FUNCTIONAL GENOMIC CHARACTERIZATION OF FRUIT QUALITY TRAITS IN APPLE (*Malus x domestica* Borkh.)

Claudius Marondedze

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor Philosophiae at the Faculty of Science, University of the Western Cape.

Supervisor: Prof. D. Jasper G. Rees

Co-supervisor: Ass. Prof. Bongani K. Ndimba

September 2009

ABSTRACT

Functional genomic characterization of fruit quality traits in apple (*Malus x domestica* Borkh.)

Claudius Marondedze

Department of Biotechnology, Faculty of Science, University of the Western Cape, South Africa.

The domesticated apple (*Malus* x *domestica* Borkh.), belonging to the *Malus* genus of the Rosaceae family, is one of the edible pomaceous fruits. Since it is one of the important commercial fruit crops worldwide, the quality of the fruit is crucial to breeders and farmers as it ultimately determines acceptance of a cultivar for consumption. Fruit quality is also a critical determinant factor that is used to estimate the potential of apples to have a long shelf life.

The introduction of marker-assisted selection (MAS) has allowed hastening of traditional breeding and selection of high-quality apple cultivars. The availability of genetic linkage maps, constructed by positioning molecular markers throughout the apple genome, enables the detection and analysis of major genes and quantitative trait loci (QTLs) contributing to the quality traits of a given genotype. Therefore, the primary aim of this study was to construct a genetic linkage map of the 'Golden Delicious' x 'Dietrich' population for the identification of QTLs associated with fruit quality traits and then to examine the apple fruit pulp proteome with a specific focus on fruit firmness.

In this regard, genomic DNA was extracted from leaves of the 'Golden Delicious' x 'Dietrich' population and used in megaplex PCR reactions. The PCR products were analysed prior to scoring of alleles. Polymorphic markers were then used to construct genetic linkage maps. The genetic linkage maps constructed in this study comprise of 167 simple sequence repeats (SSR) markers, 33 of these were newly developed markers. The 17 linkage groups of apple were constructed and aligned to existing apple genetic maps. The maps span 1,437.8 cM and 1,491.5 cM for 'Golden Delicious' and

'Dietrich', respectively.

The constructed framework genetic linkage maps and phenotypic data collected between 2005 and 2007 enabled the detection and analysis of QTLs contributing to nine fruit quality traits, namey firmness, juiciness, crispness, colour, stripness, form, acidity, size and russeting. Using the maximum likelihood based interval mapping, multiple QTL-mapping and Krustal-Wallis analysis several QTLs were identified for each of the fruit quality traits under study. The identified QTLs coding for these nine traits were then compared with the previously mapped QTLs in apples. The comparison revealed new QTLs associated with stripness, russeting and form.

With respect to fruit firmness trait, proteomic approach was applied to further confirm its polygenic nature. Proteomics can be used to map translated genes and loci controlling their expression, leading to the identification of proteins accounting for the variation of complex phenotypic traits. This part of the study was aimed first at identifying and characterising proteins from apple fruit pulp, and then investigating differential protein expression influencing fruit firmness. Total soluble proteins (TSP) were extracted from the fruit pulp and resolved using two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). For the total fruit proteome, an average of 290 protein spots were resolved on 2D-PAGE. A total of 135 spots were positively identified by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (MS), corresponding to 111 non-redundant proteins, and were classified into 11 functional categories. Proteins identified were validated against the *Malus* EST database from the NCBI. Two major functional categories consisted of proteins involved in energy metabolism and defense or detoxification and this accounted for 31% of the identified proteins. A large proportion (23%) of the identified proteins remained unclassified.

In addition, the proteome maps from the high and low firmness apple fruit pulp were comparatively analysed to identify differentially expressed proteins between the two phenotypes. The analysis detected 54 differentially expressed proteins between the two phenotypes, suggesting that important metabolic changes influence apple fruit characteristics. Five proteins were detected as expressed only in the high firmness phenotype. In addition, some proteins involved in defense/detoxification, cell

growth/division, signal transduction and cytoskeleton were identified as up-regulated in the high firmness phenotype, while some proteins involved in ethylene biosynthesis, lipid biosynthesis and photosynthesis were down-regulated. Identification of these differentially expressed proteins shed light on the events influencing fruit pulp firmness in apples.

Overall, the study allowed the identification of putative candidate markers linked to fruit quality traits in the 'Golden Delicious' x 'Dietrich' population, which could be used by apple breeders for MAS and for the isolation and characterisation of genes associated with these traits using map-based cloning. In addition to the confirmation of the polygenic nature of fruit firmness, the proteomic data provided knowledge on the proteins regulating apple fruit pulp firmness. By using both the genomics and proteomics approaches, this study related genomic data to cell metabolism and plant phenotype, which is necessary for studying plant physiological mechanisms.

September 2009.

KEY WORDS

Functional genomic characterization of fruit quality traits in apple (*Malus x domestica* Borkh.)

Claudius Marondedze

Apple

Differential expression

Fruit quality

Fruit pulp

Genetic mapping

Linkage groups

Malus x domestica

Microsatellite

Proteome map

Quantitative trait loci

To my parents (Emilio and Simbisai), wife (Ludivine), brothers (Edson, Innocent and Andrew) and sister (Priscah)

"Knowledge is Power"

DECLARATION

I declare that 'Functional genomic characterization of fruit quality traits in apple (*Malus x domestica* Borkh.)' is my own work that has not been submitted for any degree or examination in any other University and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Claudius Marondedze September 2009

Signed:……

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my supervisor Prof. D. Jasper G. Rees and co-supervisor Ass. Prof. Bongani K. Ndimba for their guidance and support throughout this study. Special thanks goes to Prof. P. Gouws, Dr. L. Holtman, Dr. L.A. Thomas, Dr. I. Labuschagné, Dr. J.-M. Celton, Mr T. Koopman and Mr M.K. Soeker for their valuable input in this study. I also wish to thank Dr. W.T. Sanyika, Dr. J. Mafofo, Dr. M. Chibi, Dr. R. Ngara, Mr K. Ngara, Mr J. Baison, Mr P. Hove, Mrs R. Ndimba, Mrs M. Madduri, Mrs X.S.F. Sanyika, Miss R. M. Mudzonga and Ms H.V.H. Mathye for their support throughout the difficult times during my studies.

Special mention goes to the members of the Biochemistry and the Proteomics research laboratories (Department of Biotechnology, University of the Western Cape) for their assistance when required. For the support provided by my family during this study, I want to convey my appreciation and gratitude to all of them. I also wish to thank THRIP (DTI) and DFPT for their financial support. Above all, I thank God for all his blessings and support granted to me throughout my studies.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

χ^2	Chi-square
∩	Intersection
%	Percentage
1D-PAGE	One-dimensional polyacrylamide gel electrophoresis
2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
A	Adenine
AAT	Aspartate transaminase
ABA	Abscisic acid
ABC	ATP-binding cassette
ABI	Applied Biosystems
ac	Acidity
ACC	1-aminocyclopropane-1-carboxylase
ADH	Alcohol dehydrogenase
ADP	Adenosine diphosphate
AFLPs	Amplified fragment length polymorphisms
ALPs	Amplicon length polymorphisms
ANOVA	Analysis of variance
APS	Ammonium persulphate
ASB	Alkylamidosulfobetaine
ATP	Adenosine triphosphate
BLASTN	Basic Local Alignment Search Tool Nucleotide
BLASTP	Basic Local Alignment Search Tool Protein
bp	Base pair
BSA	Bovine serum albumin

C	Cytosine
°C	Degree Celsius
Ca^{2+}	Calcium ion
CAPS	Cleavable amplified polymorphic sequences
CBB	Coomassie brilliant blue R-250
CC	Coiled-coil
cDNA	Complementary deoxyribonucleic acid
CHAPS	3-[(3-cholamidopropyl)dimethyl-ammonio]-1-propanesulphonate
CHCA	α-cyna-hydroxy-cinnamic
CI	Confidence interval
cm	Centimetre
cM	CentiMorgan
CO_2	Carbon dioxide
co	Colour
CP	Cross pollination
cr	Crispness
CTAB	Cetyl trimethyl ammonium bromide
cv.	Cultivar
'D'	'Dietrich'
Da	Dalton
DAPI	4'-6-diamidino-2-phenylindole
DEPC	Diethylpyrocarbonate
DFPT	Deciduous Fruit Producers Trust
dH_2O	Distilled water
DHAR	Dehydroascorbate reductase

DHCIb	Dynein heavy chain isoform
DIGE	Difference gel electrophoresis
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EDTA	Ethylenediamine tetraacetic acid
ER	Estrogen receptor
ESI	Electrospray ionisation
EST-SSR	Expressed sequence tag-simple sequence repeats
FAO	Food and Agriculture Organization
fi	Firmness
g	*g*-force
g	Gram
G	Guanine
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
'GD'	'Golden Delicious'
gDNA	Genomic deoxyribonucleic acid
GDP	Gross domestic product
GFP	Green fluorescent protein
GGT	Graphical genotyper®
G-6-P	Glucose-6-phosphate
GMP	Guanosine monophosphate
GTP	Guanosine triphosphate
GW	Genome wide

h	Hours
H^+	Hydrogen ion
HCl	Hydrochloric acid
HDMF	4-hydroxy-2,5-dimethyl-3(2H)-furanone
HFP	High firmness phenotype
HiDRAS	High-Quality Disease Resistant Sustainable Agriculture
HMMF	4-hydroxy-5-methyl-2-methylene-3(2H)-furanone
HPLC	High performance liquid chromatography
Hsp	Heat shock proteins
IDH	Isocitrate dehydrogenase
IEF	Isoelectric focusing
IM	Interval mapping
IMP	Inosine monophosphate
IPG	Immobilized pH gradient
ISA	Inter-simple sequence repeat amplification
ju	Juiciness
kDa	Kilo Daltons
kVh	Kilo volt-hours
KW	Krustal-Wallis
LC	Liquid chromatography
LDL	Low density lipoprotein
LFP	Low firmness phenotype
LG	Linkage group
LOD	Likelihood of the odds
LOX	Lipoxygenases

LRR	Leucine-rich repeat
LRS	Likelihood ratio statistic
M	Moles l^{-1}
m/z	Mass to charge ratio
MAS	Marker assisted selection
MAB	Marker assisted breeding
Mal	Major allergen
MALDI-TOF	Matrix Assisted Laser Desorption/Ionisation Time-of-Flight
MASCOT	Mass spectroscopy and proteomics
MDH	Malate dehydrogenase
mers	Nucleotides
min	Minutes
ml	Millilitre
mM	Millimolar
MOWSE	Molecular weight search
MQM	Multiple-QTL models
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MSDB	Mass spectrometry database
MudPIT	Multi-dimensional protein identification
MW	Molecular weight
N	Newton
Na^+	Sodium ion
$NaBH_4$	Sodium borohydride
NaCl	Sodium chloride

NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NaOH	Sodium hydroxide
NBS	Nucleotide binding site
NCBI	National Center for Biotechnology Information
NIRS	Near infrared spectroscopy
nm	Nanometer
p	Probability
PAGE	Polyacrylamide gel electrophoresis
PAL	Phenylalamine ammonia-lyase
PBS	Phosphate buffered saline
PC	Paper chromatography
PCR	Polymerase Chain Reaction
pH	Potential of hydrogen
p*I*	Isoelectric point
PMF	Peptide mass fingerprinting
PR	Pathogenesis related
PVP	Polyvinylpyrrolidone
PVPP	Polyvinylpolypyrrolidone
Q	Quantitative test
QTLs	Quantitative trait loci
r	Recombination frequency
RAPDs	Randomly amplified polymorphic DNAs
RFLPs	Restriction fragment length polymorphisms

RNA	Ribonucleic acid
ROS	Reactive oxygen species
RT-PCR	Reverse transcriptase-Polymerase chain reaction
ru	Russeting
RuBisCO	Ribulose-1,5-bisphosphate carboxylase/oxygenase
S-6-P	Sorbitol-6-phosphate
SAM	S-adenosyl-L-methionine
SCARs	Sequence characterised amplified regions
SDS	Sodium dodecyl sulphate
si	Size
SIMAP	Similarity matrix of proteins
SNP	Single nucleotide polymorphism
SOX	Sorbitol oxidase
SSCP	Single-stranded conformation polymorphism
SSR	Simple sequence repeat
st	Stripness
STS	Sequence-tagged sites
T	Thymine
TAF	TATA box binding protein-associated factors
Taq	*Thermus aquaticus*
TBE	Tris borate EDTA
TBP	Tributyl phosphine
TCA	Trichloroacetic acid
TE	Tris EDTA
TEMED	*N,N,N′,N′*-tetramethylethylenediamine

T-test	Student *t*-test
TIR	Toll/interleukin-1 receptor
TLC	Thin layer chromatography
Tris	Tris(hydroxymethyl)aminomethane
TSP	Total soluble proteins
TSS	Total soluble solids
µg	Microgram
$\mu g.\mu l^{-1}$	Microgram per microlitre
µl	Microlitre
UK	United Kingdom
USA	United States of America
UTP	Uridine-5′-triphosphate
UV	Ultraviolet light
V	Voltage
v/v	Volume per volume
w/v	Weight per volume

ABBREVIATIONS FOR AMINO ACIDS

Amino acid	Three-letter abbreviation	One letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

CHAPTER 1

LITERATURE REVIEW

1.1 APPLES

Apple (*Malus x domestica* Borkh.), from the Rosaceae family, belongs to the genus *Malus*, which bears pomaceous fruits. Other members of the Rosaceae family include pear (*Pyrus communis*), peach (*Prunus persica*) and cherry (*Prunus avium*). Apple is the most ubiquitous fruit in temperate regions and it has been cultivated throughout Europe and Asia since antiquity (Huxley, 1992; Janick *et al.*, 1996). This is one of the most widely cultivated fruit tree and most probably the first one to be cultivated (Janick *et al.*, 1996; Azuma *et al.*, 2001).

Apple fruits are used as a food crop and a source of pectin, which is used to thicken jams and culture media for laboratory uses (Brouk, 1975). In addition, apple fruits can be processed into sauces, slices, sweets, alcoholic beverages, vinegar or juice (Janick *et al*., 1996). They are also an excellent dentifrice, the mechanical action of eating a fruit serving to clean both teeth and gums (Grieve, 1984). Further, the seed oil can be used as an illuminant (Duke, 1983; Grieve, 1984).

In temperate regions, apples are very popular fruits as a result of their availability to consumers. Unlike other fruits, with the exception of citrus fruits such as oranges, apples can be consumed off the tree or stored for months. When appropriately stored, apples can last up to a year, without major loss of nutritive value (Janick *et al*., 1996). It is because of their long shelf life and nutritional value that 63% of total fruit production is consumed locally [South Africa; Deciduous Fruit Producers Trust (DFPT), 2009]. In

addition, about 50% of the apples produced for local consumption are processed for taste/thickener (DFPT, 2009).

In 2006, 93 countries worldwide were producing apples, with China accounting for 40% of world production (Global Crop Diversity Trust). The world apple demand is therefore highly dependent on the apple production from China. With a production of approximately 650,000 tons per annum, mainly from the Western Cape province, South Africa is the largest producer of apples in Africa, and ranked 17th worldwide (FAO, 2008). The apple industry plays a vital role for South Africa, representing approximately 28% of the total deciduous cultivated fruit area after grapes (30%). In addition, 37% of apples produced are channelled towards the export market representing 38% of deciduous fruit exportation. Apples are mostly exported to Europe, North America and Asia. This accounts for as much as 80% of the apple industry's income and contributes significantly to the Western Cape gross domestic product (DFPT, 2009).

1.1.1 Botanical origin of apples

Although the exact origin of the domesticated apple remains unknown, its wild ancestor is believed to be *M. sieversii*, a wild plant that originated from Alma-Ata (literally meaning "father of apples"), a major city in China (Sauer, 1993; Janick *et al*., 1996). Genetic analysis of wild trees found in mountainous regions of Central Asia in Southern Kazakhstan, Kyrgyzstan, Tajikistan, and China revealed that they belong to the species *M. sieversii* and display similarities to the domesticated apple (Juniper *et al.*, 2001).

Malus x domestica is suggested to result from the hybridization between *M. sieversii* and *M. prunifolia* or between *M. baccata* and *M. sieboldii* in the eastern regions and probably between *M. sieversii* and *M. turkmenorum* or between *M. sieversii* and *M. sylvestris* in the western regions (Juniper *et al.*, 2001; Robinson *et al.*, 2001; Harris *et al.*, 2002). Apples were then spread throughout Asia from Western China to the Black sea by merchants travelling from Europe using the Old Silk Road. Apple cultivation diversified and became an important part of both Greek and Roman horticulture, after the development of the grafting technique in Mesopotamia 3,800 years ago (Robinson *et al.*, 2001).

Importantly, studies using both nuclear and chloroplast DNA confirmed that the domesticated apple is closely related to *Malus* species (Harris *et al*., 2002). However, the most accurate nomenclature is disputed between *Malus* x *domestica* (Borkh.) and *Malus* x *pumila* (Mill.), the former being most commonly used (Korban and Skirvin, 1994).

1.1.2 Selection of apple cultivars

The development and selection of new apple cultivars displaying one or more desirable characteristics/traits, like improved yield, disease and chilling resistance or fruit quality, have traditionally been carried out by vegetative propagation. Some individuals of *M. sieversii* have shown some resistance to many pests and diseases affecting *Malus x domestica* and have become the subject of research to develop new disease-resistant cultivars. In addition, several wild species, like *M. baccata* and *M. sylvestris*, have been used for breeding new apple cultivars with improved cold tolerance and/or fruit quality (Sauer, 1993). Fruit quality traits important in breeding include texture, colour and

appearance of apple, among others (Stebbins, 1991; Guilford *et al.*, 1997; Galli *et al.*, 2005).

Vegetative propagation by grafting has allowed the selection of high quality apples through the centuries by propagating a single tree developed from a seed, which coincidentally had more appealing fruit and growth characteristics than other locally known varieties (Stebbins, 199l). Some of the cultivars dominating the industry include 'Granny Smith', 'Royal Gala', 'Fuji' and 'Pink Lady' (Janick *et al.*, 1996; Norton, 1997). These cultivars have been produced from successful breeding programs, with the principle objective of increasing marketability of apples.

1.2 FRUIT QUALITY

As defined by Kramer and Twigg (1970), quality represents the characteristics differentiating individual products and has significance in determining the degree of acceptability of the product by the buyer. Consumers evaluate quality on the basis of their five senses: sight, hearing, smell, taste, and touch. As a result scientists have attempted to quantify and evaluate quality using these same senses (Gliha *et al.*, 1981). Early work in the evaluation of taste and flavor of apples was based largely on the personal judgment of scientists conducting the trials rather than through scientific assessment of sensory characteristics. About 30 years ago, Hedonic tests, which are based on scientific assessment of sensory characteristics, were introduced and used for measuring consumer acceptance of fruit quality characters, like colour and flavor, on the apple cultivars 'Delicious', 'Gala' and 'McIntosh' (Liu and King, 1978; Crassweller *et al.*, 1984; Greene and Autio, 1990, 1993). Hedonic tests use a multivariate technique to reveal relevant relationships between products and sensory preferences of consumers,

based on sensory preference (Jonsson, 2003).

In addition, cultural behaviors play an important role on the evaluation of fruit quality. For example, consumers from countries in Southern Europe prefer fruits with low acidity and red or yellow skin colouration, whereas consumers from countries in Northern Europe prefer green fruits with high acidity (Alavoine *et al.*, 1990). On the other hand, Swiss apple consumers prefer crispy and juicy apples (Hoehn *et al*., 2008). Therefore, it is clear that the majority of the consumers focus their choice on texture and taste rather than aroma and appearance of the fruit (Alavoine *et al.*, 1990; Daillant-Spinnler *et al.*, 1996; Hoehn *et al*., 2008). Consumer preferences are also influenced by past memories and experiences with a given product and cultivar name (Alavoine *et al.*, 1990; Crosby *et al.*, 1992; Durner *et al.*, 1992; Greene and Autio, 1993; Greene, 1998).

Along with quality, the nutritive characteristics of apple fruits are also important to consumers (Hoehn *et al.*, 2003; Babojelic *et al.*, 2007). Apple fruits have health promoting values as they contain high levels of vitamins, organic acids, fibre and soluble sugars. They are also a source of natural antioxidants (Harker *et al.*, 2002). Antioxidants have an ability to scavenge reactive oxygen species in the cytoplasm, thus are important in limitation of oxidative lipid damage in cells (Harker *et al.*, 2002).

Thus, for the determination of fruit quality, it is necessary to measure the relevant quality trait indicators including the sensory characteristics, nutritional values, chemical compositions, and mechanical attributes (Abbott *et al.*, 2000; Jobling, 2002; Sivakumar, 2006). Fruit quality needs to be measured using standard instruments like the penetrometer for determination of firmness to ensure reproducibility of measurements

without human bias (Bramlage, 1983; DeLong *et al.*, 2000; King *et al.*, 2000; Liebhard *et al.*, 2003a; Peng and Lu, 2007).

1.2.1 Fruit quality traits

Fruit quality traits include appearance (colour, shape, stripness and size), flavor, and texture (firmness, juiciness and crispness) (Malanczuk, 2005). While fruit quality is mainly appreciated through appearances, it is texture and flavor that ultimately determine consumer satisfaction. However, appearance characters like colour are not reliable indicators for determining flavor and texture (Seymour *et al.*, 2002; Gessler and Patocchi, 2007). Currently, technologies for determining traits according to fruit appearance are being adopted, especially in the United States and other industrialised nations.

1.2.1.1 Colour

Colour is one of the most important fruit quality traits that create the image of a fruit (Dobrzanski and Rybzynski, 2002). Colour in apples can be a dominant visual trait that influences purchasing patterns of consumers (Crassweller *et al*., 1984). A good relationship has been established between anthocyanin levels, visual rating and chromaticity values (Singha *et al.*, 1991a,b). In addition to the visual appeal of fruits, anthocyanins, the pigment that confer the red colouration of the skin as well as other organs of the plant, have been shown to be beneficial to human health due to their antioxidant activity (Takos *et al.*, 2006).

Colour evaluation is also crucial to determine fruit maturity. For example as green apple fruits ripen, the skin colour changes from green to yellow. The observation of

skin colour is a valuable method to monitor the ripening process, as revealed in a study using the cultivars 'Gala', 'Fuji' and 'Braeburn' (Plotto *et al.*, 1995). Skin colour charts can also be developed and used on new apple cultivars to provide a standard method for optimum fruit harvest. However, colour evaluations are subjective due to human bias and the inability to compare data from different years and geographical locations. Automation of colour determination using colorimeters or other colour sensing devices may allow colour data to be measured in internationally accepted units and thus allow proper data comparison.

Cultivar preference depends on external colour uniformity, fruit colour consistency among individuals from the same cultivar, ground colour differences, blush and ground colour intensity (saturation of red), fruit coverage of high colour area, physical defects, bruising and the stage of maturity (Dobrzanski and Rybzynski, 2002; Bouchra *et al.*, 2008). However, poor colour development has been shown to induce reduction in profitability in various countries. In 2003, this led to about 30% apple export decline in Australia and The Netherlands (Iglesias *et al.*, 2008). An increase in consumer interest towards red apples rather than green and bluish apples was also demonstrated in The Netherlands (Pan and Shu, 2007).

1.2.1.2 Size

Fruit size refers to the extent of largeness or smallness of a fruit, which is measured using fruit diameter. Fruit size correlates positively with profits. It is therefore a commercially valuable trait in the apple as well as in other commercial fruits like pears. Importantly, size, together with shape and colour, are the most crucial fruit quality characteristics influencing consumers (Stanley *et al.*, 2001). Apple growers throughout

the world use chemicals or hand thinning techniques in order to increase fruit size (Lortze and Bergh, 2004). Elimination of some fruitlets and/or a part of the developing fruitlets makes more photosynthate available for the remaining fruitlets. This leads to an increase in fruit size (Zaragoza *et al*., 1992). Zaragoza *et al*. (1992) also showed that an increase in fruit size could be attained prior to a massive reduction in fruit number. In addition, thinning the fruit during the one year may increase fruit size, thereby increasing crop value and induce some flowering the following year (Guardiola and Garcia-Luis, 2000).

Several studies have investigated the mechanisms influencing fruit size and revealed the positive correlation between fruit size and cell numbers (Lortze and Bergh, 2004, Volz *et al.*, 2003). Volz *et al.* (2003) also determined the positive correlation between the fruit diameter and the total cell count specifically in the fruit cortical region. Fruits with greater cell production rate from the beginning of development to maturation developed into larger fruits. In addition, a study on cell cycle genes showed that a two-fold overexpression of the *Md;CycD3* gene was observed in large fruits, suggesting that this gene induced cell division, thus leading to increase in fruit size (Janssen *et al.*, 2008; Malladi, 2005).

1.2.1.3 Acidity

Fruit flavor is also one of the most crucial criteria for the selection of apple seedlings, even though it is one of the most difficult quality traits to analyse. It is determined by a combination of chemical compounds like acids, sugars and aromatic substances. Tannins may also be present but are not in sufficient quantity to influence fruit flavor, with the exception of some cultivars used in cider brewing (Lea and Drilleau, 2003).

Besides, Alavoine *et al*. (1990) suggested that sugar content and acidity (antagonist to sugar content) are the most important determinants for consumer acceptance although the acidity taste appreciation depends upon the culture of the residents. Mechanical techniques that involved the use of a near infrared (NIR) spectroscopic instrument have been employed to assess sugar content, acidity and firmness and thus, determine the internal quality and taste of apples (Figure 1.4, Liu, 2004).

The concentrations of sugars and acids and their variations throughout fruit development and storage have been analysed in only a few common cultivars (Brown and Harvey, 1971; King *et al.*, 2000; Liebhard *et al.*, 2003b). A study by Brown and Harvey (1971) revealed that the concentrations of sugar and malic acid in ripe fruits accounted for the important variation among cultivars but they were relatively constant within a given cultivar. It was further revealed that sweetness and sourness were independently and quantitatively inherited. In addition, QTLs controlling both sweetness and acidity have been detected and were shown to be located on different linkage groups (King *et al.*, 2000; Liebhard *et al.*, 2002).

1.2.1.4 Russeting

Russeting is characterized by the development of a brown layer of suberized cells that forms soon after epidermal cell damage. As cork cells develop on the epidermis, they are pushed outward and become exposed to the surface as the fruit matures (Matteson-Heidenreich *et al.*, 1997). The epidermal cell damage can be caused by stresses including cool, wet weather, pests like apple rust mite (*Aculus schlechtendali*), infection with viruses, fungi such as apple scab (*Venturia inaequalis*) and powdery mildew

(*Podosphaera leucotricha*) or bacteria such as fire blight (*Erwinia amylovora)* and *Pseudomonas* sp. during fruitlet development (Matteson-Heidenreich, 1997).

