *Sf*-allele pollinated with self-pollen than those that did not have this allele. It is now well established that the presence of this allele confers self-fertility in almond regardless of the *S*-allele of the pollen, which was also shown in this study. The mode of self-fertility has not been fully clarified and the elucidation of its control has been a popular topic of research. It has however been established that pollen tube arrest occurs in the upper section of the style in SI crosses (de Nettancourt 1977; Ben-Njima and Socias i Company 1995). No significant differences were observed in pollen tube levels in the lower section of the style (an indication of pollen tube arrest in the upper syle) when comparing self-pollination or cross-pollination in SC versus self-incompatible groups. The only significant difference was in the percentage of pollen tubes reaching the ovule in both of these treatments. Ortega and Dicenta (2003) discussed the difficulty in examining pollen tube data as a measure of self-fertility, and also observed no difference in the levels of pollen tubes in the upper parts of the style when half- and fully-compatible cross pollinations were compared. Based on the results of this study the determination of self-fertility by pollen tube analysis would be more accurate if the number of pollen tubes reaching the ovule were used instead of pollen tube arrest in the style. This does not agree with data presented by Dicenta et al. (2001) who showed an absence of difference in the number of pollen tubes in the ovary between self- and cross-pollinations of a number of cultivars. Therefore the most robust method for self-fertility

assessment is by molecular analysis, followed by fruit set and finally pollen tube penetration of the ovary.

Fertilisation of both ovules in the ovary leads to two kernels being produced within the same endocarp, a situation known as double kernel production. A predisposition to double kernel production has a genetic component since some cultivars are able to pass this on to further generations (Spiegel-Roy and Kochba 1974). However the expression of this trait is modified by the environment (Barbera et al. 1988), and the development of doubles has in most cases been linked to environmental conditions during ovule development, particularly temperature at preblossom. Work by Egea and Burgos (1995) has suggested that high preblossom temperatures cause earlier degeneration of the secondary ovule and therefore act negatively on double kernel percentage. Later work by these authors has shown cultivars which consistently show a high level of doubles, such as 'Malaguẽna' do not have a significant difference in ovule size between the primary and secondary ovule three days after anthesis (Egea and Burgos 2000) which could be a factor in double production. The degeneration of the secondary ovule in the cultivars studied in this work appeared to be under genotypic control, consistent with observations by Gradziel and Martínez-Gómez (2002) on 'Nonpareil'. In this study the highest average percentage doubles was in 2001, and the preblossom average temperature in July was 16.6°C and in August 18.0°C. The August value was 0.7°C less than 2002 and 2003. However cooler temperatures have been shown to retard pollen tube growth (Socias i Company et al. 1976). This is a complex trait and temperature effects are only one factor in determining the level of double kernels (Asensio and Socias i Company 1996). One reason for the large difference in doubles between 2001 and the other two years may be due to the juvenile

phase of the trees. In 2001, there was a much larger percentage of doubles in the 40.1-50%, 60.1-70% and 90.1-100% categories which was not observed in the other two years. In 2001 a large number of trees produced only a few nuts which may not have been a large enough sample size to accurately assess this trait. For example, a tree that produced a total of 3 kernels in 2001 produced 100% doubles, in 2002 and 2003 30 kernels were collected and the percentage doubles dropped to 6% and 0% respectively. The limited sample number and resulting sample bias in 2001 may therefore account for a large proportion of the differences observed between 2001 and the following two years data.

Percentage doubles showed a heritability of 0.686, which is nearly double the value obtained by Arteaga and Socias i Company (2002) (0.38) but close to the value of 0.51 from a study by Kester et al. (1977). The difference between these values may be a result of the differences in the calculation of heredity. As in the study by Kester et al. (1977) the data was from an average of several families and the inclusion of families with no double kernels, which has been suggested by Arteaga and Socias i Company (2002) as a source of variation between results.

Shell hardness is determined by the percentage of kernel to whole nut weight, and this population was derived from a cross between a paper and hard shelled cultivar. The inheritance of this trait in the progeny was strongly skewed, 90% of the population showed a hard or semihard shell type for all three years. This trait is controlled by a single gene, *D*, based on the segregation of this trait in the population it can be assumed that 'Nonpareil' is homozygous recessive *dd*, and 'Lauranne' is homozygous dominant, *DD*, and the majority of the progeny are heterozygous *Dd*. Shell hardness showed a lower heritability than was expected based on the inheritance and similar averages for the three years which suggested

a strong genetic control. However, this value was still within the range observed for other studies of 0.55-0.32 (Arteaga and Socias i Company 2002; Kester et al. 1977). Kester (1996) proposed a difference in the temporal expression of *D* and *d*, within the shell consequently affecting the shell hardness. Both this and the effect of modifier genes on this major gene may account for the 10% of individuals which did not show the expected heterozygous semihard to hard phenotype.

Inshell weight is a function of the combined weight of the shell and kernel. A high positive correlation was observed between inshell weight and kernel weight, where an increase in kernel weight causes a concomitant increase in inshell weight. The increase in kernel weight potentially caused the observed increase in shell hardness in 2002 since this is also influenced by kernel weight. A positive correlation was observed between kernel weight and kernel length indicating kernels increase in weight by increasing length as opposed to thickness or width. Kernel length showed a high heritability and this indicates this trait is genetically determined rather than changing in response to environmental conditions. Kernel width is a more important factor in determining kernel shape than kernel length, as shown by the positive correlation between these traits. Kernel thickness showed a very low heritability which is consistent with values obtained by Arteaga and Socias i Company (2002) but much lower than values obtained by Chandrababu and Sharma (1999) and Kester et al. (1977). The increase in kernel weight showed no obvious correlation to weather conditions, although a decrease in the mean maximum air temperature in January and February 2002 compared to 2001 and 2003 may have reduced the plants response to heat and moisture stress. Stress at the final stage of growth in nectarine has shown to be linked to a decrease in yield (Naor et al. 2001). Hawker and Buttrose (1980) stated kernels

reach their maximum weight at approximately 15 weeks after anthesis, cooler temperatures in November 2001 compared to the other years may have also reduced the stress on the plants and allowed maximal nutrient assimilation in this period resulting in an overall increased average kernel weight for this year. Kester (1996) supports this and has shown improper filling of kernels which reduces kernel weight is caused by moisture stress and other adverse conditions (Kester 1996).

Climatic conditions at harvest storage conditions and nut maturity have been proposed to affect testa colour (Kester et al. 1990). Some casual observations in this study have indicated darkening of the kernel testa occurs as the kernel undergoes its final weeks of maturity and once mature testa colour starts to darken when the nut is cracked from the shell and exposed to light. It was noticed that some trees were not as mature as other trees for each of the years harvest, as indicated by hull dehiscence and dehydration. This in combination with the storage conditions and subjective nature of the assessment may be responsible for the skewness of the results over the three years. The gradation over several colour shades makes scoring intermediately coloured nuts to a particular class difficult.

Testa pubescence is another trait that was subjectively assessed; overall there was a greater proportion of pubescent individuals than smooth individuals for all years. The physiosensory assessment of this trait is influenced by a number of different stimuli such as the level of pubescence, kernel rugosity and kernel colour. The overall pattern of inheritance indicates that this trait may be controlled by a major gene interacting with two or more modifier genes. The scoring of smooth individuals in the progeny was difficult to determine as this trait appears to be a lower level of the pubescent character rather than a distinctly different class observed for the smooth parent, as examined in Chapter 7. Scoring

of this trait is therefore a matter of degrees and a less subjective assessment of this trait such as microscopy analysis or development of a molecular marker, as outlined in Chapter 7 would be beneficial to remove subjective bias.

Presence of a crease in the kernel was observed only in 2.6% to 6.5% of the progeny over the three years. Kester et al. (1977) calculated a high heritability of 0.79 for this trait, whereas Arteaga and Socias i Company (2002) calculated a much lower heritability. It is likely that similarly to double kernels, a particular individual may have a genetic predisposition to this trait which is manifested under certain environmental conditions which have not been identified, and to date no information to date has identified this trait's cause or inheritance.

Kernel taste in this study was assessed only as the presence of a semi-bitter or marzipan like flavour or a sweet flavour (absence of any semi-bitter flavour). All other taste attributes such as degree of sweetness, oiliness, woody taste, nutty flavour and texture properties were ignored. Dicenta and García (1993) has shown a gradation between sweet and the total bitter taste in almond, which is correlated to the levels of cyanogenic glycosides in the kernel. The presence of semi-bitter individuals does not correlate with the segregation ratio proposed by Heppner (1923) of 3:1 and results from the present study agree with Dicenta and García (1993) and Ledbetter and Pyntea (2000) of a more complex mode of inheritance. It is not known if there is any environmental effect on the expression of this trait, and since bitterness is proposed to be a defense mechanism to insect attack (Ladizinsky 2000), it would be interesting to know if adverse environmental conditions could cause a plant response mechanism by increasing the levels of cyanogenic glycosides. These compounds are proposed to be produced in the mother plant and are translocated to

the seed (Frehner et al. 1990), and no data are evident indicating if the levels in the seed are static or in a constant state of flux, which may be triggered by external factors. If this was the case and the mother plant could influence the levels of the compounds in the seed, this may account for the difference observed in the increased number of semi-bitter individuals in 2001 compared to the identical levels observed in 2002 and 2003. Greater disease levels were observed in 2001, in particular the level of rust attack, which affected the majority of individuals in this year. Rust is a fungal infection that affects the foliage of the plant and can cause defoliation of the tree (Reuveni 2000). Defoliation may have been a trigger for the increase in cyanogenic compounds in the fruit of the most severely affected trees resulting in an overall increase in the number of semi-bitter individuals. Cyanogenic levels have been shown to increase in response to Al^{4+} and NH_4 in peach (Graham 2002) and to grazing levels in clover (Hayden and Parker 2002). Therefore the elucidation of the control and environmental influences of this trait warrant further investigation.

As no direct investigation was undertaken to specifically determine the physical parameters influencing these traits, it is difficult to state exact reasons for the variation over the three years. Several potential causal relationships have been identified in this study with regard to environmental and biotic factors and trait variation over the three years. The next phase of this work would be to carry out experiments with clonal material under controlled environmental conditions. This type of approach would expand the knowledge on the causes and interrelations of some of these traits, and provide a better understanding of these important characteristics in almond.

The aim of this investigation was primarily to identify traits segregating in the population which could then be incorporated into the genetic linkage map, and used for

marker assisted selection. Seven traits were selected for integration into the genetic map which is discussed in Chapter 8, these were *S*-alleles, kernel shape, kernel thickness, kernel taste, double kernel production, testa pubescence and testa colour.

CHAPTER 7

Characterisation of Factors Responsible for Almond Kernel Testa Pubescence, and Development of a Molecular Marker by Bulked Segregant Analysis

INTRODUCTION

The sensory attributes of almond kernels depend on several factors, such as flavour, size, shape, colour, rugosity, and pubescence, and therefore improvement by plant breeding is a difficult process. To add to the complexity facing breeders, European and American almond producers and consumers favour different kernel attributes, with the former preferring a dark flat nut, such as 'Marcona' and the latter a fatter paler nut, such as 'Nonpareil' (Socias i Company, pers. comm. 2003). One distinguishing difference between the cultivars grown for these two markets is kernel testa pubescence, with cultivars of European descent tending toward a pubescent testa while American cultivars, particularly 'Nonpareil' are usually smooth, or non-pubescent.

Studies on testa pubescence are limited and generally focus on cellular development or heritability. Testa pubescence has a low heritability of between 0.30 –

0.39 (Arteaga and Socias i Company 2002; Kester et al. 1977) suggesting that environmental factors can alter the expression of this trait.

Brooks (1939) and Hawker and Buttrose (1980) have published in-depth data on the development of almond fruits, which differ quite markedly from the closely related peach, apricot, cherry, and plum in which a fleshy mesocarp develops. In contrast, the almond mesocarp is dry and leathery, and the economically important part of the crop is the kernel or seed. The kernel testa is maternal and is derived from the outer and inner integuments of the ovule (Hawker and Buttrose 1980). When mature, the testa shrinks and encloses the cotyledons and embryo in a dry textured skin. At maturity, the testa is the most external tissue of the seed, has no cellular connection with the cotyledons, embryo, or endosperm, and can be easily removed from the rest of the seed (Pascual-Albero et al. 1998).

The function of the testa seems to be largely protection for the embryo, but Hawker and Buttrose (1980) suggested that the testa is an important organ for the movement of metabolites into the seed since it is supplied by vascular tissue from the pericarp. They also observed that, in the first few weeks of fruit development, the cells of the outer layer of the testa became enlarged and filled with darkly staining contents. In addition to this study, Pascual-Albero et al. (1998) also observed cellular structures that projected from the outer epidermal layer of the testa. Young et al. (2004) examined the cellular structure of the almond cotyledon and testa microscopically, and observed thin walled cells ranging from 100-300 μm in width projecting from the surface. Directly below this outer surface they observed the remaining portion of the testa to be composed of approximately 20 layers of flattended parenchyma cells. (Young et al. 2004). It is not known if the presence of these epidermal protuberances confer an adaptive advantage to the individual.

Mesocarp pubescence in peach and other *Prunus* such as apricot has been shown to be due to surface hairs that are tubular and anchored between epidermal cells (Archibald and Melton 1987; Atwood and Arnold 1982; Creller and Werner 1996). In peach, mesocarp pubescence is controlled by the G locus, where homozygous recessive individuals (gg) produce fruit that lacks pubescence, known as the nectarine type of *Prunus persica* (Creller and Werner 1996). The control of testa pubescence in almond is not well understood despite the fact that it is a potential factor for consumer preference and marketability. Therefore it would be desirable to develop a technique to identify genotypes that can express this trait. The identification of a genetic marker linked to kernel pubescence would allow faster screening of individuals in large breeding programs to be carried out at the juvenile seedling stage by marker assisted selection (MAS).

The technique known as bulked segregant analysis (BSA) developed by Michelmore et al. (1991) has become an invaluable method for identifying markers linked to genes of economic or scientific interest in a target population. BSA is a versatile and robust technique when compared with other methods, such as genetic map construction in combination with linkage association (Dirlewanger et al. 1996; Hurtado et al. 2002) and the analysis of near-isogenic lines (NILs) (Young et al. 1988), which are also used to detect sequences linked to a gene of interest. The success of BSA lies mainly in its simplistic approach, which follows three basic steps. Firstly, the segregating population is assessed and scored for the phenotype of interest, then the genomic DNA of between eight and twelve individuals of the same phenotype (at equal DNA concentrations) is pooled into a 'bulk' mix. A second bulk is produced in the same manner with individuals of the alternative phenotype, for example in the case of a marker for a particular disease, resistance versus susceptible individuals. Finally these

genomic DNA bulks are screened against a number of primers until a consistent polymorphism between the two groups is observed.

The primer type of choice for BSA analysis has generally been Randomly Amplified Polymorphic DNA (RAPDs) because these primers are random in their binding pattern, require no prior sequence knowledge, and many primers are available. Studies using RAPDs for BSA have identified markers linked to disease resistance in lettuce (Michelmore et al. 1991), resistance to leaf spot in olive (Mekuria et al. 2001b), waxiness in *Eucalyptus* (Wirthensohn et al. 1999), and aroma in rice (Lorieux et al. 1996). In *Prunus*, markers have been developed to identify the freestone character in peach (Warburton et al. 1996) peach/nectarine, red/green leaf and pollen fertility/pollen sterility in peach (Chapparo et al. 1994), self-compatibility in apricot (Badenes et al. 2000), and late bloom time in almond (Ballester et al. 2001). Therefore, not only do these examples illustrate the usefulness of BSA for identifying molecular markers linked to traits of interest of diverse origin, but also its practical application to identifying markers in *Prunus*. In almond, the two classes observed for kernel pubescence are those that have a smooth testa and those that have a pubescent testa.

The objectives of this study were firstly to examine of the cellular structure of the testa using light and scanning electron microscopy. This was to determine if a difference in the surfance of the testa is responsible for differences in mouthfeel pubescence between smooth and pubescent. The second aim was to develop a molecular marker linked to the presence of kernel pubescence using bulked segregant analysis, and the development of sequence specific primers for this trait for use in breeding programs.

MATERIALS AND METHODS

Plant Material

Almond genotypes used for this study were derived from the F_1 hybrid population of a cross between 'Nonpareil' (NP, phenotypically smooth testa) and 'Lauranne' (LA phenotypically pubescent testa). Details of this cross are provided in Chapter 2.

For the analysis of a range of cultivars, kernel samples were selected from the collection held by the Almond breeding program at The University of Adelaide, or collected from trees held at the Claremont orchard at the Waite Campus of The University of Adelaide.

Organoleptic Evaluation of Testa Pubescence

All 181 individuals comprising the F_1 hybrid population were organoleptically assessed by the author for mouthfeel. Kernels for evaluation were selected randomly from the tree, and a sample size of two kernels was deemed to be a representative sample for each tree. To ensure classifications by this panelist were veritable two kernels from 24 individuals (14 for 2001) were submitted to a sensory evaluation panel, consisting of three additional panelists. The panelists were asked to organoleptically assess the nuts from each tree and score the mouthfeel of the testa surface as either smooth or pubescent (only these two options were available) by comparison to the two parents whose phenotype was given.

Microscopy Techniques

For all microscopy sample preparations, at least two kernels of each of the parents, and two hybrids that had been organoleptically assessed as smooth or pubescent were cracked from the shell immediately prior to analysis to eliminate mechanical

damage to the testa surface by external factors. Kernels were sampled at approximately 12 months after fruit maturity, with the exception of immature kernels of 'Nonpareil', which were sampled at approximately 24 weeks after pollination, and the samples of 'Lauranne', which were sampled at approximately 30 weeks and 24 months after fruit maturity for scanning electron microscopy (SEM) and light microscopy respectively.

Light Microscopy

Surface samples of four whole kernels for each phenotype and parent (NP, LA, five hybrids with smooth testa, and five hybrids with a pubescent testa) were initially observed under a stereo dissecting microscope, as a preliminary observation of pubescence of the testa without any prior treatment, using a Leica MS 5 dissecting microscope with supplementary light supplied by Intralux 4000-1 (Volpi, Switzerland). Digital images were captured by the attachment of a Nikon 995 CoolPix digital camera (Nikon, Japan).

Two samples for light microscopy were prepared from two kernels for each individual, and treated separately for analysis. Testa samples were sectioned perpendicular to the cotyledon surface by hand with a sterile razor blade, stained with Toluidine Blue O (0.05% v/v Toluidine Blue O in benzoate buffer, pH 4.5) for approximately 1 min, and excess stain removed with RO water. Both stained and unstained sections were then mounted on a standard microscope slide in 90% v/v glycerol and observed with a Zeiss Axiophot Photomicroscope with transmitted light. Digital images were captured using an attached JVC colour video camera, using the computer program AcQuis (Syncroscopy, USA) for image manipulation.

Scanning Electron Microscopy

For SEM, samples were prepared for the parents, one hybrid with smooth testa, and one hybrid with a pubescent testa. For each individual eight samples were taken from two kernels derived from the one tree. Samples were prepared by slicing thin (*ca* 0.1 mm) parallel sections of the testa from the cotyledon tissue. Excess cotyledon tissue was removed from the underside of the section and the remaining tissue mounted outer testa surface up on aluminium SEM stubs using double-sided tape. Samples were prepared for SEM analysis by sputter coating with approximately 15 nm of carbon/gold using a vacuum evaporator, and examined using a Philips XL20 SEM at an accelerating voltage of 10 kV. The images were digitally captured.

Cell Density and Diameter Measurements

Cell density measurements were made by analysing the total number of collapsed cells per field of view from SEM images for up to eleven images and expressing this number as a percentage of the total number of cells. Measurements for cell diameter were made of the transverse axis of between 50-100 individual cells over 19 fields of view from the SEM digital images of testa surface. Measurements were only made for the samples described in the previous section.

Data Analysis

In order to determine if there was a significant difference between cell type (expanded versus collapsed) and cell type diameter in relation to classification (NP smooth, LA pubescent, F_1 smooth, F_1 pubescent) linear mixed modelling using restricted maximum likelihood (REML) was performed, since each treatment group contained an unequal number of measurements.

Diagnostic plots were performed to determine if the REML assumption, that normally distributed residuals have constant variance, was correct. If a departure from the assumption of constant variance was observed, the data for length were transformed using a natural logarithm. The Wald test statistic, which is asymptotically distributed as chi-square, was used to test the interaction between cell diameter and cell type and the independent effects of cell diameter and cell type. The tests were conducted in a hierarchal manner, whereby if the interaction effect was significant (p-value < 0.05) then the Wald test to analyse separate cell diameter and cell type effects was not applied. To detect where these differences lay, pairwise mean comparisons (on a log scale) was applied using the least significant difference (LSD) value. The LSD value was calculated from twice the standard error of the mean at the 5% level of significance.

Bulked Segregant Analysis for Testa Pubescence

The selection of individual trees for analysis by BSA was determined by analysing the sensory data from each year and panel for consistency between the assessed scores (9-12 assessments/parent or hybrid individual), given in Table 7.2.

Sixteen individuals from the F_1 population were selected to construct two DNA bulks. In addition, DNA was prepared from the parental genotypes. DNA was extracted using the modified Lamboy technique described in Chapter 3. Aliquots of the DNA prepared from each individual (300 ng), of either the smooth or pubescent phenotype, were pooled together to form two distinct bulks for each phenotype, at a final concentration of 20 ng/μL. A total of eight inter-simple sequence repeat (ISSR) primers [$(CT)_8G$, $(GA)_8T$, $(CA)_8G$, $(CA)_8T$, $(AC)_8G$, $(AG)_8YC$, $(CT)_8RC$, $(CA)_8RG$] (Proligo, Australia) were screened for polymorphism generation between the maternal and paternal DNA and both bulked samples. The PCR conditions and cycling parameters

were identical to the protocol described in Chapter 4. A negative control, containing no DNA, was included for all PCR reactions to check for DNA contamination.

The presence of an amplified fragment that was present in one parent and bulk, and absent in the other parent and bulk was used to indicate potential linkage to the trait of interest. In order to eliminate false positives, the PCR was repeated on the bulks and parents. If the parental and bulk banding pattern was consistent between the two reactions then PCR with the same primer was repeated on each individual comprising the bulks, and with the parents again.

This PCR was repeated to check for consistency, and reproducibility was high when all factors were held constant. From these three primers, $(CA)_8RG$ was selected for further investigation because of the bright segregating band and uncomplicated banding pattern.

Data Analysis

The recombination frequency, *r*, was calculated using the equation:

$$r = Ab + aB/ n$$

Where Ab is the number of pubescent individuals without the marker, aB is the number of smooth individuals with the segregating band (marker), and n is the total number of individuals. Ab + aB therefore equates to the total number of recombinants observed.

Conversion of Inter-Simple Sequence Repeat to Sequence Characterised Amplified Region

Isolation and Purification of ISSR Fragment

Another PCR was performed on the parents and all 16 individuals comprising the bulks with identical conditions as already described. Both bulked DNA samples were included to confirm band identity by direct comparison.

The segregating DNA fragment was excised from the agarose gel, using a sterile surgical blade, for four individuals: 'Lauranne', the smooth testa hybrid showing the fragment, and two hybrids from the pubescent testa bulk that showed the brightest amplification. The DNA from the excised band was extracted from the gel using the QIAquick Gel Extraction Kit with microcentrifuge directions (QIAGEN, Roche Molecular Systems, USA). After the gel slice was excised and weighed, three volumes of binding and solubilisation buffer was added and heated at 50°C for 10 min, and samples were vortexed three times during this incubation. One volume of isopropanol was added to the melted sample and mixed, and the entire solution was then applied to the QIAquick spin column. The sample was centrifuged at 14 000 rpm for 1 min and the flow through discarded. The column was washed with 750 μL of ethanol washing buffer and centrifuged at 14 000 rpm for 1 min. The flow through was discarded and the sample re-centrifuged at 14 000 rpm for 1 min. To elute the DNA from the column, it was placed into a clean sterile 1.5 mL Eppendorf tube and 30 μL buffer (10 mM Tris-Cl) was applied to the centre of the membrane and left to stand for 1 min before centrifuging at 14 000 rpm for 1 min. Eluted samples were stored at –20°C until required for further analysis.

DNA cloning

The plasmid vector pCR®-2.1-TOPO® was used in conjunction with the TOPO TA Cloning® Kit (Invitrogen, USA) using the conditions outlined in the supplied instruction manual with the following modifications. Each individual cloning reaction for transformation contained 3 μL of purified gel extract, 1 μL of supplied salt solution (1.2 M NaCl, 0.06 M $MgCl_2$), 1 μL sterile water, and 1 μL TOPO® vector in a 6 μL total reaction volume. This reaction was gently mixed and incubated at room

temperature (RT, 22-24 °C) for 15 min and then placed on ice. Chemically competent TOP10F' One Shot® *E. coli* were transformed by gently mixing 2 μL of cloning reaction and incubating on ice for 10 min. The cells were heat shocked at 42°C for 30 sec, placed back on ice until 250 μL of RT SOC medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl. 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM SO_4, 20 mM glucose, Invitrogen, USA) was added. The samples were then incubated in an orbital shaker at 37°C for 1 hr. Colour selective plates were prepared by coating Luria-Bertaini (LB) agar plates (1 g/100 mL BiTek Agar, 0.5% Bacto™ Yeast Extract (Difco Laboratories, USA), 1.0% tryptone, 1.0% NaCl, containing a final concentration of 50 μg/mL ampicillin) with 0.8 mg 5-bromo-4-chloro-3-indolyl-beta-D-galactoside (X-gal, Promega, USA) and 0.825 mg isopropylthio-β-D-galactoside (IPTG, Sigma, USA) and preheated at 37°C. For each transformation two plates were spread with either 30 μL or 50 μL of transformation mixture to ensure well spaced colonies. After overnight incubation at 37°C, eight positive colonies were selected for each transformation (four from each plate dilution). A single colony was picked from the plate using a sterile tip and transferred to a sterile 10 mL tube containing 3 mL of LB Broth (1.0 % tryptone, 0.5% yeast extract, 1.0% NaCl) with 150 mg of ampicillin added immediately prior to adding the bacteria, and incubated with shaking at 37°C overnight.

Plasmid preparation

Small-scale plasmid preparation was carried out using the UltraClean™ 6-Minute Mini Plasmid Prep Kit (MolBio Laboratories, USA). Cultures were transferred to 1.5 mL sterile Eppendorf tubes and centrifuged at 14 000 rpm for 1 min to pellet the cells. Two spins were necessary to centrifuge the total volume and the supernatant was decanted between spins. After the supernatant was decanted the samples were spun at

14 000 rpm for approximately 5 sec and the remaining liquid drawn off with a pipette. The cells were resuspended in 50 µL of lysis buffer by bump vortexing, 100 µL of alkaline SDS solution was added and inverted once, and 325 µL of neutralising buffer was then added and inverted once gently. The sample was then centrifuged at 14 000 rpm for 1 min, and, if the pellet was not firm, re-centrifuged for an additional 5 min. The supernatant was transferred to a spin column by pipette and centrifuged at 14 000 rpm for 30 sec and the flow-through discarded. The bound DNA was washed by adding 300 µL of ethanol wash buffer and centrifuged at 14 000 rpm for 30 sec. The flow-through was discarded and the sample centrifuged again for another 5 sec to remove residual ethanol. The spin column was then transferred to a fresh sterile 2 mL Eppendorf tube and the plasmid DNA eluted by adding 50 µL of elution buffer to the middle of the column and centrifuging at 14 000 rpm for 30 sec. The samples were stored at –20°C until required.

To check that all clones contained the correct insert, the plasmids were digested by restriction enzyme. Plasmids were digested with 5 U of *EcoR* I (Promega, USA), 14.7 µL of sterile MQ water, 2 µL of 10 x restriction buffer (900 mM Tris-HCl, pH 7.5, 500 mM NaCl, 100 mM $MgCl_2$), 0.2 µL of acetylated BSA and 3 µL of plasmid DNA, and incubated at 37°C for 90 min. Prior to electrophoresis, digests were mixed with 10 x loading dye (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) Electrophoresis was carried out on 1.5% agarose gels in 0.5 x TBE buffer at 80 mA constant current. A 100 bp ladder (Geneworks, Australia) was also loaded for fragment size identification. For each individual, one plasmid preparation containing an insert of the expected size was selected and prepared for sequencing.

Plasmid preparation for sequencing

To prepare the sequencing reaction, 2 μL of the digested plasmid DNA was mixed with 13.9 μL sterile MQ water, 4.0 μL of terminator ready reaction mix (BigDye® Terminator Kit V 3.1, Applied Biosystems, USA) and 0.32 pmol of M13 forward primer (5'-TGT AAA ACG ACG GCC AGT-3'). The same conditions were used for the M13 reverse primer (5'-CAG GAA ACA GCT ATG ACC-3') to sequence the antisense strand. The sequencing PCR conditions consisted of 25 cycles of 96°C for 10 sec, 50°C for 5 sec, and 60°C for 4 min.

An isopropanol precipitation was performed to purify the PCR product prior to sequencing. 80 μL of 75% isopropanol was added to the PCR mix and briefly vortexed. The samples were left to incubate for 15 min at RT and then centrifuged at 14 000 rpm for 20 min. After centrifuging, the supernatant was gently removed by pipette, the pellet was washed with 250 μL of 75% isopropanol and then re-centrifuged for 5 min at 14 000 rpm. The supernatant was removed by pipette and the pellet dried of all residual isopropanol by heating at 90°C in a heat block with the lids open for 1 min.