Several fungi, including *Aureobasidium pullulans* and *Rhodotorula glutinis*, which are responsible for powdery mildew, have been shown to induce fruit russet on cultivars like 'Jonathan' and 'McIntosh' (Figure 1.1). These two fungi are known to attack on the surfaces of apple fruits and leaves. Fruits are particularly susceptible to this kind of russet at bloom and during development, from four weeks after pollination to fruit maturity (Matteson-Heidenreich, 1997).

Pests like the apple rust mite affect both the leaves and fruits. They can limit fruit size when feeding on premature terminal buds causing them to set early. The mites can feed directly on the fruit skin around the calyx end, causing tan russeting (Manganaris *et al.*, 1994). Light and golden coloured apple cultivars like 'Golden Delicious' are more susceptible to russeting that red apples like 'Fuji', although effects vary among cultivars (Taylor and Knight, 1986).

Because of a loss in attractiveness, the russeting and cracking of apple fruit skin often results in the low grading of a considerable proportion of the crop, leading to decreased profitability. In addition, increased rotting, and possible water loss of russeted stored fruits cause further economic loss (Taylor and Knight, 1986). Apple russeting can be controlled by cultivating tolerant cultivars like 'Fuji'. However, since knowledge of the genetic factors underlying russeting susceptibility is limited, preventive chemical treatments are generally used (Eccher *et al*., 2008). Recently, marker-assisted selection

has been implemented in this study to assist in the detection of resistant cultivars as early as the seedling stage (Maharaj, 2007).

A B

Figure 1.1: Russeting effect on apple fruits (Matteson-Heidenreich, 1997).

Russeting of apple fruit caused by apple rust mite (A) and powdery mildew on an unsprayed 'Golden Delicious' fruit (B).

1.2.1.5 Fruit stripness

The apple peel colouration has a significant impact on fruit appearance, which in turn plays an important role on fruit quality. Cliff *et al*. (2002) showed that fruit peel colour patterns affect consumer preferences, although they do vary with culture. For example, New Zealand consumers prefer striped apples, whereas Nova Scotia consumers prefer non-striped apples (Cliff *et al.*, 2002). Since consumers are becoming increasingly health conscious, anthocyanins, which are proposed to lower the risk of cancer and cardiovascular disease, have made apple consumption more valuable (Brown *et al.*, 1998).

Variations in the extent of fruit stripness are observed among cultivars and trees, as shown in Figure 1.2. The specific causes of stripness and its extent on fruits are poorly understood. Takos *et al.* (2006) suggested that the varying accumulation of

anthocyanin, which causes the red colouration on apple skin (Jones *et al.*, 2003), might explain the extent of stripness. Anthocyanin accumulation may be affected by abiotic factors, particularly light and temperature, as well as plant and genetic factors. Plant factors include the crop load and canopy shape of the tree, while genetic factors include mutations that may occur during propagation. The mutation of several genes, like *MYB*10 that controls anthocyanin production has been suggested to influence anthocyanin accumulation (Takos *et al*., 2006). Though this may be the cause of stripness, no data on the variations of fruit stripness over years and the cultivar effect on fruit peel colour and pattern is currently available.

Figure 1.2: Extent of stripness on apple fruits (http://www.nyapplecountry.com).
(A). Stripped 'honeycrisp' cultivar, (B). 'Gala' red blushed stripped and (C). Red stripped 'Canada Red'.

1.2.1.6 Form

Fruit size and shape are amongst the major factors determining yield, quality and consumer acceptability for many crops. Wild apple species bear small round fruit adapted for seed production and dispersal, while in cultivated apple varieties, fruit size, form and shape vary to a great extent. The cultivated apple varieties have been selected and propagated for their utility and/or sheer novelty (Herrera, 1992).

Like most of the traits important to agriculture, both fruit size and shape are quantitatively inherited (Grandillo and Tanksley, 1999). Most of the knowledge gained on the genetic regulation of fruit morphology has been gained using tomato (*Solanum lycopersicum*), a model fruit-bearing species characterised by a wide range of fruit morphology (Grandillo and Tanksley, 1999). The detection of genes encoding form and size in tomato opened a new dimension aimed at understanding fruit form and size in other species like apple and pear.

Further, mapping studies using common, orthologous genetic markers have allowed the identification of QTLs linked to fruit size and shape in multiple fruit-bearing crops from the Solanaceae family like eggplant (*Solanum melongena*) and pepper (*Capsicum annuum*, Ben-Chaim *et al.*, 2001; Doganlar *et al.*, 2002) and the Rosaceae family like pear and apples (King *et al.*, 2000; Liebhard *et al.*, 2002). These QTLs could be a result of mutations in orthologous genes. In addition, Doganlar *et al*. (2002) identified the ortholog of *fw2.2* as being the most significant locus accounting for variations in fruit size in eggplant , while in pepper, it corresponds to a major fruit weight QTL.

Apple fruit shape has been analysed using an image analysis program, which extracts fruit caliper measurements, using Fourier descriptors (Figure 1.3). Five independent shape traits, namely fruit aspect, asymmetric-crown, fruit conicity, asymmetric-sides and fruit squareness, were identified using principal component analysis (PCA) of Fourier descriptors (Currie *et al.*, 2000; De Salvador *et al.*, 2006). Out of these, aspect, conicity and squareness were observed useful for the visual assessment of shape. Fruit aspect can be assessed by calculating the fruit length to width ratio and fruit conicity by measuring the distance of the maximum width from the base of the fruit to fruit length

ratio as well as the calyx basin width to fruit width ratio, while the squareness can be estimated by calculating the ratio of the product of calyx basin width and distance of the maximum width from the calyx end of the fruit by the product of fruit length and fruit width (De Salvador *et al.*, 2006). Therefore, aspect, conicity and squareness can be used for genetic analysis of shape.

A B C D

Figure 1.3: Common apple fruit shapes (http://www.applejournal.com)

(A) Oblong-conical such as 'Harvey', (B) round to conical such as 'Elstar', (C) plain conical shape and round such as 'Winter banana' (D) shaped apple such as 'Empress'.

1.2.1.7 Juiciness

Juiciness is an important factor influencing fruit taste and can be assessed by measuring total soluble solids using a refractometer. Tu and de Baerdenaeker (1997) determined that a minimum content of 12° Brix for total soluble solids, a minimum acidity of 3 g l^{-1}, as measured by the malate content, and a minimum firmness of 45 N are the acceptable limits for the consumption of 'Golden Delicious' apples.

Apple juiciness has been shown as affected by the growth conditions, including the soil type and its nutritional composition (Babojelic *et al.,* 2007). The same study also showed that some cultivars like 'Idared' had predominant juiciness characteristics in

comparison to others, like ‘Granny Smith’ and ‘Pink Lady’, thus revealing that juiciness is also influenced by genetic factors.

1.2.1.8 Crispness

Crispness is proposed to be the most sought-after apple trait (Kilcast and Fillion, 2001). It is a result of the cell wall strength and turgor pressure (Oey *et al*., 2007). Turgor pressure is exerted by intracellular liquids on the cellular membrane and cell wall and imparts turgidity and rigidity and thus crispness. In addition, turgor pressure has a major influence on tissue strength and macroscopic fruit firmness (Oey *et al*., 2007).

Consumers prefer crisp and juicy apples to soft or slightly mealy ones. The pulp should literally crack upon chewing, melt and disappear. In addition, a fruit must also be firm to be appealing to the consumer (Abbott *et al.*, 2004). For example, the exceptional crispness and juiciness of the ‘Honeycrisp’ apple cultivar greatly accounted for its commercial success. Therefore, research focused on improving crispness, juiciness, flavor and colour of fruits is important since these traits are important to consumers (Kilcast and Fillion, 2001).

1.2.1.9 Firmness

Fruit firmness/softness is a critical determinant to estimate the potential of a fruit to last through prolonged storage. Excessive softening is considered undesirable in apples and can lead to lower sensory values for firmness, juiciness, crispness and crunchiness, but increases mealiness (Abbott *et al.*, 1984; Jaeger *et al.*, 1998; Jaeger, 2000; Harker *et al.*, 2003) and thus reduced consumer acceptability (Liu and King, 1978). Apples, like other climacteric fruits, display an increase in ethylene production during ripening, which is

responsible for changes in texture, firmness and colour. Endogenous ethylene induces fruit ripening and promotes softening of some apple cultivars (Abeles and Biles, 1991).

Fruit firmness assessment, also known as fruit pressure test or penetrometer test is generally measured using a puncture test in the apple fruit (Johnston *et al.*, 2002). Many markets routinely use fruit firmness as a guide to ensure that apples delivered to customers have the required textural characteristics all year-round (Johnston *et al.*, 2002). The firmness test was used to determine the factors influencing the texture of 'Royal Gala' (Volz *et al.*, 2003). The study revealed that cell size, cell packing and intercellular air space influence fruit firmness as well as crispness. Firmer fruits were observed having reduced cortical cells and intracellular spaces but increased cell packaging (Volz *et al.*, 2003). Besides the use of the penetrometer, fruit firmness can be assessed using a HyperSpectral Imaging system, a more reliable and non-destructive technique, which is based on the ability to defract infrared light.

1.2.1.9.1 Fruit firmness assessment using HyperSpectral Imaging system

Fruit quality can be measured using non-destructive methods, thus delivering superior quality and consistent fresh products to the consumer. Internal fruit quality can be measured using an optical technique like the Near-infrared spectroscopy (NIRS) (Figures 1.4-1.5) in the Hyperspectral Imaging system. This technique measures diffusely reflected or transmitted light over a range of invisible wavelengths. It has been used for predicting firmness (Lu *et al.*, 2000; Servakaranpalayam, 2006) and to some extent sweetness of apples and other fresh fruits (Servakaranpalayam, 2006).

When a light beam is incident upon a fruit, part of the light is absorbed while the rest is scattered in the form of either backscattering reflectance or transmission (Figure 1.5). Light absorption is related to some of the fruit components like sugar, water, or chlorophyll. Scattering, on the other hand, is associated with the structural features of the fruit and hence, is useful for measuring the textural properties of fruit. If both absorption and scattering are measured, greater knowledge about both chemical (sugar and acid) and physical (firmness) properties of fruit can be obtained (Lu, 2004).

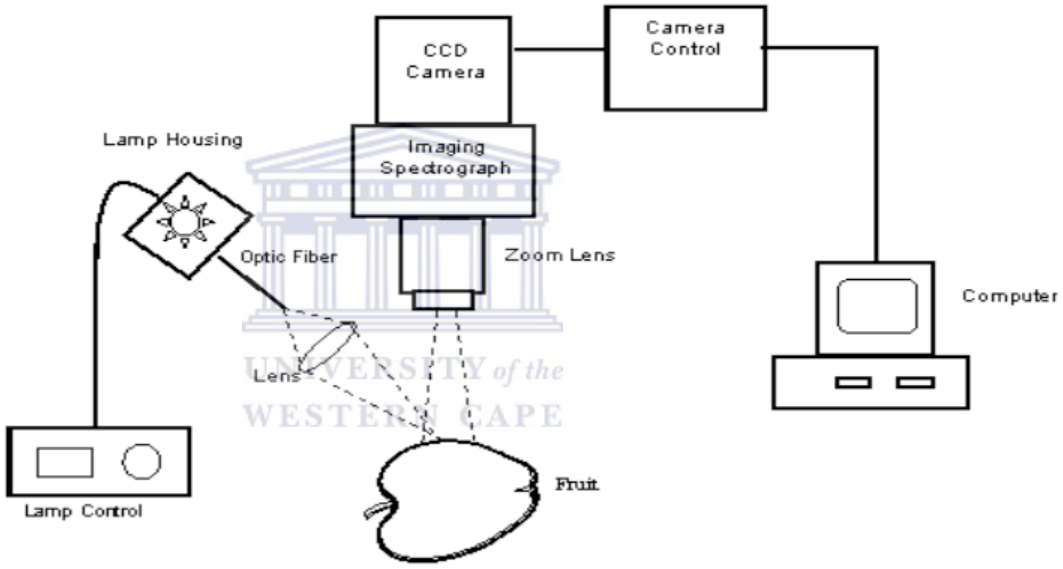


Figure 1.4: Fruit firmness assessment using the Hyperspectral imaging system.

Incident light is focused on the fruit, which is partly absorbed and partly scattered in the form of backscattering reflectance or transmission. The transmitted light is reflected onto the imaging spectrograph, attached to the camera, which will capture the spectrum and convey spectral data for analysis to the computer (Adapted from Lu *et al.*, 2000)

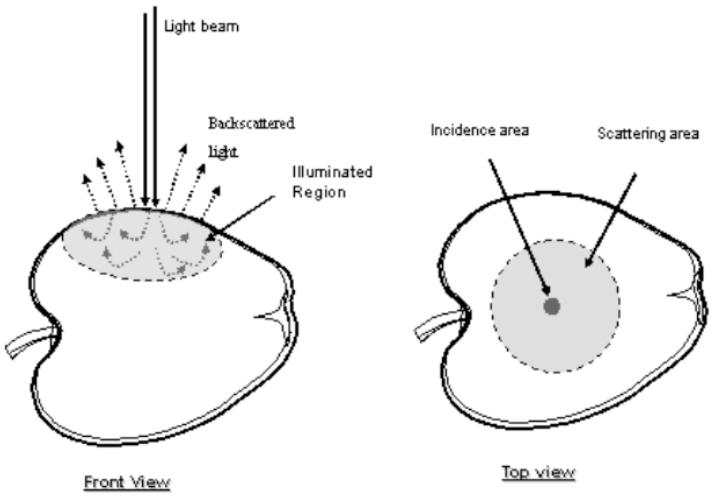


Figure 1.5: Concept of measuring Hyperspectral scattering on an apple fruit.

The front view shows the backscattering of the incident light on the fruit illumination region and the top view shows the actual incidence area and scattering region on the fruit. (Adapted from Lu, 2004).

Near infrared spectroscopy can be applied for sweetness and firmness assessment of apple and other fruits like peach (*Prunus persica*) and pear (*Pyrus pyrifolia*) or their classification prior to commercialization (Lu, 2004). The technique has been used to measure several properties in a wide range of products like soluble solids in apple juice (Ventura *et al.*, 1998) and in peaches (Kawano *et al.*, 1995), dry matter in onions (*Allium cepa*) (Gerald *et al.*, 1989) and internal quality in peaches and nectarines (*Prunus persica* var. nectarina; Slaughter *et al.*, 2003). This method however, is performed on harvested fruits prior to distribution to the market, which is important in marketing good quality product.

1.3 PLANT BREEDING

Plant breeding can be defined as the art and science of changing the genetics of plant species in order to create improved genotypes/phenotypes for specifically defined purposes (Chahal and Gosal, 2002). It leads to the production of new cultivars displaying one or several improved characters with regards to yield, disease resistance and/or quality (Chahal and Gosal, 2002). The overall objective of plant breeding is to increase the marketability, while reducing the production cost (Hrazdina, 1994). This involves traditional breeding through controlled cross-pollination, or genetic engineering, followed by artificial selection of progeny (Tobutt *et al.*, 2000).

Integrating both traditional breeding and genetic engineering is vital to genetically improve apple plants and fruits. This dual approach can be used to maximize the success of biotechnology and breeding programs. Each approach has its own advantages and disadvantages.

1.3.1 Traditional breeding

Traditional breeding has been applied to produce new varieties or lines with desirable properties using deliberate interbreeding of closely or distantly related individuals (Janick *et al*., 1996). Parents are chosen on the basis of desired traits they possess, to produce offsprings that will contain the desirable characteristics from both parents, while minimizing their negative traits. Seedlings from cross-pollination may closely resemble their parents, but never be identical to either one and therefore, loose the varietal identity. An example was the cross between ‘McIntosh’ and ‘Delicious’ that resulted in the creation of ‘Empire’, which resembles ‘McIntosh’ but is genetically distinct from it (Hrazdina, 1994). To ensure hybrid vigor, progeny from the cross (F1

generation) needs to be crossed with the high yielding line to obtain a new plant that display the desired characteristics from both parents (Chahal and Gosal, 2002).

Although apples have been cultivated for centuries and are one of the most important fruit crops worldwide, genetic studies and breeding are hampered by their long juvenile phase, space, time and the cost of screening and maintenance of populations, the high chromosome number (2n=34), high level of heterozygosity and self-incompatibility and the out-breeding mode of reproduction and vegetative development (Maliepaard *et al*., 1998). Furthermore, breeding of new apple cultivars is a very long and tedious process requiring an average of 20 years, from cross-pollination, seedling selection to field trials and new cultivar release. This is partly due to the long juvenile phase a characteristic of fruit trees. In this regard, recent modern techniques have been developed to ensure early selection of plants displaying desired characteristics and therefore reduce the cost of classical breeding (Gessler and Patocchi, 2007).

Traditional breeding can alter both simple and complex traits and at the same time losing varietal identity, while modern biotechnology techniques maintain the varietal identity but can only improve simple traits (Collard *et al.*, 2005). Modern biotechnology techniques offer a great potential for plant breeding by greatly reducing the time necessary to breed new crop varieties. The use of molecular techniques allows the introgression of desirable genes among varieties from related wild species, without undergoing the genetic engineering process. In addition, molecular markers can be used to monitor linkage drag during each backcross generation. In this regards, individuals with recombinant haplotypes can be selected to remove linkage drag. Linkage drag refers to the reduction in fitness in a cultivar due to deleterious genes introduced along

with the beneficial genes during backcrossing (Young and Tanksley, 1989; Chahal and Gosal, 2002).

1.3.2 Genetic engineering

Genetic engineering, also termed recombinant DNA technology refers to the direct manipulation of the genes of an organism. Biotechnological techniques like genetic transformation can be used to insert genes of interest or block the expression of less desirable genes. Since genes of interest from other crops can be used, genetic transformation therefore may be useful for breeding apple cultivars resistant to diseases like fire blight (Reddy and Thomas, 1996; Norelli *et al.*, 2003), apple scab or powdery mildew (Manganaris *et al.*, 1994) or to insects (Brown, 1995). Transformation has been used to generate apple plants that are less susceptible to fire blight, a bacterial disease caused by *Erwinia amylovora* (Hanke *et al.*, 2002; Peil *et al.*, 2006).

Though, genetic engineering is a powerful technology capable of providing great benefits, it also could carry risks. This technology and its resulting products have always been highly regulated in the United States, for example, and the scientific basis for the regulatory oversight is under constant review by the scientific community as well as by the relevant federal and state government agencies (McHughen, 2006). As a result, crops and foods derived from genetic engineering receive greater regulatory safety evaluation prior to commercial release than any other crops or foods in the history of agriculture. This regulatory framework ensures that the safety of genetically engineered crops is superior or at least comparable to the safety of conventionally produced crops, although regulatory costs for genetic engineering crops are much higher than for conventional ones (McHughen, 2006). Thus, this technique is barely

used nowadays because of food safety regulations (Sithole-Niang, personal communication).

1.3.3 Marker assisted breeding

Marker assisted breeding (MAB) is a form of biotechnology which uses genetic fingerprinting techniques to assist plant breeders in matching molecular profile to the physical properties of the variety. It is particularly promising for perennial tree crops, like apple, since many important traits are expressed only after several years of costly field maintenance. MAB allows the marker-assisted introgression of important and/or favorable genes from exotic donors to enhance elite breeding material (Lecomte *et al*., 2004).

1.3.4 Marker-assisted selection

Marker-assisted selection (MAS) is based on the results of DNA testing to assist in the selection of individuals that will be used as parents of the next generation (Van Eenennaam, 2004). MAS enables the detection of favorable alleles in the early stages of plant development and thus allowing a significant reduction of the breeding population (Liebhard *et al*., 2003a). In MAS, marker(s) should co-segregate or be closely linked with the desired trait. This technique has the potential to greatly increase the efficiency of plant breeding.

The developments in the DNA marker technology together with the concept of MAS provide new solutions for selecting and maintaining desirable genotypes. The molecular marker techniques are useful tools in breeding for the detection of genes coding

economically important traits and also the selection of plants displaying these traits (Janick *et al*., 1996; Maliepaard *et al.,* 1998; Liebhard *et al*., 2002, 2003a; Galli *et al*., 2005). Once molecular markers closely linked to desirable traits are identified, MAS can be performed in early segregating populations and/or at early stages of plant development, thus limit the cost involved in tree maintenance. MAS can be used to pyramid the major genes including resistance genes, with the ultimate goal of producing varieties with more desirable characters (Liebhard *et al.*, 2002; Graham *et al.*, 2009). Thus, with MAS it is now possible for the breeder to conduct several rounds of selection every year, a process called MAB.

Molecular markers have allowed the mapping and tagging of many agriculturally important genes, which forms the foundation of MAS in crop plants. Molecular tags, a prerequisite for MAS, have been developed for many crop plants using different kinds of molecular markers (Liebhard *et al*., 2003a). Molecular markers offer great scope for improving the efficiency of conventional plant breeding by carrying out selection not directly on the trait of interest but on molecular markers linked to this trait. However, this requires a molecular marker to be closely linked to the trait of interest. These markers are unaffected by the environment and the plant growth conditions and are detectable at all stages of plant growth (Hospital *et al.*, 1997).

1.4 MOLECULAR MARKERS

Molecular markers can be defined as segregating variations in known short DNA sequences (Young, 1992). Molecular markers are often referred to as "fingerprints" because they can be as unique and useful for identifying plants or plant characteristics as fingerprints are to humans. Each individual plant has a unique set of genetic markers

(Brown, 1995). The quality of a molecular marker is typically determined by its heterozygosity, which is defined as the probability of an individual to have two different alleles at a given locus, in the mapping population (Botstein *et al.*, 1980). Molecular markers like microsatellites, have become the basic tools in genetic mapping mainly because they are easily identifiable, locus-specific, highly polymorphic, easily genotyped and transferable between cultivars (Maliepaard *et al.,* 1998; Liebhard *et al*., 2002, 2003a). This makes molecular markers very useful in the generation of genetic linkage maps for different plant populations.

Molecular markers can be used to study the relationship between an inherited disease and its genetic cause like a particular mutation of a gene that results in a defective protein. It is known that closely related pieces of DNA on a chromosome tend to be inherited together (Nagaraju *et al.*, 2002). This property enables the use of a marker, which can then be used to determine the precise inheritance pattern of the gene, which has not yet been positioned. The markers can be detected directly by DNA sequencing, or indirectly by the use of isozymes (Tanksley, 1993).

Several apple genetic maps have been developed by positioning genetic markers. These markers enabled detecting location of genes of interest. Molecular markers have been found for burr knots, columnar form, and certain fruit attributes (Ames *et al.*, 1984), apple scab (Yang *et al.*, 1997a; Belfanti *et al.,* 2004; Silfverberg-Dilworth *et al*., 2005), and storage scald (Rupasinghe *et al.*, 2000).

While polygenic characters were very difficult to analyze using traditional plant breeding methods, molecular markers allow easy tagging of desirable polygenic traits

(Mullis, 1990). These markers are used for genetic mapping, localization of major genes and quantitative trait loci (QTLs) detection. Several types of markers are available, including isoenzymes (Vinterhalter and James, 1986; Weeden and Lamb, 1987; Manganaris and Alston, 1989; Weeden, 1989; Samimy and Cummins, 1992), restriction fragment length polymorphisms (RFLPs, Botstein *et al.*, 1980; Helentjaris *et al.*, 1987; Weber and Helentjaris, 1989; Roche *et al.*, 1997). Polymerase chain reaction (PCR)-based DNA markers like random-amplified polymorphic DNAs (RAPDs, Welsh and McClelland, 1990; Williams *et al.*, 1990; Williams *et al.*, 1991; Hemmat *et al.*, 1994; Conner *et al.*, 1998a), sequence characterised amplified regions (SCARs, Williams *et al.*, 1991; Markussen *et al.*, 1995; Gianfranceschi *et al.*, 1998), sequence-tagged sites (STS) and inter-simple sequence repeat amplification (ISA), amplified fragment length polymorphic DNAs (AFLPs, Becker *et al.*, 1995; Vos *et al.*, 1995), amplicon length polymorphisms (ALPs, Mullis, 1990; Mohan *et al.*, 1997), single nucleotide polymorphism (SNP, Collins *et al.*, 1998; Cho *et al.*, 1999; Brown, 2006; Celton *et al.*, 2009), diversity arrays technology (DArTs, Wenzl *et al.*, 2004; Xia *et al.*, 2005; Mace *et al.*, 2008) and microsatellites/simple sequence repeats (SSRs, Liebhard *et al.*, 2003c; Newcomb *et al.*, 2006; Silfverberg-Dilworth *et al.,*2006; Naik *et al.*, 2006 Gasic *et al.*, 2008; Celton *et al.*, 2009), are also available.

1.4.1 Microsatellites

Microsatellites, also called SSRs, are simple, di-, tri- and or tetra-nucleotide sequence motifs flanked by unique sequences. The repeat units range from one to 13-mers in length and are usually repeated five to 20 times (Brown, 2006). In nature, the most abundant microsatellite motif reported in plants is (AT)n, while (AC)n is the most abundant in the human genome (Brown, 2006).

Allelic differences (simple sequence length polymorphisms) are usually the result of the variable numbers of repeat units within the microsatellite structure. The number of repeats at any one locus can be determined by PCR using primers that will anneal at either side of the short tandem repeat. Moreover, since Guilford *et al*. (1997) characterised microsatellites in apples and Hemmat *et al.* (2003) developed a microsatellite marker containing a simple sequence repeat (SSR), linked to the Co gene for columnar tree habit, SSRs have become predominant markers used for MAS. As such, SSRs have a great potential for improving valuable traits of plants.

SSRs are valuable genetic markers since they are transferable from one cultivar to another, co-dominant and can detect a high level of allelic diversity. They are easily and economically assayed by PCR (McCouch *et al.*, 1997; Adam-Blondon *et al.*, 2004). In addition to the number of alleles detected, heterozygosity is one of the most important characteristics of a locus. From the perspective of genetic linkage map construction, a marker on a locus with heterozygosity higher than 70% is commonly considered a highly informative marker, because the segregation can easily be monitored within a given population (Gupta *et al.*, 2003; Chagné *et al.*, 2004; Oraguzie *et al.*, 2005).

Microsatellites developed in apple are easily transferable not only between cultivars but also between closely related species. Apple SSRs were successfully mapped on pears, which belong to the same subfamily of Maloideae (Yamamoto *et al.*, 2001; Yamamoto *et al.*, 2002a). Apple SSRs have also been applied to the subfamily Prunoideae using peach and cherry (Yamamoto *et al*., 2002a).

Apple EST sequencing projects have created thousands of apple EST sequences and can be used to search for SSR repeats and develop new SSR markers (Crowhurst *et al.*, 2005), thus facilitating map construction as demonstrated by Silfverberg-Dilworth *et al.* (2006) with the construction of the apple reference genetic linkage map. The apple SSR database (http://www.hindras.unimi.it), a European initiative, was also established to saturate genetic maps constructed using SSRs, thus increasing the apple genome coverage. In genetic mapping, SSRs are considered valuable when polymorphic in the nucleotide sequence. However, SSRs derived from ESTs (EST-SSRs) were reported to have lower levels of polymorphism than genomic SSRs in cereal crops (Gupta *et al.,* 2003), which is contrary to tree species like pine (Brown *et al.*, 2001). Generally, the amplification quality of SSRs derived from genes is better than genomic SSRs (Gupta *et al*., 2003). Most apple SSR markers published are predominantly derived from genomic sequences (genomic-SSRs) and are mostly used for genetic map construction (Wheeler *et al.*, 2006).

1.5 GENETIC MAPPING

Genetic mapping refers to the determination of the relative position of a gene on a chromosome. This process is achieved through constructing genetic maps using molecular markers such as SSRs or SNPs and then detecting relative position of a gene on the maps. The genetic map is composed of linkage groups, each of them representing a chromosome (Maliepaard *et al.*, 1998; Shibaike, 1998).

A genetic map is based on the frequencies of recombination between markers during crossover of homologous chromosomes. The greater the frequency of recombination (segregation) between two genetic markers, the further apart they are assumed to be.

Conversely, the higher the frequency of association between the markers, the smaller the physical distance between them (Copenhaver *et al.*, 1998). For the first genetic linkage maps, detectable phenotypes like the eye colour of *Drosophila melanogaster* were used as markers. The detectable phenotypes were obtained by genetic studies using the Mendelian principles and involved directed breeding programs for experimental organisms or pedigree analysis of plants (Copenhaver *et al.,* 1998). The mapping and sequencing of plant genomes are useful for elucidating gene or genome primary structure, while the EST sequences profile the transcriptome and can show gene regulatory events (Baldi *et al.*, 2004; Watanabe *et al.*, 2006). In this regard molecular markers are used to identify and tag desired genes.

1.5.1 Linkage analyses and genetic map construction

The introduction of molecular markers facilitates the performance of linkage analysis and map construction, even for large segregating populations. Linkage can be defined as the genetic distances between polymorphic traits, which may be recognized as differences in enzyme activities, restriction fragment lengths or nucleotide sequences at an allelic locus (Luo *et al.*, 2000). Linkage analysis refers to the determination of the probability for genes and other genetic markers to be inherited together because of their location near one another on the same chromosome. Linkage analysis is based on the co-segregation of adjacent markers and their co-transmission to the next progeny generation (Wu *et al.*, 2007; Wen and Wu, 2008).

The construction of a genetic map and the analysis of the linkage groups using molecular markers is far more complex using a full-sib family from an out breeding plant species, as up to four segregating alleles (tetraploids) for each locus can be

detected. Also, in general, the parental origin (grand parental lines) at any given locus is unknown (Maliepaard *et al.*, 1998). In this regard, Maliepaard *et al.* (1998) recommended the use of multi-allelic markers like SSRs, to integrate homologous linkage groups of the respective parents, since recombination can be estimated separately in male and female parents.

Software that can handle large numbers of segregating/polymorphic loci was developed for the generation of genetic linkage maps. This software, *JoinMap*, compute the frequencies of recombination between markers during crossover (Copenhaver *et al.,* 1998). Genetic linkage maps facilitate a better understanding of the genome structure and localization of genes through identification of quantitative trait loci.

1.5.2 Overview of genetic map construction

The construction of a genetic map involve five steps, namely the creation of a segregating mapping population, the determination of genotypes of all loci (single locus analysis), the estimation of recombination frequencies between the pairs of loci (two-locus analysis), the establishment of linkage groups and the ordering of the linkage groups.

1.5.2.1 Creating a segregating population

Mapping populations generally used for genetic mapping are generated by controlled crosses between two or more parents to create variation at the phenotypic level for a trait of interest. When the parents used in out-breeding show heterozygosity, which is the case with most apple cultivars, then genetic variation exists between the parents.

This variation is essential in order to trace recombination events in their siblings (Liebhard *et al.*, 2003b; Liebhard *et al.*, 2003c; Mahmoud *et al.*, 2008).

1.5.2.2 Determination of genotypes or single-locus analysis

Genotypes of each individual in a mapping population are determined at each locus in a process called single-locus analysis. Seedlings derived from a cross between two diploid, heterozygous parents may have up to four alleles segregating at any given locus, which are expected to be passed onto the progeny in equal ratios. The possible inheritance of alleles by the progeny and their expected allelic frequencies are presented in Table 1.1.

Any deviation from the expected segregation pattern (according to the allele frequency ratios shown in Table 1.1) might be an indication of low informative data, generally caused by missing data from the individuals in the population being mapped, non-random or insufficient sampling (Liebhard *et al.*, 2003c; Mahmoud *et al.*, 2008).

Table 1.1: Allelic frequencies expected in the progeny derived from a cross between two diploid parents.

Cross	Number of possible progeny genotypes	Frequency			
		A1	A2	A3	A4
ab x cd	4	0.25	0.25	0.25	0.25
ab x ab	3	0.25	0.25	0.5	0
ab x cc	2	0.5	0.5	0	0
aa x aa	1	1	0	0	0

1.5.2.3 Two-locus analysis

Genetic linkage map construction is based upon the recombination frequencies observed between different markers. High quality raw data files with low levels of genotyping errors and/or missing data are essential for the construction of a genetic linkage map. Liebhard *et al.* (2002) observed cases where faulty classification of a single individual in the mapping population caused an entire chromosome segment to change its orientation. Although missing observations might result in recombination frequencies estimated from smaller data sets being less accurate, this is less critical than genotypic errors (Van Ooijen and Voorrips, 2001).

The construction of a genetic linkage map requires knowledge of heritage or separation rate of the loci by genetic recombinations (Liebhard *et al.*, 2003c). Recombination frequency (r) between two loci depends on the distance that separates them. A recombination frequency of zero means that no recombination occurred at this location. Thus, it can be assumed that the two loci were inherited together. A recombination frequency of zero does however not necessarily imply that the two loci are located next to each other on the actual genome (Burr and Burr, 1991). A marker pair is linked when the marker frequencies obtained in the progeny are significantly different from the expected frequencies in the absence of linkage (Liebhard *et al.*, 2003c; Dugo *et al.*, 2005).

1.5.2.4 Establishing linkage groups

Linkage groups are calculated on the basis that alleles of markers positioned on different chromosomes segregate independently, while alleles of markers on the same chromosome pair segregate more often in the same combination. A linkage group is

thus a group of loci with a recombination frequency lower than 0.5, as expected for independent, or unlinked loci. If no false linkage is assumed, then the number of genetic linkage groups statistically obtained should correlate to the haploid number of chromosomes. The number of linkage groups can however exceed the number of chromosomes when no loci on a large segment of a particular chromosome are observed, causing a 'chromosome break' in the linkage group (Van Ooijen and Voorrips, 2001).

The *JoinMap*® software has been used for the construction of recently published apple genetic linkage maps (Igarashi *et al*., 2008; Van Dyk, 2008; Woodhead *et al*., 2008; Celton *et al*., 2009; Patocchi *et al*., 2009). It uses the "Likelihood of the odds" (LOD) score test as the statistical criterion to test linkage, identified as the base 10 logarithm of the likelihood ratio. Often an LOD value of three or above is used as the significance threshold, meaning that linkage is at least 1000 times more likely than independent segregation. As a chi-square test (χ^2), LOD value of three corresponds to a significance of probability value (*p*) of 0.0002 (Hemmat *et al.,* 1997; Gianfranceschi *et al*., 1998). In their studies, Liebhard *et al.* (2003a) and Silverberg-Dilworth *et al*. (2006) used more stringent conditions, with a LOD score of four and five, respectively, to group markers belonging to the same linkage group in comparison with a LOD score of three previously used by Hemmat *et al.* (1997) and Gianfranceschi *et al*. (1998). A high level of stringency is necessary as many pairs of markers are usually tested, while a low LOD score induce markers to influence one another, resulting in larger distances between them (Liebhard *et al*., 2003a).

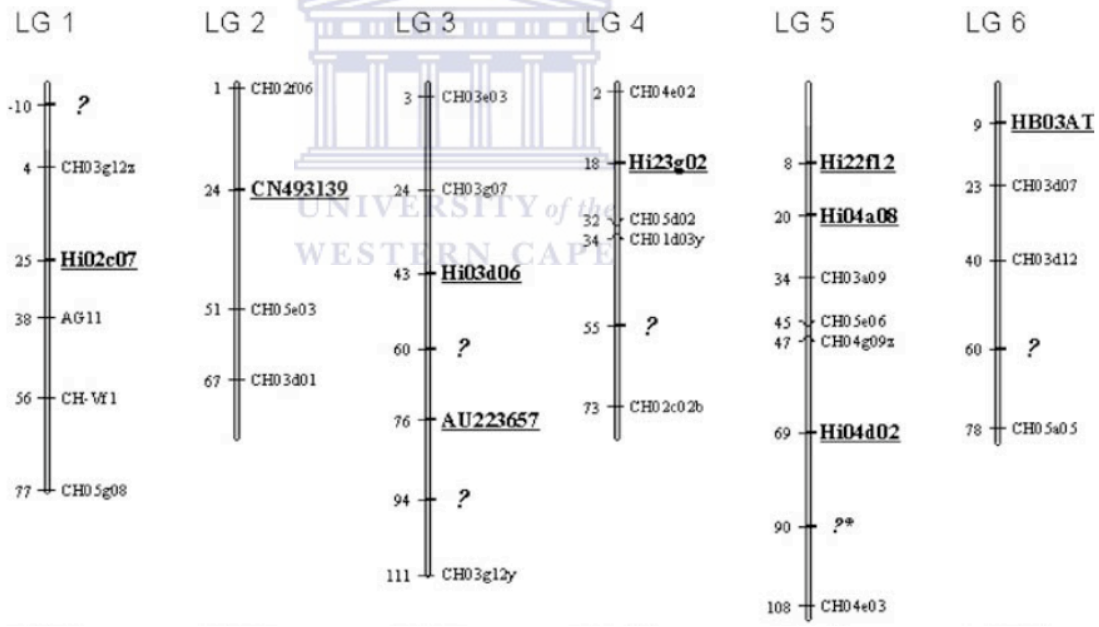
1.5.2.5 Linkage group ordering

Linkage group ordering refers to the orderly arrangement of the groups obtained during map calculation of groupings (Van Ooijen and Voorrips, 2001) using a reference standard map. The number of possible groupings for any given linkage group increases as the number of markers positioned on the map increase. This is generally performed using regression-mapping algorithm in the *JoinMap*® software, which is based on the sequential addition of markers in a systematic way (Van Ooijen and Voorrips, 2001). Recombination frequencies between markers determine their order on the genetic linkage map. Therefore, a map is calculated using recombination frequencies, after which spatial sampling of markers is used to obtain a framework map. Individual markers are added to this framework map in order to determine their position and identify markers causing map distortions.

1.5.3 Genetic linkage maps for apple

Apple genetic linkage maps were initially constructed mainly using RFLP, isozyme and RAPD markers (Maliepaard *et al.,* 1998). Elaborate parental maps have been produced using informative SSR markers (Maliepaard *et al.*, 1998; Conner *et al.,* 1998; Hemmat *et al.,* 1997, 2003; Liebhard *et al*. 2002; Silfverberg-Dilworth *et al*., 2006; Igarashi *et al.,* 2008; van Dyk, 2008; Celton *et al.,* 2009; Pattochi *et al*., 2009). These maps were all developed from scion cultivars with the exception of the study by Celton *et al*. (2009), which used apple rootstocks. These were aligned using the numbering of the 17 linkage groups based on the map published by Maliepaard *et al.* (1998). The development of these high-quality, accurate and high-density genetic linkage maps enables the discovery of genetic markers linked to desired traits and then the localization of genes controlling these traits.

A minimum of 840 molecular markers, including AFLP, RAPD, SCAR and SSR, have been used for the generation of the reference genetic linkage maps using a controlled mapping population of 'Fiesta' x 'Discovery' (Liebhard *et al.,* 2003a). The reference genetic linkage maps constructed were further saturated by Silfverberg-Dilworth *et al*. (2006) with the addition of 148 SSR markers (Figure 1.6). The transferability of SSRs across the *Maloideae* species allowed the successful transfer of 41% of the markers during the saturation of the reference map (Silfverberg-Dilworth *et al*., 2006). The SSRs covered about 85% of the genome with an average distance of 15 cM between markers (Silfverberg-Dilworth *et al.*, 2006).



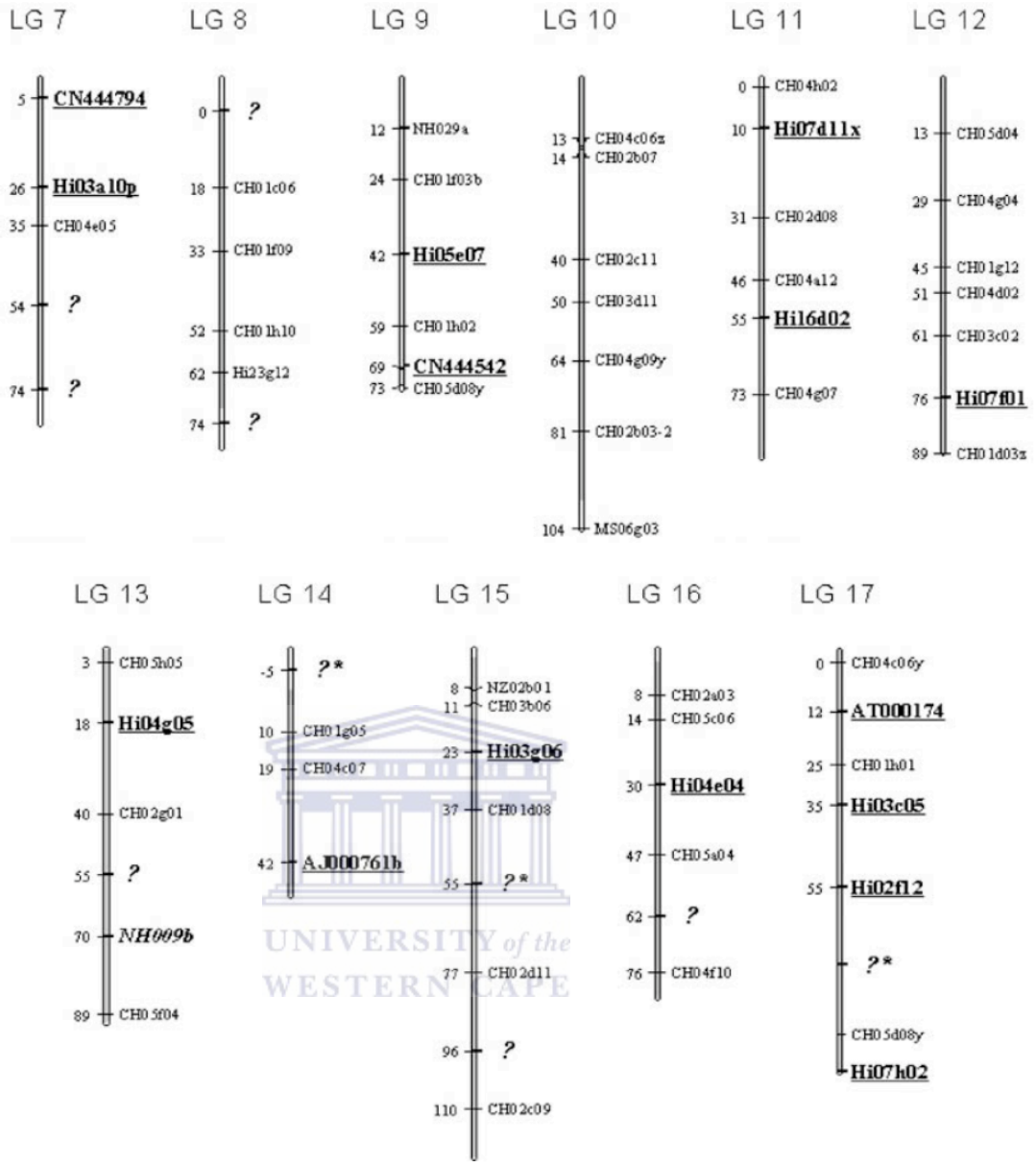


Figure 1.6: Global coverage of the apple genome using a set of 102 SSR primer pairs.

Map positions, expressed in cM, were aligned to the 'Discovery' maps. *Grey filled bar segments* indicate linkage group segments covered by the 'Fiesta' x 'Discovery' maps (Silverberg-Dilworth *et al.,* 2006). Open bar segments indicate linkage group segments not covered by the Silverberg-Dilworth *et al.* (2006) map, but revealed in other unpublished linkage maps. The symbol '?' indicate that no primer pairs were publicly available and the symbol '?^{x} marks positions of unpublished SSR markers, which are expected to become available in the near future.

1.6 QUANTITATIVE TRAIT LOCI

Quantitative trait loci (QTL) also known as polygenes, refers to individual loci controlling a quantitative trait. A quantitative trait is a genetic characteristic whose phenotypic variation is continuous and determined by the segregation of multiple loci (Tanksley, 1993; Durel *et al.*, 2003a; Calenge *et al.*, 2005). Molecular analyses of quantitative traits provide new tools, for both plant breeding, and gene cloning (Chahal and Gosal, 2002).

QTL mapping allows for the determination of the influence of individual genes or group of genes to phenotypic traits, which are complex and polygenic characters (Young, 1996). QTL mapping is based on the use of pedigrees, also known as lineages, that show segregation on phenotypic traits, and the identification of genes by linkage analysis with molecular markers (Yin, 2006). Individual traits are influenced by the environment or by the combination of the environment and genetic factors (Bradshaw and Stettler, 1995; Young, 1996; Luo *et al.,* 2000, Gardiner *et al.*, 2007).

A characteristic is simply inherited when controlled by one or several genes with a major effect and can be used to group seedlings into several discrete categories. For example, disease resistance in seedlings can be categorized as either resistant or susceptible. On the other hand, characteristics are quantitatively inherited when they are controlled by polygenes, like apple flavor, firmness, texture, fruit size or colour (Sivakumar, 2006; Zhang *et al.*, 2006)**.** For example, traits like colour are controlled by polygenes including genes coding for colour intensity (anthocyanin biosynthesis pathway) and distribution, which can lead to stripness (Sivakumar, 2006).

1.6.1 Mapping QTLs with molecular markers

Following the map construction, QTLs can be identified by testing the putative relationship between the entire set of molecular markers and quantitative trait distribution or the phenotype of interest (Tanksley, 1993). Traditionally, phenotypic data were partitioned into genetic (additive, dominant and epistatic effects) and environmental variances (Paterson *et al.*, 1991a). From this information it was then possible to estimate the heritability of a trait, predict its response to detect and estimate the minimum number of genes coding for a given trait. Nowadays, this step is carried out using the software *MapQTL*® (van Ooijen, 2004). This software has been shown to limit the statistical uncertainty observed in genetic mapping. This new approach, performed by joint analysis of segregation of marker genotypes and phenotypic values of individuals, facilitates the dissection of polygenic characters into discrete genetic loci, thus defining the roles of individual genes (Young, 1996; van Ooijen, 2004).

It is generally estimated that almost 80% of all parental chromosomes are passed to the next generation with either one or no recombination event (Singer *et al.*, 2006). Therefore, any mapping population needs to be sufficiently large in order to study accurately a large number of meiosis, resulting in enough crossovers to map QTLs. Adjacent QTLs are difficult to distinguish when a minimum of two genes are closely linked and may be mistaken for one gene. In some cases, false positive QTLs, also referred to as Type I error, may be detected. On the other hand, some QTLs, referred to as Type II error, can be missed when genes are linked in repulsion (Paterson *et al.*, 1991b; Tanksley, 1993).

While the identification of QTLs is primordial for unravelling complex traits like fruit quality, the detection of markers linked to these QTLs is necessary for MAS breeding, and the isolation and characterization of the gene(s) related to the QTLs. In apples, several candidate genes (Hough *et al.*, 1953; Knight *et al.*, 1962; Mowry and Dayton, 1964; Alston and Briggs, 1968; Dayton and Williams, 1968; Alston and Briggs, 1970; Thompson and Taylor, 1971; Aldwinckle *et al.*, 1977; Alston and Briggs, 1977; Dayton, 1977; Battle *et al.*, 1995; Bénaouf and Parisi, 2000; Bus *et al.*, 2000; Gianfranceschi and Soglio, 2004; Patocchi *et al.*, 2004; Tartarini *et al.*, 2004; Bus *et al.*, 2005a; Bus *et al.*, 2005b) and QTLs (King *et al.*, 2000; King *et al.*, 2001; Durel *et al.*, 2003b; Liebhard *et al.*, 2003b; Liebhard *et al.*, 2003c; Calenge *et al.*, 2004; Stankiewicz-Kosyl *et al.*, 2005) have been identified using various populations and cultivars (Table 1.2). The large numbers of QTLs suggest that most of the candidate genes that have been mapped play a role in determining the expression of a respective trait in the seedlings. Although QTLs controlling fruit quality traits have been identified, their genes have not been mapped to date. While the identification of QTLs is important for unravelling complex traits like fruit quality traits, the detection of markers linked with these QTLs is necessary for MAS and gene isolation and characterisation.

Table 1.2: Candidate genes for a variety of phenotypic traits mapped in *Malus x domestica* Borkh. by genetic mapping.

Trait	Linkage group														
	1	2	4	5	6	7	8	9	10	11	12	13	15	16	17
Fruit skin colour								[a,b]Rf							[b]Rs
Fruit cortex colour, red foliage								[c]MdM YB10							
Fruit acidity- malic acid														[a]Ma	
Columnar growth									[a]Co						
Dwarfing				[d]Dw1											
Ethylene production or Fruit ripening									[e]Md-AC01				[e]Md-ACS1		
Self-incompatibility															[a]SI
Rosy leaf curling resistance						[f]Sd-1 [g]Sd-2									
Non-specific lipid transfer protein or allergens			[i]Mal d 3		[h]Mal d 1			[i]Mal d 2 [i]Mal d 4			[i]Mal d 3	[h]Mal d 1		[h]Mal d 1	

Table 1.2 continued

Trait	Linkage group														
	1	2	4	5	6	7	8	9	10	11	12	13	15	16	17
Rosy apple aphid resistance							[g]Dp-fI								
Powdery mildew resistance							[j]Pl-w			[k]Pl-2	[l]Pl-d [m]Pl-I				
Major apple scab resistance	[n]Vf [o]Va [o]Vb	[p]Vr2 [q]Vh2 [q]Vh4 [q]Vh8 [r]Vbj [q]Vt57					[s]Vfh		[t]Vd		[u]Vg [v]Vb				[w]Vm
Woolly apple aphid resistance							[x]Er-I [x]Er-3								[x]Er-2

[a]Maliepaard *et al*., 1998, [b]Weeden *et al*.,1994, [c]Chagné *et al*., 2007, [d]Rusholme-Pilcher *et al*., 2008, [e]Costa *et al*., 2005, [f]Cevik and King, 2000, [g]Cevik and King, 2002, [h]Gao *et al.,* 2005a, [i]Gao *et al.,* 2005b, [j]Batlle and Alston, 1996 & Evans and James, 2003, [k]Gardiner *et al*., 2003, [l]James *et al*., 2004, [m]Lesemann and Dunemann, 2006, [n]Vinatzer *et al*., 2004, [o]Hemmat *et al*., 2003, [p]Patocchi *et al*., 2004, [q]Bus *et al*., 2005b, [r]Gygax *et al.,* 2004, [s]Bénaouf and Parisi, 2000, [t]Tartarini *et al*., 2004, [u]Durel *et al*., 2000, [v]Erdin *et al*., 2006, [w]Patocchi *et al*., 2005 and [x]Bus *et al*., 2008.

1.6.2 Mapping QTLs for fruit quality

In addition to the identification of several QTLs for growth characteristics, Liebhard *et al.* (2003b) assessed traits associated with fruit development and quality. In this study, one major QTL accounting for 16% of the variability was associated with fruit harvest date and eight QTLs associated with fruit weight were identified on LG 3 of 'Discovery'. However, since the expression of some of the traits varies with tree growth and development, it is necessary to assess phenotypic data over several years to establish true phenotypes and accurately identify the genetic loci involved (Gardiner *et al.*, 2007).

Some fruit quality traits like texture are complex characters. Texture is defined as a combination of firmness, juiciness and crispness, which are in turn determined by several other variables, like cell size or cell wall strength. These combinations of factors complicate the identification of QTLs controlling texture. Several QTLs linked to fruit quality traits were identified using a reference genetic map derived from 'Prima' x 'Fiesta' mapping population (Maliepaard *et al*., 1998). King *et al*. (2000) assessed fruit quality through a range of mechanical measurements and sensory parameters and identified QTLs positioned on seven linkage groups controlling firmness, stiffness and a number of sensory attributes (Table 1.3).

Two studies using apple populations with 'Fiesta' as one of the parents identified QTLs controlling fruit firmness (King *et al*., 2000; Liebhard *et al*., 2003b). The two studies demonstrated that the expression of 'Fiesta' alleles contributing to fruit firmness could be different in different genetic backgrounds, or that certain alleles may only be expressed in certain environments (Liebhard *et al*., 2003a; Gardiner *et al.,* 2007).

A recent study by Kenis and Keulemans (2008) assessed the inheritance of fruit quality traits using a population derived from ‘Telamon’ and ‘Braeburn’ over two successive seasons. A total of 74 QTLs were identified as controlling the major fruit physiological traits including fruit height, diameter, weight and stiffness, flesh firmness, rate of flesh browning, acidity, the ºBrix content and harvest date. As only 26 out of the 74 QTLs identified were detected in both years of investigation, the study demonstrated that QTLs are influenced by the environmental conditions and thus are not stable (Kenis and Keulemans, 2008). Although direct comparisons between studies are difficult as they use different markers and trait evaluation protocols, they provide insights on the QTLs or genes controlling fruit quality traits.

Table 1.3: Quantitative trait loci controlling fruit quality traits mapped in *Malus x domestica* Borkh. by genetic mapping.