Sequencing was performed using an ABI PRISM 3700 Capillary DNA Sequencer (Institute of Medical and Veterinary Science, Adelaide).

Sequence analysis and specific primer design

To check for sequence homology between individuals and between the forward and reverse sequences, sequences were aligned (with reverse complementation of one strand) using the 'bestfit' function of WebANGIS (www.angis.org.au). All individuals sequenced with the M13 forward primer were aligned and assessed for sequence homology using the 'eclustalw' program. The longest reading frame was determined by converting the forward sequences into protein sequences in all six reading frames using the 'map' function of WebANGIS. The whole forward nucleotide sequence was

submitted to the WebANGIS data base search, 'FastA'. Both the forward nucleotide sequence and protein sequence were submitted to the NCBI (www.ncbi.nlm.nih.gov/blast) Blast-EST 'blastn' and the protein-protein blast search, 'blastp' (Altschul et al. 1997).

Primer Selection and Optimisation

Two initial primers PubescentF and PubescentR (Table 7.1) were designed automatically using the Primer3 computer program (http://frodo.wi.mit.edu) from the cloned forward sequence of 'Lauranne'. Conflicting secondary structures and primer properties were assessed using Net Primer (www.PremierBiosoft.com). These primers were synthesised by Proligo (Australia). The PCR conditions were identical to those described for the ISSR primer, except that the concentration for both the forward and reverse primers was 0.5 μM.

Table 7.1 Primer sequences and G/C base pair composition of original ISSR primers and sequence specific primers used to identify a marker linked to the pubescent phenotype

Classification	Primer Name	Sequence*	G/C content (%)
ISSR	(CT)8 G	CTC TCT CTC TCT CTC TG	53
	(CT)8RC	CTC TCT CTC TCT CTC TRG	56
	(CA)8RG	CAC ACA CAC ACA CAC ARG	56
Sequence Specific	PubescentF	GGG TTC GGG TCG TCT TTT AT	50
	PubescentR	TGT TTT CCT TTC TTT TCC GTC T	36
	PubescentR2	CCA CCA TGC TAC AAA CAA	44
SCAR	ISSRspecificF	CAC ACA CAC ACA CAC AGG TAT CAT	46
	ISSRspecificR	CAC ACA CAC ACA CAC AAG TGG TT	48

*R = A or G

A replacement reverse primer, PubescentR2 (Table 7.1), was designed using Primer3, assessed using Net Primer, and synthesised by Proligo (Australia). This primer was used together with the original forward primer, PubescentF, under the same PCR

conditions. Optimisation of this primer pair was carried out by varying the annealing temperature, cycle number, magnesium and template concentration, the addition of T4 Gene 32 Protein (Roche Diagnostics, USA), and the inclusion of the original ISSR primer at a concentration of 0.5 μM. In order to determine the optimal annealing temperature for the primer pair, PubescentF and PubescentR2, an annealing temperature gradient was set up between 61°C to 51°C with 8 increments between these temperatures (61°C, 60.3°C, 59.1°C, 57.2°C, 54.7°C, 53.0°C, 51.8°C, 51.0°C) using a Bio-RAD icycler (Bio-RAD, Australia).

A second set of primers, ISSRspecificF and ISSRspecificR (Table 7.1), were designed and screened for secondary structures and internal binding sites using MacVector (Accelrys, USA), which incorporated the original ISSR sequence with an additional 7 bp anchor, and synthesized by Proligo (Australia). These primers were optimised by altering the annealing temperature and number of PCR cycles. The final optimised PCR conditions were 40 ng of genomic DNA, 1.5 mM $MgCl_2$, 0.5 μM Forward and Reverse primer, 0.2 μM dNTPs, 1 x cresol red dye, 1 x PCR buffer, and 1.0 U Taq Polymerase (Invitrogen, USA). PCR was carried out under the following conditions; 95°C for 3 min, followed by 23 cycles of 95°C for 30 sec, 54°C for 45 sec, and 72°C for 1 min, followed by a final extension step of 72°C for 10 min.

Assessment of Primer Efficaciousness

To test the reliability of the ISSR specific SCAR primers, they were tested on the parents and all sixteen individuals comprising the bulks. The primers were then screened on the parents and all 93 F_1 individuals. This PCR was repeated to check for reproducibility and consistency in the results.

Verification of SCAR Amplification

In order to ascertain the correct amplification of the original sequenced fragment using the SCAR primers, PCR was repeated on the three individuals for which the original ISSR sequence was obtained ('Lauranne' and two hybrids with pubescent testa). This PCR product was cloned and sequenced using the plasmid vector pCR®-2.1-TOPO® as outlined in the preceding methods (Conversion of Inter-Simple Sequence Repeat to Sequence Characterised Amplified Region). The only modification was in the initial cloning reaction where 4 μL of fresh PCR product was used in place of the gel extract, with 1 μL of salt solution and 1 μL of TOPO® vector to a total volume of 6 μL. All three forward sequences were aligned with their complementary anti-sense reverse sequence, and showed 99-100% sequence similarity using the 'besfit' function of WebANGIS.

To determine sequence similarity between all the forward sequences of the original ISSR and the SCAR amplification, the 'multialignment' function of WebANGIS was used. Base changes, expressed as a percentage similarity, were calculated by visually scoring nucleotide differences between all six sequences with the derived consensus sequence.

Application of SCAR Testa Pubescence Marker

The SCAR primers were screened on twelve almond cultivars; 'Baxendale', 'Carmel', 'Chellaston', 'Ferragnés', 'Johnsohn', 'Keanes', 'Lauranne', 'Mission', 'Nonpareil', 'Somerton', 'Strout', and 'Thompson'. Two kernels from each of these cultivars were assessed organoleptically and the testa classified as either smooth ('Mission', 'Nonpareil' and 'Strout') or pubescent ('Baxendale', 'Carmel', 'Chellaston', 'Ferragnés', 'Johnsohn', 'Keanes', 'Lauranne', 'Somerton', and 'Thompson'. The PCR

conditions and electrophoresis were as outlined above (Primer selection and optimisation).

RESULTS

Organoleptic Analysis

Table 7.2 shows the sensory data for all four panelists for each selected tree over three years. The results from two of the panelists (C and D) for 2002 and 2003 showed a strong skew toward one particular mouthfeel, pubescent for panelist C and smooth for panelist D. Consistency for all three years for all the four panelist varied from 100% agreement to 41.6%, indicating the sensory evaluation of kernels of some trees was difficult to determine.

Individuals selected for inclusion for BSA bulks showed an agreement of above 66.6% (Table 7.2). Four out of the eight hybrids selected to comprise the pubescent bulk showed 100% agreement, no hybrids in the smooth bulks showed 100% agreement, indicating that this phenotype may be more difficult to score accurately.

Stereo Microscopy

When viewed under the stereo microscope a difference was observed between the testa surfaces of the two parents (Figure 7.1). 'Lauranne' (LA) testa showed discrete protuberances arising from the epidermal surface, interdispersed with cells that collapsed (Figure 7.1B). In contrast the phenotypically smooth testa parent 'Nonpareil' (NP) showed only collapsed cells (Figure 7.1A). The F_1 hybrids were intermediate between the two parents, having both expanded and collapsed cells (Figures 7.1CD). The phenotypically smooth hybrid (Figure 7.1C) showed expanded cells that appeared

smaller and less uniform in shape than those observed in the pubescent hybrids (Figure 7.1D).

Light Microscopy

The protuberances fitted the classification for trichomes since they were epidermal cells with a high length/width ratio (Wagner et al. 2004). The trichomes arose from the epidermal layer of the testa which appeared to be composed of two layers. A section of testa prepared from an immature kernel of NP at *ca* 24 weeks of development (Figure 7.2) showed expanded epidermal trichomes, with the sub-epidermal cell layer comprising a number of expanded cell layers. Figure 7.3 shows a section of the testa and cotyledon tissue of NP at kernel maturity with collapsed trichome epidermal and sub-epidermal layers of isodiametric cells, cotyledon epidermal cells and cotyledon parenchyma cells containing lipid and protein bodies. The region of detachment of the testa from the surface of the cotyledon occurred at the junction of the testa and the isodiametric cells of the cotyledon epidermis (Figure 7.4).

At nut maturity LA trichomes were thick walled, large and mainly expanded with a few collapsed, as was also observed in the hybrids with pubescent testa (Figures 7.5B and D). NP showed only collapsed trichomes, with a crater or dish-like appearance in cross section (Figure 7.5A), the hybrids with a smooth testa showed a phenotype intermediate between the two parents (Figure 7.5C). These were both collapsed and expanded trichomes, but the expanded cells were more irregular in shape, and smaller in comparison to those of the pubescent hybrids (Figure 7.5C).

Scanning Electron Microscopy

Scanning Electron Microscopy (SEM) was used to analyse the surface structure of the testa under higher magnification and resolution than the stereo and light

Table 7.2 Raw data from sensory panel, for panelists A, B, C and D, for 2001-2003, and assignment for Bulked Segregant Analysis. Percentage of smooth scores given as a percentage of the total scores. S= Individual scored for smooth testa, P= Individual scored for pubescent testa, - = no data

Parent/Hybrid	Taster and Year												Number of smooth	Total n	% smooth	BSA Assignment
	2001				2002				2003							
	A	B	C	D	A	B	C	D	A	B	C	D				
'Nonpareil'	S	S	P	S	S	-	-	-	S	S	P	S	7	9	78	**Smooth**
'Lauranne'	P	P	P	P	P	-	-	-	P	-	-	-	0	6	0	**Pubescent**
Smooth	S	S	P	S	S	S	S	S	S	S	P	S	10	12	83	**Smooth**
Smooth	S	S	P	S	S	S	P	S	S	S	P	S	9	12	75	**Smooth**
Smooth	S	S	P	S	S	S	S	S	S	S	P	S	10	12	83	**Smooth**
Smooth	S	S	P	S	P	-	-	-	S	S	P	S	6	9	67	**Smooth**
Smooth	S	S	P	S	P	S	S	S	P	S	P	S	8	12	67	**Smooth**
Smooth	P	S	P	S	S	S	P	S	S	S	P	S	8	12	67	**Smooth**
Smooth	S	S	P	S	P	S	P	S	S	S	P	S	8	12	67	**Smooth**
Smooth	P	S	P	S	P	S	P	S	P	S	P	S	6	12	50	Not used
Smooth	S	S	P	S	S	-	-	-	P	S	P	S	6	9	67	**Smooth**
Pubescent	P	P	P	P	P	P	-	-	P	P	P	S	1	9	11	**Pubescent**
Pubescent	P	P	P	P	P	P	P	P	P	P	P	S	1	12	8	**Pubescent**
Pubescent	P	P	P	S	P	P	P	P	P	P	P	S	2	12	17	**Pubescent**
Pubescent	P	P	P	S	P	-	-	-	P	P	P	S	2	9	22	**Pubescent**
Pubescent	P	P	P	S	P	P	P	P	P	P	S	S	3	12	25	Not used
Pubescent	S	P	P	S	P	-	-	-	P	P	P	S	3	9	33	Not used
Pubescent	P	P	P	P	P	-	-	-	P	P	P	P	0	9	0	**Pubescent**
Pubescent	P	P	P	P	P	-	-	-	P	P	P	P	0	9	0	**Pubescent**
Pubescent	P	P	P	P	S	-	-	-	P	P	P	S	2	9	22	Not used
Pubescent	P	P	P	S	P	S	P	P	P	P	P	S	3	12	25	Not used
Pubescent	P	P	P	S	S	P	P	P	P	P	P	S	3	12	25	Not used
Pubescent	P	P	P	S	P	-	-	-	P	P	P	S	2	9	22	Not used
Pubescent	P	P	P	S	P	P	P	P	P	P	P	S	2	12	17	Not used
Pubescent	P	P	P	P	P	P	P	P	P	P	P	P	0	12	0	**Pubescent**
Pubescent	P	P	P	P	P	-	-	-	P	P	P	P	0	9	0	**Pubescent**

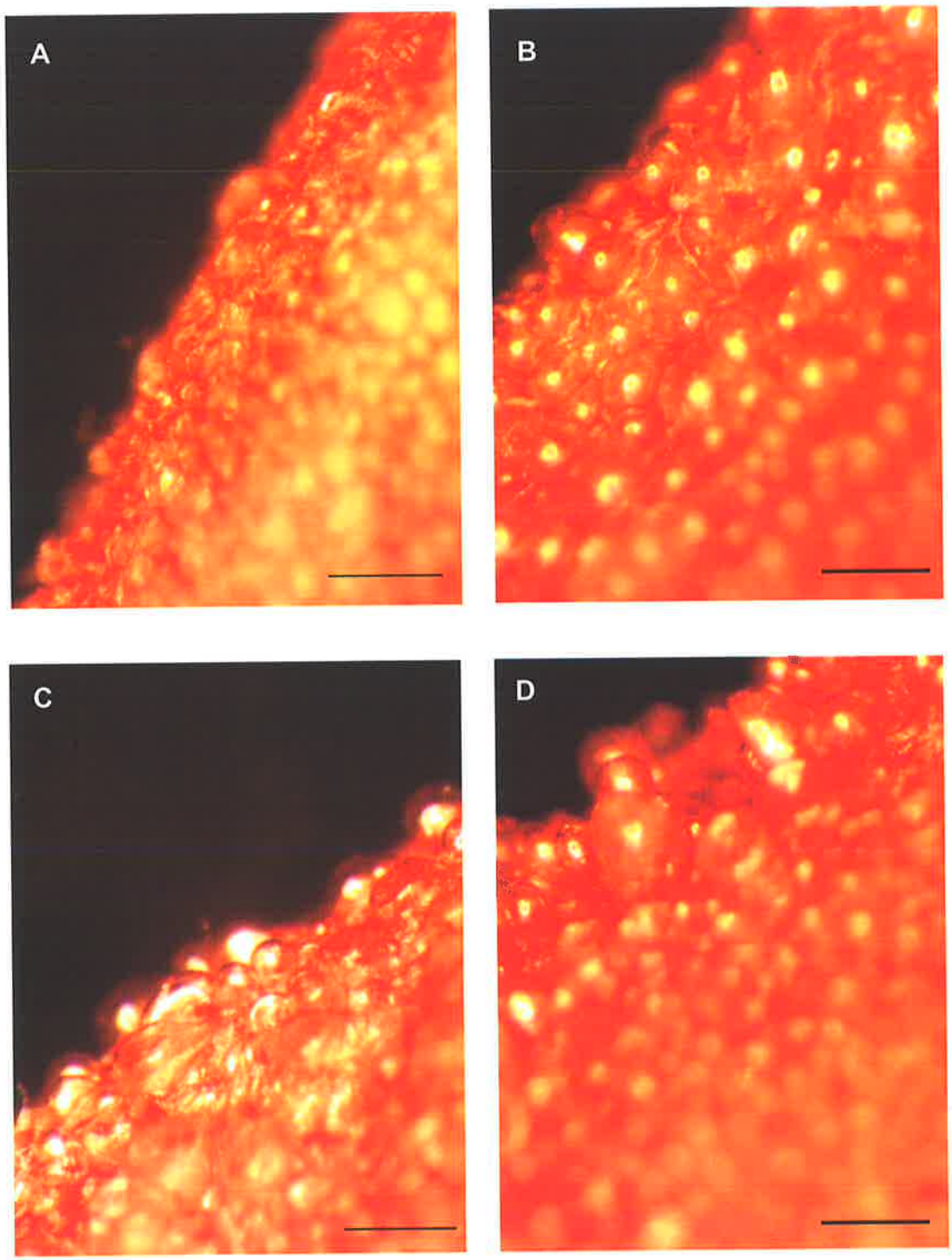

Figure 7.1 Stereo micrograph of testa surface 'Nonpareil' (A), 'Lauranne' (B), Hybrid with smooth testa (C), Hybrid with pubescent testa (D). Scale bar = 100 μm

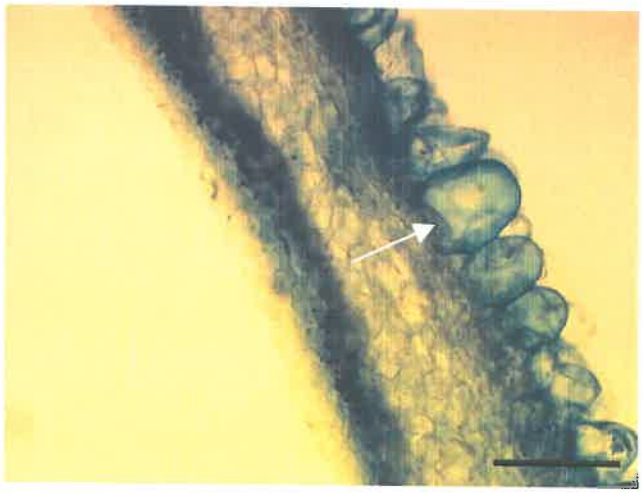

Figure 7.2 Light micrograph of transverse section of immature 'Nonpareil' almond testa (*ca* 24 weeks) arrow showing enlarged cells, which are not visible at nut maturity, stained with Toluidine Blue. Scale bar = 300 μm

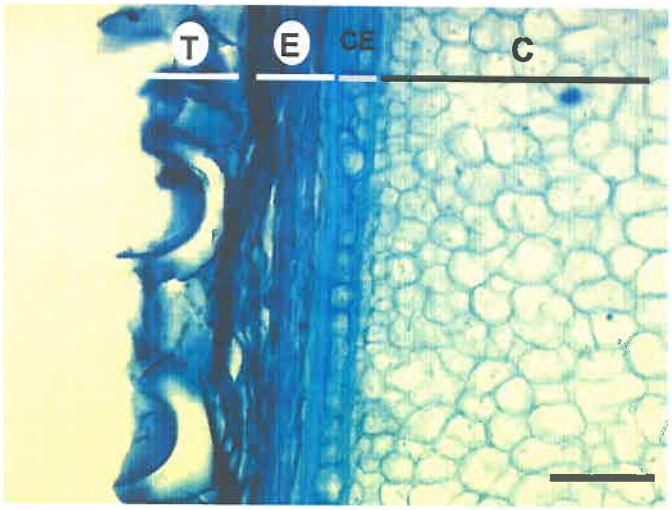

Figure 7.3 Light micrograph of cross section of 'Nonpareil' almond testa and cotyledon tissue showing collapsed trichomes (T), collapsed sub-epidermal layer (E), cotyledon epidermis (CE) and cotyledon tissue (C), stained with Toluidine Blue. Scale bar = 100 μm

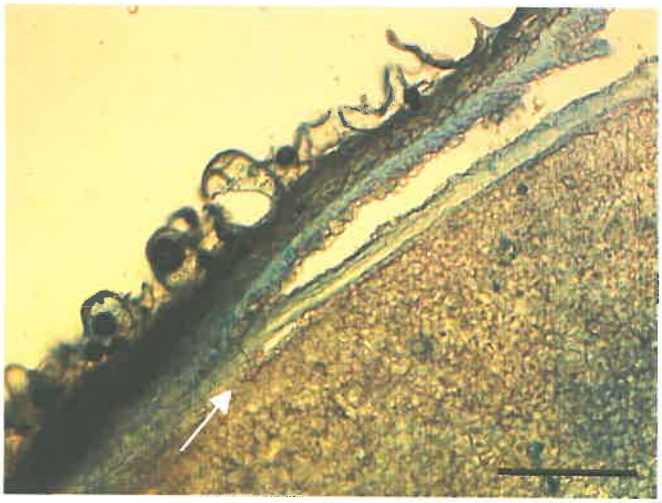

Figure 7.4 Light micrograph of section of almond testa and cotyledon tissue of 'Lauranne' showing cleaving of testa tissue from the cotyledons (arrow) at the junction of the testa and cotyledon epidermal tissues, stained with Toluidine Blue. Scale bar = 300 μm

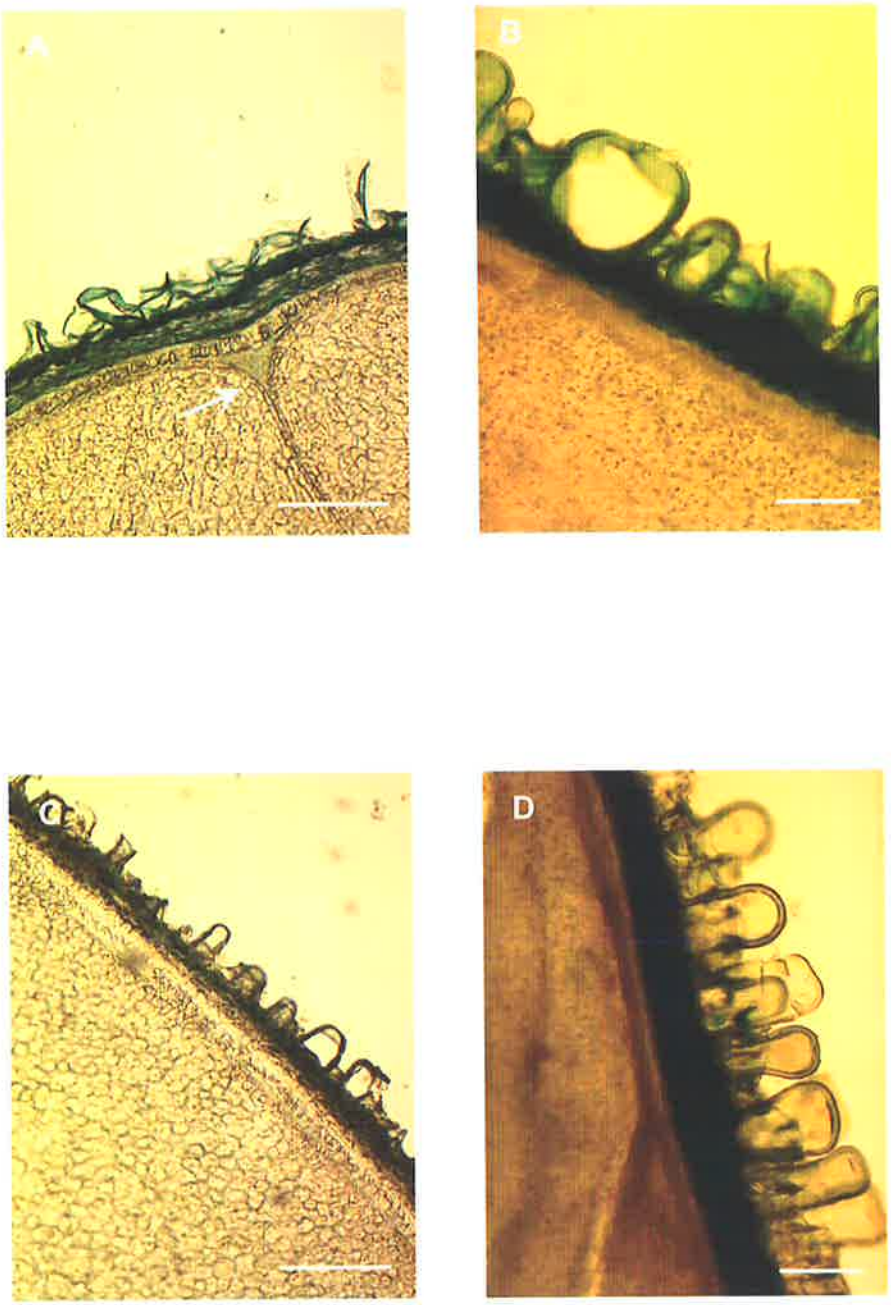

Figure 7.5 Light Micrographs of sectioned testa tissue 'Nonpareil' (A), 'Lauranne' (B), Hybrid with Smooth Testa (C), Hybrid with Pubescent Testa (D), stained with Toluidine Blue. Arrow showing vascular trace. Scale bar = 300 μm

microscope. The cellular arrangement comprised mainly expanded cells in LA and both hybrids (Figure 7.6B, C, D), and only collapsed cells observed for NP (Figure 7.6A). The similarity between the cells observed using the stereo microscope, light microscope and SEM illustrates that sample preparation is not likely to be responsible for the differences observed between the two cell types, or between individuals. The cell wall of the trichomes maintains its structural integrity when sectioned and prepared for SEM, indicating it is composed of a rigid substance at the developmental stages kernels were prepared for this study. Various stages of cellular collapse were evident, when the testa was viewed from above, as shown in Figure 7.6, both LA and the pubescent hybrid had expanded cells that were fairly uniform (Figure 7.6A and D), whereas the hybrid with a smooth testa showed expanded cells that were distorted and elongated, which was also shown in some collapsed cells (Figure 7.6C). When these individuals were viewed in side profile as shown in Figure 7.7, the differences in cell uniformity and height were more evident, collapsed cells appeared to have a cell wall that had sunken into a depression, resulting in folding and creasing of the surface (Figure 7.8). Expanded cells appeared to have a balloon shape, with a restricted pentagonal base, with some cells showing a slightly sunken top surface. Various stages of cellular collapse could be seen on the one section of the testa, as shown in Figure 7.9.

All trichomes showed a number of pits in the lower section and base, which appear to traverse the cell wall (Figure 7.10). The trichomes were unicellular, as indicated from the light microscopy work. The rigid cell wall of the LA trichome was approximately 4 μm thick.

Data Analysis

Cell diameter of both expanded and collapsed cells and density of collapsed cells were measured from SEM images for the parents and hybrids. The mean cell diameters,

are presented in Table 7.3, the LSD values are given in Table 7.4. Cell diameters of all four genotypes, both collapsed and expanded, were significantly different except the collapsed cell diameter of the pubescent hybrid and LA. NP had a significantly larger collapsed cell diameter than the others, and the smooth hybrid had the shortest collapsed cells. All values obtained for the expanded cell type were significantly different, with the hybrid with pubescent testa having the largest cell diameter, and the hybrid with the smooth testa the shortest. LA and the two hybrids show a significantly larger cell diameter when the cell was expanded than when collapsed.

The density of collapsed cells for each individual is given in Table 7.5. Both hybrids showed a collapsed cell density lower than the parents, the hybrid with the pubescent testa showed the lowest density of collapsed cells, while NP showed 100% collapsed cells, since no expanded cells were ever observed. The low sample size precludes accurate statistical analysis of these data.

Bulked Segregant Analysis

From an initial screen of eight ISSR primers, three produced a segregating band between the parents and two bulks, $(CT)_8G$, $(CT)_8RC$, and $(CA)_8RG$. Both $(CT)_8G$ and $(CA)_8RG$ produced a fragment corresponding to the pubescent phenotype of the paternal genotype 'Lauranne' and the hybrids with the pubescent testa, whereas $(CT)_8RC$ produced a fragment corresponding to the smooth phenotype of the maternal genotype 'Nonpareil' and the hybrids with the smooth testa. This PCR was repeated to check for consistency, and reproducibility was high when all factors were held constant. When these primers were screened on all the individuals comprising the bulks they showed a recombination frequency of 0.18, 0.25 and 0.18 for $(CT)_8G$, $(CT)_8RC$, and $(CA)_8RG$ respectively (Table 7.6). From this initial screen, ISSR primer $(CA)_8RG$ was selected

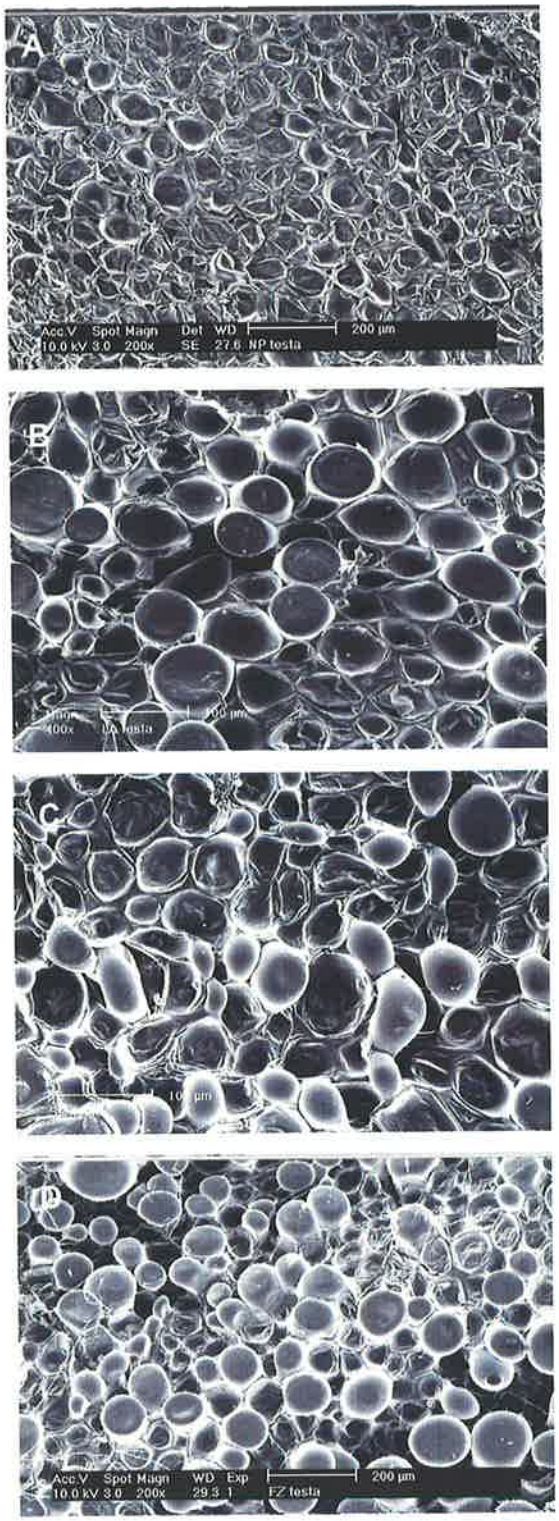

Figure 7.6 Scanning electron micrograph of almond kernel testa surface showing differences in the number and size of expanded and collapsed epidermal trichomes. For 'Nonpareil' (A) only collapsed cells are present, 'Lauranne' (B) shows both collapsed and expanded cells, which are also evident in the hybrids with smooth testa (C) and those with a pubescent testa (D).