Trait	**Linkage groups**															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Fruit flesh firmness	X[b, h]					X[a]		X[b, h]		X[b, h]	X[a]	X[a]		X[a]		
Fruit weight	X[a]		X[a]	X[c]		X[a]		X[a]		X[a]		X[a]			X[a]	X[a]
Malic acid content																X[e]
Fruit acidity								X[a]								X[a]
Sugar content			X[a]			X[a]		X[a]	X[a]						X[a]	
Number of fruit					X[a]										X[a]	X[a]
Hardness										X[b, h]						
Crispness	X[b, h]				X[b, h]					X[b, h]		X[b, h]	X[b, h]			X[b, h]
Juiciness	X[b, h]											X[b, h]				X[b]
Granularity		X[b, h]														
Sponginess	X[b, h]				X[b, h]	X[b, h]										X[b, h]
Compression	X[c]					X[c]		X[c]				X[c]			X[c]	

Table 1.3 continued

Trait	Linkage groups															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Fruit taste: overall liking												X[b]				X[b]
Fruit cortex colour, red foliage									X[f]							
Skin colour									X[e]							X[d]
Fruit ripening										X[g]					X[g]	
Slow breakdown	X[b. h]															
Allergens				X[k]		X[i]			X[j]			X[k]	X[i]			X[i]

[a]Liebhard *et al*., 2003b, [b]King *et al*., 2000, [c]King *et al*., 2001, [d]Weeden *et al*.,1994, [e]Maliepaard *et al*., 1998, [f]Chagné *et al*., 2007, [g]Costa *et al*., 2005, [h]Maliepaard *et al*., 2001, [i]Gao *et al.,* 2005a, [j]Gao *et al.,* 2005b and [k]Gao *et al.,* 2005c.

1.7 PROTEOMICS

The term ‘proteome’ first used in 1995, describes the protein complement of a genome (Blackstock and Weir, 1999). Proteomics can be defined as the large-scale study of protein expression, interaction, structure and functions (Ndimba and Thomas, 2008). Proteomics covers the systematic analysis of the proteins expressed by a genome, from the identification of their primary amino-acid sequence to the determination of their relative amounts, their state of modification and association with other proteins or molecules like carbohydrates (Barbier-Brygoo and Joyard, 2004). Therefore, a proteome study is expected to represent a comprehensive survey of all proteins expressed at a given time and at a defined physiological state in an organism of interest (Aebersold and Mann, 2003; Barbier-Brygoo and Joyard, 2004). The evaluation of the protein status of cell type, tissues, organs and whole organism, is an important strategy to address complex biological questions, like linking phenotype and genotype and identifying biomarkers (Pirondini *et al*., 2006).

Contrary to the genome, which is a static entity, the proteome is a dynamic entity changing during development and in response to internal or external stimuli. Until recently, protein expression patterns were estimated through transcriptomics, the large-scale analysis of mRNA expression (Ndimba and Thomas, 2008). However, only proteomics takes into consideration all the changes, including post-transcriptional, translational, post-translational modifications and interactions that can modify proteins, to obtain a global, integrated view of the cellular processes and networks (Aebersold and Mann, 2003; Ndimba and Thomas, 2008). Proteomics studies cover expression proteomics, which is the study of global changes in protein expression, and cell-map proteomics, which is the systematic study of protein-protein interactions through the

isolation of protein complexes (Blackstock and Weir, 1999). In plants, examples of comparative proteomic studies include the search for specific biomarkers involved in plant tolerance to water deficit (Riccardi *et al.*, 1998) or the comparison of *Arabidopsis thaliana* ecotypes during exposure to abiotic stress (Chevalier *et al.*, 2004).

1.7.1 Gel-based and non-gel based proteomics

Technically, proteomics can be divided into gel based and the non-gel based proteomics (Monteoliva and Albar, 2004). Gel-based proteomics involves the separation of complex protein mixtures with gels. Following isolation, protein mixtures can be separated using either one-dimensional polyacrylamide gel electrophoresis (1D-PAGE), which separates proteins according to molecular weight, or two-dimensional PAGE (2D-PAGE), which separates proteins according to their isoelectric point (pI) and subsequently their molecular weight (Klose, 1975; O'Farrell, 1975). pI is the pH at which a protein carries no net electrical charge (Beranova-Giorgianni, 2003; Garfin, 2003). The separation by 2D-PAGE allows for the resolution of proteins differing by a single charge, thereby allowing the separation of isoforms and identifying post-translational modifications like phosphorylation or glycosylation (Thomas, 2008).

Proteins are then stained for visualisation using for example Coomassie Brilliant Blue (CBB) or silver nitrate for total protein detection. On the other hand, fluorescent dyes like the Pro-Q Diamond or Emerald are also available for the detection of a specific subset of proteins like phosphoproteins or glycoproteins, respectively (Agrawal and Thelen, 2006). After gel staining, proteins are quantified by densitometry, whereby differences in staining intensity are expected to correlate with differences in protein abundance, for comparative analysis and the detection of differentially expressed

proteins. Further, the bands and/or spots of interest can be proteolysed using trypsin and identified by peptide mass fingerprinting using mass spectrometry, typically, Matrix-Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF, Xu *et al.*, 2006) or by protein sequencing using tandem mass spectrometry. Finally, proteins are identified through database searching using the MASCOT search engine (Hood *et al.*, 2004; Bhadauria *et al.*, 2007). A typical workflow of the general procedure of protein analysis from sample collection to protein identification is illustrated in Figure 1.7.

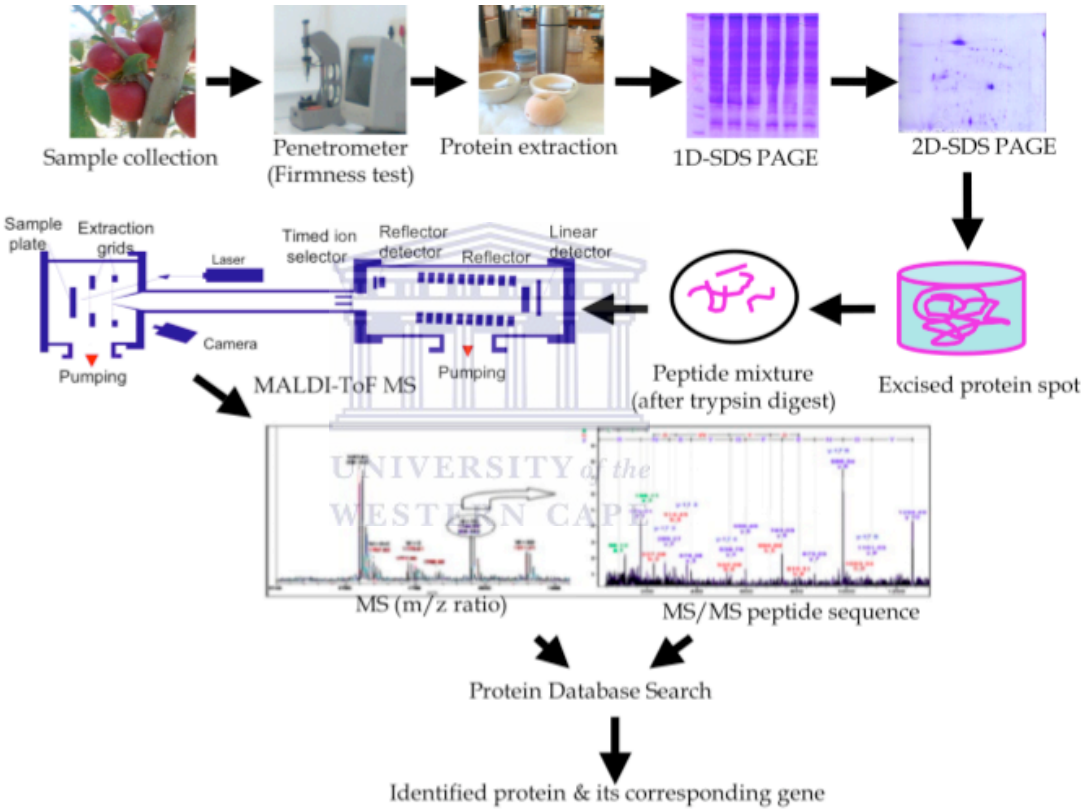


Figure 1.7: Typical gel-based proteomics workflow (modified from Bhadauria *et al.*, 2007).

Gel-based proteomics includes the collection of samples and preparation of protein extracts, the separation of proteins by 1D-PAGE and/or 2D-PAGE. After analysis, protein spots of interest are excised out of the gel and in-gel digested typically using trypsin. The digested proteins are identified either by peptide mass fingerprinting using mass spectrometry or by protein sequencing using tandem mass spectrometry. Finally, proteins are identified through database searching, typically using the MASCOT search engine.

In non-gel-based proteomics, multi-dimensional protein identification techniques (MudPIT) and liquid chromatography (LC)-MS methods are implemented, whereby complex mixtures of protein samples are first enzymatically digested, then separated using high performance separation techniques in liquid prior to their identification by MS (McCormack *et al.*, 1997; McDonald *et al.*, 2002). Non-gel based approaches have the advantage of being less affected by the size and solubility of protein extracts, while the gel-based approaches are favoured for their unsurpassed protein resolving power and relative affordability (Blackstock and Weir, 1999).

Since its introduction in 1975 (O'Farrell and Klose, 1975), 2D-PAGE has remained an unchallenged method of preference for the efficient separation of complex protein mixtures and the study of proteome dynamics. Nevertheless, this strategy is time-consuming for the determination of differentially expressed proteins between two samples (control *versus* experiment). However, the introduction of the difference gel electrophoresis (DIGE) technique, which is based on fluorescent dyes to label protein samples prior to 2D-PAGE, allows multiple samples to be co-separated and visualized on a single 2D-PAGE gel (Unlu *et al.*, 1997; Tonge *et al.*, 2001). The images can then be merged and differential protein expression between the control and the experiment(s) determined using image analysis softwares like the PDQuest software (Bio-Rad Laboratories, Ltd, UK).

1.7.2 Fruit proteomics

The characterisation of fruit, which comprises both skin and flesh tissues, is apparently an essential parameter for understanding fruit development. Thus, characterisation of proteins isolated from skin tissue is evidently essential for comprehending fruit ripening

in fruits like grapes (Okamoto *et al*., 2004; Deytieux *et al*., 2007), and colour differences between cultivars, for example in apples, 'Golden Delicious' and 'Dietrich' (Mathye, unpublished data). In addition, fruit skin proteomic analysis is crucial to understand metabolic changes occurring during the fruit ripening process such as increase in fruit size, softening and cell expansion (Coombe and McCarthy, 2000).

In recent years, proteome analysis has been successfully applied to fruit pulp tissues of banana (Peumans *et al*., 2000; Wang *et al*., 2006), ripe grape mesocarp (Sarry *et al*., 2004), tomato (Rocco *et al.*, 2006; Wang *et al*., 2006), strawberry (Hjerno *et al.*, 2006), citrus fruit (Katz *et al.,* 2007), browning induction in 'Conference' pears (Pedreschi *et al.*, 2007; 2008) and apple (Barraclough *et al.,* 2004; Guarino *et al*., 2007). In apple, the studies reported the proteomic analyses of pseudocarp tissue (Guarino *et al*., 2007). In this study, total protein samples from the fruit pseudocarp of 'Annurca' were electrophoretically separated and protein spots identified by a combined MALDI-TOF MS and liquid chromatography electrospray ionisation (LC-ESI) mass spectrometry approach. A total of 44 spots were identified and associated to 28 different genes. In addition, energy metabolism was the functional class that was most represented in their map. Besides the difficulties in the identification of pseudocarp proteins related to the poorly characterised *Malus* genome, this study illustrates the importance of proteomic studies in providing insights into the link between proteomics and functional genomics. This relation is determined through linking the identified proteins to their associated genes. This information is crucial in the present study where the quantitative nature of the genomic regions associated with fruit firmness were confirmed using proteomic approach.

AIMS AND OBJECTIVES OF THE STUDY

Increasing demands for the improvement of apple fruit quality has given rise to the research for the identification of QTLs linked to variety of fruit quality traits. Identification of these traits can then facilitate marker-assisted selection of seedlings possessing such traits early with the plant development.

While several studies for the identification of fruit quality related QTLs have been carried out using the 'Prima' x 'Fiesta' (King *et al*., 2000), 'Fiesta' x 'Discovery' (Liebhard *et al*., 2003a) and 'Telamon' x 'Braeburn' (Kenis *et al*., 2006; Kenis and Keulemans, 2008) apple population, no information has yet been generated for one of the economically valuable 'Golden Delicious' x 'Dietrich' population. In addition, only a few proteomics studies have been reported on apples using pseudocarp tissue from a cultivar 'Annurca' (Guarino *et al*., 2007) and flower bud tissues (Cao *et al.*, 2008). However, the entire fruit pulp proteome and the molecular mechanism influencing fruit firmness have not been characterised.

Accordingly, the aim of this study was to construct a genetic linkage map of 'Golden Delicious' x 'Dietrich' population for the identification of QTLs associated with fruit quality traits and then examine the apple fruit pulp proteome with specific focus on fruit firmness. This would facilitate the detection of biomarkers involved in fruit firmness.

To achieve this aim, the objectives below were formulated;

- To collect phenotypic data on the apple fruit firmness, crispness, juiciness, form, colour, stripness, russeting, acidity and size from the mapping population 'Golden Delicious' x 'Dietrich',

- To extract genomic DNA from leaves, test SSR markers and to construct a genetic linkage map of 'Golden Delicious' x 'Dietrich' using the *JoinMap*® V4.0 software,
- To identify QTLs controlling firmness, crispness, juiciness, form, colour, stripness, russeting, acidity and size using the *MapQTL*®V4.0 software,
- To identify possible candidate markers for MAS.
- To extract total soluble protein (TSP) from apple fruit pulp of 'Golden Delicious' x 'Dietrich' population,
- To generate a 2D-PAGE proteome maps for both firm (high firmness) and soft (low firmness),
- To identify proteins visualised in the proteome maps using MALDI-TOF MS and then characterise the positively identified proteins,
- To comparatively analyse the proteome maps of high and low firmness apples using the PDQuest software.
- To examine the molecular mechanisms influencing and/or influenced by fruit firmness.

CHAPTER 2

MATERIALS AND METHODS

2.1 CHEMICALS

Acetone	Merck
Acetonitrile	Merck
Acrylamide: Bis-Acrylamide (40%, 19:1)	Promega
Acrylamide: Bis-Acrylamide (40%, 37:5:1)	Bio-Rad
Agarose D1 LE	Promega
Alkylamidosulphobetaine (ASB)-14	Bio-Rad
Ampholytes (pH 3-10)	Bio-Rad
Ammonium bicarbonate	Merck
Ammonium persulphate (APS)	Merck
β-mercaptoethanol	Fermentas
Boric acid	Merck
Bovine Serum Albumin (BSA)	Roche
Bradford protein dye	Bio-Rad
Bromophenol blue	Sigma
Buffer saturated phenol	Invitrogen
3-[3-cholamidopropyl]-dimethyl-ammonio-1-propane sulphonate (CHAPS)	Sigma
Sequazyme peptide mass standard kit	Applied Biosystems
Chloroform	BDH chemicals
Coomassie brilliant blue R-250 (CBB)	Bio-Rad
Dithiothreitol (DTT)	Fermentas

Ethylenediamine tetraacetic Acid (EDTA)	Merck
Ethanol	BDH chemicals
Ethidium bromide	Sigma
Formaldehyde solution	Riedel-de Haën
Formamide	Merck
Formic acid	Merck
Glacial acetic acid	Merck
Glycerol	Merck
Glycine	Merck
Hydrochloric acid (HCl)	Merck
Iodoacetamide	Bio-Rad
Iso–propyl alcohol	BDH chemicals
LIZ500™	Applied Biosystems
Methanol	Merck
Megaplex PCR mix	Qiagen
Mineral oil	GE Healthcare
Potassium acetate	Sigma
Protein markers	Fermentas
Propan-2-ol	Merck
Proteinase K	Roche
RNase A	Roche
Silver nitrate	Merck
Sodium acetate	Merck
Sodium borohydride	Saarchem
Sodium chloride (NaCl)	Merck

Sodium dodecyl sulphate (SDS)	Bio-Rad
Sodium hydroxide (NaOH)	BDH chemicals
Sodium sulphite	Merck
Spermidine	Sigma
Tributyl phosphine (TBP) reducing agent	Bio-Rad
N,N,N′,N′-Tetra methylethylenediamine (TEMED)	Promega
Thiourea	Sigma
Trichloroacetic acid (TCA)	Merck
Tris (hydromethyl) aminomethane (Tris)	Merck
Trypsin	Promega
Urea	Merck

All the chemicals used were of AnalaR or equivalent grade. All buffers were prepared with dH_2O purified using the Maxima water purification system (Elga, Wycombe, England).

2.2 BUFFERS AND SOLUTIONS

80% Acetone	80% (v/v) acetone, 20% (v/v) dH_2O
Agarose sealing solution	0.5% (w/v) agarose, 0.002% (w/v) bromophenol blue dissolved in 1X SDS running buffer
10% APS	10% (w/v) APS in dH_2O
2X Cetyl trimethyl ammonium bromide (CTAB)	100 mM Tris-HCl, 20 mM EDTA, 4 M NaCl, 2% (w/v) PVP40, 0.6 M sodium sulphite (pH 8.0)
CBB stock solution	1.25% (w/v) CBB in dH_2O

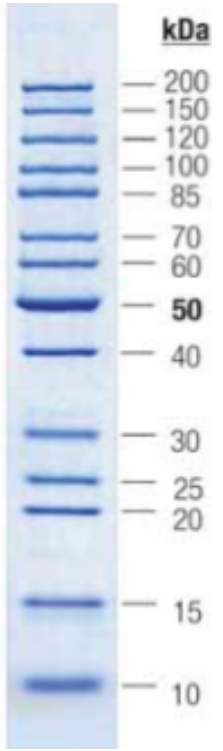
CBB staining solutions:

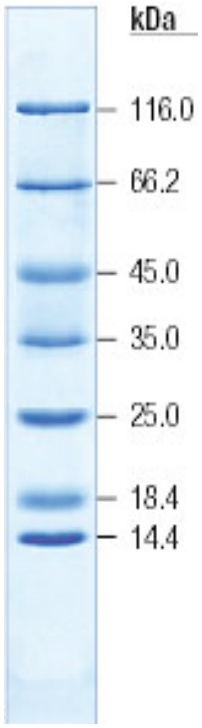
Staining solution I:	10% (v/v) glacial acetic acid, 2% (w/v) CBB stock solution, 25% (v/v) propan-2-ol
Staining solution II:	10% (v/v) glacial acetic acid, 0.25% (v/v) CBB stock solution, 10% (v/v) propan-2-ol
Staining solution III:	10% (v/v) glacial acetic acid, 0.25% (v/v) CBB stock solution
Destaining solution:	5% (v/v) methanol, 10% (v/v) glacial acetic acid, 1% (v/v) glycerol
DNA loading buffer	0.25% (w/v) bromophenol blue, 0.25% (v/v) xylene cyanol, 30% (v/v) glycerol
Equilibration buffer I	6 M urea, 50 mM Tris-HCl (pH 8.8), 2% (w/v) SDS, 30% (v/v) glycerol, 1% (w/v) DTT
Equilibration buffer II	6 M urea, 50 mM Tris-HCl (pH 8.8), 2% (w/v) SDS, 30% (v/v) glycerol, 2.5% (w/v) iodoacetamide
70% Ethanol	70% (v/v) absolute ethanol in dH_2O
Extraction/Lysis buffer	7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 40 mM Tris (pH 8.5), 0.001% (v/v) TBP
1% Formic acid	1% (v/v) formic acid in 25 mM ammonium bicarbonate
IEF rehydration buffer	7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 40 mM Tris (pH 8.5), 0.001% (v/v) TBP, 1% (v/v) ASB-14, 0.2% (w/v) DTT.
IPG strip equilibration buffer I	6 M urea, 50 mM Tris-HCl (pH 8.8), 2% (w/v)

	SDS, 30% (v/v) glycerol, 1% (w/v) DTT.
IPG strip Equilibration buffer II	6 M urea, 50 mM Tris-HCl (pH 8.8), 2% (w/v) SDS, 30% (v/v) glycerol, 2.5% (w/v) iodoacetamide.
PCR master mix	1X *Taq* polymerase buffer, 50 mM $MgCl_2$; 0.2 mM each of dATP, dCTP, dGTP, and dTTP, one unit of *Taq* DNA polymerase
Phenol:chloroform:isoamyl alcohol	25 (v/v) Tris-buffered phenol, 24 (v/v) chloroform, 1 (v/v) isoamyl alcohol
Potassium acetate	3 M potassium acetate, adjusted to pH 4.8 with glacial acetic acid
4X Resolving gel Buffer	0.75 M Tris adjusted to pH 8.8 with HCl
2X RNase Buffer	20 mM Tris-HCl, 10 mM EDTA, 0.6 M NaCl (pH 8.0), dissolved in DEPC treated H_2O
SDS-PAGE loading dye	50 mM Tris (pH 6.8), 10 mM DTT, 5% SDS (w/v), 0.1% (w/v) bromophenol blue, 20% (v/v) glycerol
1X SDS-running buffer	0.1% (w/v) SDS, 0.025 M Tris (pH 8.8), 0.192 M glycine
10% SDS solution	10% (w/v) SDS in dH_2O
Silver-staining solutions	
Solution A:	0.1% (w/v) silver nitrate in dH_2O, kept in the dark
Solution B:	1.5% (w/v) NaOH, 0.01% (w/v) $NaBH_4$, 0.004% (v/v) formaldehyde solution, kept in the dark

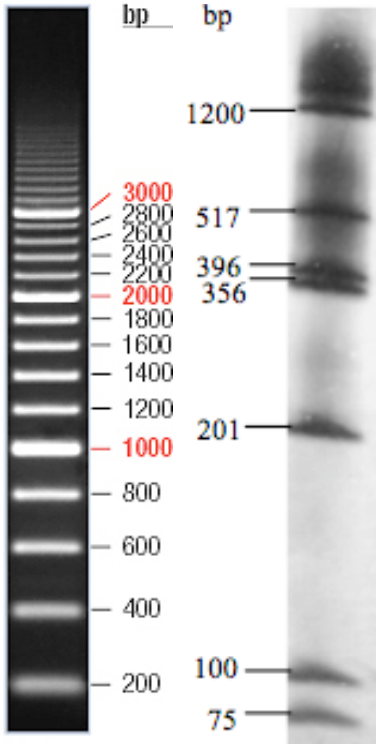
Sodium acetate	3 M sodium acetate (w/v), adjusted to pH 5.5 with glacial acetic acid
2X Stacking gel buffer	0.25 M Tris adjusted to pH 6.8 with HCl
10X TBE	0.9 M Tris, 0.89 M boric acid, 0.032 M EDTA, diluted 20 times prior to electrophoresis of agarose gels
10X TE	100 mM Tris-HCl (pH 7.5), 10 mM EDTA, diluted 10 times for resuspending DNA samples and primers
TE-RNAse	0.0625 $mg.ml^{-1}$ RNase in 1X TE
TCA	10% (w/v) TCA in acetone
Tris base	1.5 M Tris, pH 8.8
Trypsin solution	10 $\mu g.ml^{-1}$ trypsin in 25 mM ammonium bicarbonate

Protein ladders





DNA ladders (Fermentas Life Sciences, SA)



2.3 PLANT MATERIAL

2.3.1 Apple population

An apple population consisting of 248 individuals (trees) from a cross of 'Golden Delicious' x 'Dietrich' was used in this study (Figure 2.1). The plants were grown at the Drostersnes experimental farm of the Agricultural Research Council (Vyeboem, Western Cape, South Africa). Cultivar (cv.) 'Golden Delicious' is characterised by green to golden soft fruits, while cv. 'Dietrich' is characterised by red firm fruits (I. Labuschagné, personal communication). In this study, 'Golden Delicious' was used as a female parent in this cross and 'Dietrich' as the male parent. Although it is common practice to always indicate the male parent to the left of the 'x' sign, the cross, 'Golden Delicious' x 'Dietrich' was labelled the other way round throughout this study. This was done to maintain consistence with the records from the breeders. 'Golden Delicious' was referred to as parent 1 and 'Dietrich' as parent 2.

Figure 2.1: Fruits collected from the 'Golden Delicious' x 'Dietrich' population

The 'Golden Delicious' x 'Dietrich' population is generally characterised by green, greenish-red to red and sometimes deep red firm fruits.

2.3.2 Leaf sample

Five young fully developed apple leaves were harvested in February 2007 from each of the 248 individual trees of the 'Golden Delicious' x 'Dietrich' population (section

2.3.1). The leaves collected from each of the individuals were packed in plastic bags, chilled on ice for approximately 1 h during transportation and stored at -20ºC.

2.3.3 Apple fruit material

Mature palatable (ripe) fruits from the 'Golden Delicious' x 'Dietrich' population (section 2.3.1) were periodically collected (from March to April) from each of the individuals bearing fruits for three years between 2005 and 2007. Freshly harvested fruits were stored for two days at room temperature; a standard procedure followed prior to marketing the apples (I. Labuschagné, personal communication) and then stored at 4°C prior to fruit quality sensory tests (section 2.4.1). These samples were used for proteomics work.

2.4 PHYSIOLOGICAL METHODS

2.4.1 Phenotypic data collection

Once harvested (section 2.3.2), fruit firmness, juiciness, crispness, acidity, size, form, colour, stripness and russeting were assessed on every fruit. These measurements were performed every experimental year from 2005 to 2007.

2.4.2 Fruit trait assessment

Apples were harvested up to twice a week as they reached maturity in 2005, 2006 and 2007. All collected apples were assayed immediately after harvest and analysed over a period of 12 weeks. Of each tree, 10 of the collected apples were analysed at a time for different phenotypic traits. Colour was assessed visually and then measured using a colorimeter (Minolta CR-400, Minolta Co., Ltd, Japan). With the exception of colour,

all the other physiological traits were analysed biweekly. Sensory analyses, performed by trained technicians, were carried out to measure russeting, stripness, form, crispness and acidity. These were evaluated on a scale of 1-10. Size of the fruits was determined by measuring the diameter of each fruit using a caliper. Firmness was measured as described in section 2.4.3. Penetrated fruits (after firmness test) were pressed with a hand-operated squeezer and the juice collected. Sugar content (Juiciness) was determined with a digital refractometer (model PR-1, Atago Co. Ltd, Tokyo, Japan), by adding a few drops of apple juice onto the lens of the measuring device. Refractometry results were recorded in °Brix, which is equal to g sugar per 100 ml juice at 20°C.

2.4.3 Fruit pulp firmness measurement

Twenty fruiting individuals from the 'Golden Delicious' x 'Dietrich' population (section 2.3.1) were selected (Table 7.1) according to preliminary results obtained by M.K. Soeker (unpublished data) between 2005 and 2007. Fruits from the selected individuals were collected between the 10th March and 16th April 2008 for pulp firmness test using a penetrometer (Magness-Taylor pressure tester, Italy). The penetrometer is a device measuring the force that is applied while piercing a fruit. Firmness was measured taking into consideration the size and position of the fruit on the parent tree (Bramlage, 1983). Prior to measuring, the skin of the fruits was peeled off using a commercial peeler for consistency. The penetrometer was then applied on three different regions of every fruit, using three apples for every individual to limit variation. The fruits were then snap frozen in liquid nitrogen and stored at -80°C prior to their use for proteomic analyses (section 2.6).