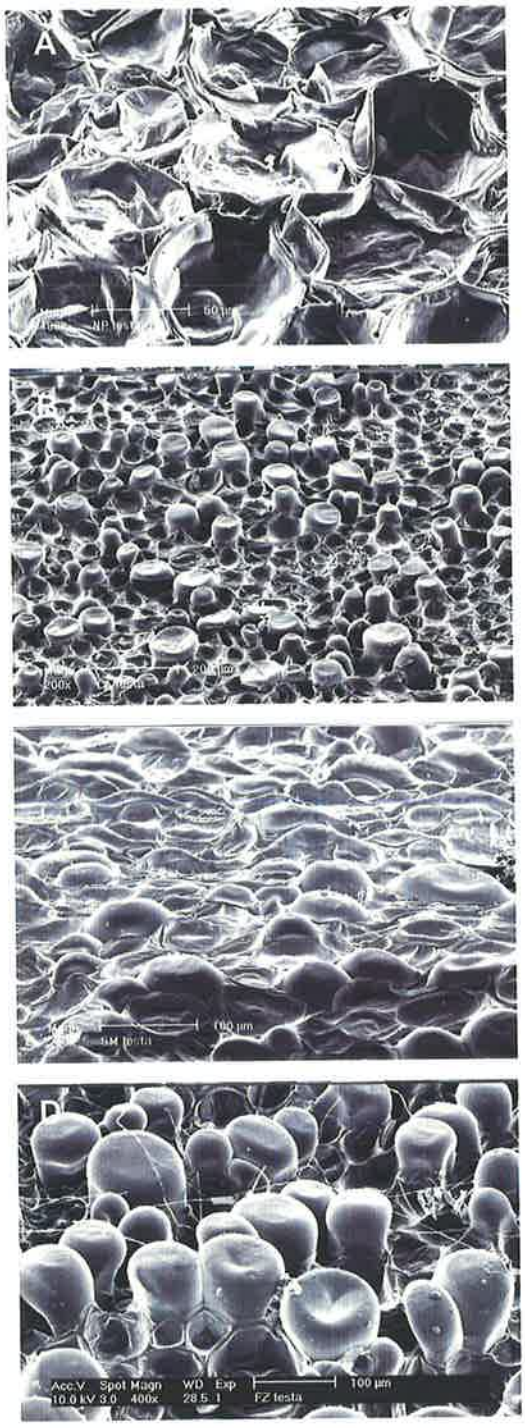

Figure 7.7 Scanning electron micrograph of almond kernel testa surface showing differences in the number and size of expanded and collapsed epidermal trichomes. For 'Nonpareil' (A) only collapsed cells are present, 'Lauranne' (B) shows both collapsed and expanded cells, which are also evident in the hybrids with smooth testa (C) and those with a pubescent testa (D).

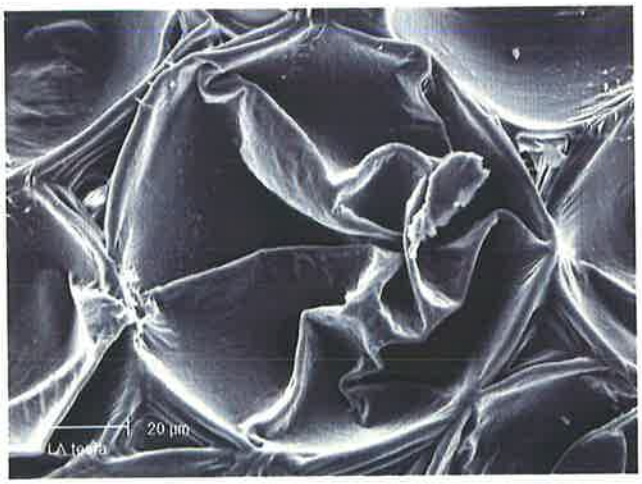

Figure 7.8 SEM micrograph of the surface of 'Lauranne' almond testa showing collapsed cell wall folded into a depression

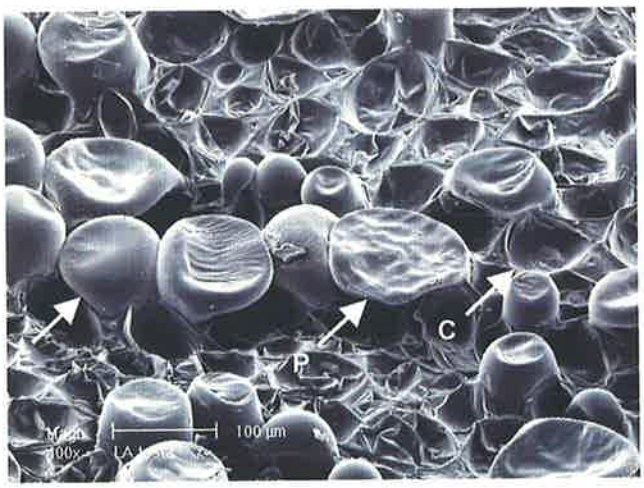

Figure 7.9 SEM micrograph of 'Lauranne' almond testa surface showing trichomes in various stages of collapse, from fully expanded (arrow E), partially collapsed (arrow P) and wholly collapsed (arrow C)

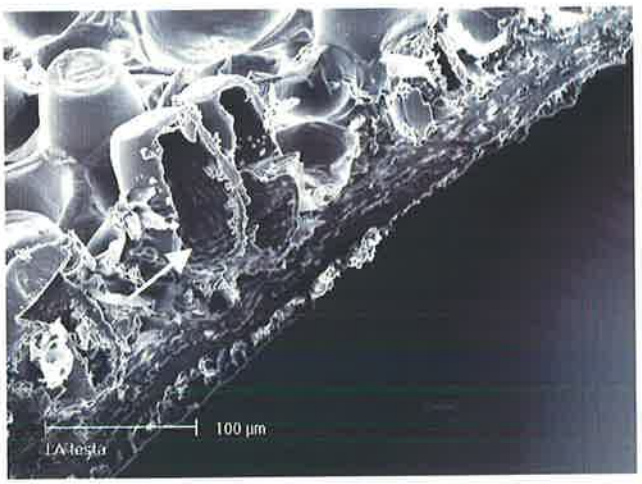

Figure 7.10 SEM micrograph of transverse section of 'Lauranne' almond testa arrow showing pitted cell wall and base

Table 7.3 Average mean maximum diameter (μm) of testa cells, significant differences between groups are given in Table 7.4.

	Cell Type	
	Collapsed	**Expanded**
'Nonpareil'	102.6	-
'Lauranne'	82.7	84.4
Hybrid with Smooth Testa	70.2	73.3
Hybrid with Pubescent Testa	87.5	111.7

Table 7.4 Absolute differences between cell diameter means of individuals (back-transformed on a log scale) for LSD analysis. Values of differences between means for collapsed cells are in open boxes, values for expanded cells are given in shaded boxes.

	'Nonpareil'	'Lauranne'	Hybrid with Smooth Testa	Hybrid with Pubescent Testa
'Nonpareil'				-
'Lauranne'	0.216*		0.142*	0.280*
Hybrid with Smooth Testa	0.379*	0.163*		0.422*
Hybrid with Pubescent Testa	0.159*	0.057	0.220*	

(LSD equals 0.0914 at 5% level, significant differences are marked by an asterisk *)

Table 7.5 Average density of collapsed cells expressed as a percentage of the total number of cells per field of view

Individual	**Collapsed Cell Density (%)**
'Nonpareil'	100
'Lauranne'	68.7
Hybrid with Smooth Testa	65.6
Hybrid with Pubescent Testa	63.9

for further investigation based on the production of a bright segregating band, the absence of other closely spaced bands, and a relatively low recombination frequency.

Conversion of ISSR to SCAR

A band of 733 bp amplified by $(CA)_8RG$ was found to segregate in the individuals comprising the DNA bulk, this band was cloned and sequenced from the parent contributing the band ('Lauranne'), two pubescent hybrids, and the smooth hybrid which showed the band from the bulked DNA samples (Figure 7.11). The fragment of 733bp, cosegregating with the pubescent testa phenotype, has been designated $(CA)_8RG$-733. The forward and reverse sequence alignments from each individual showed a percentage similarity of between 99.9% and 95.9%. When all four forward sequences were aligned they also showed a high percentage similarity. The longest forward sequence with the least sequencing errors was selected and the longest protein sequence was determined by reading in all six frames. The corresponding protein sequence of 237 amino acids was submitted to the 'blastp' protein search, but failed to show any significant homology to sequences published in the database. The whole forward nucleotide sequence was submitted to the 'FastA' and 'blastn' databases. The cloned sequence showed homology to a number of short sequences from *Homo sapiens* BAC clones but no significant homology to any known functional sequences. From the 'blastn' EST division search, nine sequences were identified with a sequence identity of between 100% (40 bp) to 90% (180 bp). The longest sequence corresponded to an EST from a half-ripe apricot fruit Lambda Zap II library (Genebank assession number: CB818522) (Mbeguie-A-Mbeguie et al. 1999).

Based on the derived sequence of the $(CA)_8RG$-733 fragment, two oligonucleotides of 20 and 22 bases were synthesised and used as sequence specific primers. The forward and reverse primers PubescentF and PubescentR that were

designed from this initial sequence failed to produce a clear banding pattern under any of the conditions trialled for optimisation. The reverse primer contained a high percentage of thymine nucleotides, 63% (Table 7.1) although Net Primer failed to detect any conflicting secondary structures such as hairpins, dimers, palindromes, or repeats. This primer was substituted with a replacement reverse primer, PubescentR2, which had a lower thymine content but showed an internal hairpin when assessed using Net Primer.

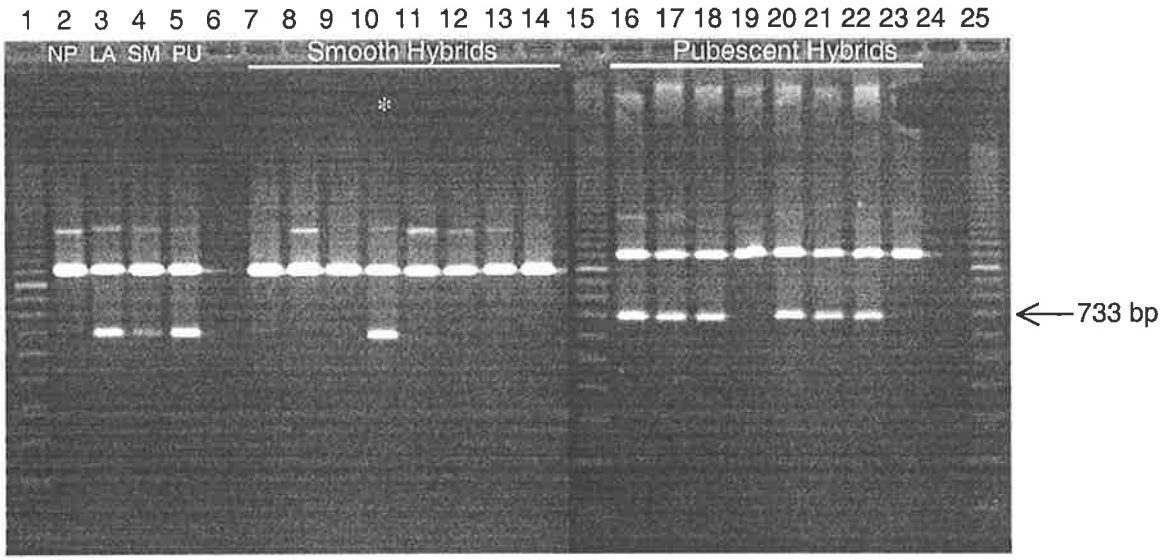

Figure 7.11 Ethidium bromide stained garose gel showing ISSR banding pattern for $(CA)_8RG$, segregating band indicated by arrow at 733 bp. Lane 1 100 bp molecular ladder, lane 2 'Nonpareil', lane 3 'Lauranne', lane 4 Smooth bulk (note faint band caused by one recombinant individual *), lane 5 Pubescent bulk, lane 6 Blank, lanes 7-14 hybrids with smooth testas, lane 15 100 bp molecular ladder, lanes 16-23 hybrids with pubescent testas, lane 24 negative control, lane 25 100 bp molecular ladder

To eliminate this problem, three bp were dropped from the 3' end of the primer selected by Primer3. The primer combination Pubescent F and Pubescent R2 also failed to produce a PCR product of the expected size under any conditions trialled, including adjusting the number of cycles, annealing temperature, magnesium and template DNA concentration.

A second set of primers was then selected that contained all the CA repeats from the original ISSR primer $(CA)_8RG$, with an additional 7 bp added to the 3' end of the sequence. These primers were designated as ISSRspecificF and ISSRspecificR, the sequence characterised amplified region (marker) is designated ISSRspecific. After optimisation, these primers produced a distinct product of approximately 700 bp which segregated in the parents and progeny comprising the bulks. The recombination frequency of the original ISSR primer and the ISSR specific primers was calculated as 0.18 (Table 7.6). When these primers were screened on the whole progeny the recombination frequency increased to 0.34 for the former and 0.51 for the latter (Table 7.7).

The application of this SCAR marker was tested by screening the primers on twelve cultivars that had been organoleptically assessed as either having a smooth or pubescent testa. Only one individual of the twelve ('Chellaston') did not show a band when it was expected to, and a recombination frequency of 0.08 was calculated for this evaluation (Figure 7.12).

DISCUSSION

The control of testa pubescence in almond has not been established and there is a paucity of information in the literature regarding the effect of surface cellular structure on organoleptic/sensory stimulation.

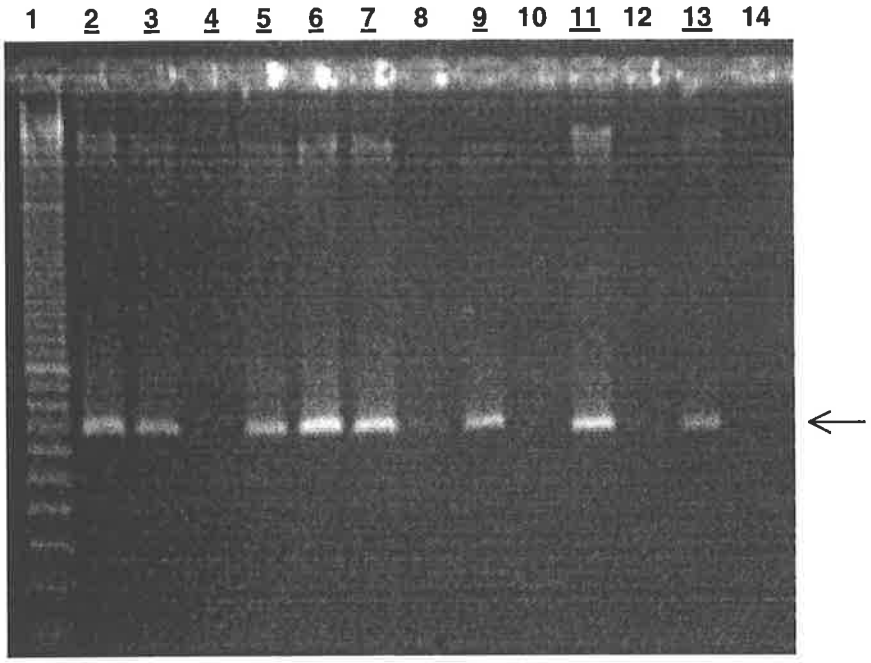

Figure 7.12 SCAR marker amplified with primers PubescentF and PubescentR2 denoted by arrow on twelve almond cultivars,. Lane 1 100 bp molecular ladder, lane 2 'Baxendale', lane 3 'Carmel', lane 4 'Chellaston', lane 5 'Ferragnés', lane 6 'Johnston', lane 7 'Keanes', lane 8 'Mission', lane 9 'Lauranne', lane 10, 'Nonpareil', lane 11 'Somerton', lane 12 'Strout', lane 13 'Thompson' lane 14 negative control. Lane numbers underlined correspond to cultivars assessed as having a pubescent phenotype

Table 7.6 *r*-values for individuals used for BSA with three ISSR primers, and SCAR primers

Locus		Phenotype*				*r*
A	B	AB (n)	Ab (n)	aB (n)	ab (n)	
Pubescent	$(CT)_8$ G	6	2	1	7	0.18
Smooth	$(CT)_8$RC	5	3	1	7	0.25
Pubescent	$(CA)_8$RG	6	2	1	7	0.18
Pubescent	SCAR	6	2	1	7	0.18

* AB = phenotype (A) with marker (B), Ab = phenotype (A) without marker (B), aB = phenotype (a) with marker (B), ab = phenotype (a) without marker (B). Ab + aB is total number of recominants

Table 7.7 *r*-values for whole progeny using SCAR primers and ISSR primer

Locus		Phenotype*				*r*
A	B	AB (n)	Ab (n)	aB (n)	ab (n)	
Pubescent	$(CA)_8$RG	59	25	8	2	0.34
Pubescent	SCAR	38	47	2	8	0.51
$(CA)_8$RG	SCAR	40	28	0	27	0.29

* AB = phenotype (A) with marker (B), Ab = phenotype (A) without marker (B), aB = phenotype (a) with marker (B), ab = phenotype (a) without marker (B). Ab + aB is total number of recominants

This preliminary study has identified a difference in testa surface cellular structure between the almond cultivars 'Nonpareil' and 'Lauranne' and their F_1 hybrids. Organoleptic evaluation of the two parents and the hybrids distinguished two classes, those that were smooth in texture/mouthfeel, and those that were pubescent. The panel used in this study was also able to detect a difference in the mouthfeel of the kernel surface, this lead to the examination of the microstructure of the testa as a proposed difference for the mouthfeel between these two types.

Investigation of the kernel surface using microscopy showed that there is a potential association between the mouthfeel of the testa and the number of cells projecting from the epidermal layer of the testa of mature nuts. These protuberances fit the classification of trichomes (Wagner et al. 2004), and they were observed to be simple and appear to be non-glandular. The perception of a difference between smooth and pubescent types by the tasting panel could be attributable to the state of the trichomes at nut maturity. Two cell states were observed, those that appeared to be expanded and those that were collapsed. A gradation in the mouthfeel perception of pubescence was reported by members of the tasting panel, which may be linked to the density of expanded cells in comparison to collapsed cells and the height and length of these cells. The smooth parent NP showed only collapsed cells on the testa surface at nut maturity indicating that all trichomes on the surface of the testa collapsed at some stage of development. The pubescent parent LA showed trichomes that were both expanded and collapsed. In the F_1 hybrids, two classes were observed corresponding to the organoleptic phenotypes of the parents. The pubescent hybrid nuts showed a cellular arrangement similar to the pubescent parent, whereas the smooth hybrid nuts had fewer and shorter trichomes in comparison. In order to make a tenable conclusion to whether expanded cell density or height is responsible for the differences in mouthfeel of

pubescent versus smooth hybrids a greater sample size needs to be utilized. However some preliminary conclusions can be drawn from the presented data.

There appeared to be some form of genetic control over this trait, which governs the perceived mouthfeel of the kernel as pubescent or smooth. This was determined by the development of a molecular marker using BSA. By comparing the ISSR banding patterns between a pooled DNA sample of eight individuals showing the pubescent phenotype to a sample of eight smooth individuals a band was identified with putative linkage to the trait of interest. Michelmore et al. (1991) calculated the probability of individuals in two bulks having the band unlinked to the marker as 2×10^{-6} for bulks of 10 individuals, which correlates to 3.5×10^{-5} for eight individuals, therefore the chances of observing false positives are relatively small using this method. The fragment derived from the pubescent parent 'Lauranne' and the converted SCAR was used to screen each individual comprising the bulks and the 93 individuals of the mapping population. Conversion of the ISSR to a SCAR marker was undertaken in order to provide a more robust marker type for this form of analysis (Paran and Mitchelmore 1993). In particular, the use of SCAR markers limits the misclassification of bands from the co-migration of equal sized bands with different sequences, which has been observed using RAPDs (Gu et al. 1999; Mekuria et al. 2002). This could potentially give rise to skewed data, and since SCAR markers are sequence specific this source of error is effectively removed.

The SCAR marker amplified a band in 75% of the pubescent individuals and this product was absent in 87.5% of the smooth individuals comprising the population used for BSA. The recombination frequency of the SCAR marker with the smooth and pubescent phenotype of the bulks was below the value for random assortment, indicating that the marker was linked to the trait of interest. This suggests that the

marker is linked to the trait for pubescence at a distance of approximately 18 cM from the gene. When this primer was tested on the total segregating population there was not a close link to the trait. Of the purported 85 individuals classified as pubescent, only 45% showed the expected banding pattern. When both the ISSR primer $(CA)_8RG$ and the SCAR marker were tested on the parents and 93 progeny the *r*-value increased to 0.34 and 0.51 respectively. The discrepancy between the *r*-values obtained from the BSA classifications and values from the whole progeny may be due to the misclassification of smooth individuals in the population as pubescent. Several hypotheses can be proposed to account for this discrepancy. Firstly texture is a sensory property, and analysis of this trait is therefore an individual sensory perception. Several factors contribute to this trait, and it is therefore a multi-parameter attribute. Texture is classified as a sensory property (Surmacka Szczeniak 2002), and gender, age and personality affect sensory perception (Stevens 1996). Colour is also a key attribute for many taste profile preconceptions (Koch and Koch 2003). In this case the smooth parent had a pale testa and the pubescent parent a dark testa. Panelists may have had a preconceived notion of the pubescence level based on the colour of the testa. The demarcation of the pubescent character is based on a number of geometrical factors such as the level, size and shape of the trichomes as observed by microscopy techniques. These temporal cues and differences in cell type may influence the ability of the panelist to determine the level of pubescence accurately as is shown in a generic study by Cascio and Sathian (2001). Therefore misclassification of the F_1 hybrids by sensory analysis alone may be due to the difficulties of determining distinct classes. In order to assess this hypothesis, the SCAR marker was tested on twelve almond cultivars that showed distinctly pubescent or smooth testas. The SCAR marker showed a recombination frequency of 0.08, which was much closer to the level observed in the BSA analysis.

Screening of this marker in other breeding populations segregating for this trait and submitting kernel samples to a panel for sensory analysis to eliminate individual bias is required to validate this markers utility for screening of this trait.

Although the function of these trichomes on almond kernel testa is unclear at this stage, a developmental sequence can be proposed based on the data collated from SEM of both phenotypes. Hawker and Buttrose (1980) suggested that early in kernel development these cells are darkly stained, and peaks in the accumulation and concentration of starch and reducing sugars occur in the testa tissue between 12-16 weeks after flowering. This accumulation corresponds to the displacement of the nucellus tissue by the endosperm and embryo (Hawker and Buttrose 1980). Therefore a potential role for the testa is the storage of sugars which are then translocated to the developing endosperm and embryo. This could occur via plasmodesmata, as pits were observed in the trichome walls, and through the linked vascular network. The cells comprising the testa layer started to collapse at week 14, and by week 28 (kernel fully mature) had wholly collapsed to form a dry, furry covering over the embryo (Hawker and Buttrose 1980). This covering was observed in the present study where immature nuts of NP at approximately 24 weeks after pollination, showed a high proportion of expanded trichomes and the underlying epidermal layer was not as compact as observed in the mature nut. NP may have a later onset of cellular collapse than the cultivars studied by Hawker and Buttrose (1980). This may be triggered by environmental cues, which may explain the low heritability observed for this trait (Arteaga and Socias i Company 2002; Kester et al. 1977).

A difference in the cell wall thickness of the protruding expanded trichomes could be a factor contributing to the rate and level of cellular collapse observed between the two phenotypes. Young et al. (2004) observed the large expanded cells had a thin

cell wall in their study of NP. It is possible that individuals with the pubescent phenotype have cell walls that are thicker, thus more resistant to cellular collapse than the thinner walled cells of smooth testa individuals. This needs further investigation to establish if this is a possible reason for the difference in cell collapse levels between the phenotypes. Another hypothesis to account for the differences observed between smooth and pubescent kernels in this study may be that an alteration in the signals initiating trichome induction and suppression affects the rate of expansion/collapse in these cells. In *Arabidopsis*, over 40 genes have been identified as having a role in trichome development (Wagner et al. 2004), and the levels of trichome inducing factor (TCIF) and trichome suppressing factor (TCSF), in combination with other regulators, control trichome development (Kragler et al. 1998). If a similar situation operated in the control of trichomes in almond testa, then in early development of the seed TCIF might initiate trichome development. At approximately 12 weeks these cells fill with sugars and between 12-16 weeks the sugars move out of the cells into the developing endosperm and embryo. Once the cell is void of contents, it may undergo programmed cell death at which time another regulator may be expressed causing the cells to collapse. Since cells of the hybrid F_1 smooth type may have fewer and smaller trichomes it may be that a regulator such as TCSF lowers the number and size compared to the pubescent type.

Because the smooth testa phenotype is fairly scarce in comparison to the pubescent type in almond germplasm it is likely that the pubescent phenotype is the wild type. The presence of a completely smooth testa may be a result of over-expression of the cellular collapse regulator since it appears that these cells were once fully expanded, as observed in the immature nuts of 'NP'.

In addition to, or as an alternative to the function of the testa as an energy storage organ, other roles for these cells could be in reducing mechanical abrasion and

damage to the underlying embryo, water and nutrient absorption, and perhaps in seed germination after the seed is mature. Similar roles have been established for other simple non-glandular trichomes (Wagner et al. 2004).

The homology of the sequence derived from the initial ISSR primer and SCAR marker to a sequence published from an EST of developing apricot fruit (Mbeguie-A-Mbeguie et al. 1999) merits further attention. Chromosome walking to obtain the full gene sequence, provided that the SCAR marker is close to the gene of interest in almond, and work on the expression of this gene in the fruit could potentially provide information on the control of trichome development in almond. Physiological studies using microscopy techniques at varying intervals of kernel development are required to determine the mode of cellular development and factors controlling the formation and collapse of testa trichomes, in particular the correlation between solute movement in the testa and cellular collapse. The combination of physiological and genetic studies will then provide a better understanding in the role and function of these cells in seed development of almond. The development of a molecular marker will be beneficial to the Australian almond breeding program, as it has the potential to screen individuals for this trait which may be important in the development of cultivars for kernel export to different markets.

CHAPTER 8

Construction of a Genetic Linkage Map for the F_1 Hybrid Mapping Population Using RAPD, ISSR, SCAR, SSR, and Morphological Markers

INTRODUCTION

Traditional breeding methods have made significant contributions to the improvement of modern crop species. However, in the majority of long lived tree crop species, targeting important agronomic traits, such as disease resistance, bloom time, organoleptic qualities, and self-fertility by classical breeding has been slow. Almond breeders are hampered by the limited knowledge of the genetics of important agronomic traits and the inability to promptly select promising individuals from large breeding programs. The use of molecular techniques has the capacity to enable the breeder to improve the efficiency of almond breeding by providing markers for marker assisted selection (MAS) and the identifcation genes controlling traits of interest.

The development of genetic linkage maps offers a method for early selection strategies in many tree crops (Staub et al. 1996). The goal of genetic mapping is to produce a map saturated with markers within a short distance of loci of interest. This

technique is a powerful tool for identifying regions of genes controlling both simple and complex traits and offers a more direct method for selection of these genes than by classical techniques. By combining the use of MAS with genetic linkage maps, the long lag time between developing the initial cross and evaluating the progeny in long-lived perennial tree crop species can be significantly reduced. This circumvents the need for plants to become sexually mature or to be exposed to a disease inoculum before screening for the desired phenotype can be carried out, and therefore huge savings can be made to the breeder in time, space, and money.

The development of genetic linkage maps for woody perennials follows the same basic premise regardless of the species under investigation. Firstly a population is selected with enough genetic diversity between the parents to ensure adequate polymorphisms between the offspring (Staub et al. 1996; Weeden 1994; Weeden et al. 1994). Populations used for mapping are generally F_1, F_2, backcross (BC), recombinant inbred lines (RIL), or double haploid (DH). A popular method for the development of mapping data for species where F_2, BC, RIL, or DH are difficult to produce because of self-incompatibility and/or a long juvenile phase, as is the case for many perennial crops, is a double pseudo-testcross (DPSTC) situation (Grattapaglia and Sederoff 1994). The term DPSTC is used because the procedure is essentially analogous to a testcross, except that the configuration of markers is not known *a priori*, rather it is determined after analysing the segregation of the markers in the parents and progeny. In a DPSTC situation, a separate map is created for both parents, commonly based on dominant markers, where the segregation of these markers is heterozygous in one parent and homozygous null in the other parent, segregating 1:1 in the progeny as would be the case in a testcross situation. The DPSTC strategy has been successfully applied for genetic mapping in tree-crop species such as apple (Conner et al. 1997), chestnut

(Casasoli et al. 2001), eucalyptus (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996), larch (Arcade et al. 2000), oak (Barreneche et al. 1998), olive (la Rosa et al. 2003; Wu et al. 2004), rose (Debener and Mattiesch 1999), and tea (Hackett et al. 2000).

Genome mapping in fruit trees has become well advanced in the past decade, with extensive research in *Prunus* genus being carried out by the European *Prunus* mapping projects (Baird et al. 1996; Joobeur et al. 1998), with the most popular species for study being peach. Among the *Prunus*, almond is the only nut crop, and although mapping is not as advanced as peach, compared to other nut crops it is well studied. A number of molecular marker systems are available to detect DNA polymorphisms that vary in their method of detection, multiplex ratio, and ability to detect heterozygous genotypes. In almond, a variety of DNA marker types have been used to develop genetic maps. Randomly amplified polymorphic DNA (RAPD) markers are generated by random PCR amplification of the genome by single arbitrary primers (usually decamers). These markers are popular for the production of preliminary maps since they are cost effective, and have the capacity to generate a large number of fragments per amplification. These markers can be converted to sequence characterised regions (SCAR) markers, which gives a higher level of reproducibility (Paran and Michelmore, 1993). Simple sequence repeats (SSRs), also known as microsatellites, are comprised of tandem arrays of 2-5 repeated nucleotides, where the variation in length gives rise to polymorphisms. SSRs are currently the method of choice for the development of molecular markers due to their co-dominant inheritance, high reliability, and potential for cross-species transferability (Rossetto 2001). In addition to RAPDs and SSRs, restriction fragment length polymorphisms (RFLPs) and isozymes have also been used for mapping in almond (Viruel et al. 1995).

Viruel et al. (1995) published the first genetic linkage map for almond, using an F_1 cross between the cultivars 'Ferragnès' x 'Tuono'. Joobeur et al. (2000) published a second generation map for the same population, based on RAPD and SSR markers, but to date no morphological traits have been placed on this map. Two smaller targeted mapping projects have placed self-incompatibility on this framework almond map (Ballester et al. 1998), and late bloom on another almond population map derived from the cross of 'Felisia' x 'Bertina' (Ballester et al. 2001). Genetic maps have also been constructed for interspecific crosses of peach and almond, and the map for the almond 'Texas' x peach 'Earlygold' (T x E) is assumed to be the most saturated map for *Prunus* (Aranzana et al. 2003; Foulongne et al. 2003; Lambert et al. 2003). The localisation of morphological traits on *Prunus* maps is most advanced for peach (Chapparo et al. 1994; Dirlewanger et al. 1996; Dirlewanger et al. 1998; Warburton et al. 1996).

For MAS to be an effective method for plant breeding in *Prunus* more morphological traits need to be added to these maps, particularly for species other than peach. The development of additional molecular maps for almond, and the construction of a more saturated consensus map for *Prunus*, containing markers derived from peach, almond, apricot, plum, cherry and wild species, also has the capacity to greatly improve stone fruit breeding efficiency.

The aim of this study was to produce a genetic linkage map composed of a framework of RAPD, ISSR, SCAR, and SSR, and morphological markers. The ultimate aim is to use this genetic map as a tool to detect the presence of morphological traits of interest using MAS.

MATERIALS AND METHODS

Linkage Map Pedigree

An intraspecific almond F_1 population was derived from a controlled cross between two divergent cultivars. 'Nonpareil' (NP) was used as the female parent, and 'Lauranne' (LA) as the male parent (Chapter 2). From the 181 progeny derived from this cross, 93 individuals were randomly selected to comprise the population for which the genetic map was constructed.

Molecular Techniques

DNA extraction

Total genomic DNA was extracted from leaves using the optimised protocol of Lamboy and Alpha (1998) described in Chapter 3. DNA quantity and quality was assessed by spectrophometric analysis and working samples were diluted to either 20 ng/μL or 10 ng/μL depending on the technique used for molecular marker derivation.

Molecular Marker Generation

A total of 60 RAPD, 23 ISSR, 1 SCAR, and 36 SSR primers were screened for polymorphism generation as outlined in Chapter 4. Polymorphic markers were selected for consistent and reproducible fragments based on visual scores of agarose gels, autoradiograms, and fluorescent detection of SSRs (Chapter 4).

Morphological Markers

Twelve quantitative traits were evaluated for three years (2001, 2002, and 2003), and χ^2 analyses were performed to test for deviations from the expected Mendelian ratios.

Measurements for the total number of traits are described in more detail in Chapter 6. The seven segregating traits that were selected for inclusion in the map are listed in Table 8.1, and these were scored for the 2003 harvest (third fruiting year) and treated as dominant marker types. *S*-alleles were scored using allele specific primers as described in Chapter 5. Each allele was treated independently and entered into the mapping matrix as a single entity. The double kernel character was calculated based on the number of double kernelled fruit in a sample of 20-50 nuts, expressed as a percentage of the total. The segregation class was selected as either absent (0%) or present (<0%). Kernel shape and thickness were measured using digital calipers at the widest, thickest, and longest points. Kernel shape was determined by the length/width ratio classified into five groups, which were compressed into two classes for linkage analysis, broad (<0.56 cm) and medium/narrow (>0.55 cm). Kernel testa colour was scored visually in comparison to the two parents and scored as pale or dark. Both kernel taste and testa pubescence were assessed by sensory analysis and scored on the basis of consistency over three years of data (2001, 2002 and 2003).

Marker Nomenclature

RAPD markers were coded by the primer name followed by the fragment size, for example a fragment of 456 bp amplified by primer OPA-02 was coded OPA02-456. For ISSR primers, the primer sequence was used as the primer name followed by the fragment size, for example a fragment of 1125 bp amplified by primer $(CA)_8G$ was coded $(CA)_8G$-1125. SSR and SCAR markers were designated by their primer name, based on their derivation which is outlined in Chapter 7.

Morphological markers were coded with a truncation of the trait description as shown in Table 8.1.

Linkage Analysis and Map Construction

Only markers that segregated in the parents and the F_1 generation were selected to construct the mapping matrix based on a three-tier selection criterion (Chapter 4).

For dominant markers (RAPD, ISSR, and SCAR), bands were scored as present or absent and checked for segregation to a 3:1 or 1:1 ratio. For codominant markers (SSR) segregating 1:1:1:1 or 1:2:1, alleles were coded based on ascending band size. For all marker types, polymorphisms were coding according to the Joinmap 3.0 data coding system (Van Ooijen and Voorrips 2001).

Two data sets were constructed, one for each parent based on the parental origin of the marker in accordance with the DPSTC strategy. Markers were sub-divided into three data sets, (1) testcross markers segregating with a ratio of 1:1, heterozygous in the female parent NP and homozygous in male parent LA, (2) testcross markers homozygous in NP and heterozygous in LA, and (3) intercross markers heterozygous in both parents segregating with ratios of 3:1, 1:2:1, or 1:1:1:1.

Co-dominant markers were included in both parental data sets to identify homologies between parental maps. Segregation distortion of markers was evaluated using the LINKEM program (Vowden and Ridout 1994) and the internal χ^2 analysis of the Joinmap 3.0 program. Joinmap 3.0 was used for linkage analysis and map construction using the cross-pollination (CP) format. For construction of the initial parental map, all segregating markers were included regardless of any distortion detected. A LOD score range of 2.0 to 10.0 was set as a linkage threshold and the maximal number of linkage groups corresponding to the chromosome number (n = 8) were selected for conversion to map format. The LOD score was assessed manually when selecting mapping nodes (predominantly LOD 3.0 - 4.0) selected automatically by the program. Markers that interfered with linkage group determination were checked for phase and segregation distortion and omitted by selecting the exclude marker function in the loci node. All

other thresholds were set by default of the program and were held constant. Map distances in centiMorgans were calculated by converting the recombination frequencies using Kosambis mapping function (Kosambi 1944).

To produce a pictorial representation of the mapping data, the output map data 'map text' was copied into the data file of Map Chart 2.1 program (Voorrips 2002).

Generation of the Integrated Parental Map

Inspection of the two parental maps was carried out to identify markers heterozygous in both parents that were localised on the maps. In order to produce an integrated map, groups were joined where corresponding intercross markers were present in both parental data files, and the groups merged using the combine groups for map integration function. Markers were merged into linkage groups with a LOD score of between 2.0 – 15.0.

RESULTS

In order to minimise experimental artifacts, primers selected for mapping were subjected to a three-step screening procedure. Firstly, primers were tested for polymorphisms between both parents, secondly, primers producing a polymorphic band were screened on the parents again and with 10-12 progeny, and finally, primers that produced a clear, unaltered banding pattern in the first two stages of selection were screened on the parents and the total 93 mapping progeny. RAPD and ISSR primer screening was repeated on the total population again to ensure correct scoring of the data. The second and third screens for the SCAR and SSR markers were compared for consistency and reproducibility. This three-step screening was used to limit the impact.

Table 8.1 Segregation of morphological traits scored in the parents and F_1 progeny used in linkage analysis. Mapping data nomenclature is given in italics and brackets

Phenotype	Class	Parental Phenotype 'Nonpareil'	Parental Phenotype 'Lauranne'	Segregation ratio in F_1	Chi-Square value
S-allele *(S7S8)* *(S3Sf)*	S3 S7	S7 S8		1:1	1.82
	S8 Sf		S3 Sf	1:1	0.01
Kernel Shape *(Shape)*	Broad Medium/Narrow	Medium/Narrow	Broad	3:1	1.58
Kernel Thickness *(Thick)*	Thick Medium/Thin	Thick	Medium/Thin	1:1	0.10
Kernel Taste *(Taste)*	Sweet Semi-bitter	Sweet	Sweet	3:1	5.57
Double Kernels *(Double)*	Absent (0%) Present (<0%)	Absent	Present*	1:1	1.33
Testa Pubescence	Smooth Pubescent	Smooth	Pubescent	1:1	59.52
Testa Colour *(Colour)*	Pale Dark	Pale	Dark	1:1	0.54

* Doubles observed in other environments (Socias i Company pers. com.)

of scoring error for genetic map construction. Morphological markers were scored over three harvest years to check for consistency.

From the 119 primers screened, 33 were polymorphic resulting in the production 73 segregating markers. From these 73 segregating markers, 19 were inherited from the female parent 'NP', 30 from the male parent 'LA', and 24 from both parents, the segregation types observed and the number of distorted loci are summarised in Table 8.2. Overall 74% of markers analysed in the progeny segregated according to the expected mendelian inheritance, and at $\alpha = 0.05$, considering 73 marker loci, approximately four markers could be expected to display distortion by chance alone. For the female parent 30% of markers were distorted, for the male parent only 21% of markers showed a distorted ratio. For the paternally inherited distorted markers, the distortion was due to an excess of heterozygotes in the majority of cases (78%). No specific correlation was observed for the four distorted markers in the maternal group. There were fewer distorted loci among the SSR markers for both parental types than for RAPD and ISSR loci and the majority of distorted loci in the male parent were of the ISSR type (56%).

For map construction, 43 markers from NP and 54 markers from LA were selected. Of these 19 from NP and 30 from LA were grouped with a LOD score of between 2.0 to 4.0, and of these, 77% were mapped with a LOD of 4.0. Two parental maps were constructed according to the DPSTC mapping strategy as described by Grattapaglia and Sederoff (1994), resulting in 7 linkage groups for NP (Figure 8.1) and 8 for LA (Figure 8.2). Linkage groups were composed of between 2 - 11 loci, and were between 10.5 – 86.6 cM in length. The average marker density was 9.4 cM/marker for 'NP' and 9.6 cM/marker for LA. On a single linkage group basis, linkage group NP-C showed the highest marker density of 3.5 cM/marker, while NP-B showed the lowest

marker density of 24.6 cM/marker, which is probably due to the presence of only two markers on the latter linkage group. The longest linkage group for both parents was G4 (86.6 and 84.1 cM), which also had the highest number of linked markers (8 and 11). The summarised data for linkage group length, loci type, number of skewed loci, average density of markers, and anchor loci are given in Table 8.3. No particular evidence of clustering of the loci was observed, although for six primers $(AC)_8G$, $(AG)_8YT$, $(CA)_8G$, $GCA(CA)_9$, OPC-10, and OPC-11 two fragments of different size derived from the same primer were localised on the same linkage group with a distance of 1.9 – 24.9 cM between them (Table 8.4). Most of these co-localisations of fragments derived from the same primer were observed in the male parent map 'LA'.

Of the 97 markers used for parental map construction, 24 were heterozygous in both parents, and 18 common loci were identified to be used as loci bridges between the maps. Of these markers, 14 were used to produce an integrated parental map, representing approximately 19% of all markers studied. The integrated map consisted of 6 linkage groups, with between 2 and 4 markers, varying in length from 6.0 to 49.1 cM in length (Figure 8.3), and with a marker density of 11.5 cM/marker (Table 8.3).

Of the eight morphological markers selected for mapping, three loci associated with phenotypic characters could be placed on the maps. According to the morphological assessment NP has a pale testa, whereas LA has a dark testa, and this character segregated with a ratio of 1:1 in the progeny suggesting monogenic inheritance. Linkage analysis placed this locus at the top of linkage group G4-LA, between the markers PS12e2 and $(AG)_8YC$-1150 (Figure 8.2). Both of these molecular markers are present in the parental maps and have been placed on the integrated map. The locus for taste was loosely linked to the ISSR marker $(AG)_8YC$-1786 and was placed at the bottom of B-NP and B-IN, 49.1 cM from the marker. Taste was assessed

as sweet in both parents and therefore expected to segregate 3:1 for the sweet versus semi-bitter phenotype. It showed distortion from this segregation at $\alpha = 0.05$, indicating either that it is not likely to be inherited in a simple mendelian fashion, or that some other factor associated with scoring this trait distorted the ratio. The loci for the self-incompatibility alleles, *S*7 and *S*8, were placed at the distal region of G6-NP, 6.3 cM from the marker OPA08-467 and segregated 1:1 in the population.

DISCUSSION

A total of 73 polymorphic markers were used to construct two low to moderate density genetic linkage maps for almond using a F_1 full-sib hybrid population, derived from a cross between the cultivars 'Nonpareil' and 'Lauranne'. This population was used to construct these genetic linkage maps as the parents were assumed to have a high level of heterozygosity between them leading to a high frequency of gene recombination in the progeny, which is a prerequisite of the DPSTC strategy (Grattapaglia and Sederoff 1994). The allelic transmission from parent to offspring via recombination during meiosis enabled the relative position of genetic loci in relation to each other to be inferred. Using this strategy, seven and eight separate linkage groups were constructed for the female and male parents respectively, putatively corresponding to the haploid chromosome number of almond (n = 8). Markers that were in an intercross segregation (heterozygous in both parents) were used to merge the two parental maps into a preliminary consensus map for this cross, consisting of six linkage groups.

The discrepancy between the number of linkage groups expected to coalesce and the number observed for the female parent and the integrated map is most likely due to the high number of unincorporated markers and the low number of homologous markers for the integrated map. The inclusion of more co-dominant markers would be expected to alleviate this situation and increase the number of linkage groups. The incorporation

Table 8.2 Analysis of marker segregation and distortion in the F_1 population and number of distorted markers by marker type

	Marker inherited from		
	NP	LA	Both
Total number of markers analysed in segregating population	19	30	18
Expected segregation	1:1	1:1	3:1 1:1:1:1
Number of markers with deviation from the expected ratio at $\alpha = 0.05$	4	9	6
Percentage of distorted markers as a percentage of total (%)	30	21	25
Number of distorted markers based on marker type			
RAPD	2	2	1
ISSR	2	5	1
SSR	-	1	2
Morphological	-	1	2

Table 8.3 Summary of linkage group composition for 'Nonpareil' (NP), 'Lauranne' (LA), and Integrated (IN) Maps

	Average			G1	G3			G4			G6			G8*		A			B			C		1
	NP	LA	IN	LA	NP	LA	IN	NP	LA	IN	NP	LA	IN	NP	LA	NP	LA	IN	NP	LA	IN	NP	IN	LA
Total number of markers	36	35	14	2	7	7	2	8	11	4	7	3	2	6	6	2	2	2	2	2	2	3	2	3
RAPD	9	6		-	1	2	-	3	3	1	1	-	-	2	-	2	2	2	-	-	-	-	-	-
ISSR	15	18		1	4	3	1	2	4	1	3	1	-	2	5	-	-	-	1	2	1	3	2	2
SCAR	1	1		-	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SSR	8	9		1	1	1	1	3	3	2	2	2	2	2	1	-	-	-	-	-	-	-	-	1
Morphological	2	1		-	-	-	-	-	1	-	1	-	-	-	-	-	-	-	1	-	1	-	-	-
Length (cM)	337.3	336.3	161.9	32.5	47.7	67.6	10.1	86.6	84.1	48.2	49.2	38.2	24.9	71.0	63.6	23.6	23.6	23.6	49.1	18.7	49.1	10.5	6.0	40.0
Density of markers /cM	9.4	9.6	11.5	32.5	6.8	9.6	10.1	10.8	7.6	12.1	7.0	12.7	12.5	11.8	10.6	11.8	11.8	11.8	24.6	9.4	24.6	3.5	3.0	13.3
Number of distorted markers	4	9	6	1	2	2	0	2	2	1	2	1	1	1	1	0	0	0	1	0	1	0	0	1
Number anchored to integrated map				-	2	2	-	4	4	-	2	2	-	-	-	2	2	-	2	1	-	2	-	-
Number anchored to TxE map				1	1			2			1			2										

* For G8, G8a and G8b grouped together for analysis

Table 8.4 Summary of primers producing two fragments co-localised on the same linkage group for a cross between 'Nonpareil' x 'Lauranne'

Primer	Fragment Size (bp)	Linkage Group	Distance between markers (cM)
$(AC)_8G$	775 820	G3LA	19.3
$(AG)_8YT$	960 1957	G4LA	19.9
$(CA)_8G$	492 1125	G8aLA	24.9
$GCA(CA)_9$	376 535	G8aLA	1.9
OPC-10	625 763	G4LA	23.7
OPC-11	850 1150	ANP, ALA, AIN	23.6

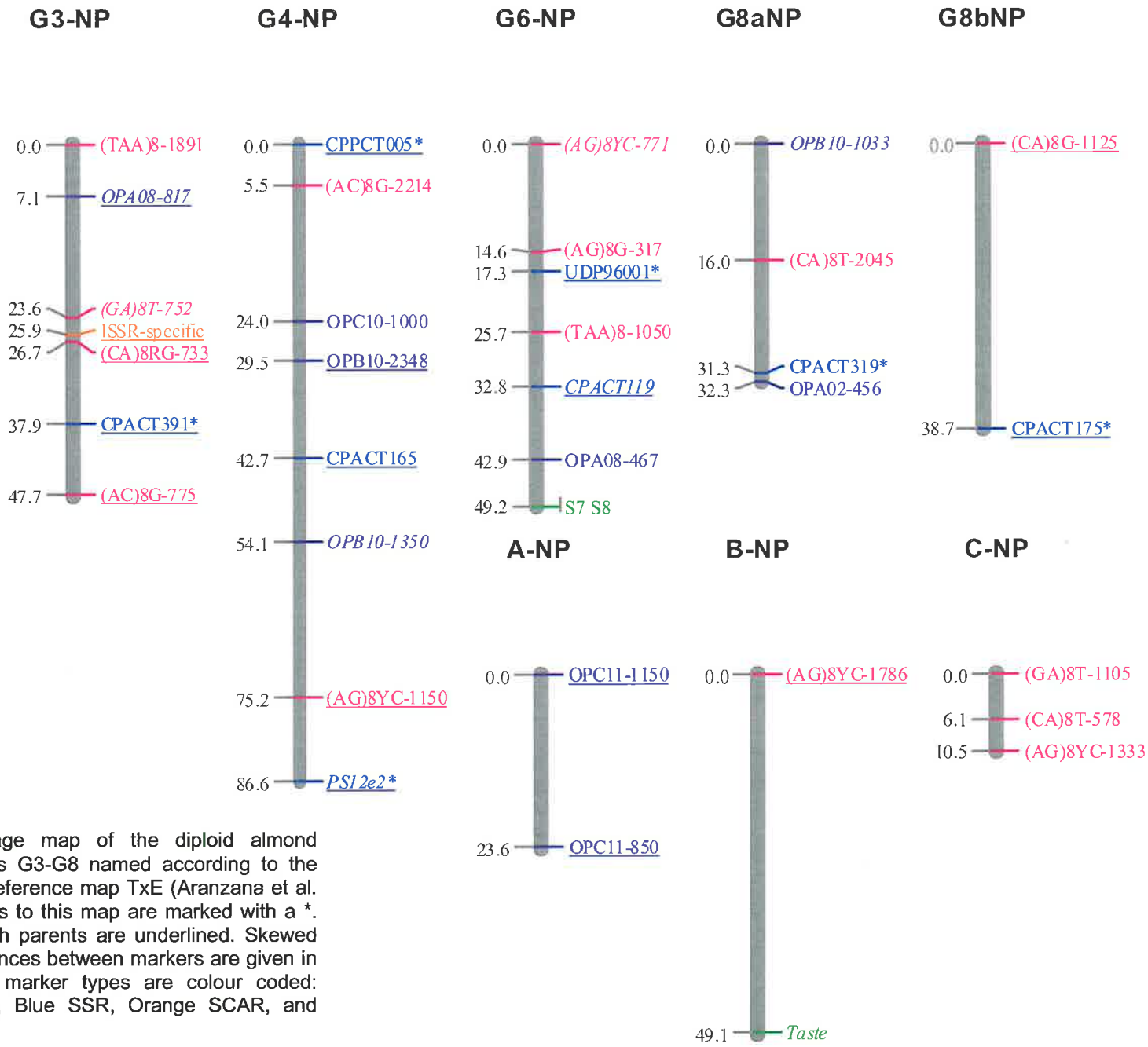

Figure 8.1 Genetic linkage map of the diploid almond 'Nonpareil'. Linkage groups G3-G8 named according to the convention of the *Prunus* reference map TxE (Aranzana et al. 2003). Markers homologous to this map are marked with a *. Markers inherited from both parents are underlined. Skewed markers are in italics. Distances between markers are given in centi Morgans. Individual marker types are colour coded: Purple RAPD, Pink ISSR, Blue SSR, Orange SCAR, and Green Morphological

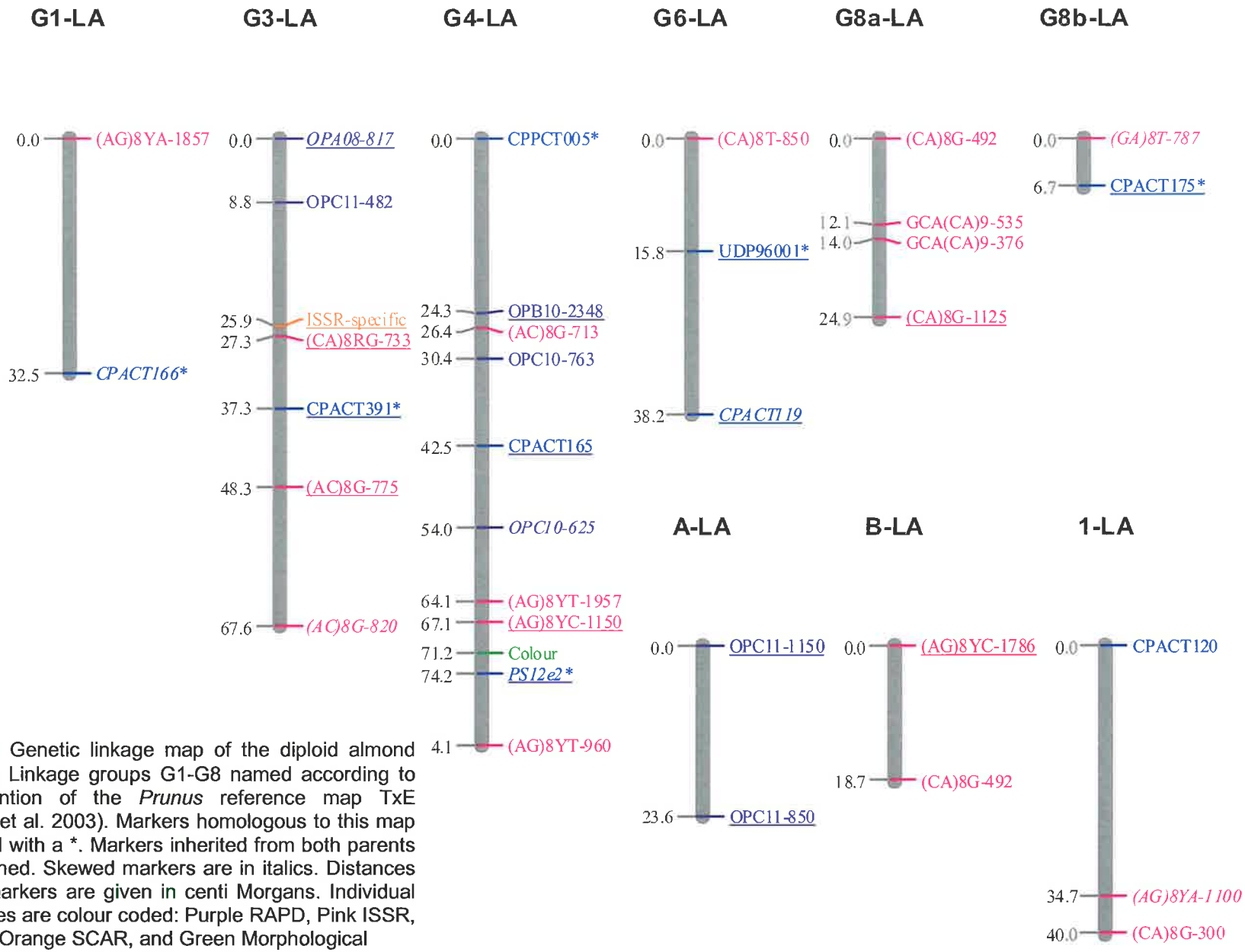

Figure 8.2 Genetic linkage map of the diploid almond 'Lauranne'. Linkage groups G1-G8 named according to the convention of the *Prunus* reference map TxE (Aranzana et al. 2003). Markers homologous to this map are marked with a *. Markers inherited from both parents are underlined. Skewed markers are in italics. Distances between markers are given in centi Morgans. Individual marker types are colour coded: Purple RAPD, Pink ISSR, Blue SSR, Orange SCAR, and Green Morphological

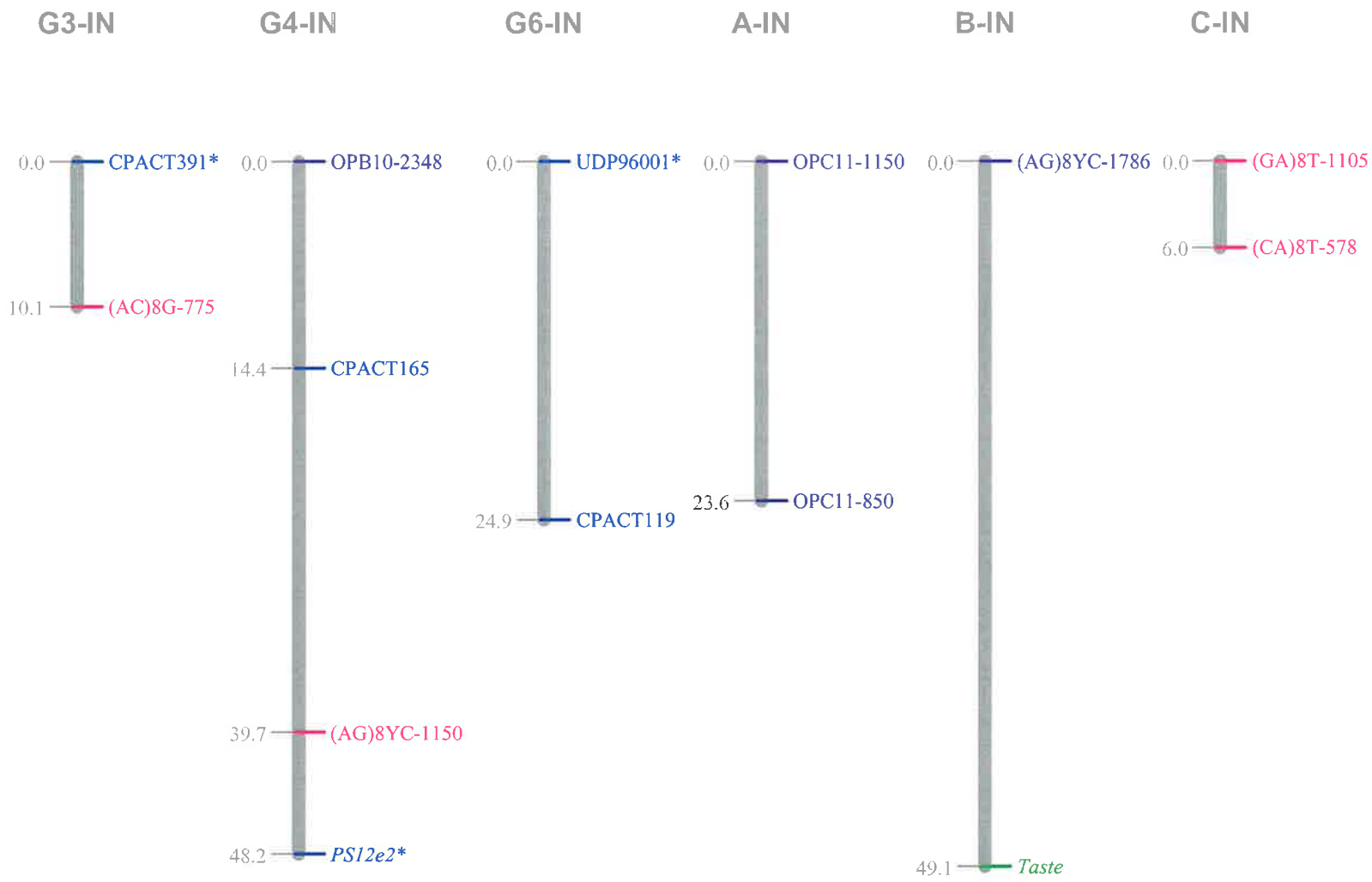

Figure 8.3 Integrated genetic linkage map of the diploid almond cross 'Nonpareil' x 'Lauranne'. Linkage groups G3-G6 named according to the convention of the *Prunus* reference map TxE (Aranzana et al. 2003). Distances between markers are given in centi Morgans. Individual marker types are colour coded: Purple RAPD, Pink ISSR, Blue SSR, Orange SCAR, and Green Morphological

of more markers would also be expected to merge single markers that are unlinked so far to the major linkage groups that may presently be in regions of very low marker frequency.