2.5 GENOMICS METHODS

2.5.1 DNA extraction

Genomic DNA was extracted from leaves (section 2.3.2) using a modified CTAB DNA extraction protocol (Murray and Thompson, 1980). Prior to extraction, the 2X CTAB solution was incubated for 20 min at 62ºC. Approximately 0.1 g of leaf material was ground to a fine powder in liquid nitrogen with a sterile mortar and pestle and transferred into a 2 ml tube. The sample was homogenised by mixing 1 ml of pre-warmed 2X CTAB containing 0.2% (v/v) β-mercaptoethanol and was incubated for 30 min at 62ºC. Then, 10 $\mu g.\mu l^{-1}$ of proteinase K was added and the sample was incubated for 30 min at 37ºC. A 1 ml of chloroform–isoamyl alcohol solution (24:1) was added, mixed by inversion and centrifuged at 10,000 × *g* for 10 min at room temperature. Once transferred to a fresh 2 ml tube, the aqueous phase was mixed with 2.5 volumes of ice-cold absolute ethanol and 0.1 volume of 3 M ammonium acetate (pH 5.2), mixed by inversion, incubated for 30 min at -20ºC and centrifuged at 16,000 × *g* for 10 min at room temperature. After discarding the supernatant, the pellet was washed with 500 µl of 70% (v/v) ice-cold ethanol, and collected by centrifugation at 16,000 × *g* for 2 min at room temperature. This step was repeated once more. The pellet was then air-dried, resuspended in 200 µl TE-RNase (section 2.2) and incubated for 30 min at 37ºC. The DNA was precipitated with 2.5 volumes of ice-cold absolute ethanol and 0.1 volume of 3 M ammonium acetate (pH 5.2), mixed by inversion, incubated for 60 min at -20ºC and collected by centrifugation at 16,000 × *g* for 10 min at room temperature. The DNA pellet was washed with 500 µl ice-cold 70% (v/v) ethanol and collected by centrifugation at 16,000 × *g* for 2 min at room temperature. The pellet was then air-dried and finally resuspended in 100 µl of TE (section 2.2). The DNA samples were

then quantified using a Nanodrop™ spectrophotometer (GlobalSpec, Wilmington, DE, USA) and diluted to the final concentration of 5 ng.µl^{-1}.

2.5.2 Microsatellite primer design

Apple microsatellites (SSRs) were identified from the apple EST data (Korban *et al.*, 2005; Naik *et al.*, 2006; Newcomb *et al.*, 2006) and designed using the Tandem Repeats Finder database system (Benson, 1999) software (http://tandem.bu.edu/). Both published apple and pear SSRs primer pairs (Guilford *et al.*, 1997; Gianfranceschi *et al.*, 1998; Hokanson *et al.*, 1998; Liebhard *et al.*, 2002; Yamamoto *et al.*, 2002a; Yamamoto *et al.*, 2002b; Liebhard *et al.*, 2003b; Vinatzer *et al.*, 2004; Silfverberg-Dilworth *et al.*, 2006) and those newly designed primers (Table 2.1) were synthesised at Applied Biosystems (ABI, Forster City, California, USA). The primer derived from the sequence closest to the SSR was fluorescently labelled using either 6-Fam (blue), Ned (yellow), Vic (green) or Pet (red) dye for detection using the Genetic Analyser.

Table 2.1: List of primers and megaplexes grouped to the megaplex mix.

The table lists the primer number, accession number, primer nucleotide sequence (forward and reverse nucleotide sequence), expected amplicon size range, predicted linkage group of the marker and the references of the primer design.

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 1						
A341	CO723148	PET-CGG TGG TGA CTA GTA TCA GC	TAT GGA GGA AGA AAC TGA GGC	131	No position	**
A236	CH01e09b	GAG AAA CCG TTT GAT TAC AGC	PET-CTC CAT CCC CAA TCA CAC C	242	17	**
A425	CO756781	PET-ATA AGT TTA GGC TCA TCT GCC	AAA CCC ATC CCA CTT AAG GC	331	No position	***
A166	CH04g12	PET-CAC CGA TGG TGT CAA CTT GT	CAA CAA AAT GTG ATC GCC AC	141-186	No position	Liebhard *et al.*, 2002
A335	CO052033	NED-TTG CCA ATC CGC ATT CGC C	TGA GGT TCC CGC CCT TGC	192	5	**
A329	CN496002	NED-AGC AGC AGC TAG GCT AGA GC	AAA TTG CCT TGC CAG ATT AGC	227	5	**
A318	CN580227	NED-GAC GTA AAA TCC CTA ATT CCC	TCA TCC CAG TCG TCT TCC C	246	No position	***
A227	CN493171	NED-TCT TAC TTC GTC GGT GGA CC	TGT GTG GCT ATT ACC TGA GG	345	No position	**
A107	CH04d02	NED-CGT ACG CTG CTT CTT TTG CT	CTA TCC ACC ACC CGT CAA CT	118-146	12	Liebhard *et al.*, 2002
A320	CN580637	6-FAM-ACA ACA GCT GAC GAA CAA GC	CTA CTC GTC GAA GTA CGC C	213	15	***
A114	CH03b06	FAM-GCA TCC TTG AAT GAG GTT CAC T	CCA ATC ACC AAA TCA ATG TCA C	111-131	15	Liebhard *et al.*, 2002
A120	CH05e04	FAM-AAG GAG AAG ACC GTG TGA AAT C	CAT GGA TAA GGC ATA GTC AGG A	153-234	16	Liebhard *et al.*, 2002
A288	CN909118	6-FAM-CTG AGG ACT CTT CTA CCC C	CAG CAG CCA CAG AAT CAG C	258	No position	***
A281	CN870040	VIC-CCT CAG CAT CAT CAA CCC C	GGA AAT GCG ATT TCG AAC CC	310	16	***
A294	CN946851	VIC-AAT GAC TCA AGC GAT CAG GG	CCG ATC CAA GTA GTT AAC GG	361	No position	***
A93	CH05f06	VIC-TTA GAT CCG GTC ACT CTC CAC T	TGG AGG AAG ACG AAG AAG AAA G	166-184	5	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 2						
A417	CH04f04	PET-GTC GGT ACA AAC TCA GGA CC	CGA CGT TCG ATC TTC CTC TC	229	4	Liebhard *et al.*, 2002
A346	CO753022	PET-CTG AGT CTT TGT TTT TGC TCG	GCT CCG CCT CTC TGT ACC	355	13	**
A139	CH02a04	PET-GAA ACA GGC GCC ATT ATT TG	AAA GGA GAC GTT GCA AGT GG	66-112	2, 7	Liebhard *et al.*, 2002
A128	CH01b121	PET-CGC ATG CTG ACA TGT TGA AT	CGG TGA GCC CTC TTA TGT GA	125-178	4, 12	Liebhard *et al.*, 2002
A244	CN947446	VIC-CCG TTA CAG CTA TCC AAA CC	ATA ATG GCC ATT CTG TTC AGC	186	15	**
A298	CN943252	TCC CAC TGA CAC TAT CAC C	VIC-TGC AGG AAA TGA GAA TGC GC	198	No position	***
A327	CN490324	ATA GAG AGG TAG AGG ACT GG	VIC-TTC GCC CAG TGT AAC ATT GG	230	17	**
A60	CH03g04	VIC-ATG TCC AAT GTA GAC ACG CAA C	TTG AAG ATG GCC TAA CCT TGT T	122-144	14	Liebhard *et al.*, 2002
A56	CH05f04	VIC-GAT GAT GGT GCT CTC GGT TAT T	TTA TGT TGG GTA ATG TCT TCC G	160-172	13	Liebhard *et al.*, 2002
A15	28f4	VIC-TGC CTC CCT TAT ATA GCT AC	TGA GGA CGG TGA GAT TTG	112	No position	Celton, pers comm.
A279	CN887525	NED-TAG TAG CTA CAC ACT CTT TCC	GCA TTG CCT TGA GCT CCA G	217	5	***
A398	CN490644	NED-ATC TCA CAC CTC AGC AGT GA	CTT CTG CCC AAT TCA AGA CC	264	10	**
A293	CN944444	NED-TAG TGC AAG TAC TGG GGC C	CAT CGA TAG AAT AGG ACG GC	383	No position	***
A80	MS02a01	NED-CTC CTA CAT TGA CAT TGC AT	TAG ACA TTT GAT GAG ACT G	170-194	10	Liebhard *et al.*, 2002
A219	CN580620	6-FAM-TGC GGT CAA CGA TGT CTT CG	AAG GTA CAA GCC CGC AAA GG	383	12	**
A59	CH03d08	6-FAM-CAT CAG TCT CTT GCA CTG GAA A	TAG GGC TAG GGA GAG ATG ATG A	129-161	14	Liebhard *et al.*, 2002
A43	CH04e03	6-FAM-TTG AAG ATG TTT GGC TGT GC	TGC ATG TCT GTC TCC TCC AT	179-222	5	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 3						
A369	CO865608	CAA CAA GTG TGC CTC TGT GG	PET-AGC AAG CAA CAG ATC AAG CC	159	1	Lerato, 2004
A177	CH05h05	PET-ACA TGT CAC TCC TAC GCG G	GTG CAG TGA TTA GCA TTG CTG T	174	13	Liebhard *et al.*, 2002
A235	CN881550	ATC CAA ACA ACC CCA TTG CG	PET-AGT CGA TGT TGA ACG CTC CA	355	No position	**
A125	CH02g01	PET-GAT GAC GTC GGC AGG TAA AG	CAA CCA ACA GCT CTG CAA TC	91-121	No position	Liebhard *et al.*, 2002
A207	CN495433	VIC-ACA AGA GCA GCA GCA TTT CG	GTA GCG TGT TTC AGG CAG TC	263	No position	**
A187	CN490566	VIC-AGC GCA ATG GCG TTC TAG G	AGC TGC GCT ATC TTC TCA GC	336	No position	**
A66	MS01a05	VIC-GGA AGG AAC ATG CAG ACT	TGA TGT TTC ATC TTT ACA	158-176	14	Liebhard *et al.*, 2002
A112	CH05c04	VIC-CCT TCG TTA TCT TCC TTG CAT T	GAG CTT AAG AAT AAG AGA AGG GG	186-258	13	Liebhard *et al.*, 2002
A234	CN938125	NED-GCC TTC ATC CCC CCT TGA	GGT GTA TAG GAA TCT TGG AG	353	17	**
A452	CO900827	NED-ACC TTG GTG GCC AAG TAG C	CTT GCG TAT CAA AGC TGC CG	446	No position	***
A71	CH01h011	NED-GAA AGA CTT GCA GTG GGA GC	GGA GTG GGT TTG AGA AGG TT	114-134	17	Liebhard *et al.*, 2002
A74	CH02a10	NED-ATG CCA ATG CAT GAG ACA AA	ACA CGC AGC TGA AAC ACT TG	143-177	10	Liebhard *et al.*, 2002
A113	CH01d08	NED-CTC CGC CGC TAT AAC ACT TC	TAC TCT GGA GGG TAT GTC AAA G	238-290	15	Liebhard *et al.*, 2002
A345	CO755814	6-FAM-AAC ATC AAG ACA GAG AAG AGC	CGT CTT CTT CAC AAA CTC CG	261	No position	**
A119	CH05a04	NED-GAA GCG AAT TTT GCA CGA AT	GCT TTT GTT TCA TTG AAT CCC C	159-189	16	Liebhard *et al.*, 2002
A182	CN445253	6-FAM-TGC AAG AAT CAT CCA CTT CC	TTG GAC CTG TGA GGA CTC C	315	No position	**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 4						
A390	CH05a03	PET-CGG CTG AGC ATG GTT ACT TC	TGA TCG TTG TGA AAG CTC CA	300	No position	Liebhard *et al.*, 2002
A158	CH04d08	PET-AAT TCC ACA TTC ACG CAT CT	TTG AAA GAC GGA AAC GAT CA	116-142	No position	Liebhard *et al.*, 2002
A148	CH03a03	PET-CCA CTT GGC AAT GAC TCC TC	ACC TTA CCG CCA ATG TGA AG	154-182	No position	Liebhard *et al.*, 2002
A137	CH01e121	PET-AAA CTG AAG CCA TGA GGG C	TTC CAA TTC ACA TGA GGC TG	210-224	No position	Liebhard *et al.*, 2002
A231	CN580271	VIC-TCT GGC TCT CAT CGG TTT GC	TCG ATG CCC TTG TAA CGC C	206	No position	**
A29	AT000141	VIC-GAA ATA AAC ACC GAG TAA ACA G	TGC TAT CTG GTT TTC TTT TAG C	56-100	No position	Liebhard *et al.*, 2002
A92	CH03a09	VIC-GCC AGG TGT GAC TCC TTC TC	CTG CAG CTG CTG AAA CTG G	125-143	5	Liebhard *et al.*, 2002
A122	CH04c06	VIC-GCT GCT GCT GCT TCT AGG TT	GCT TGG AAA AGG TCA CTT GC	155-186	17	Liebhard *et al.*, 2002
A200	CN493925	NED-TCT CCT TCA CTT CCC ATT CC	TGG TGA TGG CAT ACA CAT CC	406	No position	**
A64	CH05e05	NED-TCC TAG CGA TAG CTT GTG AGA G	GAA ACC ACC AAA CCG TTA CAA T	138-160	14	Liebhard *et al.*, 2002
A109	CH05d11	NED-CAC AAC CTG ATA TCC GGG AC	GAG AAG GTC GTA CAT TCC TCA A	171-211	12	Liebhard *et al.*, 2002
A67	CH02c09	NED-TTA TGT ACC AAC TTT GCT AAC CTC	AGA AGC AGC AGA GGA GGA TG	233-257	15	Liebhard *et al.*, 2002
A377	CO068842	NED-TGG TTG GAG ATG TTC CAT GG	ACC AGC TAG ATT ATC TTC TGC	333	13	**
A380	CO866737	6-FAM-AGC AGC TTC CGT TTC CCT G	AAA CAA CCC ACG CTC GGA G	242	13	**
A106	CH03c02	6-FAM-TCA CTA TTT ACG GGA TCA AGC A	GTG CAG AGT CTT TGA CAA GGC	116-136	12	Liebhard *et al.*, 2002
A63	CH05d03	6-FAM-TAC CTG AAA GAG GAA GCC CT	TCA TTC CTT CTC ACA TCC ACT	152-187	14	Liebhard *et al.*, 2002
A99	CH03d02	6-FAM-AAA CTT TCA CTT TCA CCC ACG	ACT ACA TTT TTA GAT TTG TGC GTC	201-223	11	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 5						
A412	CH05g01	PET-TTT CAT TCA ACT TCA CCT CTC	CTC CTT TCC GAT TCT TCT ATT TCA	215	No position	Liebhard *et al.*, 2002
A172	CN492999	PET-ATG AGA GAG AGC TAC CTC AC	GTA CAA GTT CAG CAG TGA CC	91-143	9, 17	**
A171	CH04g09	PET-TTG TCG CAC AAG CCA GTT TA	GAA GAC TCA TGG GTG CCA TT	168-200	10	Liebhard *et al.*, 2002
A173	CH05a02	PET-GTT GCA AGA GTT GCA TGT TAG C	TTT TGA CCC CAT AAA ACC CAC	236-276	No position	Liebhard *et al.*, 2002
A401	CN544835	VIC-AGG AGA GCT TTC TGC ATT CC	AGC GCT ATC CCC AGC TGC	187	5	**
A260	CN935817	VIC-GCC TTC CAA GCG TCT TGG	TTA TCA ACA AGC GCC GTT CC	239	12	**
A331	AB162040	VIC-GGA GTG CTA TTA GCT CCT CC	TCC TTG AAT CTC AAC TCT AGG	294	12	**
A94	CH03d12	VIC-GCC CAG AAG CAA TAA GTA AAC C	ATT GCT CCA TGC ATA AAG GG	108-154	6	Liebhard *et al.*, 2002
A10	02b1	NED-CCG TGA TGA CAA AGT GCA TGA	ATG AGT TTG ATG CCC TTG GA	238	No position	Celton, pers comm.
A253	CO540769	NED-TCC TAG GGT CGG AGA GCA G	CTC AAG AAT CAC CAA CAA TGC	263	15	**
A376	CO867345	NED-TAC ATC CAC CAT GGA AAG ATC	CTG GTC GGA CAG GTT AAC G	368	No position	**
A49	CH05c07	NED-TGA TGC ATT AGG GCT TGT ACT T	GGG ATG CAT TGC TAA ATA GGA T	111-149	9	Liebhard *et al.*, 2002
A62	CH04f06	NED-GGC TCA GAG TAC TTG CAG AGG	ATC CTT AAG CGC TCT CCA CA	159-179	14	Liebhard *et al.*, 2002
A14	23g4	6-FAM-TTT CTC TCT CTT TCC CAA CTC	AGC CGC CTT GCA TTA AAT AC	88	No position	Celton, pers comm.
A217	CN579502	6-FAM-TCG TGA AGT GCC AAG TAT CG	TGG CGG ACT GCT CAA TTG C	280	No position	**
A238	CN865016	6-FAM-TTC TTC ACA CCC TTC AAT CC	AAA GCG CCT GCG ATT GCG	344	15	**
A87	CH03e03	6-FAM-GCA CAT TCT GCC TTA TCT TGG	AAA ACC CAC AAA TAG CGC C	106-216	3	Liebhard *et al.*, 2002
A73	CH01f12	6-FAM-CTC CTC CAA GCT TCA ACC AC	GCA AAA ACC ACA GGC ATA AC	145-162	10	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 6						
A323	CH04d08	PET-AAT TCC ACA TTC ACG CAT CT	TTG AAA GAC GGA AAC GAT CA	246	No position	Liebhard *et al.*, 2002
A167	CN544851	PET-TTG TCG GAT TTG TAA CCC TAG	TTC CAT ATC AGT TTG GAC ACC	111-135	8, 15	**
A165	CH05c02	PET-TTA AAC TGT CAC CAA ATC CAC A	GCG AAG CTT TAG AGA GAC ATC C	141-177	10	Liebhard *et al.*, 2002
A168	CH05d08	PET-TCA TGG ATG GGA AAA AGA GG	TGA TTG CCA CAT GTC AGT GTT	182-220	No position	Liebhard *et al.*, 2002
A202	CN494248	VIC-ACC TCT CTT CAT TCT TCT CC	GAA GAG CAT AGA AGA ACA CC	316	No position	**
A57	CH01g05	VIC-CAT CAG TCT CTT GCA CTG GAA A	GAC AGA GTA AGC TAG GGC TAG GG	140-188	14	Liebhard *et al.*, 2002
A115	CH03b10	NED-CCC TCC AAA ATA TCT CCT CCT C	CGT TGT CCT GCT CAT CAT ACT C	99-121	15	Liebhard *et al.*, 2002
A72	CH05g03	NED-GCT TTG AAT GGA TAC AGG AAC C	CCT GTC TCA TGG CAT TGT TG	135-192	17	Liebhard *et al.*, 2002
A76	CH02c11	NED-TGA AGG CAA TCA CTC TGT GC	TTC CGA GAA TCC TCT TCG AC	219-239	10	Liebhard *et al.*, 2002
A196	CN492626	NED-TGC AGG TTG AGA TGG TTT GG	GAC CCA AGA ACA ACA AAA CC	310	No position	**
A222	CN581649	NED-AGC CCT GAT CTT CCT CTA GC	ACG AAC TAC CAC CTC AAA CC	382	14	**
A213	CN496756	NED-TCG GTG GAA GAC CAA GCA G	CAT GAT CAT GTG GCG CCG T	473	No position	**
A193	CN492206	6-FAM-ACA TAC TGG AGT CTG CGA GC	CAA TAC GCT AGT GAA GAC GC	379	13	**
A85	CH03d01	6-FAM-CGC ACC ACA AAT CCA ACT C	AGA GTC AGA AGC ACA GCC TC	95-115	2	Liebhard *et al.*, 2002
A44	CH05e06	6-FAM-ACA CGC ACA GAG ACA GAG ACA T	GTT GAA TAG CAT CCC AAA TGG T	125-222	5	Liebhard *et al.*, 2002
A48	CH01h02	6-FAM-AGA GCT TCG AGC TTC GTT TG	ATC TTT TGG TGC TCC CAC AC	236-256	9	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 7						
A133	CH01d03	PET-CCA CTT GGC AAT GAC TCC TC	ACC TTA CCG CCA ATG TGA AG	136-160	No position	Liebhard *et al.*, 2002
A352	CO865608	CAA CAA GTG TGC CTC TGT GG	PET-AGC AAG CAA CAG ATC AAG CC	168-184	15	Lerato, 2004
A147	CH02h11b	PET-GGG ACG TAA ACA GGT ATT CTC TC	ATG GTT AGG CCA AGC ACA TC	214-240	No position	Liebhard *et al.*, 2002
A136	CH01e121	PET-AAA CTG AAG CCA TGA GGG C	TTC CAA TTC ACA TGA GGC TG	246-278	No position	Liebhard *et al.*, 2002
A215	CN496844	VIC-GGA TCA ACA GCA ACA GCA GC	CTT GGA CCG GAG CAT GTC C	293	No position	**
A90	CH02b121	VIC-GGC AGG CTT TAC GAT TAT GC	CCC ACT AAA AGT TCA CAG GC	101-143	5	Liebhard *et al.*, 2002
A81	MS06g03	VIC-CGG AGG GTG TGC TGC CGA AG	GCC CAG CCC ATA TCT GCT	154-190	10	Liebhard *et al.*, 2002
A79	MS01a03	VIC-AGC AGT ATA GGT CTT CAG	TGC GTA GAT AAC ACT CGA T	235-249	10	Liebhard *et al.*, 2002
A36	CH02g09	NED-TCA GAC AGA AGA GGA ACT GTA TTT G	CAA ACA AAC CAG TAC CGC AA	98-138	8	Liebhard *et al.*, 2002
A38	CH05e03	NED-CGA ATA TTT TCA CTC TGA CTG GG	CAA GTT GTT GTA CTG CTC CGA C	158-190	2	Liebhard *et al.*, 2002
A118	CH02d10a	NED-TGA TTT CCT TTT TCG CAA GG	TTC ATC GTT CCC TCT CCA AC	215-229	16	Liebhard *et al.*, 2002
A274	CN925672	NED-ACA CGG TAA ACA CTA CCA CC	GCG AAC TTC ACC TTC GCA AA	264	16	***
A12	NZ05g08	6-FAM-CGG CCA TCG ATT ATC TTA CTC TT	GGA TCA ATG CAC TGA AAT AAA CG	121	No position	Celton, pers comm.
A283	CN921216	6-FAM-CGC ACA CCC CCA AAT GCG	AGA GCT TGT CGC CCT CGG	379	No position	***
A89	CH04e02	6-FAM-GGC GAT GAC TAC CAG GAA AA	ATG TAG CCA AGC CAG CGT AT	143-163	4	Liebhard *et al.*, 2002
A98	CH02d121	6-FAM-AAC CAG ATT TGC TTG CCA TC	GCT GGT GGT AAA CGT GGT G	177-199	11	Liebhard *et al.*, 2002
A78	COLa	6-FAM-AGG AGA AAG GCG TTT ACC TG	GAC TCA TTC TTC GTC GTC ACT G	220-240	10	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 8						
A365	CO903680	PET-CAG CAG TTG CAA CAA GTC C	GTG GAA ATG GCT AAG CAA GC	244	11	Lerato, 2004
A162	CV128959	PET-AAA TAG TGT GGA AGA CGC GG	CAA TAT ACT AAT GAG TCC TTC G	144-166	No position	***
A126	CH01b09b	PET-TTA TAG CAG CAA CAG GAG CG	TAT TCG GGA GGC ATG GTA TG	172-182	No position	Liebhard *et al.*, 2002
A145	CH02g01	PET-GAT GAC GTC GGC AGG TAA AG	CAA CCA ACA GCT CTG CAA TC	198-238	13	Liebhard *et al.*, 2002
A381	CO751676	VIC-TGT GGC TCT GGA TGG TTC C	TAC CAG TCC ATC CGT ATA GC	220	10	**
A91	CH03a04	VIC-GAC GCA TAA CTT CTC TTC CAC C	TCA AGG TGT GCT AGA CAA GGA G	92-124	5	Liebhard *et al.*, 2002
A84	CH02f061	VIC-CCC TCT TCA GAC CTG CAT ATG	ACT GTT TCC AAG CGA TCA GG	135-158	2	Liebhard *et al.*, 2002
A35	CH01f021	VIC-ACC ACA TTA GAG CAG TTG AGG	CTG GTT TGT TTT CCT CCA GC	174-206	12	Liebhard *et al.*, 2002
A61	CH04c07	NED-GGC CTT CCA TGT CTC AGA AG	CCT CAT GCC CTC CAC TAA CA	98-135	14	Liebhard *et al.*, 2002
A34	CH01c06	NED-TTC CCC ATC ATC GAT CTC TC	AAA CTG AAG CCA TGA GGG C	146-188	8	Liebhard *et al.*, 2002
A42	CH05d02	NED-AAA CTC CCT CAC CTC ACA TCA C	AAT AGT CCA ATG GTG TGG ATG G	203-225	4	Liebhard *et al.*, 2002
A428	CO902639	NED-CTC CTT TAT CTC TTT CCT CCC	TTG TCG TCC CAA ATC AAG CC	343	No position	***
A75	CH02b03b1	6-FAM-ATA AGG ATA CAA AAA CCC TAC ACA G	GAC ATG TTT GGT TGA AAA CTT G	77-109	10	Liebhard *et al.*, 2002
A95	CH01f091	6-FAM-ATG TAC ATC AAA GTG TGG ATT G	AAT TCC AAT TTC AGA ACA GG	125-160	8	Liebhard *et al.*, 2002
A108	CH04g04	6-FAM-AGT GGC TGA TGA GGA TGA GG	GCT AGT TGC ACC AAG TTC ACA	170-186	12	Liebhard *et al.*, 2002
A65	CH05g11	6-FAM-GCA AAC CAA CCT CTG GTG AT	AAA CTG TTC CAA CGA CGC TA	201-255	14	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 9						
A130	CH04g12	PET-CAC CGA TGG TGT CAA CTT GT	CAA CAA AAT GTG ATC GCC AC	92-108	No position	Liebhard *et al.*, 2002
A135	CN910036	GAG AAA CCG TTT GAT TAC AGC	PET-CTC CAT CCC CAA TCA CAC C	118-140	10	**
A169	CO723148	PET-CGG TGG TGA CTA GTA TCA GC	TAT GGA GGA AGA AAC TGA GGC	152-200	No position	**
A146	CO756781	PET-ATA AGT TTA GGC TCA TCT GCC	AAA CCC ATC CCA CTT AAG GC	214-236	No position	***
A319	AF527800	VIC-TAT TCT GAA GCC TCA TGG CC	ATC TCC ACT ACC TAA CTG CC	340	17	Liebhard *et al.*, 2002
A41	CH02c02b	VIC-TGC ATG CAT GGA AAC GAC	TGG AAA AAG TCA CAC TGC TCC	78-126	4	Liebhard *et al.*, 2002
A97	CH01f03b	VIC-GAG AAG CAA ATG CAA AAC CC	CTC CCC GGC TCC TAT TCT AC	139-183	9	Liebhard *et al.*, 2002
A37	CH02c061	VIC-TGA CGA AAT CCA CTA CTA ATG CA	GAT TGC GCG CTT TTT AAC AT	216-254	2	Liebhard *et al.*, 2002
A96	CH01h101	NED-TGC AAA GAT AGG TAG ATA TAT GCC A	AGG AGG GAT TGT TTG TGC AC	94-114	8	Liebhard *et al.*, 2002
A116	CH04g10	NED-CAA AGA TGT GGT GTG AAG AGG A	GGA GGC AAA AAG AGT GAA CCT	127-168	15	Liebhard *et al.*, 2002
A45	CH03d07	NED-CAA ATC AAT GCA AAA CTG TCA	GGC TTC TGG CCA TGA TTT TA	186-226	6	Liebhard *et al.*, 2002
A266	CN851624	NED-AAC TGT AGA AAA AAC ACT CCC	GGT CCT CCT TTC ACA AAT GC	409	No position	***
A111	CH03h03	6-FAM-AAG AAA TCG GAT CCA AAA CAA C	TCC CTC AAA GAT TGC TCC TG	72-120	13	Liebhard *et al.*, 2002
A121	CH02g04	6-FAM-TTT TAC CTT TTT ACG TAC TTG AGC G	AGG CAA AAC TCT GCA AGT CC	132-197	17	Liebhard *et al.*, 2002
A52	CH02d08	6-FAM-TCC AAA ATG GCG TAC CTC TC	GCA GAC ACT CAC TCA CTA TCT CTC	210-254	11	Liebhard *et al.*, 2002
A422	CV627191	CTT AAT CAC CCA TCA TTC CCC	6-FAM-CTC TGT CGG CTA ACT AAC CC	300	No position	***