Debener and Mattiesch (1999) proposed increasing the mapping population size as a means of introducing unlinked single and linked marker pairs into the main linkage groups. Analysis of genetic maps produced for other *Prunus* species (Table 8.5), and for the T x E population used as the reference map for *Prunus* (Aranzana et al. 2003; Joobeur et al. 1998), indicates that the population size used for this study is unlikely to be the reason for the discrepancy of the number of derived linkage groups in comparison to the haploid chromosome number. Therefore increasing the marker number and using more informative marker types such as RFLP and SSR would be more beneficial for further map construction than increasing the population size.

No obvious clustering was observed in either the parental or integrated maps, although this could be masked by the low marker density and may be observed if more markers were introduced into the map. The colocalisation of two markers derived from the same primer on the same linkage group was more prevalent in the male parent, particularly for ISSR primers. This phenomenon has also been reported using AFLP markers in *Populus deltoides* (Wu et al. 2000) and salmonid-specific small interspersed nuclear elements in rainbow trout (Young et al. 1998). The presence of colocalised ISSR markers on the same linkage group may be due to the primer amplifying the same locus at different binding sites, resulting in the different fragment size. The dispersal of distorted loci did not appear to be random, rather they tended to be localised toward the distal ends of the linkage groups. Eleven linkage groups, five in the female, and six in the male had distorted markers present. Maliepaard et al. (1998) observed that markers with skewed segregation patterns could not usually be mapped. In this study, 80% of

skewed markers in 'NP' and 53% in 'LA' were able to be mapped, indicating that this is not the case in almond. The percentage of skewed loci is comparable to levels of 11-23% observed in rose (Debener and Mattiesch 1999), 18% in oak (Barreneche et al. 1998), 15% in *Eucalyptus* (Marques et al. 1998), 11.3% in apricot (Lambert et al. 2003), and 10% in almond (Joobeur et al. 2000). However the number of skewed loci was approximately half of the 43% of distorted markers obtained in intraspecific peach x almond mapping programs (Aranzana et al. 2003), 46% (Joobeur et al. 1998) and 37% (Foolad et al. 1995). Segregation distortion was observed in approximately one third of markers mapped in a peach x *Prunus davidiana* cross, and although several mechanisms were proposed to account for this, one mechanism alone could not account for the total number of distorted loci observed (Foulongne et al. 2003). In intraspecific hybrids, segregation distortion is proposed to be a consequence of hybrid breakdown (Foolad et al. 1995). Several reasons have been proposed for segregation distortion observed in tree species, including genetic load (Bradshaw and Stettler 1994), segregation with a pollen lethal gene or SI locus (Ballester et al. 1998; Gebhardt et al. 1991), statistical bias or genotyping and scoring errors (Plomion et al. 1995). Several authors have removed distorted markers from the first round of map construction and then observed that the inclusion of these markers in a later round made no significant difference to gene order or linkage map order (Conner et al. 1997; Debener and Mattiesch 1999; Kuang et al. 1999; la Rosa et al. 2003). Lambert et al. (2003) proposed the discrepancy between marker order in their apricot F_1 map compared to the T x E map is a result of the high level of distorted markers included in the latter map. The inclusion of more markers on the presented maps will determine if this is the case for almond, or if genetic distance between species is responsible for the change in loci order between less closely related individuals.

The localisation of three morphological traits on the map is the first step toward using this map as a tool for marker assisted selection. Both testa colour and the sweet/semi-bitter taste character have not been mapped prior to this study. The gene for self-incompatibility (SI) has been mapped to the end region of linkage group G6 from a 'Ferragnés' x 'Tuono' cross (Ballester et al. 1998). Foulongne et al. (2003) also proposed the same region for the location of a SI locus, the end of linkage group LG6, for a *P. persica* x *P. davidiana* cross. The alleles, *S7* and *S8*, were mapped at the same position in this study, indicating that the mapping of locations for morphological traits is consistent at the species level. It would be interesting to determine if this same homology existed at the genus level in *Prunus*. Markers for self-incompatibility have been identified in both apricot (Badenes et al. 2000) and cherry (Wiersma et al. 2001), and these could be placed on maps developed by Lambert et al. (2003) and Wang et al. (1998). In this study, morphological traits were selected for analysis on the basis of their importance for release in new commercial cultivars developed by the Australian almond breeding program.

The integration of the parental maps was accomplished using heterozygous and co-dominant markers, which also allowed the identification of homologous linkage groups between the parents. No change in marker order was observed between the parents and the integrated maps although a change in the distances between markers was observed. This was expected, since although the values were relatively close, in the integrated map the combined recombination frequency estimate is the average over the male and female meioses. The integrated map is therefore a statistical compromise between the maternal and paternal recombination values rather than a biological reality. However, integrated maps are the ideal situation for mapping

Table 8.5 Summary of data from genetic linkage maps produced for domesticated *Prunus* species. For map lengths female and male parent map length is denoted by (F) and (M) respectively.

Species	Population Type	Population Number	Marker Used	Number of Markers	Map Length (cM)	Reference
P. dulcis	F_1	60	Isozyme RFLP	120	393 (F) 394 (M)	Viruel et al. (1995)
	F_1	60	RAPD SSR	60	415 (F) 416 (M)	Joobeur et al. (2000)
P. dulcis x P. persica	F_2	64	Morphological Isozyme RFLP	117	800	Foolad et al. (1995)
	F_2	75	Isozyme RFLP	246	491	Joobeur et al. (1998)
	F_2	82	SSR	109	522	Aranzana et al. (2003)
P. persica	F_2	1027	Morphological Isozyme RAPD	103	396	Chapparo et al. (1994)
	F_2	71	Morphological RAPD RFLP	65	332	Rajapaske et al. (1995)
	F_2	63	Morphological Isozyme RAPD AFLP ISSR RFLP	270	712	Dirlewanger et al. (1998)
P. armeniaca	F_1	81	AFLP RAPD RFLP SSR	132	511 (F) 467.2 (M)	Hurtado et al. (2002)
	F_2	76	Morphological AFLP SSR	211	602	Vilanova et al. (2003)
	F_1	142	AFLP RFLP SSR	141	538 (F) 699 (M)	Lambert et al. (2003)
P. ceracus	F_1	86	RFLP	126	461.6 (F) 279.2 (M)	Wang et al. (1998)

programs, since they permit the identification of conserved loci in related species and their use as reference points to transfer linkage information. SSRs are transportable across *Prunus*, as shown by a number of studies (Aranzana et al. 2003; Lambert et al. 2003; Martínez-Gómez et al. 2003ab; Mnejja et al. 2003; Rossetto 2001; Vilanova et al. 2003; Zhebentyayeva et al. 2003). In addition, peach and cherry SSRs have been used for construction of genetic maps in pear (Yamamoto et al. 2002b), indicating that cross-genus transfer is possible for these markers in the Rosaceae. Among the ten SSRs used in this study, six were developed from almond, three from peach, and one from cherry. These markers can be regarded as the preferential marker type for mapping, particularly for the development of genera specific consensus maps. Seven of these SSRs are anchor loci to the T x E map (Aranzana et al. 2003; P. Arús, unpublished). Aranzana et al. (2003) have proposed a set of SSRs, based on the T x E map, which could be used as a 'genotyping set' for this genera., two SSRs used in this study are part of this set.

The ultimate application of this map is to improve almond-breeding efficiency through the use of molecular marker technologies. MAS is the most common objective of genetic map construction, which can be used either for accelerated backcrossing by tracking the segregation and inheritance of traits through a number of backcrosses (Foulongne et al. 2003), or as is the case in most woody perennials, for the direct selection of desirable traits in breeding programs. It is desirable that the selected marker is closely linked to the trait of interest to ensure linkage is not lost through successive breeding cycles. Therefore marker saturation is important to reduce linkage disequilibria and to diminish the effectiveness of selection. The incorporation of more molecular markers to integrate the remaining five unlinked morphological traits is the next step for improving this map for further use in MAS. In the future, when additional markers are added to the map to improve genome coverage, marker saturation, and accuracy in some

regions, gene pyramiding can be employed. Screening for a number of traits at the same time makes MAS more cost effective since the majority of costs are associated with sampling and DNA extraction (Luby and Shaw 2001). Thus with further saturation this map will become a valuable tool for use in breeding programs in Australia and strengthen research on genetic mapping of stone fruit worldwide.

CHAPTER 9

General Discussion

Early settlers first introduced the almond to Australia nearly 200 years ago, and these first varieties were hard-shelled 'Jordan' types of European descent (Quinn 1928). The second major introduction of almonds into Australia was of cultivars derived in North America of paper/soft shelled varieties such as 'Nonpareil' (Moss 1965).

The most widely planted cultivars in Australian commercial orchards are two American cultivars 'Nonpareil' and 'Carmel' (Bennett and Johns 2002), with 'Nonpareil' the single most planted cultivar (Bennett 1999). In the world market, Australia is a relatively small almond producer, accounting for only 1.1% of the total world production (FAO 2003). In order to gain a greater market share, particularly for export into Mediterranean countries and to remain competitive in the world market, Australia needs to increase its production. In combination with an increase in yield, Australian producers need to maintain kernel quality, relative disease free status, and promote better marketing for this product, both domestically and internationally. This can be achieved in two ways, firstly more trees can be planted, or alternately better cultivars can be developed.

This research program was initiated with the aim of developing a genetic linkage map for use in almond breeding, using a variety of molecular techniques. Secondly the population used for map construction was examined in order to establish the control and genetic inheritance of several morphological traits which are key breeding objectives. In particular, the physiological control of testa pubescence was examined by microscopy techniques and a molecular marker was identified by bulked segregant analysis.

Breeding programs have been undertaken worldwide in order to develop new cultivars, by ameliorating desired, and excluding unfavourable traits in the resulting progeny. Breeding objectives in Australian programs have a similar goal to established breeding programs in a number of countries including Spain (Socias i Company et al. 1997), Portugal (Oliveira 1999) and America (Gradziel and Kester 1999). In 1997, the University of Adelaide established the Australian almond breeding program in partnership with the Almond Board of Australia (Bertozzi et al. 1998). The aim of this program was to develop a variety of cultivars for eventual commercial release which showed improved yield, bloom time, and kernel characteristics, as well as incorporating the *Sf*-allele, to eliminate the reliance on cross-pollinator varieties in the orchard (Wirthensohn and Sedgley 2002). Germplasm for this program is largely imported from overseas since there are no wild species in Australia, and no survey has been undertaken to establish potential germplasm pools in abandoned orchards or isolated individual trees, such has been established and utilised in Australian olive breeding programs (Mekuria et al. 2001; Sedgley 2000).

The development of breeding programs is a huge investment in time, space and money, and therefore prospects for removing some of these costs would be advantageous to the breeder. The use of molecular techniques has the capacity to greatly improve the breeding efficiency of long-lived woody perennials such as almond

(Weeden et al. 1994). The construction of genetic linkage maps is a popular technique for marker assisted selection (MAS) in breeding programs (Weeden 1994). MAS is the indirect selection for the trait of interest by selection of a linked molecular marker phenotype. The use of MAS for fruit tree breeding programs is more cost effective than for annual crops because of the long juvenile phase and large size of many of these trees. In almond the major focus of crop improvement is in regard to fruit quality enhancement, however selection for traits associated with fruiting can only be assessed once the tree becomes sexually mature, which may take between 3-5 years in almond (Kester and Gradziel 1996). Therefore the use of MAS for screening at the juvenile phase is proposed to be as cost effective as maintaining trees in the orchard until maturity for assessment (Luby and Shaw 2001). The effectiveness of MAS versus phenotypic selection depends on a number of criteria being fulfilled. Firstly in order to produce a robust molecular map, high quality DNA must be extracted from the species under investigation. This was examined in Chapter 3, where the comparison of four different DNA extraction techniques from almond leaves was carried out. Results indicated that not all protocols are equally suited for extraction of high quality intact DNA in this species. The modified technique of Lamboy and Alpha (1998) fulfilled all criteria for a suitable extraction method with respect to DNA yield, integrity, purity and suitability for RAPD analysis.

Secondly the individuals used to derive the mapping population must also be considered prior to developing a genetic linkage map. As outlined in Chapter 2, the cultivars 'Nonpareil' and 'Lauranne' were used to derive the F_1 population. These two cultivars were selected for various reasons, firstly although almond has a high level of heterozygosity compared to peach (Byrne 1990), cultivars of American and Russian descent have a rather low level of heterozygosity (Bartolozzi et al. 1998; Lansari et al.

1994). In order to maximise genetic diversity in the F_1 an American and European cultivar were selected. 'Nonpareil' was selected owing to its prevalence in the Australian (Bennett 1999) and American industry (Hauagge et al. 1987), and its highly valued nut characteristics such as colour, general appearance, size and very low percentage of doubles. 'Lauranne' was chosen as the other parent because it has the self-fertile character derived from its male parent 'Tuono' (Grasselly 1972). In addition, this population was chosen because despite its economic importance, no known map based on 'Nonpareil' has been established, and genetic mapping in almond has mainly focused on European cultivars. Of the European cultivar maps, several well-developed maps have been produced for a 'Ferragnés' x 'Tuono' population (Ballester et al. 1998; Joobeur et al. 2000; Viruel et al. 1995), which is the parental combination of 'Lauranne'. Given the close genetic similarity it was assumed this could be valuable for comparison between markers on both maps, to check for homologous linkage groups.

Prior to map production the population was tested for hybridity, since the inclusion of non-hybrid mapping data can greatly skew the results (Kearsey and Pooni 1996). This also applies to the misclassification of molecular data (Skroch and Nienhuis 1995), which can be influenced by the use of low quality DNA (Reiss and Rutz 1999). The determination of hybridity was carried out by analysing the *S*-allele complement of each individual, in combination with molecular markers and cluster analysis, as outlined in Chapter 5. *S*-alleles are a robust technique for verifying hybridity, provided the *S*-alleles of the parents are known. One *S*-allele is inherited from both parents, and in all reported cases follows a 1:1 segregation ratio in the progeny (Dicenta and García 1993; Ortega and Dicenta 2003; Socias i Company 1991; Socias i Company 2001; Vargas et al. 1997). Therefore progeny with aberrant or an absent *S*-allele complement can be assumed not to be true hybrids of the parental cross. *S*-allele inheritance is widely used

for establishing cross-compatibility between cultivars, especially newly derived cultivars for which the *S*-allele combination is not known, and to track self-fertility inheritance in breeding programs (Batlle et al. 1998; Bŏsković et al. 1997, 1999, 2003; Boskovic et al. 1999; Channuntapipat et al. 2002; Kester et al. 1994; Kester and Micke 1994; Tamura et al. 2000; Tao et al. 1999). This methodology has also been used to detect self fertility in cherry (Hauck et al. 2001; Wiersma et al. 2001) and Japanese apricot (Tao et al. 2000). This population segregated for several morphological traits, which were examined in Chapter 6, which included self-fertility.

In this study Randomly Amplified Polymorphic DNA (RAPD), Inter-Simple Sequence Repeats (ISSR) and Simple Sequence Repeats (SSR) were used to generate molecular markers, which is described in Chapter 4. All these marker types were useful for molecular marker generation and map construction. The use of co-dominant marker types such as SSR and Restriction Fragment Length Polymorphisms (RFLP) allows the establishment of anchor loci between various maps. The use of SSRs in this project allowed the identification of identical marker loci present on the *Prunus* reference map 'Texas' x 'Earlygold' (Aranzana et al. 2003; Arús et al. 1994) as shown in Chapter 8. This map is the most saturated for *Prunus* and several other studies have used anchor loci to establish homology between linkage groups in apricot (Hurtado et al. 2002; Lambert et al. 2003; Vilanova et al. 2003a), plum (Claverie et al. 2004), sour cherry (Wang et al. 1998), a peach x *P. davidiana* cross (Foulongne et al. 2003), another peach x almond cross (Jáuregui et al. 2001), and a distantly related apple x pear study (Yamamoto et al. 2001). Syntenic relationships among the *Prunus* may allow the presence of a gene on the linkage group or chromosome of one species to be predicted on another species. Therefore the genetic gain obtained in one species may be applied to a wide spectrum of related species.

The genetic map produced in this study provides an excellent backbone onto which further SSRs from the *Prunus* reference set (Aranzana et al. 2003) can be incorporated for additional map saturation. This approach will also provide more anchor loci for the production of syntenic maps for the genus, in addition to enabling the unlinked morphological markers identified in Chapter 8 to be incorporated into the map. Three morphological markers have been placed on the map, with four unlinked at the present time. The addition of auxiliary molecular markers will allow the identification and localisation of quantitative trait loci (QTL), such as bloom time, inshell and fruit weight, and kernel shape on the map.

A collaborative approach to mapping in *Prunus* has several advantages other than only comparing homologous linkage groups. Identification of disease resistance in almond is lagging behind other stone fruit species. Several bacterial, fungal and viroid pathogens attack stone fruit and are often multi-host, causing significant losses to the grower and industry as a whole. Although disease resistance in *Prunus* is multigenic in most cases, single dominant genes can be identified for MAS (Claverie et al. 2004). Markers have been identified for root-knot nematode (RKN) resistance in myrobolan plum (Lecouls et al. 1999) and the localisation of resistance gene clusters has been identified on LG2 and LG7 of the 'T x E' map (Claverie et al. 2004). Of serious threat to the Australian stone fruit industry is the disease Sharka which is not present in this country. The determination of Sharka resistance by classical means is very time consuming and prone to the production of false positives because of the difficulty scoring the trait accurately (Hurtado et al. 2002). In order to preempt a devastating blow to the industry should this disease breach the strict Australian quarantine defense, markers which have already been developed could be used to screen breeding programs for this trait. Sharka resistance genes have been localised to LG1 of the 'T x E' map

(Hurtado et al. 2002; Vilanova et al. 2003a), and this could be applied for assessing resistance in almond progenies in this country. Other projects at the University of Adelaide are assessing resistance to bacterial spot (Li et al. 2004) and anthracnose (Colmagro et al. 2002) in almond which could be placed on the genetic linkage map once the mapping population was screened for these traits.

Genetic markers are not only used to produce genetic maps, and as described in Chapters 1 and 4, they have also been applied for DNA fingerprinting and assessment of phylogenetic relationships, paternity analysis and for bulked segregant analysis (BSA). BSA is a useful technique for identifying traits of interest without the need to produce an entire genetic map, or for targeting regions on maps with a low marker density. This technique has been used to identify markers linked to late bloom in almond (Ballester et al. 2001), several fruit traits in peach (Chapparo et al. 1994; Warburton et al. 1996) disease resistance in plum (Claverie et al. 2004) and self-fertility in apricot (Badenes et al. 2000). In this study BSA was used to identify a marker linked to the pubescent testa character, as presented in Chapter 7. BSA also offers another technique to obtain markers flanking traits of interest for use in map based cloning. Obtaining the full genomic sequence of the pubescent phenotype and comparison to homologous sequences in other species would allow the elucidation of how this trait might be controlled in almond. Map based cloning is only possible with markers closely linked to the target gene and the availability of a large insert library. The development of a library requires the extraction of quality high molecular weight DNA, which was established in Chapter 3. Libraries have been developed for a number of *Prunus* species including almond (Ushijima et al. 2001), peach (Georgi et al. 2002; Wang et al. 2001, 2002) and apricot (Vilanova et al. 2003b). These libraries are now being used for screening for

genes, and in the future can be used for physical mapping and sequencing of the genome.

The production of genetic maps, which include important agronomic loci, can also be used for mapping expressed sequence tagged sites (EST). EST databases can be screened for sequences homologous to traits of interest such as disease resistance, fruit weight, and biochemical pathway regulation (Martin 1998). These sequences can then be placed on the genetic map to determine if they are located and linked to traits of interest. This approach has been used to identify *R*-homologues, which are disease resistance genes, in soybean (Kanazin et al. 1996), cacao (Kuhn et al 2003) and *Arabidopsis* (Botella et al. 1997), and this technique has been employed for identifying orthologous genes in related species such as *Arabidopsis* and tomato (Fulton et al. 2002), and *Arabidopsis* and potato (Gebhardt et al. 2004). It is foreseeable that the same approach can be used for identifying homologous genes in *Prunus*, from already generated ESTs in almond (Jiang et al. 2003) and apricot (Decroocq et al. 2003).

The approaches discussed above are all viable options for MAS in almond breeding programs. The use of MAS for fruit breeders offers the opportunity to improve the selection efficiency of breeding programs while limiting the risk associated with culling potentially genetically superior individuals by phenotypic selection alone. The adoption of any new selection technique must be both cost and benefit efficient. A study in pigs has shown that the use of MAS for QTL selection is a cost effective and viable method in a commercial enterprise (Hayes and Goddard 2003). In the case of MAS for this project it has been implemented in three of the four stages, the first of which was the production of molecular markers. Secondly the establishment of a genetic map based on these markers precluded the detection of linkage between markers and three traits of agronomic importance and scientific interest. Finally the use of this genetic

map and identified linked markers in breeding programs can be undertaken. This study has provided a huge step forward for MAS in almond, and the tighter the linkage between the trait and marker, the more efficient selection based on this marker will be. Further saturation of this map will allow it be used for step four of MAS, and as a valuable tool in the almond breeding program at the University of Adelaide, as well as providing data for other stone fruit mapping programs worldwide.

The progress made by the utilisation of DNA marker technology for crop improvement has been greatly accelerated in the last decade by improved techniques and the adoption of automated methodologies. The development of high density molecular maps now offers the possibility to map and tag virtually all traits of interest, from simply inherited to quantitative trait loci (Gibson and Sommerville 1993; Kumar 1999; Remington et al. 2001). The application of this technology can also be applied to map based cloning of candidate genes. This approach has been used for identifying clones linked to *S*-loci in apricot (Vilanova et al. 2003) and almond (Ushijima et al. 1999) and the *evergrowing* gene in peach (Wang et al. 2002). The availability of techniques for producing BAC libraries in peach (Georgi et al. 2002; Wang et al. 2001) and apricot (Vilanova et al. 2003) can be applied to almond for map base cloning and gene isolation, in combination with high density maps. The production of genetic and physical maps for members of the *Prunus* will allow cross-referencing of conserved gene order and expede the transfer of orthologous genes among related species (Säll et al. 1993). These approaches will further advance the understanding of plant breeding and characterization of previously unknown genes in *Prunus*.

REFERENCES

Ahn S, Tanksley S (1993) Comparative linkage maps of the rice and maize genomes. Proceedings of the National Academy of Science 90:7980-7984

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

Anderson D, Sweeny D, Williams T (1994) Introduction to Statistics, Concepts and Applications. West Publishing Company, St Paul, pp Chapters 3 and 13

Anon (2002) Agricultural Commodities. Australian Bureau of Statistics, Australia, pp 16

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

Aranzana M, Carbo J, Arús P (2003a) Using amplified fragment length polymorphisms (AFLPs) to identify peach cultivars. Journal of the American Society for Horticultural Science 128:672-677

Aranzana MJ, Pineda A, Cosson P, Dirlewanger, Ascasibar J, Cipriani G, Ryder CD, Testolin R, Abbott A, King GJ, Iezzoni AF, Arus P (2003b) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. Theoretical and Applied Genetics 106:819-825

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

Archibald R, Melton L (1987) The anatomy of the fleshy pericarp of maturing Moorpark apricots, *Prunus armeniaca*. New Zealand Journal of Botany 25:181-184

Arteaga N, Socias i Company R (2002) Heritability of fruit and kernel traits in almond. Acta Horticulturae 591:269-274

Arulsekar S, Parfitt D, Kester D (1989) Comparison of isozyme variability in peach and almond cultivars. Journal of Heredity 77:272-274

Arumuganathan K, Earle E (1991) Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter 9:208-218

Arús P (1996) Developing tools for a more efficient fruit breeding: The European *Prunus* Mapping Project. FAO- Nucis-Newsletter. Vol. 5, pp 6-8

Arús P, Messeguer R, Viruel M, Tobutt K, Dirlewanger E, Santi F, Quarta R, Ritter E (1994a) The European *Prunus* mapping project. Progress in the almond linkage map. Euphytica 77:97-100

Arús P, Olarte C, Romero M, Vargus F (1994b) Linkage analysis of ten isozyme genes in F_1 segregating almond progenies. Journal of the American Society for Horticultural Science 119:339-344

Asensio M, Socias i Company R (1996) Double kernels in almond: an open question. FAO - Nucis Newsletter 5:8-9

Asíns MJ, Mestre P, García JE, Dicenta F, Carbonell EA (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

Atwood J, Arnold C (1982) Changes in trichome length and density during growth of peach fruits. HortScience 17:228-230

Ayliffe M, Lawrence G, Ellis J, Pryor A (1994) Heteroduplex molecules formed between allelic sequences cause nonparental RAPD bands. Nucleic Acids Research 22:1632-1636

Badenes M, Hurtado M, Sanz F, Archelos D, Burgos L, Egea J, Llácer G (2000) Searching for molecular markers linked to male sterility and self-compatibility in apricot. Plant Breeding 119:157-160

Baird W, Ballard R, Rajapakse S, Abbott A (1996) Progress in *Prunus* mapping and application to molecular markers to germplasm improvement. HortScience 31:1099-1106

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

Ballester J (1998) Almonds Genetica Vegetal. Universitat Autonoma, Barcelona

Ballester J, Bŏsković R, Batlle I, Arús P, Vargus F, de Vincente M (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 Vincente M (2001) Genetic mapping of a major gene delaying bloom time in almond. Plant Breeding 120:268-270

Barbera G, Di Marco L, Fatta Del Bosco G, Inglese P (1988) Behaviour of 26 almond cultivars growing under rainfed and semiarid conditions in Sicily. In: Grasselly C (ed) GREMPA XII (groupe de recherche d'etude Mediterraneen pour le pistachier et l'amandier). Agriculture: program de recherche Agrimed, Luxembourg, Rapport

Barreneche T, Bodenes C, Lexer C, Trontin J-F, Fluch S, Streiff R, Plomion C, Gössell Kremer A (1998) A genetic map of *Quercus robur* L. (pendulate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5s rDNA markers. Theoretical and Applied Genetics 97:1090-1103

Bartolozzi F, Warburton M, Arulsekar S, Gradziel T (1998) Genetic characterisation and relatedness among California almond cultivars and breeding lines detected by randomly amplified polymorphic DNA (RAPD) analysis. Journal of the American Society for Horticultural Science 123:381-387

Batlle I, Tanriver E, Romero M, Vargas F (1998) Stylar ribonucleases applied to almond breeding. *Prunus* Genetic Resources Newsletter:5-6

Baxter P (1997) Growing fruit in Australia. For profit or pleasure. Macmillan, Sydney

Beckmann J, Soller M (1983) Restriction fragment length polymorphisms in genetic improvement: methodologies, mapping and costs. Theoretical and Applied Genetics 67:35-43

Beckmann J, Soller M (1986) Restriction fragment length polymorphisms and genetic improvement of agricultural species. Euphytica 35:111-124

Belthoff L, Ballard R, Abbott A, Baird W, Morgens P, Callahan A, Scorza R, et al (1993) Development of a saturated linkage map of *Prunus persica* using molecular based marker systems. Acta Horticulturae 336:51-56

Bennett C (1999) Almond industry profile in Australia. NUCIS 8:17-18

Bennett C, Johns G (2002) The Australian almond industry - committed to research. Acta Horticulturae 591:159

Bennett M, Smith J, Heslop-Harrison J (1982) Nuclear DNA amounts in angiosperms. Proceedings of the Royal Society London 216:179-199

Ben-Njima N, Socias i Company R (1995) Characterisation of some self compatible almonds. I. Pollen tube growth. HortScience 30:318-320

Bernad D, Socias i Company R (1995) Characterisation of some self-compatible almonds. II. Flower phenology and morphology. HortScience 30:321-324

Bertozzi T, Bennett C, Sedgley M (1998) Almond improvement in Australia. NUCIS 7:10-11

Blair M, Panaud O, McCouch S (1999) Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.). Theoretical and Applied Genetics 98:780-792

Boiteux L, Fonseca M, Simon P (1999) Effects of plant tissue and DNA purification method on randomly amplified polymorphic DNA-based genetic fingerprinting analysis in carrot. Journal of the American Society for Horticultural Science 124:32-38

Boonprakob U, Byrne D, Graham C, Okie W, Beckman T, Smith B (2001) Genetic relationships among cultivated diploid plums and their progenitors as determined by RAPD markers. Journal of the American Society for Horticultural Science 126:451-461

Bornet B, Branchard M (2001) Nonanchored Inter Simple Sequence Repeat (ISSR) Markers: Reproducible and specific tools for genome fingerprinting. Plant Molecular Biology Reporter 19:209-215

Bŏsković R, Tobutt K, Batlle I, Duval H (1997) Correlation of ribonuclease zymograms and incompatibility genotypes in almond. Euphytica 97:167-176

Bŏsković R, Tobutt K, Batlle I, Duval H, Martínez-Gómez P, Gradziel T (2003) Stylar ribonucleases in almond: correlation with and prediction of incompatibility genotypes. Plant Breeding 122:70-76

Bŏsković R, Tobutt K, Duval H, Batlle I, Dicenta F, Vargas F (1999) A stylar ribonuclease assay to detect self-compatible seedlings in almond progenies. Theoretical and Applied Genetics 99:800-810

Botella M, Coleman M, Hughes D, Nishimura M, Jones J, Sommerville S (1997) Map positions of 47 *Arabidopsis* sequences with similarity to disease resistance genes. Plant Journal 12:1197-1211

Botstein D, White R, Skonick M, Davis R (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American Journal of Human Genetics 32:314-331

Bradshaw H, Stettler R (1994) Molecular genetics of growth and development in *Populus*. II. Segregation distortion due to genetic load. Theoretical and Applied Genetics 89:551-558

Brady J, Scott N, Thomas M (1996) DNA tying of hops (*Humulus lupulus*) through application of RAPD and microsatellite marker sequences converted to sequence tagged sites (STS). Euphytica 91:277-284

Brooks R (1939) A growth study of the almond fruit. Journal of the American Society for Horticultural Science 37:193-197

Browicz K, Zohary D (1996) The genus *Amygdalus* L. (*Rosaceae*): Species relationships, distribution and evolution under domestication. Genetic Resources and Crop Evolution 43:229-247

Burgos L, Egea J, Guerriero R, Viti R, Monteleone P, Audergon J (1997) The self-compatibility trait of the main apricot cultivars of new selections from breeding programs. Journal of Horticultural Science 72:147-154

Butler D, Cullis B, Gilmour A, Gogel B (2000) SAMM Reference Manual. Queensland Department of Primary Industries

Byrne D (1990) Isozyme variability in four diploid stone fruits compared with other woody perennial plants. Journal of Heredity 81:68-71

Byrne M, Murrell J, Allen B, Moran G (1995) An integrated genetic linkage map for *Eucalyptus* using RFLP, RAPD and isozyme markers. Theoretical and Applied Genetics 91:869-875

Cai Q, Guy C, Moore G (1994) Extension of the linkage map in *Citrus* using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. Theoretical and Applied Genetics 89:606-614

Cantini C, Iezzoni A, Lamboy W, Boritzki M, 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, 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

Cascio C, Sathian K (2001) Temporal cues contribute to tactile perception of roughness. The Journal of Neuroscience 21:5289-5296

Cekic C, Battey N, Wilkinsohn M (2001) The potential of ISSR-PCR primer-pair combinations for genetic linkage analysis using the seasonal flowering locus in *Fragaria* as a model. Theoretical and Applied Genetics 103:540-546

Cerezo M, Socias i Company R, Arús P (1989) Identification of almond cultivars by pollen isozymes. Journal of the American Society for Horticultural Science 114:164-169

Certal A, Sanchez A, Kokko H, Broothaerts W, Oliveira M, Feijo J (1999) *S*-RNases in apple are expressed in the pistil along the pollen tube growth path. Sexual Plant Reproduction 12:94-98

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

Chang C, Meyerowitz E (1991) Plant genome studies: restriction fragment length polymorphism and chromosome mapping information. Current Opinions in Biotechnology 1:178-183

Channuntapipat C (2002) Almond improvement via micropropagation, cryopreservation and *S*-allele identification. PhD Thesis. Department of Horticulture, Viticulture and Oenology. Adelaide University, Adelaide, pp 223

Channuntapipat C, Collins G, Sedgley M (2002a) Self-incompatibility of almond varieties. Australian Nutgrower 16:40-41

Channuntapipat C, Sedgley M, Batlle I, Arús P, Collins G (2002b) Sequences of the genomic DNAs encoding the *S2*, *S9*, *S10*, and *S23* alleles from almond, *Prunus dulcis*. Journal of Horticultural Science and Biotechnology 77:387-392

Channuntapipat C, Sedgley M, Collins G (2001) Sequences of the cDNAs and genomic DNAs encoding the *S1*, *S7*, *S8*, and *Sf* alleles from almond, *Prunus dulcis*. Theoretical and Applied Genetics 103:1115-1122

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

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

Charlesworth D (2000) Unlocking the secrets of self-incompatibility. Current Biology 10:184-186

Che K, Zhan Q, Xing Q, Wang Z, Jin D, He D, Wang B (2003) Tagging and mapping of rice sheath blight resistance gene. Theoretical and Applied Genetics 106:293-297

Cheng H-Y, Yang W-C, Hsiao J-Y (2001) Genetic diversity and relationship among peach cultivars based on Random Amplified Microsatellite Polymorphism (RAMP). Botanical Bulletin of Academia Sinica 42:201-206

Cho Y, Eun M, McCouch S, Chae Y (1994) The semidwarf gene, *sd-1*, of rice (*Oryza sativa* L.). II. Molecular mapping and marker-assisted selection. Theoretical and Applied Genetics 89:54-59

Cipriani G, Frazza G, Peterlunger E, Testolin R (1994) Grapevine fingerprinting using microsatellite repeats. Vitis 33:211-215

Cipriani G, Lot G, Huang W-G, Marrazzo M, Peterlunger E, 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

Claverie M, Bosselut N, Lecouls A, Voisin R, Lafargue B, Poizat C, Kleinhentz M, et al (2004) Location of independent root-knot nematode resistance genes in plum and peach. Theoretical and Applied Genetics 108:765-773

Collins G, Symons R (1993) Polymorphisms in grapevine DNA detected by the RAPD PCR technique. Plant Molecular Biology Reporter 11:105-112

Colmagro S, Collins, G, Scott E (2002) Anthracnose disease in almond variation of the causal pathogen *Colletrichum actuatum*. Nutgrower 16:17-19

Condit R, Hubbell S (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Theoretical and Applied Genetics 34:66-71

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

Conner P, Brown S, Weeden N (1997) Randomly amplified polymorphic DNA based genetic linkage maps of three apple cultivars. Theoretical and Applied Genetics 122:350-359

Cordeiro V, Oliveira M, Ventura J, Monteiro A (2001) Study of some physical and nutritive composition of the Portuguese's (local) almond varieties XI GREMPA Seminar on Pistachio and Almond. Vol. 56. Cahiers OPTIONS mediterraneennes, Zaragoza, pp 333-337

Corp I (2002) S-PLUS 6 for Windows Guide to Statistics, Volume 1, Insighful Corp, Wa

Cousin M, Maataoui M (1998) Female reproductive organs in self-compatible almond (*Prunus dulcis* (Mill.) D.A. Webb) Lauranne and fertilization patterns. Scientia Horticulturae 72:287-297

Creller M, Werner D (1996) Characterizing the novel fruit surface morphology of 'Marina' peach using scanning electron microscopy. Journal of the American Society for Horticultural Science 121:198-203

Crowe L (1964) The evolution of outbreeding in plants. I. The angiosperms. Heredity 19:435-457

Crowe JF, Dove WF (1990) Mapping functions. Genetics 125:669-671

Dayanandan S, Bawa K, Kesseli R (1997) Conservation of microsatellites among tropical trees (Leguminosae). American Journal of Botany 84:1658-1663

De Benedetti L, Burchi G, Mercuri A, Pecchioni N, Faccioli P, Schiva T (2000) Random amplified polymorphic DNA (RAPD) analysis for the verification of hybridity in interspecific crosses of *Alstroemeria*. Plant Breeding 119:443-445

de Nettancourt D (1977) Incompatibility in angiosperms. Springer, Berlin, Heidelberg, New York

de Vienne D (2003) Molecular markers in plant genetics and biotechnology. Science Publishers, Inc, Plymouth, UK, pp 231

Debener T, Mattiesch L (1999) Construction of a genetic linkage map for roses using RAPD and AFLP markers. Theoretical and Applied Genetics 99:891-899

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

Demeke T, Adams R (1992) The effects of plant polysaccharides and buffer additives on PCR. BioTechniques 12:332-334

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

Di Gaspero G, Peterlunger E, Testolin R, Edwards K, Cipriani G (2000) Conservation of microsatellite loci within the genus *Vitis*. Theoretical and Applied Genetics 101:301-308

Dicenta F, García J (1993a) Inheritance of self-incompatibility in almond. Heredity 70:313-317

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

Dicenta F, García J, Carbonell E (1993) Heritability of fruit characters in almond. Journal of Horticultural Science 68:121-126

Dicenta F, Martínez-Gómez P, Ortega E (2000) Cultivar pollinizer does not affect almond flavour. HortScience 35:1153-1154

Dicenta F, Ortega E, Canovas JA, Egea J (2002) Self-pollination vs. cross-pollination in almond: pollen tube growth, fruit set and fruit characteristics. Plant-Breeding 121:163-167

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

Dicenta F, Ortega E, Cánovas J, Egea J (2002b) Self-pollination versus cross-pollination of six self-compatible almond cultivars: Pollen tube growth and fruit set XI GREMPA Seminar on Pistachios and Almonds. Cahiers options mediterraneennes, Zaragoza, pp 369-372

Dickson H (1994) Simply a social disease ? Nature 367:517-518

DiGrandi-Hoffman G, Thorp R, Lopez G, Eisikowitch D (1996) Describing the progression of almond bloom using accumulated heat units. Journal of Applied Ecology 33:812-818

Dirlewanger E, Bodo C (1994) Molecular genetic mapping of peach. Euphytica 77:101-103

Dirlewanger E, Pascal T, Zuger C, Kervella J (1996) Analysis of molecular markers associated with powdery mildew resistance genes in peach (*Prunus persica* (L.) Batsch) x *Prunus davidiana* hybrids. Theoretical and Applied Genetics 93:909-919

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

Do N, Adams R (1991) A simple technique for removing plant polysaccharide contaminants from DNA. BioTechniques 10:162-166

Dodge F (1946) A study of the degree of filling of pecan nuts as affected by the number produced by the tree. Proceedings of the American Society of Horticultural Science 47:139-146

Donini P, Elias M, Bougourd S, Koebner R (1997) AFLP fingerprinting reveals pattern differences between template DNA extracted from different plant organs. Genome 40:521-526

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

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

Ebert P, Anderson M, Bernatzky R, Altschuler M, Clarke A (1989) Genetic polymorphism of self-incompatibility in flowering plants. Cell 56:255-262

Edelman M (1975) Purification of DNA by affinity chromatography: Removal of polysaccharide contaminants. Analytical Biochemistry 65:293-297

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

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

Egea J, Dicenta F, Berenguer T, García J (2000) 'Antoneta' and 'Marta' almonds. HortScience 35:1358-1359

Eldredge L, Ballard R, Baird V, Abbott A, Morgens P, Callahan A, Scorza R, Monet, R (1992) Application of RFLP analysis to genetic linkage mapping in peaches. HortScience 27:160-163

Fang D, Roose M (1999) Inheritance of Intersimple Sequence Repeat markers in *Citrus*. Journal of Heredity 90:247-249

Fang G, Hammer S, Grumet R (1992) A quick and inexpensive method for removing polysaccharides from plant genomic DNA. BioTechniques 13:52-55

FAO (2003) Almond production (Mt) - World. Vol. 2004. FAO (www.faostat.fao.org)

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

Foolad M, Jones R, Rodriguez R (1993) RAPD markers for constructing intraspecific tomato genetic maps. Plant Cell Reports 12:293-297

Foulongne M, Pascal T, Arús P, Kervella J (2003) The potential of *Prunus davidiana* for introgression into peach [*Prunus persica* (L.) Batsch] assessed by comparative mapping. Theoretical and Applied Genetics 107:227-238

Frehner M, Scalet M, Conn E (1990) Pattern of the cyanide-potential in developing fruits. Plant Physiology 94:28-34

Fritsch P, Hanson M, Spore C, Pack P, Rieseberg L (1993) Consistency of RAPD primer amplification strength among distantly related taxa of flowering plants. Plant Molecular Biology Reporter 11:10-20

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

Garland S, Henry R (2001) Applications of molecular markers to rice breeding in Australia. Rural Industries Research and Development Corporation, Barton, pp 21

Garland S, Lewin L, Blakeney A, Reinke R, Henry R (2000) PCR based molecular markers for the fragrance gene in rice (*Oryza sativa* L.). Theoretical and Applied Genetics 101:364-371

Gebhardt C, Ritter E, Barone A, Debener T, Walkemeier B, Schachtschabel U, Kaufmann H, Thompson RD, Bonierbale MW, Ganal MW, Tanksley SD, Salamini F (1991) RFLP maps of potato and their alignment with the homologous tomato genome. Theoretical and Applied Genetics 83:49-57

Gebhardt C, Walkemeier B, Henselewski H, Barakat A, Delseny M, Stüber K (2004) Comparitive mapping between potato (*Solanum tuberosum*) and *Arabidopsis thaliana* reveals structurally conserved domains and ancient duplications in the potato genome. The Plant Journal 34:529-541

Georgi L, Wang Y, Yvergniaux D, Ormsbee T, Iñigo M, Reighard G, Abbott A (2002) Construction of a BAC library and its application to the identification of simple sequence repeats in peach [*Prunus persica* (L.) Batsch]. Theoretical and Applied Genetics 105:1151-1158

Gibson S, Somerville C (1993) Isolating plant genes. TIBTECH 11:306-313

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

Godwin I, Aitken E, Smith L (1997) Application of inter simple sequence repeat (ISSR) markers to plant genetics. Electrophoresis 18:1524-1528

Gomez E, Burgos L, Soriano C, Marin J (1998) Amygdalin content in the seeds of several apricot cultivars. Journal of Science and Food Agriculture 77:184-186

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

Gothilf S (1984) Biology of *Spectrobates ceratoniae* on almonds in Israel. Phytoparasitica 12:77-87

Goulão L, Monte-Corvo L, Oliveira C (2001a) Phenetic characterisation of plum cultivars by high multiplex ratio markers: Amplified Fragment Length Polymorphisms and Inter-Simple Sequence Repeats. Journal of the American Society for Horticultural Science 126:72-77

Goulão L, Valdiviesso T, Santana C, Moniz Oliveira C (2001b) Comparison between phenetic characterisation using RAPD and ISSR markers and phenotypic data of cultivated chestnut (*Castanea sativa* Mill.). Genetic Resources and Crop Evolution 48:329-338

Gradziel T, Kester D (1997) Breeding for self-fertility in California almond cultivars. Acta Horticulturae 470:109-117

Gradziel T, Kester D (1999) The University of California almond breeding programme: II. Breeding objectives and projects. FAO - Nucis Newsletter 8:10-14

Gradziel T, Martínez-Gómez P (2002) Shell seal breakdown in almond is associated with the site of secondary ovule abortion. Journal of the American Society of Horticultural Science 127:69-74

Graham CJ (2002) Nonstructural carbohydrate and prunasin composition of peach seedling fertilized with different nitrogen sources and aluminium Scientia Horticulturae 94:21-32

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

Grattapaglia D, Bertolucci F, Penchel R, Sederoff R (1996) Genetic mapping of quantitative trait loci controlling growth and wood quality traits in *Eucalyptus grandis* using a maternal half-sib family and RAPD markers. Genetics 144 :1205-1214

Grattapaglia D, Bertolucci F, Sederoff R (1995) Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross strategy and RAPD markers. Theoretical and Applied Genetics 90:933-947

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

Graziano E, Wirthensohn M, Ballester J, Batlle I, Arús P (2000) Development of PCR markers for the selection of self compatibility and hard shell in almond Poster presented at Primer Seminario de Mejora Genetica Vegetal, Lugo, Galicia

Gu W, Post C, Aguirre G, Ray K (1999) Individual DNA bands obtained by RAPD analysis of canine genomic DNA often contain multiple DNA sequences. Journal of Heredity 90:96-98

Gulcan R (1985) Descriptor list for almond (*Prunus dulcis*). In: Gulcan R (ed). IBPGR Secretariat, Rome

Hackett C, Wachira F, Paul S, Powell W, Waugh R (2000) Construction of a genetic linkage map for *Camellia sinensis* (tea). Heredity 85:346-355

Hagen L, Khadari B, Lambert P, Audergon J (2002) Genetic diversity in apricot revealed by AFLP markers: species and cultivar comparisons. Theoretical and Applied Genetics 105:298-305

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

Hattori J, Gottlob-McHugh S, Johnson D (1987) The isolation of high-molecular weight DNA from plants. Analytical Biochemistry 165:70-74

Hauagge R, Kester D, Arulsekar S, Parfitt D, Liu L (1987) Isozyme variation among California almond cultivars. II. Cultivar characterization and origins. Journal of the American Society for Horticultural Science 112:693-698

Hauck N, Iezzoni A, Yamane H, Tao R (2001) Revisiting the *S*-allele nomenclature in sweet cherry (*Prunus avium*) using RFLP profiles. Journal of the American Society for Horticultural Science 126:654-660

Hauck N, Yamane H, Tao R, Iezzoni A (2002) Self-compatibility and incompatibility in tetraploid sour cherry (*Prunus cerasus* L.). Sexual Plant Reproduction 15:39-46

Hausner G, Rashid K, Kenaschuk E, Procunier J (1999) The identification of a cleaved amplified polymorphic sequence (CAPS) marker for the flax rust resistance gene M^3. Canadian Journal of Plant Pathology 21:187-192

Hawker J, Buttrose M (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

Hayden KJ, Parker IM (2002) Plasticity in cyanogenesis of *Trifolium repens* L: inducibility, fitness costs and variable expression. Evolutionary and Ecology Research 4:155-168

Hayes B, Goddard M (2003) Evaluation of marker assisted selection in pig enterprises. Livestock Production Science 81:197-211

He P, Li S, Qian Q, Ma Y, Li J, Wang W, Chen Y, Zhu LH (1999) Genetic analysis of rice grain quality. Theoretical and Applied Genetics 98:502-508

Helentjaris T, King G, Slocum M, Siedenstrang C, Wegman S (1985) Restriction fragment length polymorphisms as probes for plant diversity and their development as tools for applied plant breeding. Plant Molecular Biology 5:109-118

Henry R (1997) Practical applications of plant molecular biology. Chapman and Hall, London, pp 101-133

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

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

Hill S, Stephenson D, Taylor B (1987) Almond yield in relation to tree size. Scientia Horticulturae 33:97-111

Hiramoto K, Li X, Makimoto M, Kato T, Kikugawa K (1998) Identification of hydroxyhydroquinone in coffee as a generator of reactive oxygen species that break DNA single strands. Mutation Research 419:43-51

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

Hormaza J (1999) Early selection in cherry combining RAPDS with embryo culture. Scientia Horticulturae 79:121-126

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

Hurtado M, Romero C, Vilanova S, Abbott A, Llácer G, 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

Ishimizu T, Endo T, Yamaguchi-Kabata Y, Nakamura T, Sakiyama F, Norioka S (1998) Identification of regions in which positive selection may operate in *S*-RNase of Rosaceae: Implication for *S*-allele specific recognition sites in *S*-RNase. FEBS Letters 440:337-342

Ishimizu T, Inoe K, Shimonaka M, Saito T, Terai O, Norioka S (1999) PCR-based method for identifying the *S*-genotypes of Japanese pear cultivars. Theoretical and Applied Genetics 98:961-967

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

Jackson J, Clarke G (1991) Gene flow in an almond orchard. Theoretical and Applied Genetics 82:169-173

Janssens G, Goderis I, Broekaert W, Broothaerts W (1995) A molecular method for *S*-allele identification in apple based on allele-specific PCR. Theoretical and Applied Genetics 91:691-698

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

Jiang Y, Cai M, Jiang Y, Ma R (2003) Generation and analysis of expressed sequence tags from almond (*Prunus dulcis* Mill.) pistils. Sexual Plant Reproduction 16:197-207

Jobes D, Hurley D, Thein L (1995) Plant DNA isolation: a method to efficiently remove polyphenolics, polysaccharides and RNA. Taxon 44:379-386

John M (1992) An efficient method for isolation of RNA and DNA from plants containing polyphenolics. Nucleic Acids Research 20:2381

Jones C, Edwards K, Castaglione S, Winfield M, Sala F, Vandewiel 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, Karp A (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Molecular Breeding 3:381-390

Jones D (1998) Why are so many plants cyanogenic? Phytochemistry 47:155-162

Joobeur T (1996) Cartographie genetique chez l'amandier a l'aide de marqueurs RAPD PhD Thesis. Instituto Agronomico Mediterraneo de Zaragoza, Spain

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

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

Kanazin V, Marek L, Shoemaker R (1996) Resistance gene analogues are conserved and clustered in soybean. Proceedings of the National Academy of Sciences 93:11746-11750

Kao T-H, McCubbin A (1996) How flowering plants discriminate between self and non-self pollen to prevent inbreeding. Proceedings of the National Academy of Sciences USA 93:12059-12065

Kearsey M, Pooni H (1996) Genetical Analysis of Quantitative Traits. Chapman and Hall, London, UK

Keil M, Griffin A (1994) Use of random amplified polymorphic DNA (RAPD) markers in the discrimination and verification of genotypes in *Eucalyptus*. Theoretical and Applied Genetics 89:442-450

Kester D (1965a) Inheritance of time of bloom in certain progenies of almond. Journal of the American Society for Horticultural Science 87:214-221

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

Kester D, Gradziel T (1996) Almonds. In: Janick JM and Moore JJ (ed) Fruit Breeding. Vol. III Nuts. John Wiley and Sons, New York, pp 1-97

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

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

Kester D, Gradziel T, Micke W (1994) Identifying pollen incompatibility groups in California almond cultivars. Journal of the American Society for Horticultural Science 119:106-109

Kester D, Micke W (1994) A mutation in 'Nonpareil' almond conferring unilateral incompatibility. Journal of the American Society for Horticultural Science 119:1289-1292

Kim C, Lee C, Shin J, Chung Y, Hyung N (1997) A simple and rapid method for isolation of high quality genomic DNA from fruit trees and conifers using PVP. Nucleic Acids Research 25:1085-1086

Klein I, Esparza G, Weinbaum S, DeJong T (2001) Effects of irrigation deprivation during the harvest period on leaf persistence and function in mature almond trees. Tree Physiology 21:1063-1072

Koch C, Koch E (2003) Preconceptions of taste based on colour. The Journal of Psychology 137:233-242

Kosambi D (1944) The estimation of map distance from recombination values. Annals of Eugenetics 12:172-175

Kowyama Y, Kunz C, Lewis I, Newbigin E, Clarke A, Anderson M (1994) Self-compatibility in a *Lycopersicon peruvianum* variant (LA2157) is associated with a lack of style S-RNase activity. Theoretical and Applied Genetics 88:859-864

Kragler F, Lucas W, Monzer J (1998) Plasmodesmata: dynamics, domains and patterning. Annals of Botany 81:1-10

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

Kuhn DN, Heath M, Wisser RJ, Meerow A, Brown JS, Lopes U, Schnell RJ (2003) Resistance homologues in *Theobroma cacao* as useful genetic markers. Theoretical and Applied Genetics 107:191-202

Kumar K, Das B (1996) Studies on xenia in almond (*Prunus dulcis* (Miller) D.A. Webb). Journal of Horticultural Science 71:545-549

Kumar LS (1999) DNA markers in plant improvement: An overview. Biotechnology Advances 17:143-182

la Rosa R, Angiolillo A, Guerrero C, Pellegrini M, Rallo L, Besnard G, Berville A, Martin A, Baldoni L, la Rosa R (2003) A first genetic linkage map of olive (*Olea europaea* L.) cultivars using RAPD, AFLP, RFLP and SSR markers. Theoretical and Applied Genetics 106:1273-1282

Ladizinsky G (1999) On the origin of almond. Genetic Resources and Crop Evolution 46:143-147

Lambert P, Hagen L, Arús P, Audergon J (2003) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas x peach Earlygold reference map for *Prunus*. Theoretical and Applied Genetics 108:1120-1130

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

Landry B, Kesseli R, Farrara B, Michelmore R (1987) A genetic map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics 116:331-337

Lansari A, Kester D, Iezzoni A (1994) Inbreeding, coancestry, and founding clones of almond of California, Mediterranean shores and Russia. Journal of the American Society for Horticultural Science 119:1279-1285

Lecouls A, Rubio-Cabetas M, Minot J, Voisin R, Bonnet A, Salesses G, Dirlewanger E, et al (1999) RAPD and SCAR markers linked to the *Ma1* root-knot nematode resistance gene in Myrobolan plum (*Prunus cerasifera* Ehr.). Theoretical and Applied Genetics 99:328-335

Ledbetter C, Pyntea M (2000) The significance of kernel bitterness in the context of breeding almonds for self-incompatibility. Journal of Genetics and Breeding 54:221-225

Lee H-S, Shihshieh H, Kao T-H (1994) S proteins control rejection of incompatible pollen in *Petunia inflata*. Nature 367:560-563

Levi M, Thomas C, Zhang X, Joobeur T, Dean R, Wehner T, Carle B (2001) A genetic linkage map for watermelon based on randomly amplified polymorphic DNA markers. Journal of the American Society for Horticultural Science 126:730-737

Lewis D, Crowe L (1954) The induction of self-fertility in tree fruits. Journal of Horticultural Science 29:221-225

Li W, Wirthensohn MG, Sedgley M (2004) Testing virulence of two pathovars of *Xanthomonas campestris* pv. *pruni* on three almond cultivars (*Prunus dulcis* Mill.) using detatched leaf assay. Nutgrower 18: in press

Lian C, Zhou Z, Hogetsu T (2001) A simple method for developing microsatellite markers using amplified fragments of inter-simple sequence repeat (ISSR). Journal of Plant Research 114:381-385

Lin J-J, Kuo J, Ma J (1996a) Effect of different primer combinations on the resolution of AFLP in plants with small genomes. Focus 18:68-69

Lin J-J, Kuo J, Ma J, Saunders J, Beard H, MacDonald M, Kenworthy W, et al (1996b) Identification of molecular markers in soybean comparing RFLP, RAPD and AFLP DNA mapping techniques. Plant Molecular Biology Reporter 14:156-169

Lincoln S, Daly M, Lander E (1992) Constructing genetic maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, Cambridge, Massachusettes

Litt M, Luty J (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American Journal of Human Genetics 4:397-401

Loomis W, Battalie J (1966) Plant phenolic compounds in the isolation of plant enzymes. Phytochemistry 5:423-438

Lorieux M, Petrov M, Huang N, Guiderdoni E, Ghesquiere A (1996) Aroma in rice: genetic analysis of a quantitative trait. Theoretical and Applied Genetics 93:1145-1151

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

Lui B, Knapp S (1990) GMENDEL: A program for mendelian segregation and linkage analysis of individual or multiple progeny populations using log-likelihood ratios. Journal of Heredity 81:407

Lui B, Wendel J (2001) Intersimple sequence repeat (ISSR) polymorphisms as a genetic marker system in cotton. Molecular Ecology Notes 1:205-208

Ma RC, Oliveira M (2001) Molecular cloning of the self-incompatibility genes *S1* and *S3* from almond (*Prunus dulcis* cv. 'Ferragnés'). Sexual Plant Reproduction. 14:163-167

Ma R-C, Xu Y, Xie J, Ma Y, Cao M-Q (2003) Molecular analysis of almond genetic variability in China using EST derived SSR and self-incompatibility typing. 7th International Congress of Plant Molecular Biology. (ISPMB) Barcelona, Spain

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

Manly KF, Cudmore R (1994) MapManager. A program for genetic mapping manual. Rosewell Park Cancer Institute. New York, USA. pp 64-69

Manning K (1990) Isolation of nucleic acids from plants by differential solvent precipitation. Analytical Biochemistry 195:45-50

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

Martin G (1998) Gene discovery for crop improvement. Current Opinion in Biotechnology 9:220-226

Martín-Carratalá M, Llorens-Jordá C, Berenguer-Navarro V, Grané-Teruel N (1999) Comparative study on the triglyceride composition of almond kernel oil. A new basis for cultivar chemometric characterization. Journal of Agricultural Food Chemistry 47:3688-3692

Martínez-Gómez P, Arulsekar S, Potter D, Gradziel T (2003a) An extended interspecific gene pool available to peach and almond breeding as characterized using simple sequence repeat (SSR) markers. Euphytica 131:313-322

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

Martins M, Tenreiro R, Oliveira M (2003) Genetic relatedness of Portuguese almond cultivars assessed by RAPD and ISSR markers. Plant Cell Reports 22:71-78

Mbeguie-A-Mbeguie D, Chahine H, R-M G, Gouble B, Reich M, Audergon J, Souty M, Albagnac G, Fils-Lycaon B (1999) Molecular cloning and expression of a cDNA encoding 1-aminocyclopropane-1-carboxylate (ACC) from apricot fruit (*Prunus armeniaca*). Physiologia Plantarum 105:294-303

McMurray J (1994) Fundamentals of organic chemistry. Brooks/Cole Publishing Company, California, pp 247

Mekuria G, Collins G, Sedgley M (1999) Genetic variability between different accessions of some common commercial olive cultivars. Journal of Horticultural Science and Biotechnology 74:309-314

Mekuria G, Collins G, Sedgley M (2001a) Genetic diversity of a feral population of olive trees surrounding an abandoned grove. Acta Horticulturae 546:249-251

Mekuria G, Collins G, Sedgley M (2001b) 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

Mekuria G, Sedgley M, Collins G, Lavee S (2002) Development of a sequence-tagged site for the RAPD marker linked to leaf spot resistance in olive. Journal of the American Society for Horticultural Science 127:673-676