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 10						
A277	CN866018	PET-TTC CTC TCA TCT ATC CTT TCG	GAG GTG ACA GAC AAA TTC GG	223	15	Liebhard *et al.*, 2002
A316	CN496913	GAA AGG ATG GTA CAC TCT TCG	PET-TTA GAT GCC TTA AAT ACT TCC G	290	No position	***
A346	CO753022	PET-CTG AGT CTT TGT TTT TGC TCG	GCT CCG CCT CTC TGT ACC	355	13	**
A603	EB114458	PET-TAT GAT CCA TCA CCC GAA GG	AGT CAT ACA GCT TCA CAT TCG	169	No position	*,**
A11	04h11	VIC-CTT CCA TCG AGA TTG CAT CAT A	CGA ATT GAG AGG TCG TCG TT	225	No position	Celton, pers comm.
A208	CN495651	VIC-CTT CTC CCA GAA CTG ACT GC	TCT ACA ACC GCA AAC ACG AG	398	No position	**
A88	CH02h11a	VIC-CGT GGC ATG CCT ATC ATT TG	CTG TTT GAA CCG CTT CCT TC	104-132	4	Liebhard *et al.*, 2002
A53	CH04g07	VIC-CCC TAA CCT CAA TCC CCA AT	ATG AGG CAG GTG AAG AAG GA	149-211	11	Liebhard *et al.*, 2002
A416	CO168103	NED-CTC AAA ACA AGA ACA ATG AGC C	CCC AAA AGG TTT TCC ACA CG	191	No position	***
A310	AU301301	NED-GGCATAGCAATGCTTGAAGG	GAATAGCACAAAGGAGGTTGC	232	No position	***
A300	CN939907	ATC CGC AGA ACT GAA GGC G	NED-ACT GGT CGG TTA TCG ACG G	307	No position	***
A307	CN445290	NED-TCA CTT TCT CAG TTG CTC TGG	ATG GAA GCT TAC TCT TTT CCG	348	No position	***
A383	CO903298	6FAM-TTGAGAAGCAATGCTGCCTC	TGCCACAGTTGGAAGGTGG	358	9	**
A101	CH04d07	6-FAM-TGT CCT CCA ATC TTA ACC CG	CAC ACA GAC GAC ACA TTC ACC	119-142	11	Liebhard *et al.*, 2002
A32	CH05g08	6-FAM-CCA AGA CCA AGG CAA CAT TT	CCC TTC ACC TCA TTC TCA CC	161-179	1	Liebhard *et al.*, 2002
A46	CH05a05	6-FAM-TGT ATC AGT GGT TTG CAT GAA C	GCA ACT CCC AAC TCT TCT TTC T	198-230	6	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 11						
A161	CH04f03	PET-CTT GCC CTA GCT TCA AAT GC	TCG ATC CGG TTA GGT TTC TG	175-191	No position	Liebhard *et al.*, 2002
A559	Hi03e04	PET-CTT CAC ACC GTT TGG ACC TC	GTT TCA TAT CCC ACC ACC ACA GAA G	132-160	13	Silfverberg-Dilworth *et al.*, 2006
A179	MS06c09	PET-ACT ATT GGA GTA AGT CGA	AAT ATA AGA GCC AGA GGC	102-118	No position	Liebhard *et al.*, 2002
A584	Hi06b06	PET-GGT GGG ATT GTG GTT ACT GG	GTT TCA TCG TCG GCA AGA ACT AGA G	236-262	11	Silfverberg-Dilworth *et al.*, 2006
A186	CN490349	NED-GTA CTA TCA GCA GAA ACT GG	GAT TTG AGC ACA ACA TAC GG	207	15	**
A372	CO052555	NED-GAA GTT CTC ATC AAG TCT TGC	GCT TCT GCA CAA TGG CTG G	238	13	**
A340	CO416051	NED-CCT CAC TAA ACG CAT TGC AC	CGG TAC GAT GAG GAT CAT CC	317	5	**
A424	CO415353	NED-ATG AAC AGT CAC AGA CTA TGC	AAC GAA GCA AAG GAA GAC GG	330	15	***
A180	CN444111	NED-TGA GGC CAC CTA AAT ATC AC	CAG GAT GAG AGT TCT TGA GC	409	No position	**
A188	CN490740	6-FAM-AGG ATC CTT CCT CGA TTT GC	GGA GCG CAT GAA ATT ACT GC	213	10	**
A343	CV084260	6-FAM-CAA AGC AAA ACA GAG GAT TTG	GGA GCG CAT GAA ATT ACT GC	265	16	**
A105	CH01g121	6-FAM-CCC ACC AAT CAA AAA TCA CC	TGA AGT ATG GTG GTG CGT TC	112-186	12	Liebhard *et al.*, 2002
A536	Hi02c07	VIC-AGA GCT ACG GGG ATC CAA AT	GTT TAA GCA TCC CGA TTG AAA GG	108-149	1	Silfverberg-Dilworth *et al.*, 2006
A781	Hi02h08	VIC-GCC ACT CAT ACC CAT CGT ATT G	GTT TGG CTG GGA ATA TAT GAT CAG GTG	170-200	16	Silfverberg-Dilworth *et al.*, 2006
A110	MS14b04	VIC-CCT TAA GAA TCA TGT GAT	ACT AAT GGC ACA AAG ATT GT	230-292	12	Liebhard *et al.*, 2002
A701	EG631386	ACA ACC TCT TCT TCC TCA GC	VIC-GAT ATC AGA AGG TAC ACT GAA G	389	No position	*,**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex mix 12						
A50	CH01f07a	6-FAM-CCC TAC ACA GTT TCT CAA CCC	CGT TTT TGG AGC GTA GGA AC	174-206	10	Liebhard *et al.*, 2002
A70	CH05c06	6-FAM-ATT GGA ACT CTC CGT ATT GTG C	ATC AAC AGT AGT GGT AGC CGG T	104-149	16	Liebhard *et al.*, 2002
A579	Hi07b06	6-FAM-AGC TGC AGG TAG AGT TCC AAG	GTT TCA TTA CCA TTA CAC GTA CAG C	216-222	6	Silfverberg-Dilworth *et al.*, 2006
A311	AU301254	6-FAM-TCC CGG AAA TTT TTC AAC GC	AAC GCT AGG GAT TGG TCG C	232-244	No position	***
A100	CH04a12	NED-CAG CCT GCA ACT GCA CTT AT	ATC CAT GGT CCC ATA AAC CA	158-196	11	Liebhard *et al.*, 2002
A413	CN492417	NED-TAC CAT GTT TTA GCA CCA TGG	GGC CAA GTT AGG TCA AGA CG	116-145	No position	**
A395	CN495393	NED-TCC CAA GCT CCC AAC AAA CC	CTA TCT GGG TCG GCC AGG	200-219	10	**
A725	SAA725	TGG TGG TTC TCA GTC CAG G	NED-CCA ATA GTG ATA AGC AGT TC	228-353	No position	*,**
A259	CN904905	GTT CAA TGA CTT GAA CAA GAG G	PET-TTC TGA TGA ATG AAA GCA CCT	114-138	No position	*,**
A385	CO865258	PET-CTC CTG TGA ATC TGC CAC C	AGA AGC AGC TCT GGC AGG	170-190	No position	**
A555	Hi02d04	PET-TGC TGA GTT GGC TAG AAG AGC	GTT TAA GTT CGC CAA CAT CGT CTC	217-239	10	Silfverberg-Dilworth *et al.*, 2006
A662	EB138222	PET-TGG AAG ATT GTG AAG GCA GC	TTG TGG GTG GTT CTT CAT CC	264-266	No position	*,**
A221	CN580954	VIC-TCT CTT GTC AAG GAT GGA CC	GAA TCC GAA GCA ACG GAA GC	106-118	No position	**
A540	Hi16d02	VIC-AAC CCA ACT GCC TCC TTT TC	GTT TCG ACA TGA TCT GCC TTG	141-160	11	Silfverberg-Dilworth *et al.*, 2006
A813	NZmsCO754252	VIC-CTG CCC TCA AGG AGA ATG TC	ACA GGT GCA GCA AAG GCT AT	195-197	6	Celton, pers comm.

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A550	Hi04e04	VIC-GAC CAC GAA GCG CTG TTA AG	GTT TCG GTA ATT CCT TCC ATC TTG	224-242	16	Silfverberg-Dilworth *et al.*, 2006
MonsterPlex 13						
A665	EB132264	CTC ATT GCT ACT CAC TAA TCC	6-FAM-GTT CAG AAA AGA GAG AGA GAG	119-148	No position	*,**
A491	DT041234	GCA ACT GCA AGT GAG AGG G	6-FAM-AGA AGA AGC CAT GGC CAC C	158-176	No position	***
A284	CO752155	6-FAM-TGC CTA AGA ATC CAT CTG GC	TCT CGA ACT TAC TAA CTA GGC	189-192	No position	***
A512	CN944528	6-FAM-GAC GAC GGA AAG GAA GAC G	ATT ACG CTG TTG CAG AGA GC	205-214	No position	**
A551	Hi23g02	6-FAM-TTT TCC AGG ATA TAC TAC CCT TCC	GTT TCT TCG AGG TCA GGG TTT G	229-250	4	Silfverberg-Dilworth *et al.*, 2006
A742	SAA742	TGA CAA CTA TGA TCG AAG TGG	6-FAM-TTT CAT ATC ACA TGA CGT GGC	266-275	10	*,**
A626	EB135470	6-FAM-CAT CTT TAT ATG AGC CAC TTC C	GTT GAT GCT ATT GGT AGT AGG	291-301	No position	*,**
A593	EB138715	6-FAM-GCG CGA TGC CAT CTC TGC	GGG ATC GCA GCT CAC TCC	315-338	No position	*,**
A594	EB151342	6-FAM-GCT GAA AGA TGT CAC CTA CC	CGT GGA TCC AGC CTT AGG G	359-376	No position	*,**
A344	CO905375	6FAM-AGT CTC TGT TTT TGC TCG TTC	GAA CGC CGG GTC CCT GC	407-435	No position	**
A189	CN490897	6-FAM-GCG GAG ATA AGG ATG CTT CG	CCT CAG TAC CAA ACT AGG CT	458-462	13	**
A134	CH01d07	NED-AAA ATC CAG TTT TCC ACC TC	AGT CGA AAT CCC GAA CAA TC	97-102	No position	Liebhard *et al.*, 2002
A379	CO865207	NED-TGC ACC AAA TAA GCC GAT CC	CAA GAA GTG CAA CCA GTC GA	120-138	9	**
A181	CN444846	NED-CTA GTT TCC TCC GTG GTT TCT	CGG AAA GTT TGT AGT GGT GG	150-152	No position	**
A229	CN496966	NED-GGA GGA GAA TAT GTG ATT TTG AG	GAT TGC GAC AGC ATT TAT GG	167-171	No position	**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A639	EB149851	NED-GAA CAG AGG GAA GCA GAC G	AGA AGT GGC AAC CAT GTT GC	187-202	No position	*,**
A440	CO416477	NED-CCA CAC AAC ACA AAC CAA CC	TGT GGT CAT TTG GTG AGT CC	218-224	13	***
A241	CN887787	NED-CAC TTT AGC TTA GTA CAC AGC	TGA GGT AGT AAG AGT AGA AGG	254-257	No position	**
A629	EB149808	NED-TTA AAG CTC GAG CCG AGC C	TCC AAC CCA CTA AGA TTA TCC	269-286	No position	*,**
A243	CN907588	NED-TCC AAC CCA CTA AGA TTA TCC	GGT ACT TGT TGG TGA TCT CG	304-307	No position	**
A534	DR997824	NED-GAC TGG TGA GAT AGA GAG G	ATG AGC ATC GGA TAG CTG G	319-330	No position	**
A763	SAA763	CAC CGA ACC AAT CCG TAG C	NED-AGA GAG TAT GAA AGG TGT TCC	344-355	No position	*,**
A525	CV186968	NED-ACG TAC ATG CAT GCC TTT GG	AGT CAA GAG GCA CTA TGA GC	389-397	No position	**
A647	EB146894	AAG GAA GGA GCC ATG GAG G	NED-ATA TGG AAT CTA CAA GCC ACC	422-438	No position	*,**
A397	CN491038	NED-GCT CTG TCT CGT TGA TCG G	AGC TGC TTC ACC CTC TTG C	498-510	No position	**
A163	CH04f07	PET-CAG ATC ATG AAT GAT TGA AA	GAA AAT CAC ACC CTC AAA CCA T	82-113	No position	Liebhard *et al.*, 2002
A174	CH05g02	PET-AGT GCA GCT TTC AGC TCA GAT T	AGT CAG ACA CAC CAA AAT CCC T	133-155	No position	Liebhard *et al.*, 2002
A178	CH05h12	PET-TTG CGG AGT AGG TTT GCT TT	TCA ATC CTC ATC TGT GCC AA	164-192	No position	Liebhard *et al.*, 2002
A545	Hi07f01	PET-GGA GGG CTT TAG TTG GGA AC	GTT TGA GCT CCA CTT CCA ACT CC	207-215	12	Silfverberg-Dilworth *et al.*, 2006
A686	EB106592	PET-CTT GGA AGC CCA ACG AAC C	AGA GGA GCT TGT TGT TGA GG	234-237	No position	*,**
A533	DR993168	PET-ACT TCC CTG CCG CAG AGG	CAC TTG AAG CAG ACC GAG G	249-253	13	**
A535	DR997862	PET-CAC AAT CAT ATT CCC GCA CG	TTC TTC TCC GAT GAG CAA GC	275-283	15	**
A472	DR995122	PET-CGA GGC CTT TTT TTA CTC GG	ATT GCT CTC CTG TGG TGC C	296-328	14	***

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A688	EB142061	PET-TCG ACC AGC CAG ACA AAG C	AAG AGT TGC AGG TGG GTC G	339-341	No position	*,**
A516	CO900034	PET-AAA GTC CGT TTT GGG CTG AG	GCT CTC TGC TGC CAT TTC C	353-367	15	**
A272	CN942512	PET-ATC CAT CAT CGG AAA CCT GC	AAA GAA ACT GGA GGA CCG C	389-397	14	***
A638	EB147667	PET-AGG TCT CAG GAC TCT CAG G	ATT GTT AAT GTC GGC GAA TCG	411-420	No position	*,**
A661	EB126773	PET-GTT TGT GTT TGA ACA ACG ACC	GTG GTT GTT GAG GTC GTG G	442-470	No position	*,**
A610	EB133782	PET-CTC CCA GCT CAC TTT CTCC	CAG AGG ATG CAC CAC TTG G	508-543	No position	*,**
A536	Hi02c07	VIC-AGA GCT ACG GGG ATC CAA AT	GTT TAA GCA TCC CGA TTG AAA GG	107-119	1	Silfverberg-Dilworth *et al.*, 2006
A104	CH01d09	VIC-GCC ATC TGA ACA GAA TGT GC	CCC TTC ATT CAC ATT TCC AG	131-172	12	Liebhard *et al.*, 2002
A630	DY255319	ATC GAA TTC CGT TGC TGT CG	VIC-ATC AAT CAG CAG GCT CTT CC	181-211	No position	*,**
A514	CX025465	VIC-TGC TAG AGC TGC GTT CTC C	TCG CAG ACT GCT CGC TGC	227-235	9	**
A592	EB149750	VIC-ATC AAG GTG TGA GTG TGT GC	AAG CTT GCA TCT CTA GGT CC	246-265	No position	*,**
A724	SAA724	CTC TTC ATC TGA GAA TAC ACC	VIC-AGA CTC GAG TCA TCC ATA CC	282-288	No position	*,**
A490	DR995748	VIC-TAC ACC AGC GCC ACA CCG	TGG CGA GCA CGA TGA GCG	315-338	No position	***
A498	DR992457	VIC-TCT CCA AGT GGA CGA ATC AG	TCC TCA GTG AAG ACA AAC CC	356-375	No position	**
A443	CO903797	VIC-ATT GAT ATC ACA GCT AAG CC	CCA AAA TCT CAG AAA CGG GG	399-413	16	***
A226	CN444745	VIC-AGG AAA TAA ACA CCG AGT AAA C	CAC AAG CAT CTC GAG CAC C	455-480	No position	**
A597	EB109450	VIC-GTT GAT ATC GGT ACG CTA GC	GAG GCA TCT CTG TTG GTG G	527-539	No position	*,**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex 14a						
A508	DT041145	6-FAM-TGG CTG TGA TGT CAT GAT GG	TCT AGA GTT CAT CAC AAA GAA G	63-131	No position	**
A602	EB144676	6-FAM-CAT CAG CCA TCT TCT TCT CC	CCG ATG GAA ATG CAG AAG C	161-197	No position	*,**
A418	CV150384	6-FAM-ACA AAC CAC CAC CAA TTC CC	CCT GAG AGA GCC AAT TGA GC	235-250	No position	***
A254	CN933736	TGG CAG CTC CAC CAC AAT C	6-FAM-GCC AGA TTC ACA CGA AAG C	291-334	No position	**
A262	CO865955	6-FAM-TAC TCA TGG CGG CAA CTC C	GCG GAC GGT GAT TTC TTG G	200-214	No position	**
A220	CN580732	6-FAM-ATG GGG CCA GTT ACA GGA G	CTG AAG AAA TCG CAG GTT CC	352-400	No position	**
A461	DT000945	6-FAM-AGT TGA CTA CCT CCT CCG C	GTA AGC GAT GAA ACT GAT GC	370-421	No position	***
A460	NZ26c06	NED-GAC GAA GAA CTC GCC GGA GC	CGA GGA CCA ACC CAC ACA CAA	102-165	No position	Celton, pers comm.
A400	CN578608	NED-CTT CGC CTC AGT TTC AAA CC	GAA GCC AGA GTC TGT TGC C	192-196	No position	**
A601	EB154700	NED-TTT GTT GGG ATT GTG GGT CG	GTT GCT GAG AGT GAT GAT GG	229-236	No position	*,**
A502	DR990381	NED-AAA CAC TAC TGT GCT GGT GG	AGT CCA CTT ACT ACT CCT CC	264-300	10	**
A615	EB153928	NED-CTC AAA TCC CAG AAG ATT ATC C	GTC CTC GGA ATC GTC CTC C	348-358	No position	*,**
A448	CV150002	NED-AGT TCG ATC TTT AAT GCC CC	GAA AGA GCA AGA GAG ACT GG	426-456	No position	***
A30	AT000400.1	NED-CGT ATC GAA GTA GAA CGA CG	CAG GGT TGT ACG GAT TCA CG	175-181	No position	***
A574	Hi02b07	NED-TGT GAG CCT CTC CTA TTG GG	TGG CAG TCA TCT AAC CTC CC	204-216	12	Silfverberg-Dilworth *et al.*, 2006
A531	CN943946	NED-GTC TAC TTC CAG AAC TTG CC	GAT CTC ACC ACA AAA TGC ACT	327-341	No position	**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' – 3')	Reverse primer (5' – 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex 14b						
A448	CV150002	NED-AGT TCG ATC TTT AAT GCC CC	GAA AGA GCA AGA GAG ACT GG	426-456	No position	***
A159	CH04d11	PET-ATT AGG CAA TAC ACA GCA C	GCT GCT TTG CTT CTC ACT CC	85-152	No position	Liebhard *et al.*, 2002
A265	CO723438	TCC GAT TCT CTA TCA GAT CCA T	PET-TGG ATC GGG ACA TGG AAG G	182-202	No position	**
A4	GD 100	PET-ACA GCA AGG TGT TGG GTA AGA AGG T	TGC GGA CAA AGG AAA AAA AAA AGT G	223-238	No position	(Rose and Falush, 2005)
A617	EB114260	TCA TCC TCA TCG TTT CCT CG	PET-TGT AGT TGC CTG CGA CAC C	274-290	No position	*,**
A496	DR992457	PET-TCT CCA AGT GGA CGA ATC AG	TCC TCA GTG AAG ACA AAC CC	319-330	No position	**
A13	22c6	VIC-GAC CTT TCC CTC TCC TGA	CTG GAT ATG ATT ATT GCA GA	63-142	No position	Celton, pers comm.
A833	NzmsEB137525	VIC-TCT TTC GCT GGT GTC CTC TT	GTG CTG CTT GCT GTT GTT GT	172-192	17	Celton, pers comm.
A583	Hi04f09	VIC-ACT GGG TGG CTT GAT TTG AG	GTT TCA ACT CAC ACC CTC TAC ATG C	222-258	13	Silfverberg-Dilworth *et al.*, 2006
A414	CN489062	VIC-ACA ACT TGG TTA CGC GAC AC	GAA CAG ATT AGG GTC GCT GG	284-306	No position	**
A484	DT041144	VIC-AAA TGC TGC AGT GAG GCC C	GAA TTC CAT CTA AAC GAG AGC	335-396	No position	***
A419	CO755991	VIC-AAT CTC TCG TCT GCA AAC CC	GTA TGA GTA TCC AGC ACC CG	150-154	No position	***
A598	EB138859	VIC-TAC GCT AGT GCT ACA GAA GC	AAA CTC CAT AGC AGT AGT TCG	162-169	No position	*,**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
Megaplex 15a						
A5	GD 103	6-FAM-CGG CGA GAA AAA AAA ACA ATG	GGA TAA CCG TCC CCC TCT TC	78-130	No position	Rose and Falush, 2005
A245	CN943613	6-FAM-TAG CAG AAA CCA GCA GAT GG	TGA GGC CTC GAA GAA GTG C	165-174	12	**
A567	Hi03a03	6-FAM-ACA CTT CCG GAT TTC TGC TC	GTT TGT TGC TGT TGG ATT ATG CC	205-223	6	Silfverberg-Dilworth *et al.*, 2006
A201	CN493973	6-FAM-TAC TCT CTG ATC TTC TGA TTG C	CAG TGC ACC ACC AAG TTG C	252-329	No position	**
A214	CN496821	NED-AAT GCC ACT GAA ATG ACT GC	AGC TTC GTC TAT GGA GTG C	358-410	No position	**
A82	CH02b101	NED-CAA GGA AAT CAT CAA AGA TTC AAG	CAA GTG GCT TCG GAT AGT TG	121-159	2	Liebhard *et al.*, 2002
A822	NZmsDR033893	NED-CAC TTA GGG TGT ATG GGT GTG A	TCA TTT TGG GCA GGC ACT	194-225	11	Celton, pers comm.
A826	NZmsEB111793	NED-TTG AGG GCT GCT TTC CAG	GGA GAC ATA CAA GAT TTC CAA TGA G	275-281	No position	Celton, pers comm.
A466	DT040421	GGC AGA GCA GAT GCA GAT AA	NED-TAT AAG ATG GAA GCC AAT GCC	325-350	No position	***
A676	EB106537	6-FAM-GTA CAG ATC TCG TTT CAT CAC	TGA TTG AAG GGC AGT CTT GG	178-188	No position	*,**
A738	SAA738	CGA AAC TGG TCG AAG AAC CT	6-FAM-AAA CTA CAC AGA GCA AGA TGG	332-351	No position	*,**
A195	CN492475	NED-ACT CAC CCC CTT CCT TTC C	GAA GAA AGG TAG GGG TCA GC	175-185	5	**
Megaplex 15b						
A444	CO752447	NED-AAC CCG CAA ACA AAA ATC CAG	TCG GTG ATC CGT TTC GCC	439-453	No position	***
A131	CH01c11	PET-AAA TCC TAA AAC ACA AGC AAA ACC	TGA ACC AAG TCC TCC ACT CC	109-155	No position	Liebhard *et al.*, 2002
A759	SAA759	PET-AGT TGA CTA CCT CCT CCG C	GTG GTT CTC ACG GTA CAC G	187-239	No position	*,**
A506	DR997517	PET-TCT ACA CCA CCC CGC CTC	CGA ATT CGT CAT TGG AGA GG	287-324	No position	**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A664	EB153442	PET-GGT TCA CAA GGC CAA CTT TG	ATG GTT CGA TCG GTT TAA TGC	365-373	No position	*,**
A529	CN443900	PET-AGC AAT TTT GCC TAA AAC CGA A	GCT CAT GAG GTG CGA TTG G	418-498	No position	**
A40	MS14h03	VIC-CGC TCA CCT CGT AGA CGT	ATG CAA TGG CTA AGC ATA	114-140	3	Liebhard *et al.*, 2002
A515	CV657225	VIC-TCC CTG TCA TCG AAT GAT GC	GCA AAC CCA ATC AGA AGG AC	173-200	No position	**
A510	CN881550	VIC-TCG CGG GAA GTT CCG CAG	GGC CTC AAG GAC CCA TCG	241-253	No position	**
A359	CO756752	VIC-CTC TCT GCT TTC TTT CCA GC	GGT GGC TCC GCT TTC TCC	293-345	No position	Lerato, 2004
A339	CO066563	VIC-ACA AAG GAA CAG TGA AGA CTC	TAC TTG CTC TGC ATA GTT TGG	420-438	No position	**
Megaplex 16a						
A9	NZ01a6	6-FAM-AGG ATT GCT GGA AAA GGA GG	TTA GAC GAC GCT ACT TGT CCT	87-155	No position	Celton, pers comm.
A828	NZmsCN914822	6-FAM-GAC GAT GAT CAG GCC ATT CT	TGT TCA TGT CGG TGC TCA AT	190-193	14	Celton, pers comm.
A361	CO903775	6FAM-CAT CGA TCC TTC ATG AAA GGC	GGT GGT CTG ATA TGA TTG GCG	239-251	No position	Lerato, 2004
A656	EB139609	6-FAM-ACC ATA TAC ATC TCT CTC TGC	TTC AGA AGC TGT TGT TGT TGG	311-351	No position	*,**
A336	CO168310	6-FAM-GTC GAC TTC GCC CGA AGC	ACG ACC AGG TTC ATG AAC TG	386-474	No position	**
A565	Hi08h12	NED-GAA GGA AAT CAT CAT CAA GAC G	GTT TCA AGA CCA TGG AAC AAC TTG G	191-202	10	Silfverberg-Dilworth *et al.*, 2006
A671	EB149428	NED-GTT AAT TCC GCT CCC CTC C	ATG CTT CTG GGC TCG AAC C	255-281	No position	*,**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A678	EB128431	ACG TAG TGA TAC CGG ATT CG	NED-AGA GCT AGC TAG AGA TAT TCC	322-342	No position	*,**
A370	CO052793	6-FAM-CCA TCC CTT CCT CCT ACA TC	TGG GCC TCT TGT TCA TTA GG	200-230	No position	**
A558	Hi01e10	6-FAM-TGG GCT TGT TTA GTG TGT CAG	GTT TGG CTA GTG ATG GTG GAG GTG	126-224	4	Silfverberg-Dilworth *et al.*, 2006
A473	DR996674	NED-CAA GCA GAG TAG CAA CTG C	GAG GCC TCT TGC AAT TGC G	424-428	No position	***
A328	CN489396	NED-TGG GTC TGC TGA GTA ATT AGG	TTG GGC TTG GTC GAA ACA CC	448-540	No position	**
Megaplex 16b						
A140	CH02a08	PET-GAG GAG CTG AAG CAG CAG AG	ATG CCA ACA AAA GCA TAG CC	128-177	5, 11	Liebhard *et al.*, 200
A386	CO901343	CAC CTC TTC CCT CAT CAG TC	PET-CGA CAA AGG AGA CTG AGA GG	208-233	No position	**
A760	SAA760	PET-GTC TTT GGA AGC TTG GTT GG	AAG TTA CTC TTT GTT GCT C	274-301	No position	*,**
A429	CO905285	GTT GAT TCT TAT GGC ACC GG	PET-ACC CAA ATG GCG CAA TGC C	344-382	No position	***
A445	CO068219	PET-ATT GCT TGC ACC GCA ACG C	GGA CTG ATC AAT GAC ACT CG	433-437	No position	***
A459	NZ17e06	VIC-AAC ACG CCA TCA CAC ATC	CTG TTT GCT AGA AGA GAA GTC	60-158	No position	Celton, pers comm.
A204	CN494928	VIC-AAT TAT ATC CGT CCG ACT CCA	TTA GAG TAG TCA CGA TAA TGG	209-229	15	**
A732	SAA732	TAT CGT AGA GCA GGT TGC TG	VIC-TAT CAG TAT GCA TCA CCT AC	269-309	No position	*,**
A368	CO723511	VIC-CTG TCG GGA TTC ATT GTT GC	CCG AGT AGA AGG CTG AAG C	356-434	No position	Lerato, 2004
A673	SAA673	CAC CGA ACC AAT CCG TAG C	NED-AGA GAG TAT GAA AGG TGT TCC	476-494	9	*,**
A411	CN581642	CAA GAA TAC GTT GGG CAT GG	VIC-ACA ACG ACA TAA CAA ACA CG	170-180	No position	**