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

Michelmore R, Paran I, Kesseli R (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. Proceedings of the National Academy of Sciences USA 88:9828-9832

Mnejja M, Howad W, Garcia-Mas J, Arús P (2003) Development, polymorphism and map position of almond (*Prunus amygdalis*) and Japanese plum (*P. salicina*) SSR (simple-sequence repeat) markers. 7th International Congress of Plant Molecular Biology. (ISPMB), Barcelona, Spain

Monte-Corvo L, Goulão L, Oliveira C (2001) ISSR analysis of cultivars of pear and suitability of molecular markers for clone discrimination. Journal of the American Society for Horticultural Science 126:517-522

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

Moss D (1965) The almond industry in South Australia. Journal of Agriculture South Australia 69:38-45

Muralidharan K, Wakeland EK (1993) Concentration of primer and template qualitatively affects products in random-amplified polymorphic DNA PCR. BioTechnology 14:362-364

Murfett J, Atherton T, Mou B, Gasser C, McClure B (1994) S-RNase expressed in transgenic *Nicotiana* causes *S*-allele-specific pollen rejection. Nature 367:563-566

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

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

Nagaoka T, Ogihara Y (1997) Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. Theoretical and Applied Genetics 94:597-602

Nair S, Bentur J, Prasada Rao U, Mohan M (1995) DNA markers tightly linked to a gall midge resistance gene (*Gm2*) are potentially useful for marker-aided selection in rice breeding. Theoretical and Applied Genetics 91:68-73

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

Naor A, Hupert H, Greenblat Y, Peres M, Kaufman A, Klein I (2001) The response of nectarine fruit size and midday stem water potential to irrigation levels in stage III and crop load. Journal of the American Society for Horticultural Science 126:140-143

Neaylon K, Delaporte K, Sedgley M, Collins G, Conran J (2001) Molecular analysis of hybrids among the ornamental eucalypts *Eucalyptus macrocarpa*, *E. pyriformis* and *E. youngiana*. Journal of the American Society for Horticultural Science 126:336-339

Newbigin E, Anderson M, Clarke A (1993) Gametophytic self-incompatibility systems. The Plant Cell 5:1315-1324

Nicese F, Hormaza J, McGranahan G (1998) Molecular characterization and genetic relatedness among walnut (*Juglans regia* L.) genotypes based on RAPD markers. Euphytica 101:199-206

Nillson NO, Säll T, Bengtsson B (1994) Reply to Ellis. Trends in Genetics. 10:114-115

Oliveira M (1999) Biotechnological improvement of almond in Portugal. FAO - Nucis Newsletter 8:14-16

Olson M, Hood L, Cantor C, Botstein D (1989) A common language for physical mapping of the human genome. Science 245:1434-1435

Ortega E, Dicenta F (2003) Inheritance of self-compatibility in almond: breeding strategies to assure self-compatibility in the progeny. Theoretical and Applied Genetics 106:904-911

Panaud O, Anissa-Chaib, Sarr A (2002) Dynamic conservation of apricot *Prunus armeniaca* in Saharian oases: use of AFLP markers to assess genetic diversity in traditional orchards. Euphytica 128:301-305

Pandey R, Adams R, Flournoy L (1996) Inhibition of random amplified polymorphic DNAs (RAPDs) by plant polysaccharides. Plant Molecular Biology Reporter 14:17-22

Paran I, Mitchelmore R (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theoretical and Applied Genetics 85:985-993

Pascual-Albero M, Perez-Munuera I, Lluch M (1998) Cotyledon structure of raw, soaked and roasted almond (*Prunus amygdalus* L.). Food Science and Technology International 4:189-197

Penner G, Bush A, Wise R, Kim W, Domier L, Kasha K, Laroche A, Scoles, G, Molnar, SJ, Fedak, G (1993) Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories PCR methods and applications, Cold Springs Harbor Laboratories Press. pp 341-345

Piagnani C, Iacona C, Intrieri M, Muleo R (2002) A new somaclone of *Prunus avium* shows diverse growth pattern under different spectral quality of radiation. Biologia Plantarum 45:11-17

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

Plieske J, Struss D (2001) Microsatellite markers for genome analysis in *Brassica*. Development in *Brassica napus* and abundance in *Brassicaceae* species. Theoretical and Applied Genetics 102:680-694

Plomion C, O'Malley D, Durel C (1995) Genomic analysis in maritime pine (*Pinus pinaster*). Comparison of two RAPD maps using selfed and open-pollinated seeds of the same individual. Theoretical and Applied Genetics 90:1028-1034

Potter D, Gao F, Aiello G, Leslie C, McGranahan G (2002) Intersimple sequence repeat markers for fingerprinting and determination genetic relationships of Walnut (*Juglans regia*) cultivars. Journal of the American Society for Horticultural Science 127:75-81

Pound, LM, Collins, GG (2003) The carob moth. A literature review with particular reference to its effect on almond. Nutgrower 17 :30-33

Qu L, Hancock J (1997) Randomly amplified polymorphic DNA (RAPD-) based genetic linkage map of blueberry derived from an interspecific cross between diploid *Vaccinium darrowi* and tetraploid *V. corymbosum*. Journal of the American Society for Horticultural Science 122:69-73

Quinn N (1928) The almond in South Australia. Department of Agriculture South Australia Bulletin 220:1-19

Quinn N (1941) Almond culture in South Australia. Department of Agriculture South Australia Bulletin 367:1-24

Rafalski J, Tingey S (1993) Genetic diagnosis in plant breeding: RAPDs, microsatellites and machines. Trends in Genetics 9:275-279

Rajapakse S, Belthoff LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R, Abbott AG (1995) Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. Theoretical and Applied Genetics 90:503-510

Rao K, Lakshiminarasu M, Jena K (2002) DNA markers and marker assisted breeding for durable resistance to bacterial blight disease in rice. Biotechnology Advances 20:33-47

Rattigan K, Hill S (1986) Relationship between temperature and flowering in almond. Australian Journal of Experimental Agriculture 26:399-404

Reddy M, Sarla N, Siddiq E (2002) Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica 128:9-17

Reiss R, Rutz B (1999) Quality control PCR: A method for detecting inhibitors of *Taq* DNA polymerase. BioTechniques 27:920-926

Remington DL, Ungerer MC, Purugganan MD (2001) Map-based cloning of quantitative trait loci: progress and prospects. Genetical Research Cambridge 78:213-218

Reuveni M (2000) Efficacy of trioxystrobin (Flint), a new strobilurin fungicide, in controlling powdery mildews on apple, mango and nectarine, and rust on prune trees. Crop Protection 19:335-341

RHS (1986) Royal Horticultural Society Colour Chart (2nd Edition). London

Rogers S, Bendich A (1994) Extraction of total cellular DNA from plants, algae and fungi. In: Gelvin S, Schilperoort R (eds) Plant molecular biology manual. Kluwer Academic Publishers, Dordrecht

Rohlf F (1998) NTSYS-pc: (Numerical Taxonomy and Multivariate Analysis System). Exeter Software, Setauket, N.Y.

Rossetto M (2001) Sourcing of SSR markers from related plant species. In: Henry R (ed) Plant Genotyping. The DNA fingerprinting of plants. CABI Publishing, Oxon, pp 211-224

Rovira M, Clavé J, Romero M, Santos J, Vargas F (1997) Self-incompatibility in almond progenies. Acta Horticulturae 470:66-71

Rubeena, Ford R, Taylor P (2003) Construction of an intraspecific linkage map of lentil (*Lens culinaris* ssp. *culinaris*). Theoretical and Applied Genetics 107:910-916

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

Ryder E (1971) Genetic studies in lettuce (*Lactuca sativa* L.). Journal of the American Society for Horticultural Science 96:826-828

Säll T, Nilsson N-O, Bengtsson, BO (1993) When everyone's map is different 3:631-633

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning : A Laboratory Manual. Vol. 3. Cold Spring Harbor Laboratory Press, USA

Samish R (1954) Dormancy in woody plants. Annual Review of Plant Physiology 5:183-204

Sankar A, Moore G (2001) Evaluation of inter-simple sequence repeat analysis for mapping in *Citrus* and extension of the genetic linkage map. Theoretical and Applied Genetics 102:206-214

Sax K (1923) The association of size differences with seed coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8:552-560

Schirra M (1997) Postharvest technology and utilization of almonds. Horticultural Reviews 20:267-311

Scott K (2001) Microsatellites derived from ESTs, and their comparison with those derived by other methods. In: Henry R (ed) Plant Genotyping. The DNA fingerprinting of plants. CABI Publishing, Oxon, pp 225-237

Sebastian R, Howell E, King G, Marshall D, Kearsey M (2000) An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct double-haploid mapping populations. Theoretical and Applied Genetics 100:75-81

Sedgley M (1994) Self-incompatibility in woody horticultural species. In: Williams E, Clarke A (eds) Genetic control of self-incompatibility and reproductive development in flowering plants. Kluwer Academic Publishers, London, pp 141-163

Sedgley M (2003) Wild olive selection for quality oil production. RIRDC, 19. Report number UA-54A

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

Segovia Z, Sokol F, Graves I, Ackermann W (1965) Some properties of nucleic acids extracted with phenol. Biochimica et Biophysica Acta 95:329-340

Selvi A, Shanmugasundaram P, Kumar S, Raja J (2002) Molecular markers for yellow stem borer *Scirpophaga incertulas* (Walker) resistance in rice. Euphytica 124:371-377

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

Sharma S, Knox M, Ellis T (1996) AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. Theoretical and Applied Genetics 93:751-758

Sharon D, Hillel J, Mhameed S, Cregan P, Lahav E, Lavi U (1998) Association between DNA markers and loci controlling avocado traits. Journal of the American Society for Horticultural Science 123:1016-1022

Shields R (1993) Pastoral synteny. Nature 365:297-298

Shimada T, Yamamoto T, Hayama H, Yamaguchi M, Hayashi T (1999) Application of AFLP to molecular genetic analysis in peach. Journal of the Japanese Society for Horticultural Science 68:536-542

Shimada T, Yamamoto T, Hayama H, Yamaguchi M, Hayashi T (2000) A genetic linkage map constructed by using an intraspecific cross between peach cultivars grown in Japan. Journal of the Japanese Society for Horticultural Science 69:536-542

Shioda M, Murakami-Murofushi K (1987) Selective inhibition of DNA polymerase a by a polysaccharide purified from slime of *Physarum polycephalum*. Biochemical and Biophysical Research Communcation 146:61-66

Skroch P, Nienhuis J (1995) Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. Theoretical and Applied Genetics 91:1086-1091

Sneath P, Sokal R (1973) Numerical Taxonomy. The principles and practice of numerical classification. W.H. Freeman and Company, San Francisco, pp 573

Socias i Company R (1991) Breeding self-compatible almonds. Plant Breeding Reviews 8:313-338

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 (1999) Qualitative traits in almond trees. FAO - Nucis Newsletter 8:18-20

Socias i Company R (2001) Latest advances in almond self-compatibility. Acta Horticulturae 591:205-212

Socias i Company R, Felipe A, Gómez-Aparisi J (1999a) Genetics of late blooming in almond. Acta Horticulturae 484:261-265

Socias i Company R, Felipe A, Gómez-Aparisi J (1999b) A major gene for flowering in almond. Plant Breeding 118:443-448

Socias i Company R, Felipe AJ, Gómez-Aparisi J, García JE, Dicenta F (1997) The ideotype concept in almond. Acta Horticulturae 470:51-56

Socias i Company R, Kester D, Bradley M (1976) Effects of temperature and genotype on pollen tube growth in some self-incompatible and self-compatible almond cultivars. Journal of the American Society for Horticultural Science 101:490-493

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, Beck L, King G, Ryder C, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000) Characterisation of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. Theoretical and Applied Genetics 101:421-428

Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology 98:503-517

Spada A, Caporali E, Portaluppi P, Restivo F, Tassi F, Falavigna A (1998) A genetic map of *Asparagus officinalis* based on integrated RFLP, RAPD, and AFLP molecular markers. Theoretical and Applied Genetics 97:1083-1089

Spiegel-Roy P (1986) Domestication of fruit trees. In: Barigozzi C (ed) The origin and domestication of cultivated plants. Elsevier, Amsterdam, pp 201-211

Spiegel-Roy P, Kochba J (1974) The inheritance of bitter and double kernel characters in the almond (*Prunus amygdalus* Batsch). Zeitschrift fur Pflanzenzuchtung 71:319-329

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

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

Staub J, Kuhns L, May B, Grun P (1982) Stability of potato tuber isozymes under different storage regimes. Journal of the American Society for Horticultural Science 107:405-408

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

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

Stevens D (1996) Individual differences in taste perception. Food Chemistry 56:303-311

Stone S, Goring D (2001) The molecular biology of self-incompatibility systems in flowering plants. Plant Cell, Tissue and Organ Culture 67:93-114

Stuber C (1992) Biochemical and molecular markers in plant breeding. Plant Breeding Reviews 9:37-61

Stuber C, Edwards M, Wendel J (1987) Molecular-marker facilitated investigations of quantitative trait loci in maize. II. Factors influencing yield and its component traits. Crop Science 27:639-648

Suiter K, Wendel J, Case J (1983) Linkage-1: A pascal computer program for the detection and analysis of genetic linkage. Journal of Heredity 74:203-204

Surmacka Szczeniak A (2002) Texture is a sensory property. Food Quality and Preference 13:215-225

Tamura M, Ushijima K, Sassa H, Hirano H, Tao R, Gradziel T, Dandekar A (2000) Identification of self-incompatibility genotypes of almond by allele-specific PCR analysis. Theoretical and Applied Genetics 101:344-349

Tanksley S, Young N, Paterson A, Bonierbale M (1989) RFLP mapping in plant breeding: new tools for an old science. Biotechnology 7:257-264

Tao R, Habu T, Yamane H, Sugiura A, Iwamoto K (2000) Molecular markers for self-compatibility in Japanese apricot (*Prunus mume*). HortScience 35:1121-1123

Tao R, Yamane H, Sassa H, Mori H, Gradziel T, Dandekar A, Sugiura A (1997) Identification of stylar RNases associated with gametophytic self-incompatibility in almond (*Prunus dulcis*). Plant Cell Physiology 38:304-311

Tao R, Yamane H, Sugiura A, Murayama H, Sassa H, Mori H (1999) Molecular typing of *S*-alleles through identification, characterization and cDNA cloning for S-RNases in sweet cherry. Journal of the American Society for Horticultural Science 124:224-233

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

Testolin R, Marrazzo T, Cipriani G, Quarta R, Verde I, Dettori M, Pancaldi M, 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

Tikunov YM, Khrustaleva L, Karlov G (2003) Application of ISSR markers in the genus *Lycopersicon*. Euphytica 131:71-80

Torres A, Weeden N, Martin A (1993) Linkage among isozyme, RFLP and RAPD markers in *Vicia faba*. Theoretical and Applied Genetics 85:937-945

Ushijima K, Sassa H, Tamura M, Kusaba M, Tao R, Gradziel T, Dandekar A, Hirano H (2001) Characterization of the S-locus region of almond (*Prunus dulcis*): Analysis of a somaclonal mutant and a cosmid contig for an *S* haplotype. Genetics 158:379-386

Ushijima K, Sassa H, Tao R, Yamane H, Dandekar A, Gradziel T, Hirano H (1998) Cloning and characterization of cDNAs encoding *S*-RNases from almond

(*Prunus dulcis*): primary structural features and sequence diversity of the S-RNases in Rosaceae. Molecular and General Genetics 260:261-268

Uyemoto J, Asai W, Kirkpatrick B (1999) Etiology of almond shriveled kernel disease. New Zealand Journal of Crop and Horticultural Science 27:225-228

van der Nest M, Steenkamp E, Wingfield B, Wingfield M (2000) Development of simple sequence repeat (SSR) markers in *Eucalyptus* from amplified inter-simple sequence repeats (ISSR). Plant Breeding 119:433-436

Van Ooijen J, Sandbrink J, Vrielink M, Verkerk R, Zabel P, Lindhout P (1994) An RFLP linkage map of *Lycopersicon peruvianum*. Theoretical and Applied Genetics 89:1007-1013

Van Ooijen J, Voorrips R (2001) Joinmap ® 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands

Vargas F, Clavé J, Romero M, Batlle I, Rovira M (1997) Autogamy studies on almond progenies. Acta Horticulturae 470:75-81

Vargas F, Romero M (2001) Blooming time in almond progenies. XI GREMPA Seminar on Pistachios and Almonds. Vol. 56. Cahiers OPTIONS mediterraeennes, Zaragoza, pp 29-34

Vargas F, Romero M, Batlle I (2001) Kernel taste inheritance in almond. XI GREMPA Seminar on Pistachios and Almonds. Cahiers options mediterraneennes, Zaragoza, pp 129-134

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

Vetter J (2000) Plant cyanogenic glycosides. Toxicon 38:11-36

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

Vezvaei A, Jackson J (1995) Effect of pollen parent and stages of flower development on almond nut production. Australian Journal of Experimental Agriculture 35:109-113

Vijayan K, Chatterjee S (2003) ISSR profiling of Indian cultivars of mulberry (*Morus spp.*) and its relevance to breeding programs. Euphytica 131:53-63

Vilanova S, Romero C, Abbott A, Llácer G, Badenes M (2003a) An apricot (*Prunus armeniaca* L.) F_2 progeny linkage map based on SSR and AFLP markers, mapping plum pox virus resistance and self-incompatibility traits. Theoretical and Applied Genetics 107:239-247

Vilanova S, Romero C, Abernathy D, Abbott A, Burgos L, Llácer G, Badenes M (2003b) Construction and application of a bacterial artificial chromosome

(BAC) library of *Prunus armeniaca* L. for the identification of clones linked to the self-incompatibility locus. Molecular and General Genetics 269:685-691

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

Voorrips R (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93:77-78

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

Vowden C, Ridout M (1994) LINKEM. A program for genetic linkage analysis. Horticultural Research International, East Malling, England

Wagner G, Wang E, Shepherd R (2004) New approaches for studying and exploiting an old protuberance, the plant trichome. Annals of Botany 93:3-11

Walker J (1975) The biology of plant phenolics. Edward Arnold Limited, London

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

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

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, Reighard G, Scorza R, Abbott A (2002) Genetic mapping of the *evergrowing* gene in peach [*Prunus persica* (L.) Batsch]. Journal of Heredity 93:352-358

Wang Z, Weber J, Zhong G, Tanksley S (1994) Survey of plant short tandem DNA repeats. Theoretical and Applied Genetics 88:1-6

Warburton M, Becerra-Velásquez V, Goffreda J, Bliss F (1996) Utility of RAPD markers in identifying genetic linkages to genes of economic importance in peach. Theoretical and Applied Genetics 93:920-925

Watkins R (1979) Cherry, plum, peach, apricot and almond: *Prunus spp.*, In: Simmonds N (ed) Evolution of crop plants. Longman, London, pp 242-247

Weeden N (1994) Approaches to mapping in horticultural crops. In: Gresshoff P (ed) Plant Genome Analysis. CRC Press, Boca Raton, pp 57-68

Weeden N, Hemmat M, Lawson D, Lodhi M, Bell R, Manganaris A, Reisch B, Brown, SK, Ye G-N (1994) Development and application of molecular marker linkage maps in woody fruit crops. Euphytica 77:71-75

Weissenbach J, Gyapay G, Dib C, Virgnal A, Morissette J, Millasseau P, Vaysseix G, Lathrop, M (1992) A second-generation linkage map of the human genome. Nature 359:794-801

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research 18:7213-7218

Werner D, Creller M (1997) Genetic studies in peach: Inheritance of sweet kernel and male sterility. Journal of the American Society for Horticultural Science 122:215-217

Wiersma P, Wu Z, Zhou L, Hampson C, Kappel F (2001) Identification of new self-incompatibility alleles in sweet cherry (*Prunus avium* L.) and clarification of incompatibility groups by PCR and sequencing analysis. Theoretical and Applied Genetics 102:700-708

Williams J, Hanafey M, Rafalski J, Tingey S (1993) Genetic analysis using random amplified polymorphic DNA markers. Methods in Enzymology 218:704-740

Williams J, Kubelik A, Livak K, Rafalski J, Tingey S (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18:6531-6535

Wing R, Zhang H-B, Tanksley S (1994) Map-based cloning in crop plants tomato as a model system: I. Genetic and physical mapping of jointless. Molecular and General Genetics 242:681-688

Wirthensohn M, Collins G, Jones G, Sedgley M (1999) Variability in waxiness of *Eucalyptus gunnii* foliage for floriculture. Scientia Horticulturae 82:279-288

Wirthensohn M, Collins G, Sedgley M (2001) International collaboration in almond genome mapping. Nutgrower 15:26-27

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

Woodroof J (1979) Almonds Tree Nuts. AVI Publishing Company Inc, Connecticut, USA, pp 139-180

Woolley F, Collins 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 K-S, Tanksley S (1993) Abundance, polymorphism and genetic mapping of microsatellites in rice. Molecular and General Genetics 241:225-235

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

Wu S-B, Collins G, Sedgley M (2004) A molecular linkage map of olive (*Olea europea* L.) based on RAPD, microsatellite, and SCAR markers. Genome 47:26-35

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

Wunsch A, Hormaza J (2002b) Molecular characterisation of sweet cherry (*Prunus avium* L.) genotypes using peach [*Prunus persica* (L.) Batsch] SSR sequences. Heredity 89:56-63

Wyman A, White R (1980) A highly polymorphic locus in human DNA. Proceedings of the National Academy of Sciences USA 77:6754-6758

Yamamoto T, Kimura T, Sawamura Y, Kotobuki K, Ban Y, Hayashi T, Matsuta N (2001) SSRs isolated from apple can identify polymorphism and genetic diversity in pear. Theoretical and Applied Genetics 102:865-870

Yamamoto T, Kimura T, Sawamura Y, Manabe T, Kotobuki K, Hayashi T, Ban Y, Matsuta N (2002a) Simple sequence repeats for genetic analysis in pear. Euphytica 124:129-137

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

Yamane H, Tao R, Sugiura A, Hauck N, Iezzoni A (2001) Identification and characterization of *S*-RNases in tetraploid sour cherry (*Prunus cerasus*). Journal of the American Society for Horticultural Science 126:661-667

Yoon C-S, Glawe D (1993) Pretreatment with RNase to improve PCR amplification of DNA using 10-mer primers. BioTechniques 14:908-910

Young CT, Schadel WE, Pattee HE, Sanders TH (2004) The microstructure of almond (*Prunus dulcis* (Mill.) D.A. Webb cv. 'Nonpareil') cotyledon. Lebensmittel-Wissenschaft + Technologie 37:317-322

Young N (1992) Restriction fragment length polymorphisms (RFLPs) and crop improvement. Experimental Agriculture 28:385-397

Young N, Zamir D, Ganal M, Tanksley S (1988) Use of isogenic lines and simultaneously probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. Genetics 120:579-585

Young WP, Wheeler PA, Coryell VH, Keim P, Thorgaard GH (1998) A detailed linkage map of rainbow trout produced using doubled haploids. Genetics 148:839-850

Yu Z, Mackill D, Bonman J, Tanksley S (1991) Tagging genes for blast resistance in rice via linkage to RFLP markers. Theoretical and Applied Genetics 81:471-476

Zakarya D, Yahiaoui M, Fkih-Tetouani S (1993) Structure-odour relations for bitter almond odorants. Journal of Physical Organic Chemistry 6:627-633

Zhao H, Speed T (1996) On genetic map functions. Genetics 142:1369-1377

Zhebentyayeva T, Reighard G, Gorina V, Abbott A (2003) Simple sequence repeat (SSR) analysis for assessment of genetic variability in apricot germplasm. Theoretical and Applied Genetics 106:435-444

Zietkiewicz E, Rafalski J, Labuda D (1994) Genomic fingerprinting by simple sequence repeats (SSR)-anchored polymerase chain reaction amplification. Genomics 20:176-183

Zohary M, Spiegel-Roy P (1975) Beginnings of fruit growing in the old world. Science 187:319-327

APPENDIX A

RAPD Primer Sequences

Appendix A Sequence of RAPD Primers (Operon Technologies, Qiagen, Australia) used in this study

	Primer Sequence 5' → 3'		
Primer Number	**OPA**	**OPB**	**OPC**
01	CAG GCC CTT C	GTT TCG CTC C	TTC GAG CCA G
02	TGC CGA GCT G	TGA TCC CTG G	GTG AGG CGT C
03	AGT CAG CCA C	CAT CCC CCT G	GGG GGT CTT T
04	AAT CGC GCT G	GGA CTG GAG T	CCG CAT CTA C
05	AGG GGT CTT G	TGC GCC CTT C	GAT GAC CGC C
06	GGT CCC TGA C	TGC TCT GCC C	GAA CGG ACT C
07	GAA ACG GGT G	GGT GAC GCA G	GTC CCG ACG A
08	GTG ACG TAG G	GTC CAC ACG G	TGG ACC GGT G
09	GGG TAA CGC C	TGG GGG ACT C	CTC ACC GTC C
10	GTG ATC GCA G	CTG CTG GGA C	TGT CTG GGT G
11	CAA TCG CCG T	GTA GAC CCG T	AAA GCT GCG G
12	TCG GCG ATA G	CCT TGA CGC A	TGT CAT CCC C
13	CAG CAC CCA C	TTC CCC CGC T	AAG CCT CGT C
14	TCT GTG CTG G	TCC GCT CTG G	TGC GTG CTT G
15	TTC CGA ACC C	GGA GGG TGT T	GAC GGA TCA G
16	AGC CAG GCA A	TTT GCC CGG A	CAC ACT CCA G
17	GAC CGC TTG T	AGG GAA CGA G	TTC CCC CCA G
18	AGG TGA CCG T	CCA CAG CAG T	TGA GTG GGT G
19	CAA ACG TCG G	ACC CCC GAA G	GTT GCC AGC C
20	GTT GCG ATC C	GGA CCC TTA C	ACT TCG CCA C

APPENDIX B

Weather Station Data 2000-2003

Appendix B Weather data for Renmark weather station number 2484, average mean air temperature (A) and average rainfall data (B), for 2000-2003. Collected by Bureau of Meterology, Australia

A. Temperature Data														
Year	Measurement and Unit	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2000	Mean Maximum Air Temperature (°C)	32.4	34.9	28.6	24.5	18.6	16.4	16.6	18.0	22.3	24.1	31.0	32.7	25.0
2001		37.3	34.7	28.4	24.8	20.2	17.5	15.9	18.7	22.8	21.6	26.4	28.5	24.7
2002		31.8	30.6	29.3	26.7	21.3	17.7	17.9	18.7	22.9	25.6	30.2	32.5	25.4
2003		34.8	32.0	27.2	25.1	19.4	17.1	16.7	16.9	21.2	-	-	-	-
2000	Mean Minimum Air Temperature (°C)	15.9	19.1	14.3	10.0	5.5	4.2	4.3	3.5	7.9	8.8	15.0	14.8	10.3
2001		19.2	18.1	11.7	7.9	6.4	4.9	3.8	5.2	8.1	7.8	10.9	12.3	9.7
2002		14.1	13.6	12.1	9.4	7.4	5.0	3.2	3.3	6.4	7.7	12.6	14.7	9.1
2003		16.3	16.9	11.8	9.4	7.5	5.8	3.1	4.5	5.9	-	-	-	-
B. Rainfall Data														
2000	Total Monthly Precipitation (mm)	1.0	77.2	13.4	44.2	16.2	11.2	22.4	21.8	33.6	19.6	42.6	22.4	325.6
2001		3.8	12.8	14.2	1.4	19.8	32.4	20.0	20.6	34.2	46.2	12.8	1.4	219.6
2002		3.8	0.6	1.0	6.4	14.2	24.0	9.0	7.2	8.0	0.6	8.4	12.0	95.2
2003		1.2	21.6	0.0	9.4	25.2	31.4	7.0	47.0	10.6	19.8	12.2	28.0	213.4

PAPERS

ACTA HORTICULTURAE

September 1-3, 2004, Coolum (Australia): Regional Conference: Harnessing the Potential of Horticulture in the Asian Pacific Region (Submitted)

An Integrated Genetic Linkage Map for Almond Based on RAPD, ISSR, SSR and Morphological Markers

[1]D. Gregory, [1*]M. Sedgley, [1]M.G. Wirthensohn, [2]P. Arús, [1]B. Kaiser and [1]G.G. Collins
[1]School of Agriculture and Wine, The University of Adelaide
PMB 1, Glen Osmond, 5064, South Australia, Australia
[2] Institut de Recerca i Tecnologia Agroalimentàries, Departament de Genètica Vegetal, Cabrils, Spain
[*]Author for correspondence

Keywords: *Prunus dulcis*, genome mapping, random amplified polymorphic DNA, inter-simple sequence repeats, simple sequence repeats, pseudo-testcross

Abstract

An integrated genetic linkage map of almond (*Prunus dulcis*) based on RAPD, ISSR, SSR and morphological markers was constructed using a pseudo-testcross mapping format and Joinmap 3.0. A total of 93 individuals from a F_1 full-sib family produced from a cross between 'Nonpareil' and 'Lauranne' was genotyped with 120 molecular markers (60 RAPD, 23 ISSR, 1 SCAR, and 36 SSR) to produce two parental maps. Hybridity of the mapping population was confirmed by DNA fingerprinting and cluster analysis using RAPD and ISSR markers, and SSR inheritance. The two parental maps were aligned using 12 molecular markers (2 RAPD, 4 ISSR, and 5 SSR) that were segregating in both parents (intercross markers). Two distorted markers were observed at the distal region of two linkage groups. A map position was provided for the kernel taste loci, on linkage group B, 49.1 cM from the ISSR marker $(AG)_8$YC-1786.