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A636	EB121159	VIC-GGA TCA GAG AGC TCT CAG C	TGT GTA GAG CAG TCA TGT GG	184-195	No position	*,**
A753	SAA753	VIC-TCT TTG CTT TGC CCT TGT GG	AGT CCA ATT CTT CCT CTT CAC	249-261	No position	*,**
Megaplex 17						
A458	04f3	6-FAM-CAA AAC CAC CCT CAT CCT CGA A	CCC CAA GCA GAC CTG AAG AAA	93-143	No position	Celton, pers comm.
A451	AF429983	6-FAM-TAC ACA GAC CAG TAC TCT GC	GGA GTC CCA TTT CAA TGT GG	174-219	No position	***
A192	CN491993	6-FAM-AAG CAG TCG CAG CAG GTG	AAC AAC CGT TCG GAT TCT CG	245-284	No position	**
A505	DR995002	6-FAM-ATC TGA TGG TGC ATC GGT AG	TTA GGG TCT TCT TGT CAC GC	324-334	No position	**
A595	EB148060	6-FAM-ACT CTC ATT TCT CCA CCT CC	CTC CTC TGT CTT CCT CTG G	374-441	No position	*,**
A774	Hi04e05	NED-AAG GGT GTT TGC GGA GTT AG	GGT GCG CTG TCT TCC ATA AA	116-179	8	Silfverberg-Dilworth *et al.*, 2006
A546	Hi22f12	NED-GGC CTC ACC CAG TCT ACA TT	GTT TGG TGT GAT GGG GTA CTT TGC	207-212	5	Silfverberg-Dilworth *et al.*, 2006
A308	CN444942	NED-GCT CTC AAA GTC TCT CCA GC	TAC GGA CTC TCT TTG GGG C	260-273	No position	***
A740	SAA740	TTC ACC CAA TTC CAC AAC CG	NED-TCA CTG TCG TCC AAA TCA GG	305-325		*,**
A228	CN496055	NED-CCA CAC AGA AAC GAG TCC TC	ATT TTG GTC CTC CTT GCT GG	360-364	No position	**
A563	Hi04b12	PET-CCC AAA CTC CCA ACA AAG C	GTT TGA GCA GAG GTT GCT GTT GC	138-154	8	Silfverberg-Dilworth *et al.*, 2006
A170	CH05b06	PET-ACA AGC AAA CCT AAT ACC ACC G	GAG ACT GGA AGA GTT GCA GAG G	185-215	5, 16	Liebhard *et al.*, 2002

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A403	CN494091	PET-CTT CAA CTT CTC AAA TCG ACG	CTT CTG GAA CTC AGC CTC C	253-289	No position	**
A717	SAA717	AGT TAC AAG GCG CAT TGA GG	PET-TTT CGA GTA GCT AAA GAG TCG	351-361	No position	*,**
A421	CO865954	PET-AAC ACC GTC CAG GAA TGC G	ACA CAC AGG TCT TCG CAG G	452-455	No position	***
A561	Hi05b09	VIC-AAA CCC AAC CCA AAG AGT GG	GTT TCT AAC GTG CGC CTA ACG TG	123-140	7	Silfverberg-Dilworth *et al.*, 2006
A47	CH04e05	VIC-AGG CTA ACA GAA ATG TGG TTT G	ATG GCT CCT ATT GCC ATC AT	174-227	7	Liebhard *et al.*, 2002
A378	CO753033	VIC-ACA CAG TCA TTG CTT CCT CC	ACC CAG CAT GTG GTC GAA G	273-296	No position	**
A645	EB156254	VIC-TAT TGA TTG TGT GTG TGT GCG	TAA GAG AAG ACG ACA TTG TCG	329-358	No position	*,**
A623	EB149589	VIC-TCT TTA CCT TCT TCT CCA TCC	CGG TAC GCT GTG GAC TCG	401-404	No position	*,**
Megaplex 18a						
A218	CN580519	6-FAM-TCC CCA CAC CA TTG ATT TGC	ACC TTG GAA GCT CCC TTC C	120-135	No position	**
A209	CN495857	6-FAM-TCA AAA CCC ACC TCA TAT TGC	TAG GAA GGA GAT GAG ATT TGG	145-155	3	**
A824	NZmsEB153947	GGG AGA GTT AGG GGA AAA GG	6-FAM-GGG AGA GTT AGG GGA AAA GG	166-180	11	Celton, pers comm.
A716	Contig4839	CTG TGC CGT CAT CTA TAT GC	6-FAM-AAC CAA AGA GGG AAG AGA CG	185-200	No position	*,**
A804	NZmsEB106592	6-FAM-CTC CCA CTA CTA GCC AAA CG	TTG GGA TTT GAA GGA CAG G	240-243	2	Celton, pers comm.
A612	EB1155894	6-FAM-TTT GCG ACA CGT CTC CAC C	TTG CAC CGA GCT CCT AGT C	250-280	No position	*,**
A422	CV627191	CTT AAT CAC CCA TCA TTC CCC	6-FAM-CTC TGT CGG CTA ACT AAC CC	290-310	No position	***

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A296	CN880881	6-FAM-ATA GCT CAT ACC GCT TCT CC	GTG ACG AAA ACC AAG AAC CC	390-420	No position	***
A302	CN581539	6-FAM-ACA ACA GCT GAC GAC CAA GC	GTC TCC ATG ACT TTT CTG TCC	450-500	No position	***
A7	GD 147	NED-TCC CGC CAT TTC TCT GC	AAA CCG CTG CTG CTG AAC	120-150	No position	(Rose and Falush, 2005)
A402	AT000420	NED-GTT GGA CCA ATT ATC TCT GC	ATA TAC TGG GGA GGT TGA GG	162-174	No position	**
A680	EB106034	AGA AGA AGC CCA TCC CAG C	NED-TTC ACC TTC GTC GGC ATG G	189-196	No position	*,**
A206	CN495278	NED-CCC AGA ATC ATT CAG AGA CC	GCA GGC TCC ATG CAG TTC G	240-258	No position	**,
A420	CO903145	NED-GGG CAC TGA ACG GTT CGC	CTT TAT GCA GAG ACA TGG TCC	261-263	No position	***
Megaplex 18b						
A635	EB149433	NED-CTG CAA CGT ATA CTC TAA TCC	GAA AGT AAC AAA GTA CCA GGC	285-309	No position	*,**
A614	EB155789	NED-CCC CGT TCC CTT GAA TTG TA	CCA GTG GAA CGA TGA CTG C	323-358	No position	*,**
A419	CO755991	VIC-AAT CTC TCG TCT GCA AAC CC	GTA TGA GTA TCC AGC ACC CG	150-154	No position	***
A151	CH03g06	PET-ATC CCA CAG CTT CTG TTT TTG	TCA CAG AGA ATC ACA AGG TGG A	137-171	No position	Liebhard *et al.*, 2002
A549	Hi05e07	PET-CCC AAG TCC CTA TCC CTC TC	GTT TAT GGT GAT GGT GTG AAC GTG	214-234	9	Silfverberg-Dilworth *et al.*, 2006
A485	DR993043	PET-CAC GAG GGT AAG CTC CCC	TTG GGG TTA TTG CTC TGA CG	280-300	No position	***
A301	Z71981/MDN1GN	CTT GCA CTA GTG TGC TTT GG	PET-CTT GTT GGG ATT AAA TCC GGC	331-345	12	***
A290	CN864595	CTC TGC AAA CTA CCA CCG C	PET-TCC TCC TCA ACA GCG GGG	350-360	No position	***

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A261	CO541090	PET-CCT CGG CAT CCA CAA ATC G	GAG AAG ACA AAC AGA CAC CA	400-410	No position	**
A538	CH-Vf1	VIC-ATC ACC ACC AGC AGC AAA G	CAT ACA AAT CAA AGC ACA ACC C	129-174	1	Vinatzer *et al.*, 2004
A305	CN491050	VIC-AAT CAA TGG AGA AAC GTC TGC	AAA GGA AAC CGA CTT CAC CC	220-230	No position	***
A735	SAA735	TAT CAG ATT CGT GCC ACA GC	VIC-CTT TGA CAT AGA CCC TGT CC	284-295	No position	*,**
A212	CN496144	VIC-CTC AGA CTC CTG CTG CAC C	TAC TGC CTG GTG TTT CTT CC	303-349	No position	**
A435	CO867454	VIC-ACCGCTAAATGCTGTTCAGG	CTTCACTGTGTTAGCATTGGG	360-400	No position	***
A462	DR994153	VIC-CAC GAG GCG AAA CCG ATC	AGG TCC TCA GAA CCT GAG C	450-470	No position	***
Daleen A						
A768	Hi04g11	6-FAM-CAG AGG ATT ATC AAT TGG ACG C	AAA CTA TCT CCA GTT ATC CTG CTT C	108-150	11	Silfverberg-Dilworth *et al.*, 2006
A788	Hi04a05	6-FAM-GGC AGC AGG GAT GTA TTC TG	GTT TCA TGT CAA ATC CGA TCA TCA C	180-220	9	Silfverberg-Dilworth *et al.*, 2006
A722	SAA722	AAC TCG TTT GTC AGC AGA GG	6-FAM-GTG GAA TAT GAA CAA ATC ACG	240-320	No position	f*,**
A718	SAA718	TTA AAC TGC CAA ATT GCA CGG	6-FAM-GTT GGG TAT TTG CAT GGT GG	400-470	No position	*,**
A769	Hi22d06	VIC-CCC CGA GCT CTA CCT CAA A	CAT TAT GTT TCC GGT TTT TGG	115-140	2	Silfverberg-Dilworth *et al.*, 2006
A829	NZmsCO905522	VIC-CAG GGC ACT GAC AAA GAC AG	AAT TGG AGA TTT GCG GTG TC	155-172	16	Celton, pers comm.

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A785	Hi09a01	VIC-GAA GCA ACC ACC AGA AGA GC	GTT TCC CAT TCG CTG GTA CTT GAG	174-199	11	Silfverberg-Dilworth *et al.*, 2006
A793	Hi23g08	VIC-AGC CGT TTC CCT CCG TTT	GTT TGT GGA TGA GAA GCA CAG TCA	200-230	4	Silfverberg-Dilworth *et al.*, 2006
A724	SAA724	CTC TTC ATC TGA GAA TAC ACC	VIC-AGA CTC GAG TCA TCC ATA CC	240-315	No position	*,**
A736	SAA736	TTT GAT TGG ACC TGC AGT GG	VIC-TTA GCA GCT GCT TCA GTG TG	325-380	17	*,**
A802	NZmsCN879773	NED-CCC TCT GTT ACT TTG ACT CTT CTC	TGG TTT GGG TTG AAA ATG GT	125-195	1	Celton, pers comm.
A827	NZmsEB146613	PET-AGA GTT CCG TTC CCC TCT CT	GTG GAT TCG GAA ATG CAC TC	140-210	14	Celton, pers comm.
Daleen B						
A772	Hi02a09	6-FAM-ATC TCT AAG GGC AGG CAG AC	CTG ACT CTT TGG GAA GGG C	110-195	11	Silfverberg-Dilworth *et al.*, 2006
A687	EB132187	6-FAM-TCT CCC TCA CTC GAC GTT G	GTT GCA GGA AGG AGT GTC G	220-275	No position	*,**
A726	SAA726	AAC TTG CTG AGA GAG TAA TGG	6-FAM-CAA CCA AAG GGC CTG AAG C	460-510	No position	*,**
A773	Hi23b12	VIC-TGA GCG CAA TGA CGT TTT AG	GTT TCA GGC TTT CCC TTC AGT GTC	125-175	14	Silfverberg-Dilworth *et al.*, 2006
A797	Hi02d11	VIC-GCA ATG TTG TGG GTG ACA AG	GTT TGC AGA ATC AAA ACC AAG CAA G	176-285	14	Silfverberg-Dilworth *et al.*, 2006
A714	SAA714	GTC GAT GAT CTC TGC GAG G	VIC-AGC AAG CAA AGC ATC AGA TTG	325-395	No position	*,**
A810	NZmsEB142980	NED-CCA GTT GGT TAT ACA AAT CGC AAA G	CCT GAT CCT CAA AAT TAC AGC A	80-140	4	Celton, pers comm.

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A775	Hi08e06	PET-GCA ATG GCG TTC TAG GAT TC	GGT GGT GAA CCC TTA ATT GG	120-164	13	Silfverberg-Dilworth *et al.*, 2006
A783	Hi23d11b	PET-GAC AGC CAG AAG AAC CCA AC	GTT TAT TGG TCC ATT TCC CAG GAG	165-205	4	Silfverberg-Dilworth *et al.*, 2006
A728	SAA728	TTG CTG CTG TCT GTG TTT GC	PET-GTC TCG TCG AAA TCT TAA AGG	210-285	16	*,**
A712	EB 112897	NED-CAA ATC CAG TTC GAA GTT TGG	GTC TCC GCG TCC TTA AAC G	330-390	No position	*,**
Daleen C						
A776	Hi23d02	6-FAM-CCG GCA TAT CAA AGT CTT CC	GTT TGA TGG TCT GAG GCA ATG GAG	100-155	11	Silfverberg-Dilworth *et al.*, 2006
A796	Hi08c05	6-FAM-TCA TAT AGC CGA CCC CAC TTA G	GTT TCA CAC TCC AAG ATT GCA TAC G	180-260	14	Silfverberg-Dilworth *et al.*, 2006
A777	Hi23d06	VIC-TTG AAA CCC GTA CAT TCA ACT C	GTT TCA AGA ACC GTG CGA AAT G	140-175	9	Silfverberg-Dilworth *et al.*, 2006
A789a	Hi02b10	VIC-TGT CTC AAG AAC ACA GCT ATC ACC	GTT TCT TGG AGG CAG TAG TGC AG	177-270	16	Silfverberg-Dilworth *et al.*, 2006
A789b	Hi02b10	VIC-TGT CTC AAG AAC ACA GCT ATC ACC	GTT TCT TGG AGG CAG TAG TGC AG	200-260	16	Silfverberg-Dilworth *et al.*, 2006
A715	SAA715	AGC ATC AAG CCA ATC TTT AAG C	VIC-GTA TGC TCT TCT TCT TCA TGG	320-380	No position	*,**
A778	Hi15g11	NED-TGA CAT GCA TAG GGT TAC ATG C	GTT TGG GTT CGT AAT CGT TCT TGT G	80-192	16	Silfverberg-Dilworth *et al.*, 2006

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A794	Hi01c09	NED-AAA GGC GAG GGA TAA GAA GC	GTT TGC ACA TTT GAG CTG TCA AGC	193-250	14	Silfverberg-Dilworth *et al.*, 2006
A744	SAA744	TCT ACC AAT CGT TCA AAG TCC	NED-TTA TCA GCT TTC CGA ACC TTC	260-320	16	*,**
A771	Hi21e04	PET-TGG AAA CCT GTT GTG GGA TT	TGC AGA GCG GAT GTA AGT TG	110-160	14	Silfverberg-Dilworth *et al.*, 2006
A791	Hi02c06	PET-AGC AAG CGG TTG GAG AGA	GTT TGC AAC AGG TGG ACT TGC TCT	180-270	11	Silfverberg-Dilworth *et al.*, 2006
A754	SAA754	AGC TGA TGG CCA GAA CTG C	PET-GAG GGT CCA AGT TAC AAA GG	380-510	14	*,**
Daleen D						
A820	NZmsEB116209	6-FAM-AAA ATC CCA ATT CCA AAA CC	TTG GAG CAG TGA AAG ATT GG	100-140	9	Celton, pers comm.
A780	Hi08f05	6-FAM-GTG TGG GCG ATT CTA ACT GC	GTT TCC TTT ATT CTA AAC ATG CCA CGT C	142-170	2	Silfverberg-Dilworth *et al.*, 2006
A784	Hi08d09	6-FAM-AAC GGC TTC TTG TCA ACA CC	GTT TAC TGC ATC CCT TAC CAC CAC	171-220	16	Silfverberg-Dilworth *et al.*, 2006
A800	Hi12a02	6-FAM-GCA AGT CGT AGG GTG AAG CTC	GTT TAG TAT GTT CCC TCG GTG ACG	223-280	16	Silfverberg-Dilworth *et al.*, 2006
A781	Hi02h08	VIC-GCC ACT CAT ACC CAT CGT ATT G	GTT TGG CTG GGA ATA TAT GAT CAG GTG	140-185	16	Silfverberg-Dilworth *et al.*, 2006
A801	Hi02a07	VIC-TTG AAG CTA GCA TTT GCC TGT	TAG ATT GCC CAA AGA CTG GG	210-320	2	Silfverberg-Dilworth *et al.*, 2006
A756	SAA756	ACG CTA GGA GAG AGG AAC G	VIC-GAG CAT TCC GTA TTA AAT CCG	480-550	11	*,**
A806	NZmsEB107305	NED-AAC TTC CAA ACC CCA TCT CC	AGA GCA ACC TCA CCA TCT TCA	110-190	2	Celton, pers comm.

Table 2.1 continued

Primer number	Accession number	Forward primer (5' - 3')	Reverse primer (5' - 3')	Expected amplicon Size range (bp)	Linkage group	Primers references
A766	AG11	NED-CAG ACA ACC TCC TCA CCT CA	AGT GCC CTG AAA TCT GGA TG	195-220	1	Silfverberg-Dilworth *et al.*, 2006
A779	Hi04d10	PET-AAA TTC CCA CTC CTC CCT GT	GTT TGA GAC GGA TTG GGG TAG	140-200	6	Silfverberg-Dilworth *et al.*, 2006

A/number: New coding used for easy referral for apple primer pairs.

Predicted primer pairs designed by P. Hove*, M.K. Soeker** or M.M. van Dyk*** (Apple research group, University of the Western Cape).

pers comm.- personal communication

The map positions of published primers was based on Maliepaard *et al*. (1998), Liebhard *et al*. (2002), Liebhard *et al*. (2003a), Vinatzer *et al*. (2004), Silfverberg-Dilworth *et al*. (2006) and Celton *et al.* (2009)

2.5.3 Primer optimisation for megaplex PCR

2.5.3.1 Gradient PCR

Prior to amplifying DNA from the 'Golden Delicious' x Dietrich' population (section 2.5.1), the primer annealing temperature was optimised by testing a gradient of temperatures on a Mastercycler Gradient PCR® (Eppendorf, Hamburg, Germany). During the PCR run, the annealing temperature was varied from one tube to the next by 1ºC for 10 cycles. The obtained PCR products were then electrophoresed by 6% PAGE (section 2.5.3) to determine the tube with the best temperature conditions.

2.5.3.2 Megaplex PCR

Following optimisation (section 2.5.3.1), all PCR reactions (Table 2.2) were performed using a thermal cycler GeneAmp PCR System 2700 (ABI). The cycling profile used was as follows: initial denaturation for 15 min at 95ºC, followed by 40 cycles of {60s/94ºC, 60s/58ºC and 60s/72ºC}, and final extension for 30 min at 72ºC.

Table 2.2: Composition of a megaplex PCR reaction

Reagents	Volume (µl)
PCR master mix	2.5
DNA template	1
Megaplex primer mix (0.2 µM of each primer)	0.5
PCR grade H_2O	1
Total	5

2.5.4 DNA Electrophoresis

2.5.4.1 Agarose gel electrophoresis

The quality of the isolated DNA sample was assessed by resolution on a 0.8% agarose gel. The gel was prepared by adding 0.8 g of agarose to 100 ml of 0.5X TBE and dissolved by melting using heat. The molten gel was cooled down to 56°C. Prior to pouring the molten gel into a casting plate, 0.5 μg.ml^{-1} ethidium bromide was added and mixed, and the gel was allowed to solidify.

A total of 25 ng of the DNA sample was mixed by pipetting with 2 μl of 6X DNA loading buffer (section 2.2) and loaded onto the gel. The electrophoresis was conducted in 0.5X TBE (section 2.2) at 10 V cm^{-1} for 90 min. The gel was then visualised with UV transillumination using the Gel-Doc UV Transilluminator (Bio-Rad, Hercules, California, USA).

2.5.4.2 Polyacrylamide gel electrophoresis (PAGE)

The success of the DNA amplification was assessed by resolution on a 6% PAGE. The polyacrylamide gel was prepared by mixing 6.3 g of urea, 1.5 ml of 10X TBE, 2.25 ml of acrylamide: bis-acrylamide (40% 19:1), 120 μl of APS and 9 μl of TEMED. The volume was adjusted to 15 ml with dH_20.

The gel casters were assembled using the Mini-Protean III© Cell (Bio-Rad) apparatus according to the manufacturer's instruction. The 6% (v/v) polyacrylamide gel was allowed to polymerise for 10 min at room temperature. The DNA sample was mixed in six volumes of DNA loading dye (section 2.2), denatured for 5 min at 95ºC, cooled

down on ice and loaded into the wells. Electrophoresis was conducted at 15 V cm^{-1} in 1X TBE (section 2.2) for 60 min.

2.5.5 Silver staining

The electrophoresed polyacrylamide gel was first immersed in solution A (section 2.2) for 10 min with shaking, rinsed in dH_2O and then developed in solution B (section 2.2) until an appropriate visualization of the PCR product *versus* background was achieved. Images were acquired using a Gel-Doc UV transilluminator system (Bio-Rad).

2.5.6 Automated electrophoresis using ABI PRISM® 3130xl Genetic Analyzer

A total of 2 µl of the PCR product was mixed with 10 µl of Hi-Di formamide and 0.25 µl of internal standard (GeneScan™ LIZ500 size standard, ABI) and loaded into a 96 well plate specific for ABI prism analyses. The mixture was denatured for 2 min at 95ºC, snap-cooled on ice and analysed with the ABI PRISM® 3130xl Genetic Analyzer machine (ABI). Arrays of length 36 cm and polymer (POP7) were used for electrophoresis in this study. Electrophoresis was conducted at 15 kV and 165 µA for 3 h per plate. The laser power and current were set at 15 mW and 4.7 A, respectively. In addition, the oven and cell heater temperature were set at 60°C and 50°C, respectively.

2.5.7 Allele scoring using the *GeneMapper*® V4 software

Data from section 2.5.6 were analysed using the *GeneMapper*® V4 software (ABI). The expected PCR product (allele) size data for each primer (marker) were loaded into the panel manager. The microsatellite default (table setting) was used for sizing and

genotyping microsatellite data for each megaplex. Individual seedlings were genotyped for each specific locus using the *JoinMap*® V4 coding system (van Ooijen, 2006).