Six linkage groups were obtained for the integrated map with a marker density of 11.5 cM/marker, covering 161.9 cM (Kosambi), representing a 31% coverage of the T x E *Prunus* reference map. This map provides an intial step for

producing a SSR saturated integrated genetic linkage map of the almond genome, including morphological markers. The use of both peach and cherry SSRs illustrates the use of interspecific derived primers as a source of polymorphism generation in the *Prunus*, further enhancing the collaborative effort to produce a genus wide reference map for application in breeding programs and marker assisted selection.

INTRODUCTION

Almond, *Prunus dulcis* (Miller) belongs to the family Rosaceae and is an important horticultural crop in Europe, North America, and Australia. The genus *Prunus* also contains a number of economically valuable fruit tree species including apricot (*P. armeniaca*), cherry (*P. avium* and *P. cerasus*), plum (*P. domestica* and *P. japonica*) and the closely related peach (*P. persica*). Selection of improved cultivars adapted to local climatic conditions and disease and pest resistance has been a common goal for all stone fruit breeders and several thousand new cultivars have been developed. In species with a long generation time, large investments in time, space and money for breeding programs is required. The use of molecular techniques has the capacity to greatly improve the breeding efficiency of long-lived woody perennials such as almond (Weeden, 1994).

The development of genetic linkage maps offers a method for early selection strategies in many tree crops (Staub et al. 1996) by the application of marker assisted selection (MAS). In order to use genetic mapping for MAS a map saturated with markers within a short distance of the agronomic loci of interest must be produced. By combining the use of MAS with genetic linkage maps, the long lag time between developing the initial cross and evaluating the progeny in tree crop species can be significantly reduced. Where F_2, backcross, recombinant inbred lines, or double haploid populations are difficult to produce because of self-incompatibility and/or a long juvenile phase, as is the case for many perennial crops including almond, a double pseudo-testcross (DPSTC) situation is used to produce the mapping population (Grattapaglia and Sederoff 1994). The DPSTC strategy has been successfully applied for genetic mapping in tree-crop species such as *Malus x domestica* (Conner et al. 1997), *Castanea sativa* (Casasoli et al. 2001), *Eucalyptus sp.* (Grattapaglia and Sederoff 1994; Verhaegen and Plomion 1996), *Larix* (Arcade et al. 2000), *Quercus robur* (Barreneche et al. 1998), *Olea europaea* (la Rosa et al. 2003; Wu et al. 2004), *Rosa* (Debener and Mattiesch 1999), and *Camellia sinensis* (Hackett et al. 2000).

Genome mapping in fruit trees has become well advanced in the past decade, with extensive research in the *Prunus* genus being carried out by the European *Prunus* mapping projects (Baird et al. 1996; Joobeur et al. 1998). Viruel et al. (1995) published the first genetic linkage map for almond, using an F_1 cross between the cultivars 'Ferragnès' x 'Tuono'. Joobeur et al. (2000) published a second generation map for the same population, based on random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) markers, but to date no morphological traits have been placed on this map. Two smaller targeted mapping projects have placed self-incompatibility on this framework almond map (Ballester et al. 1998), and late bloom on another almond population map derived from the cross of 'Felisia' x 'Bertina', (Ballester et al. 2001). Genetic maps have also been constructed for interspecific crosses of *Prunus*, and the map for almond 'Texas' x peach 'Earlygold' (T x E) is assumed to be the most saturated map for *Prunus* (Aranzana et al. 2003; Foulongne et al. 2003; Lambert et al. 2003). The localisation of morphological traits on *Prunus* maps is largely restricted to peach maps (Chapparo et al. 1994; Dirlewanger et al. 1996; Dirlewanger et al. 1998; Warburton et

al. 1996). In order to further utilize these maps for practical applications more morphological markers are required to be incorporated in other species maps.

The objective of this study was to produce an integrated genetic linkage map for a F_1 hybrid population of the cultivars 'Nonpareil' x 'Lauranne', constructed using RAPD, inter-simple sequence repeat (ISSR) SSR and morphological markers. This research is part of the almond breeding program at the University of Adelaide, aimed at the genetic improvement of this nut crop and the application of MAS for improved breeding efficiency.

MATERIALS AND METHODS

Plant Material and Morphological Characters

A progeny set of 93 individuals from a cross between the American self-incompatible cultivar 'Nonpareil' (female parent) and the French self-compatible cultivar 'Lauranne' (male parent) in a pseudo-testcross configuration was selected for linkage map construction. This population is maintained in the field under standard orchard conditions for phenotypic evaluation.

Kernel taste was organoleptically assessed for each individual over three growing seasons (2001-2003). For each hybrid individual two kernels from each year were tasted and scored as either sweet or semi-bitter. Final scores for inclusion into the mapping matrix was based on a consistent score over the three harvest years.

DNA Extraction

Total DNA was extracted from frozen (-80°C) leaf tissue according to the method of Lamboy and Alpha (1998), and adapted to almond as follows. Between 1.5 – 2 g of young leaves with midribs removed were ground in liquid nitrogen to a fine powder in a prechilled mortar and pestle and extracted in 10 mL of buffer [250 mM Tris-HCl (pH 8.0), 250 mM NaCl, 50 mM EDTA, 4.0% (w/v), PVP-40T, 1.5% SDS] with 50 mM diethyldithiocarbamate (DETC), 100 mM sodium ascorbate and 1.0% (v/v) β-mercaptoethanol added immediately preceding addition of ground sample. The mixture was incubated at 37°C for 30 min and inverted every 10 min, followed by extraction with chloroform : isoamyl alcohol [24 : 1 (v/v)], and centrifuged at 11 000 rpm for 15 min at 4°C. The upper aqueous phase was mixed with a 0.7 volume of 5 M NaCl, and centrifuged at 11 000 rpm for 30 min at 4°C. Cold isopropanol (-20°C) was used to precipitate the nucleic acids by adding a 0.66 volume to the supernatant, and placing at –20°C for 30 min, followed by centrifuging at 11 000 rpm for 15 min at 4°C. The DNA pellet was twice washed, firstly with 10 mL of cold wash buffer (76% ethanol, 10 mM NH_4Ac) overnight, followed by a 5 min wash with 70% ethanol, and the pellet air dried and dissolved in 600 µL of TE buffer (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). 3 µL of DNAse-free RNase A (10 mg/mL) was added to each sample and incubated at 37°C for 30 min. Remaining impurities were removed by the addition of 240 µL of 7.5 M NH_4Ac, placed at 4°C for 30 min followed by centrifugation at 11 000 rpm for 15 min at 4°C. The DNA was precipitated from the supernatant using 2 volumes of cold absolute ethanol and standing for 30 min at –20°C. DNA was recovered by centrifugation at 11 500 rpm for 15 min at 4°C, and the pellet washed for 15 min with cold wash buffer, decanted and washed with cold 70% ethanol. The pellet was air dried and dissolved in 400 µL of TE buffer. The purified DNA was checked for purity and quantity by spectrophotometry, analysing the A_{260}/A_{230} and A_{260}/A_{280} ratios, and A_{260} respectively (Sambrook et al. 1989).

RAPD and ISSR Markers

A three stage screen of primers was carried out, initially sixty RAPD primers (OPA, OPB and OPC kits, Operon Technologies, USA) and 23 ISSR primers [$(CT)_8A$, $(CT)_8G$, $(GA)_8T$, $(CA)_8G$, $(CA)_8T$, $(AG)_8G$, $(AC)_8G$, $(TG)_8G$, $(AG)_8YC$, $(AG)_8YT$, $(AG)_8YA$, $(GA)_8YC$, $(CT)_8RC$, $(CA)_8RG$, $GCA(CA)_9$, $CAG(CA)_9$, $GTC(CA)_9$, $GAC(CA)_9$, $CTG(CA)_9$, $CGA(CA)_9$, $(ACC)_6$, $(TAA)_8$, and $BDB(ACA)_5$, Geneworks, Australia] were screened on both parents. From this intial screen, a second stage screen using 11 RAPD and 13 ISSR primers was carried out on both parents and 10 F_1 progeny. A total of six RAPD and 13 ISSR primers were then used for polymorphism generation on the whole 95 F_1 progeny and parents.

The PCR protocol for RAPD analysis was performed in a 20 µL volume containing 40 ng genomic DNA, 3 mM $MgCl_2$, 0.25 µM primer, 200 µM of each dNTP, 1 x cresol red dye [2% (w/v) sucrose, 0.1 mM cresol red sodium salt, Aldrich Chemicals, USA], 1.1 U Taq polymerase and 1 x PCR buffer (Invitrogen, Australia). The amplification reaction consisted of an initial denaturation of 2 min at 95°C, followed by 45 cycles of 30 sec at 95°C, 45 sec at 48°C, and 1 min at 72°C, and a final extension of 10 min at 72°C.

ISSR amplifications were performed in a 20 µL volume containing 50 ng genomic DNA, 1.5 mM $MgCl_2$, 0.5 µM primer, 200 µM of each dNTP, 1 x cresol red dye, 1.0 U Taq polymerase and 1 x PCR buffer (Invitrogen, Australia). For primer $(AG)_8G$ 10 mM NH_4SO_4 was also included in the PCR mix to aid in amplification. Two PCR amplification cycles were used, for primers $(CT)_8G$, $(GA)_8T$, $(CA)_8G$, $(CA)_8T$, $(AC)_8G$, $(AG)_8YC$, $(CT)_8RC$, $(CA)_8RG$, the temperature profile was the same as given for the RAPD primers. For primers $(AG)_8G$, $(AG)_8YT$, $(GA)_8YC$, $GCA(CA)_9$, and $(TAA)_8$, the amplification reaction consisted of an initial denaturation of 7 min at 94°C, followed by 41 cycles of 1 min at 93°C, 1 min at 55°C [60°C for $GCA(CA)_9$, 42°C for $(TAA)_8$], and 2 min at 72°C, and a final extension of 8 min at 72°C.

Amplifications were performed in a PCT-100 96 well hot bonnet thermocycler (MJ Research Inc, USA) for both marker types. PCR products were visualised on 1.5% agarose gel for RAPD, and 1.75% for ISSR (LE, Promega, Australia) in 1 x TBE buffer (89 mM Tris HCl, pH 8.0, 89 mM boric acid, 5 mM EDTA).

SSR Markers

From the 98 SSR primers screened only on the parents, 11 were selected for polymorphism generation in the progeny [CPACT119, CPACT120, CPACT165, CPACT166, CPACT175, CPACT319, CPACT391 (P. Arús, unpublished), CPPCT005 (Aranzana et al. 2000), UDP96001, UDP96005 (Cipriani et al. 1999), PS12e2 (Downey and Iezzoni 2000)]. Primers with the prefix CPACT were amplified according to the protocol of Aranzana et al (2003) with the following modifications. For each PCR, a total of 10 ng genomic DNA was amplified with 1 mM $MgCl_2$, 0.4 µM each of forward and reverse primer, 0.4 µM of dATP, dTTP, dGTP, 1 x Taq buffer (75 mM Tris-HCl, 20 mM $(NH_4)_2SO_4$, 0.01% Tween-20), 1.0 U Taq polymerase and 0.16 µM α-^{33}P dCTP (Amersham, USA) in a total volume of 15 µL. For the remaining four primers, the forward primers were fluorescently labelled with HEX or FAM (Geneworks, Australia) or NED (Applied Biosystems, USA) and amplified with 40 ng of DNA, 2 mM $MgCl_2$, 0.2 mM each of forward and reverse primer, 0.36 mM of each dNTP, 1.0 U Taq polymerase and 1 x PCR buffer (Invitrogen, Australia) in a 20 µL reaction volume. PCR amplification was according to Joobeur et al. (2000).

PCR products were visualised by polyacrylamide gel electrophoresis with autoradiogram detection for CPACT primers according to Aranzana et al. (2003), PCR

products using fluorescent primers were scored using a 377 ABI Prism DNA sequencer and Genotyper software (Applied Biosystems, USA).

Confirmation of F_1 Hybridity

In order to verify the hybridity of the F_1 mapping population, the genetic similarities between individuals including the parents, was estimated using NTSYS-pc version 2 (Applied Biostatistics Inc, Setauket NY, USA) (Rohlf, 1998). Data from polymorphic bands scored as present or absent from six RAPD (OPA-02, OPA-08, OPB-10, OPC-07, OPC-10, OPC-11) and ten ISSR primers [$(CT)_8G$, $(GA)_8T$, $(CA)_8G$, $(CA)_8T$, $(AG)_8G$, $(AC)_8G$, $(AG)_8YT$, $(AG)_8YA$, $GCA(CA)_9$, and $(TAA)_8$] were used. A dendrogram of these data was produced using the simple matching coefficient, SAHN and UPGMA algorithms.

Linkage Analysis

The parents and all progeny were scored for presence or absence of bright, reproducible, segregating bands (or peaks for fluorescent detection). All fragment sizes were calculated using GelPro Analyser (version 3.1, Media Cybernetics, USA), segregating bands were identified by the primer code followed by the fragment size, for example marker OPA02-456 is a fragment of 456 bp amplified by the RAPD primer OPA-02.

Two data sets were constructed, one for each parent based on the parental origin of the band in accordance to the pseudo-testcross mapping strategy. Chi-squared tests were performed to check for segregation distortion of the markers. The two parental maps were produced using the software Joinmap 3.0 (Van Ooijen and Voorrips 2001), using the cross-pollination format, at a LOD of 3.0 - 4.0. An integrated map was produced from the corresponding anchor intercross markers common to both parents using the 'combine groups for mapping' function of the program. Map distances in centiMorgans were calculated by converting recombination frequencies using the Kosambi mapping function (Kosambi, 1944). These data were presented pictorially with the map drawing software MapChart 2.1 (Voorips 2002).

RESULTS AND DISCUSSION

From the 2-D dendrogram (Figure 1), three clusters are evident at the 70% similarity level, cluster A shows those most similar to the male parent 'Lauranne', cluster B those most similar to the female parent 'Nonpareil', 91% of hybrid individuals were grouped between the parents. Cluster C included individuals that were not aligned between the two parents, however analysis of the SSR profile of these individuals did not show any aberrant band transmission. Based on both these data the genotypes of all individuals indicate they are true hybrids of a cross between 'Nonpareil' and 'Lauranne'. Confirmation of correct parentage of individuals in mapping populations is important to eliminate statistical bias and skewing of recombination values in the data and analysis.

Of the 97 markers used for parental map construction, 24 were heterozygous in both parents, and 18 common loci were identified to be used as loci bridges between the maps. Of the 18 common markers, 14 were used to produce an integrated parental map, representing approximately 19% of all markers studied. The integrated map consisted of 6 linkage groups, with between 2 and 4 markers, varying in length from 6.0 to 49.1 cM (Figure 2), and with a marker density of 11.5 cM/marker.

The discrepancy between the number of linkage groups expected to coalesce and the number observed for the female parent and the integrated map is most likely due to

the high number of unincorporated markers and the low number of homologous markers common to both parents used to produce the integrated map. The inclusion of more markers of the intercross co-dominant type would alleviate this situation and potentially increase the number of linkage groups. The incorporation of more markers would also be expected to merge single markers that are so far unlinked and may presently be in regions of very low marker frequency to the major linkage groups. No change in marker order was observed between the parents and the integrated maps although a change in the distances between markers was observed. This was expected, since although the values were relatively close, in the integrated map the combined recombination frequency estimate is the average over the male and female meioses. The integrated map is therefore a statistical compromise between the maternal and paternal recombination values rather than a biological reality. However, integrated maps are the ideal situation for mapping programs, since they permit the identification of conserved loci in related species and can be used as reference points to transfer linkage information (Lambert et al. 2004).

One morphological trait of the seven studied on the whole population was localised at the distal region of the integrated linkage group B at a distance of 49.1 cM from the ISR marker $(AG)_8$YC-1786 (Figure 2). A marker for the bitter versus sweet kernel has been placed on the G5 linkage group of the T x E map at a distance of 0 cM from the RFLP marker FG202 (P. Arús unpublished, Joobeur et al. 1998). Further saturation of markers on this linkage group and inclusion of T x E anchor loci will establish if this marker for taste is part of a complex localised on this linkage group, and offers a possibility of map based cloning of this gene.

From the SSR markers used in this study, three placed on the integrated parental map can be used as anchor loci to the *Prunus* reference map T x E (Aranzana et al. 2003; P. Arús unpublished). The marker UDP96001 is part of the 24 'genotyping set' proposed by Aranzana et al. (2003) which can be used as reference markers for covering the *Prunus* genome.

Presently the integrated map covers 161.9 cM, which is approximately 48% of the two parental maps and 31% of the T x E map. It is the aim of future research to include more molecular markers, particularly SSRs belonging to the *Prunus* reference set, in order to further saturate the integrated parental map for this cross. This will coalesce the remaining two linkage groups which are outstanding at this time, corresponding to the number of linkage groups to the haploid chromosome number.

Thus with further work this map will become a valuable tool for use in breeding programs in Australia and strengthen research on genetic mapping of stonefruit worldwide.

ACKNOWLEDGEMENTS

Literature Cited

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

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

Arcade, A., Anselin, F., Fanre Rampant, P., Lesage, M., Paques, L. and Prat, D. 2000. Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. Theoretical and Applied Genetics, 100: 299-307.

Baird, W., Ballard, R., Rajapakse, S. and Abbott, A. 1996. Progress in *Prunus* mapping and application to molecular markers to germplasm improvement. HortScience, 31: 1099-1106.

Ballester, J., Bŏsković, R., Batlle, I., Arús, P., Vargus, F. and de Vincente, M. 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. and de Vincente, M. 2001. Genetic mapping of a major gene delaying bloom time in almond. Plant Breeding, 120: 268-270.

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.

Chapparo, J., Werner, D., O'Malley, D. and Sederoff, R. 1994. Targeted mapping and linkage analysis of morphological isozyme, and RAPD markers in peach. Theoretical and Applied Genetics, 87: 805-815.

Cipriani, G., Lot, G., Huang, W.-G., Marrazzo, M., 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.

Conner, P., Brown, S. and Weeden, N. 1997. Randomly amplified polymorphic DNA based genetic linkage maps of three apple cultivars. 122, 3: 350-359.

Debener, T. and Mattiesch, L. 1999. Construction of a genetic linkage map for roses using RAPD and AFLP markers. Theoretical and Applied Genetics, 99: 891-899.

Dirlewanger, E., Pascal, T., Zuger, C. and Kervella, J. 1996. Analysis of molecular markers associated with powdery mildew resistance genes in peach (*Prunus persica* (L.) Batsch) x *Prunus davidiana* hybrids. Theoretical and Applied Genetics, 93: 909-919.

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.

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

Foulongne, M., Pascal, T., Arús, P. and Kervella, J. 2003. The potential of Prunus davidiana for introgression into peach [Prunus persica (L.) Batsch] assessed by comparitive mapping. Theoretical and Applied Genetics, 107: 227-238.

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

Joobeur, T., Periam, N., de Vincente, M., King, G. and Arús, P. 2000. Development of a second generation linkage map for almond using RAPD and SSR markers. Genome, 43: 649-655.

Joobeur, T., Viruel, M., deVincente, M., Jauregui, B., Ballester, J., Dettori, M., Verde, I., Truco, M., 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.

Kosambi, D. 1944. The estimation of map distance from recombination values. Annals of Eugenetics, 12: 172-175.

la Rosa, R., Angiolillo, A., Guerrero, C., Pellegrini, M., Rallo, L., Besnard, G., Berville, A., Martin, A. and Baldoni, L. 2003. A first genetic linkage map of olive (Olea europaea L.) cultivars using RAPD, AFLP, RFLP and SSR markers. Theoretical and Applied Genetics, 106: 1273-1282.

Lambert, P., Hagen, L., Arús, P. and Audergon, J. 2003. Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) compared with the almond Texas x peach Earlygold reference map for *Prunus*. Theoretical and Applied Genetics, 108: 1120-1130.

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

Rohlf, F. 1998. NTSYS-pc: (Numerical Taxonomy and Multivariate Analysis System). Setauket, N.Y.: Exeter Software.

Sambrook, J., Fritsch, E. F. and Maniatis, T. 1989. Molecular Cloning : A Laboratory Manual. USA: Cold Spring Harbor Laboratory Press.

Staub, J., Serquen, F. and Gupta, M. 1996. Genetic markers, map construction, and their application in plant breeding. HortScience, 31: 729-741.

Van Ooijen, J. and Voorrips, R. 2001. Joinmap ® 3.0, Software for the calculation of genetic linkage maps. the Netherlands: Plant Research International, Wageningen.

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

Viruel, M., Messeguer, R., deVicente, M., Garcia-Mas, J., Puigdomenech, 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. 2002. MapChart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity, 93: 77-78.

Warburton, M., Becerra-Velasquez, V., Goffreda, J. and Bliss, F. 1996. Utility of RAPD markers in identifying genetic linkages to genes of economic importance in peach. Theoretical and Applied Genetics, 93: 920-925.

Weeden, N., Hemmat, M., Lawson, D., Lodhi, M., Bell, R., Manganaris, A., Reisch, B., Brown, S. and Ye, G.-N. 1994. Development and application of molecular marker linkage maps in woody fruit crops. Euphytica, 77: 71-75.

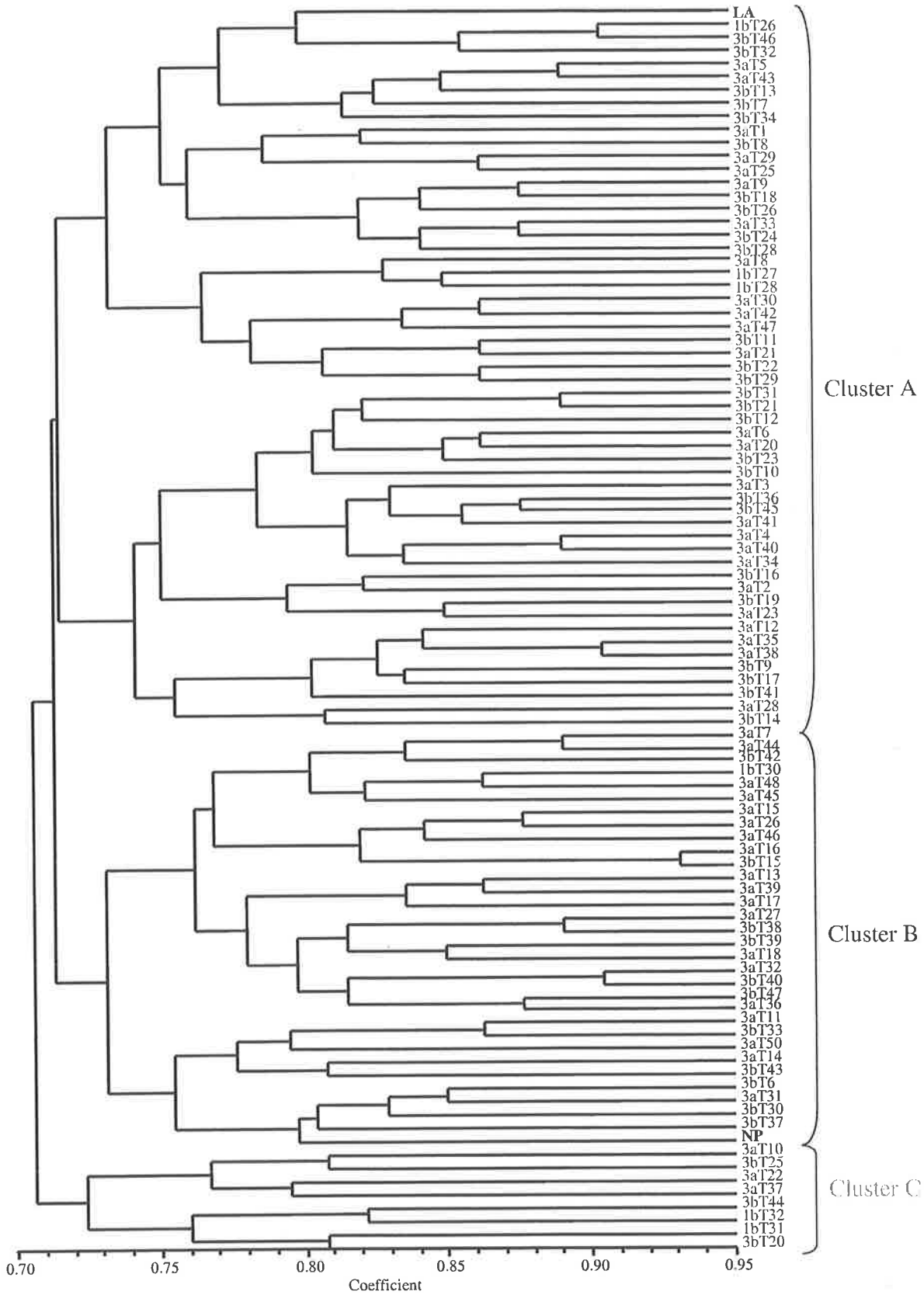

Figure 1 Dendrogram of 93 F_1 hybrid progeny and the two parents 'Nonpareil' (NP) and 'Lauranne' (LA), generated by UPMGA cluster analysis using the simple matching coefficient. Cluster A includes hybrids most similar to the male parent LA, cluster B those with similarity to the female parent NP, and cluster C those hybrids grouped as outliers.

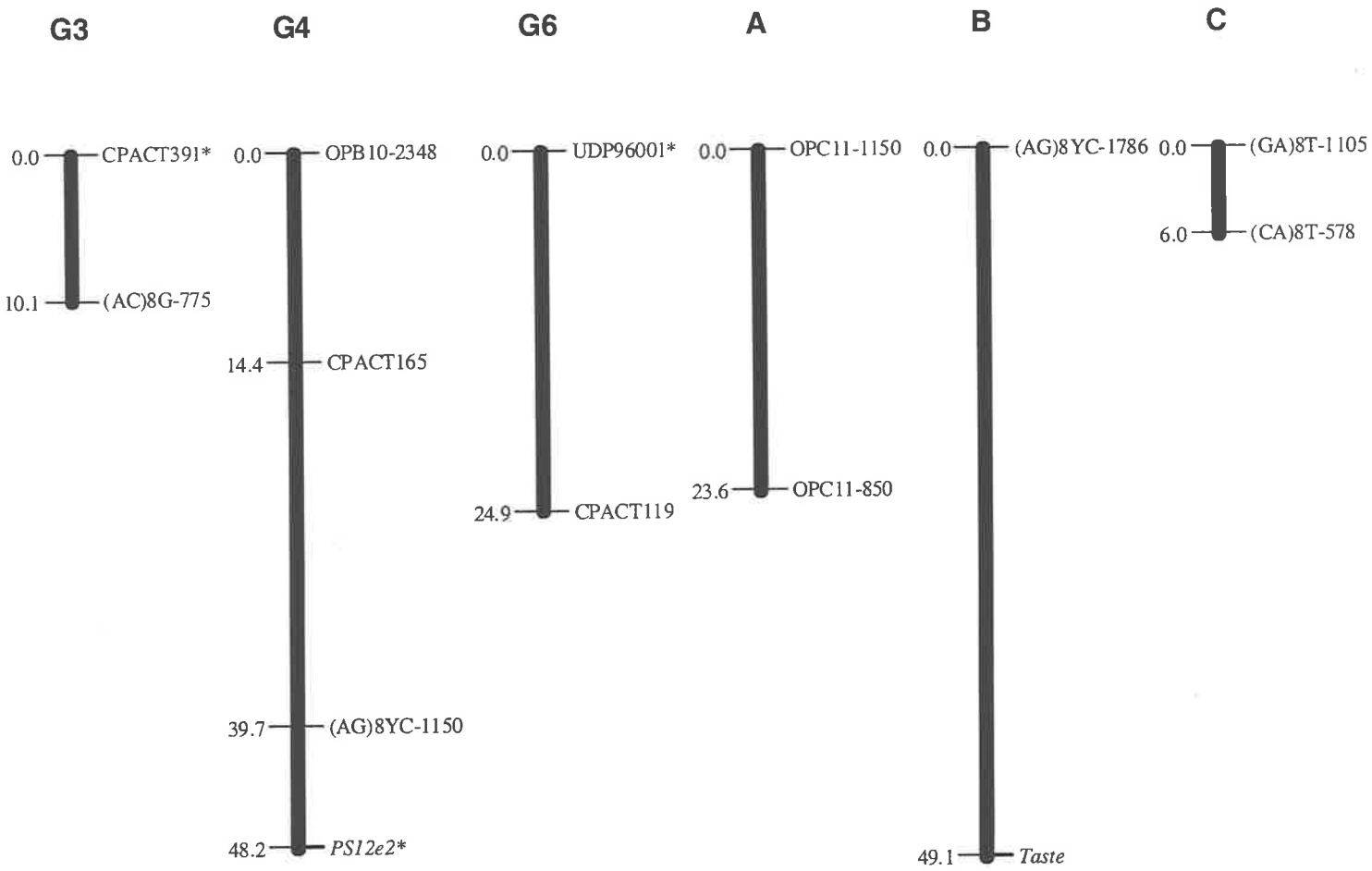

Figure 2 Integrated genetic linkage map of the diploid almond cross 'Nonpareil' x 'Lauranne'. Linkage groups G3, G4 and G6 named according to the convention of the *Prunus* reference map T x E (Aranzana et al. 2003). Markers homologous to this map are marked with a *. Skewed markers are in italics. Distances between markers are given in centiMorgans.