2.5.8 Genetic linkage map construction using the *JoinMap*® V4.0 software

Molecular markers were linked to linkage groups using the *JoinMap*® V4 software (Van Ooijen and Voorrips, 2006). Tables containing seedling genotypes (allele scoring data), as inferred by the fragment detection process described in section 2.5.7, were exported to Excel (Microsoft Office). Seedlings that had missing data points at 25% of loci tested were excluded from further analysis. The data was transposed and then imported directly into *JoinMap*® for the construction of genetic linkage maps for each parent.

Linkage groups were defined based on the recombination frequencies observed between marker pairs. A likelihood of the odds (LOD) threshold value of 6.0 was set to assign markers to the same linkage group. Markers that were excluded and/or with insufficient linkage to other marker(s) were assigned to separate linkage groups based on the strongest cross linkage information (SCL values).

The determination of marker order and distance between markers were performed separately for each group using Kosambi and regression mapping algorithms. At this point, markers causing insufficient linkage among markers within a group, and/or their incorporation into the linkage group resulted in a high mean chi-square (χ^2) value, were excluded. Their presence may indicate the occurrence of double crossover events, which are generally limited to one or two per chromosome during meiosis.

Integrated genetic linkage maps, combining segregating loci from two parentals, were constructed for the first filial (F1) progeny derived from a cross between the 'Golden Delicious' and 'Dietrich' cultivars, as well as separate parental genetic linkage maps. Graphical representation of the genetic linkage maps as well as the alignment of the maps was achieved using *MapChart*© 2.2 (Voorrips, 2002) and the reference maps from Silfverberg Dilworth *et al*. (2006), respectively. The linkage group numbering was performed in accordance with the published maps from Maliepaard *et al*. (1998). Further distortion analysis within the individual linkage groups was analysed and graphically illustrated using the Graphical Genotyping beta version (GGT V2.0).

2.5.9 Quantitative trait loci identification

2.5.9.1 Genetic linkage maps

Parental genetic linkage maps as well as the integrated genetic linkage map were construction as described in section 2.5.8. These linkage maps obtained for the mapping population derived from a cross between the 'Golden Delicious' and 'Dietrich' cultivars, were used for the identification of QTLs associated with fruit quality traits.

2.5.9.2 Phenotypic traits data

Phenotypic data described in section 2.4.2 were used for the identification of putative QTLs controlling firmness, juiciness, crispness, colour, stripness, size, acidity, form and russeting fruit quality traits. The data obtained during year I, II and III, and the overall proportion of the population representation over the three years were treated separately in order to identify QTLs that were consistent over the period of investigation.

2.5.9.3 Identification of QTLs using *MapQTL*® V5 software

QTLs were identified using the *MapQTL*® V5 software (van Ooijen, 2004). QTL analyses were first performed separately for each of the parental maps and then for the integrated linkage map. The phenotypic, genotypic and genetic linkage map data were imported into the *MapQTL*® program. Interval mapping (IM) was performed for each year the 'Golden Delicious' x 'Dietrich' population was subjected to phenotypic assessment. Due to the relatively small number of individuals both genotyped and phenotyped, as well as data inconsistencies, a putative QTL was declared significant only when the LOD threshold was equivalent or above to the genome wide (GW) in at least one year.

The IM was performed after a permutation test, a method for establishing the significance of the likelihood ratio statistic (LRS) generated by single-locus association (section 1.5.2.2; Churchill and Doerge, 1994; Doerge and Rebaï, 1996). In this test, a total of 1,000 permutations were performed to obtain a distribution of LRS values expected when no QTL linked to any of the marker loci. The LRS observed was statistically compared to the LRS expected. The genome wide threshold at 95th percentile (p-value at $\propto = 0.05$) points of the empirical distribution was used rather than individual linkage group thresholds (Manly and Olson, 1999). Therefore, the upper boundary value was considered as the significance threshold during interval and multiple-QTL models (MQM) mapping.

Interval mapping was performed for every year of phenotypic data collection. Because of the relatively small number of genotyped individuals used for phenotypic assessment and data inconsistencies of individuals in this population, a putative QTL was declared

significant only when the LOD threshold was equivalent or above to the genome wide in at least one year of investigation.

MQM mapping was performed in order to detect possible QTLs not identified with IM and to eliminate false positive QTLs. Therefore, markers with LOD values close to GW were used as cofactors for MQM mapping analysis (van Ooijen, 2004).

QTLs were graphically depicted on integrated maps as bars indicating the confidence interval (CI) of 5%, using the *MapChart*© 2.2 software. The 5% CI corresponded to a LOD score of +/-0.5 of the likelihood peak. QTLs that were identified in at least two out of the three years and/or at least on two maps were considered similar when overlapping at 5% CI. The size of a confidence interval is expected to be inversely proportional to the number of progeny in the mapping population and inversely proportional to the square of the strength of the QTL (Darvasi *et al.*, 1993).

2.5.9.4 Nonparametric mapping using Krustal-Wallis

The nonparametric mapping function (the Kruskal-Wallis test) of *MapQTL*® 5 was used to identify markers in which different genotypes can be associated with the fruit quality traits under investigation. The Krustal-Wallis test identified markers by comparing differences between genotypes and phenotypes and then ranking individuals according to their quantitative trait as well as their marker genotype. SSR markers and more specifically alleles at these loci that are good candidates for MAS were identified.

2.6 PROTEOMICS METHODS

2.6.1 Protein extraction from fruit pulp

2.6.1.1 Extraction by TCA/acetone precipitation

Prior to extracting total soluble proteins (TSP) from mature apple fruit pulp (section 2.4.2), the fruit skin was peeled off using a commercial fruit peeler. Approximately 10 g of frozen fruit pulp were ground to a fine powder in liquid nitrogen using a sterile mortar and pestle and suspended in 10 ml of 10% (w/v) TCA (section 2.2) containing 0.06% sodium sulphite. The suspension was vortexed for 1 min, incubated for 4 h on ice and centrifuged at 4,000 × *g* for 10 min at 4°C. The supernatant was discarded and the pellet washed with 10 ml of 80% (v/v) ice-cold acetone and collected by centrifugation at 4,000 × *g* for 10 min at 4°C. This washing step was repeated two more times to obtain a firm pellet. The pellet was then air-dried for 10 min at room temperature, before resuspending the pellet in approximately 3 ml of extraction buffer (section 2.2) by vortexing overnight at room temperature. The following day, the debris from the TSP was removed by centrifuging at 4,000 × *g* for 20 min at room temperature and used immediately for further analyses or stored as aliquots at -20°C.

2.6.1.2 Extraction by phenol precipitation

The TSP were also extracted from mature apple fruit pulp following a modified SDS/phenol extraction protocol (Wang *et al.*, 2006). Approximately 10 g of frozen fruit pulp were ground to a fine powder in liquid nitrogen with a sterile mortar and pestle and suspended in 10 ml of 10% (w/v) TCA (section 2.2) containing 50 mM PVPP. The suspension was vortexed and centrifuged at 4,000 × *g* for 10 min at 4°C. The supernatant was decanted and the pellet was washed twice with 80% (v/v) methanol containing 0.1 M ammonium acetate and then once with 80% (v/v) acetone. The protein

pellet was collected by centrifugation at 4,000 × *g* for 10 min at 4°C, air-dried for 10 min and resuspended in 6 ml of phenol/SDS buffer by vortexing. After incubation for 10 min at room temperature, the homogenate was centrifuged at 4,000 × *g* for 10 min at room temperature. The upper phenol phase was collected and precipitated with five volumes of 0.1 M ammonium acetate in acetone for 2 h at -20°C. Proteins were collected by centrifugation at 4,000 × *g* for 20 min at 4°C and washed first with 5 ml of 100% (v/v) methanol and subsequently with 5 ml of 80% (v/v) acetone and collected by centrifugation at 4,000 × *g* for 10 min at 4°C. After air-drying, the proteins were resuspended in 200 µl of extraction buffer (section 2.2) and used immediately for further analyses or stored as aliquots at -20°C.

2.6.2 Protein quantification

A standard curve (Figure 2.1) was generated using increasing concentrations of BSA ranging from 0 to 50 µg. A stock solution of 5 $mg.ml^{-1}$ was mixed with 10 µl of 0.1 M HCl, 80 µl of dH_2O, 900 µl of Bradford reagent and extraction buffer to a final volume of 1 ml. The absorbance of the mixture was measured at 595 nm, using a Genesis 5 Spectrophotometer (Milton Roy, Groton, CT, USA).

The concentrations of TSP extracts (sections 2.6.1.1 and 2.6.1.2) were determined using the modified Bradford assay (Bradford, 1976), as described by Ndimba *et al*. (2003). A total of 5 µl of sample (TSP extracts) was mixed with 10 µl of 0.1 M HCl, 80 µl of dH_20, 5 µl of extraction buffer and 900 µl of 20% (v/v) Bradford protein dye (section 2.2). The absorbance of the mixture was measured at 595 nm. The concentrations of the TSP extracts were then obtained by extrapolating from the bovine serum albumin (BSA) standard curve.

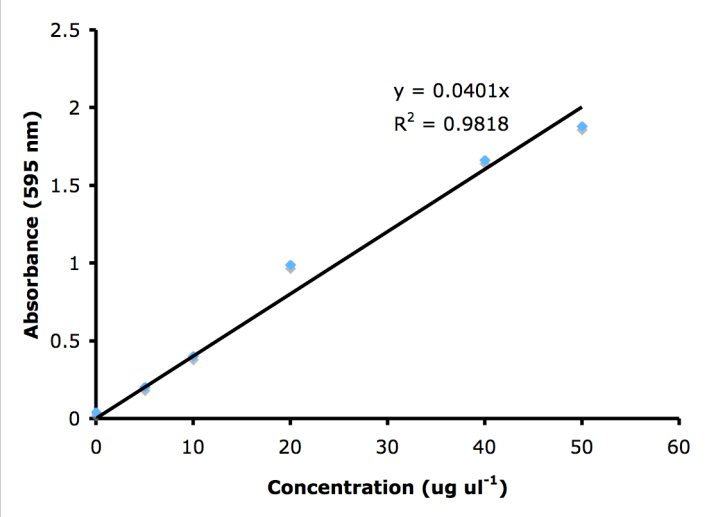


Figure 2.2: A typical Bovine serum albumin (BSA) standard curve for protein concentration determination.

2.6.3 One-dimensional polyacrylamide gel electrophoresis (1D-PAGE)

2.6.3.1 Sample preparation

Protein samples were mixed with an equal volume of SDS loading dye (section 2.2), heated for 5 min at 95°C and centrifuged at 16,000 × *g* for 2 min at room temperature before loading onto the gel (section 2.6.3.2). A total of 20 µg of protein were loaded per lane. One lane was loaded with the pre-stained protein molecular weight marker (section 2.2).

2.6.3.2 Linear sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated according to their molecular weight by 1D-PAGE as described by Laemmli (1970). The method involved preparation of stacking and resolving gels using 19:1 acrylamide: bis-acrylamide. The 12% resolving gel and the 4% stacking gel

were prepared as shown in Table 2.3 using the Mini-Protean III® Cell gel casting system (Bio-Rad). The resolving gel was allowed to polymerize for 20 min and the stacking gel for 10 min, both at room temperature. The running buffer was prepared as described in section 2.2. Protein samples were prepared as described in section 2.6.3.1 and 20 µg total protein was loaded per lane. Electrophoresis was conducted at 10 V cm^{-1}, at room temperature until the dye reached the bottom of the gel (usually 90 min).

Table 2.3: Composition of 12% resolving and 4% stacking gels for 1D-PAGE

Reagents	12% Resolving gel (ml)	4% Stacking gel (ml)
dH_2O	2.25	1
4X Resolving gel buffer	1.25	-
2X Staking gel buffer	-	1.25
19:1 acrylamide: bis-acrylamide	1.5	0.25
10% (w/v) ammonium persulphate	0.05	0.025
TEMED	0.005	0.0025

2.6.4 Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE)

2.6.4.1 Sample preparation

Total soluble proteins extracted from mature apple fruit pulp (section 2.6.1) stored at -20°C were thawed on ice and protein concentrations were determined by Bradford assay (section 2.6.2). Proteins (200 µg) were diluted in isoelectric focusing (IEF) rehydration buffer (section 2.2) and mixed with 2 µl of ampholytes (pH 3-10) to a total volume of 125 µl. The sample was then vortexed and centrifuged at 16,000 × *g* for 2 min at room temperature.

2.6.4.2 Rehydration of immobilized pH gradient strips

Dry polyacrylamide immobilized pH gradient (IPG) strips that are 7 cm in length with a pH gradient of 4-7 were used. The rehydration buffer (125 µl) containing the protein sample was loaded onto the rehydration tray and after careful removal of the protecting cover foil, the IPG strip was gently positioned on top of it, avoiding trapping air bubbles. Each strip was covered with mineral oil to prevent dehydration and allowed to passively rehydrate to their original thickness (0.5 mm) overnight at room temperature.

2.6.4.3 Isoelectric focusing (IEF)

During the first dimension, proteins are separated according to their p*I*, the pH at which a protein carries no net charge and will not migrate any further in an electrical field (O' Farrell, 1975; Klose, 1975). After overnight rehydration (section 2.6.4.2), the IPG strips were removed from the rehydration tray, rinsed with dH_2O and loaded onto an Ettan™ IPGphor II IEF machine (GE Healthcare, Bio-Sciences AB, Uppsala, Sweden), gel side facing upward. Prior to mounting the electrode pads, wicks, pre-damped with dH_2O were placed on each end of the strip to absorb excess salt. The IPG strips were overlaid with mineral oil and the focusing was performed at 20°C with the running conditions described in Table 2.6. At the end of the run, the strips were placed in an equilibration tray for equilibration (section 2.6.4.4) prior to the second dimension resolution by SDS-PAGE.

Table 2.6: Running conditions for protein focusing using IPG strips

Step and voltage mode	Voltage (V)	Duration
1: Step and Hold	250	0:15 h
2: Step and Hold	3,800	1:00 h
3: Step and Hold	4,000	10,000 vhr

2.6.4.4 Equilibration

The IPG strips were conditioned for the second dimension SDS-PAGE in a two-step equilibration in order to reduce disulphide bonds and alkylate the resultant sulfhydryl groups of the cysteine residues. Concurrently, the proteins were coated with SDS for separation on the basis of their mass. During the first equilibration step, strips were placed in an equilibration tank gel side up and incubated with 2 ml of equilibration buffer I (section 2.2) for 20 min with gentle shaking. Buffer I was discarded prior to addition of 2 ml of equilibration buffer II (section 2.2) and incubation for 20 min with gentle shaking.

2.6.4.5 Second dimension of 2D-PAGE by SDS-PAGE

In the second dimension, proteins resolved on IPG strips were applied to a 12% SDS-PAGE and separated by molecular mass. The gels were prepared in a similar manner as described in section 2.6.3.2, but without stacking gel. The strips were briefly rinsed with SDS-PAGE running buffer (section 2.2), carefully loaded on top of the resolving gels and overlaid with molten agarose sealing solution (section 2.2) to secure their position. Electrophoresis was conducted at 5 V cm^{-1} for 10 min and then at 10 V cm^{-1} at room temperature until the tracking dye reached the bottom of the gel (usually 90 min). The gels were then stained and imaged as described in section 2.6.5.

2.6.5 Detection of proteins by gel staining with Coomassie brilliant blue (CBB)

The 1D-PAGE (section 2.6.3.2) and 2D-PAGE (section 2.6.4.5) separated proteins were detected by CBB staining. Gels were immersed first in CBB staining solution I (section 2.2) for 1 h with gentle shaking at room temperature. The gel was then incubated in staining solution II (section 2.2) and finally in CBB staining solution III (section 2.2) for 1 h each with gentle shaking at room temperature. The gels were then destained in CBB destaining solution (section 2.2) until an appropriate visualization of protein against background was achieved, thus leaving only the stain linked to the protein spots. Images were acquired using the PharosFX™ Plus molecular imager scanner (Bio-Rad).

2.6.6 Comparative analysis with PDQuest™

Comparative analysis of 2D-PAGE from the high and the low firmness phenotypes was carried out using the PDQuest™ V8.1 software. Four well-separated gels of each physiological state were used to create 'replicate groups' and a 'master gel' consisting of all the spots from the gels combined. To minimize experimental variations, spot intensity was normalized on the basis of the total integrated optical density for the master gel. After automated detection and matching of spots, manual editing was performed for every identified spot. Statistic, quantitative and qualitative 'analysis sets' were created by comparing the two phenotypes. In the statistic set, the Student *t*-test with the significance level of 95% was selected. In the quantitative set, the upper and lower limits were set to 2.0 and 0.5, respectively. In the qualitative set, the detection limit of a spot versus background was set to 10-fold. Then, the Boolean analysis sets were created between the statistic sets and the quantitative or qualitative sets. The spots from the Boolean sets were compared among the four biological replicates. Only spots

displaying reproducible change patterns in at least three of the replicate gels were considered as differentially expressed.

2.6.7 In-gel trypsin digestion

Spots of interest were excised from the gel manually and transferred into sterile microcentrifuge tubes. The gel pieces were washed twice with 500 µl of 50 mM ammonium bicarbonate for 5 min each time and a third time for 30 min with occasional vortexing. The gel pieces were then destained twice with 500 µl of 50% (v/v) of 50 mM ammonium bicarbonate and 50% (v/v) acetonitrile for 30 min with occasional vortexing. These were then dehydrated with 100 µl of 100% (v/v) acetonitrile for 5 min, and completely dessicated using the Speed Vac SC100 (ThermoSavant, Waltham, MA, USA). The protein in the gel was resuspended and in-gel digested with 120 ng sequencing grade modified trypsin solution for 6 h at 37°C. The digested proteins were then stored at 4°C.

2.6.8 Protein identification using MALDI-TOF mass spectrometry

Prior to identification of digested proteins, 1 µl of each sample was mixed with the same volume of α-cyna-hydroxy-cinnamic (CHCA) matrix and spotted onto a MALDI target plate for analysis by matrix assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry (MS) using a Voyager DE Pro Biospectrometry workstation (ABI) to generate a peptide mass fingerprint (PMF). The MALDI-TOF was operated in the positive ion delayed extraction reflector mode for highest resolution and mass accuracy. Peptides were ionized with a 337 nm laser and spectra were acquired at 20 kV acceleration potential with optimised parameters. Close external calibration was

employed using the Sequazyme calibration mixture II (ABI) containing angiotensin I, ACTH (1-17 clip), ACTH (18-39 clip) and bovine insulin. This calibration method typically provided mass accuracy of 100 to 200 ppm across the mass range 900 to 5,000 Da. Peptide spectra of accumulated 1,200 shots each were automatically processed for baseline correction, noise removal, and peak deisotoping. The threshold was manually adjusted to 2-8% base peak intensity. All searches were performed against the National Center for Biotechnology Information (NCBI) and Mass Spectrometry DataBase (MSDB) peptide mass databases using Mass Spectroscopy and Proteomics (MASCOT) (http://www.matrixscience.com/search_form_select.html). Candidate identifications with molecular weight search (MOWSE) scores higher than 85 were automatically considered as positive assignments. For all other assignments of protein spots, the MOWSE score cut off threshold was set to 64. Additionally, positive protein assignments required greater than 10% sequence coverage. If more than one protein satisfied mentioned threshold criteria, the entry with the highest MOWSE score was assigned. Identified proteins were subsequently validated against the *Malus* EST database from NCBI using full protein sequences and matched peptide sequences to determine similarities. The search criteria required match of at least four peptides from PMF for successful validation.

CHAPTER 3

PHENOTYPIC DATA ANALYSIS

3.1 INTRODUCTION

The phenotype is the descriptor of the 'phenome', which is the manifestation of all the physical properties of a given organism, its physiology, morphology and behaviour (Mahner and Kary, 1997). Phenotypic analyses are necessary to link the 'visible characteristic', which is a reflection of the gene function, to the genome structures, and thus to identify genes influencing the trait of interest (Williams and Conner, 2001). They are a vital component for evolution by natural selection, and together with genetic variation, they contribute to the characteristics of the next generation. Phenotypic variations form an important adaptation mechanism for survival and are promoted in a population (Williams and Conner, 2001).

The collection of phenotypic data is an essential step towards quantitative trait analysis. Phenotypic data is descriptive and vary with the range of mutant phenotypes presented by the population under study (Drysdale, 2001; Clare, 2005). Sample tracking, quality control and data entry with a rational database, which are preferred in the case of large-scale studies, are crucial factors to consider during phenotypic data collection. Changes in environmental conditions, data collection technique and data entry errors often result in uncontrolled variability. However, this can be minimized by developing and adhering to standardised procedures, like routine monitoring of the electroconductivity of penetrometer when assessing fruit firmness. Also proper technical training prior to data collection and monitoring the effectiveness of training program by duplicating data sets are necessary to ensure comparable results (Kjemtrup *et al.,* 2003).

This chapter, thus relate to the analysis of the phenotypic data in order to assess the influence of the environment, the phenotype and the combined effect of both the environment and the genotype on the quality traits under investigation. The variation of the data for each trait was assessed statistically using either the student *t*-test or analysis of variance (ANOVA) test, where appropriate. Further, post hoc test using Turkey's Honestly Significant Difference (HSD test) was used to validate the significance of the ANOVA test. It is important to note that HSD test was only performed on ANOVA results found to have a significant effect.

3.2 RESULTS

Sensory tests were performed to determine fruit quality, using mature fruits from the 'Golden Delicious' x 'Dietrich' population. Nine traits were investigated and these include firmness, juiciness, crispness, colour, acidity, russeting, stripness, form and size. For every trait three apples were used per individual tree and three measurements were taken per apple (section 2.4.2). These tests were carried out in three consecutive years.

3.2.1 Analysis of phenotypic data representation

In order to assess if the population was well represented and thus the genetic variability, the proportion of individuals that fruited was determined for every year of investigation. A proportion of 43.5%, 34.1% and 50.4% out of the 248 individuals of the population did bear fruits in year I, II and III, respectively (Figure 3.1). A total of 71.6% of trees did bear fruits at least in one of the three years of phenotypic investigation.

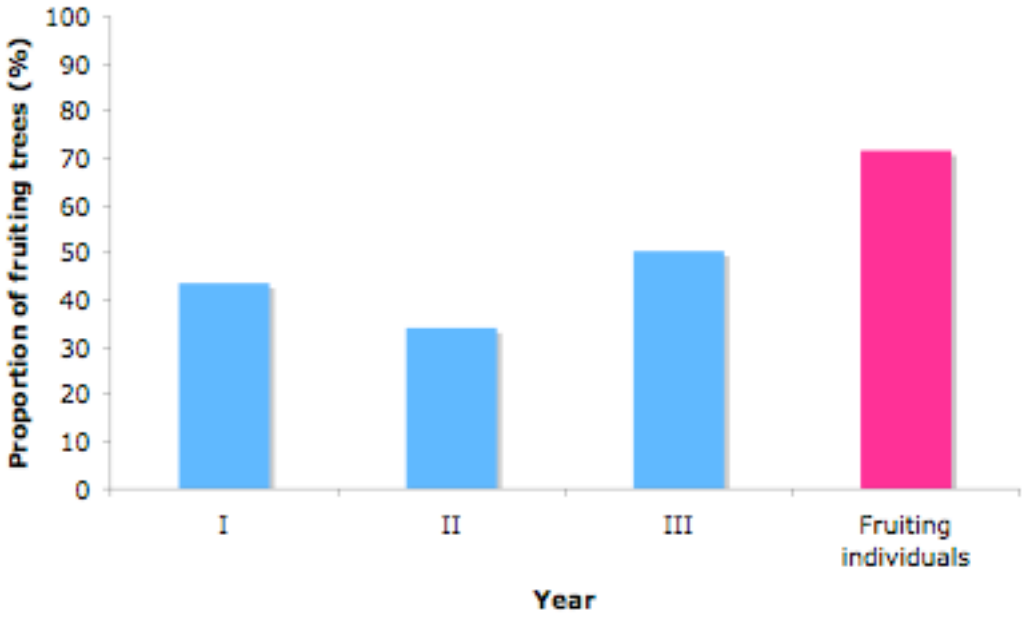


Figure 3.1: Proportion of fruiting individuals in the 'Golden Delicious' x 'Dietrich' mapping population.

Mature fruits were harvested from individuals of the 'Golden Delicious' x 'Dietrich' mapping population over three consecutive years and used for phenotypic analysis. The proportion of fruiting individuals during the first, second and third year of investigation is represented by the bars labelled I, II and III, respectively. The pink bar represents the proportion of individuals that fruited at least once over the three-year period.

3.2.2 Investigation of the effect of the environment on the quality traits

In order to assess the influence of the environment on the fruit quality traits, the raw data variations for each trait among the three years of investigation were analysed as shown on Table 3.1. Table 3.1 shows a representation of how data was analysed for all the traits under investigation. The means of the phenotypic data collected over a period of three years, for each seedling and trait were computed and summarised

Table 3.1: Representative subset of phenotypic raw data for stripness, collected from mature apples of the 'Golden Delicious' x 'Dietrich' population over a period of three years. Std dev stands for standard deviation.

Sensory trait	Stripness					
Seedling	Year I	Year II	Year III	Mean	Std dev	Std error
1-2	-	-	-	-	-	-
1-3	3	-	12.5	5.2	6.7	3.9
1-4	-	5	31	12	18.4	10.6
1-5	-	-	63	21	-	-
1-6	-	4.5	-	1.5	-	-
1-7	1.5	-	-	0.5	-	-
1-8	2	-	83.5	28.5	57.6	33.3
1-9	97	-	14	37	58.7	33.9
1-10	2.5	-	3.5	2	0.7	0.4
1-11	2	3	-	1.7	0.7	0.4
1-12	-	-	-	-	-	-
1-14	70.5	-	-	23.5	-	-
1-15	47.5	-	70	39.2	15.9	9.2
1-17	45.5	-	28	24.5	12.4	7.1
1-18	4	-	87	30.3	58.7	33.9
1-19	-	3	81	28	55.2	31.8
1-20	89.5	2.5	19.5	37.2	46.1	26.6
1-21	-	-	73	24.3	-	-
1-22	-	3	12	5	6.4	3.7
1-23	-	-	88	29.3	-	-
1-24	38	86	-	41.3	33.9	19.6
1-25	-	38	23	20.3	10.6	6.1
1-26	50	-	89	46.3	27.6	15.9
1-28	-	-	-	0	-	-
1-29	-	84	-	28	-	-
1-31	2.5	-	68	23.5	46.3	26.7
1-32	-	-	89.5	29.8	-	-
1-33	6.5	85.5	-	30.7	55.9	32.3
1-35	86.5	-	-	28.8	-	-
1-38	-	65	72.5	45.8	5.3	3.1
1-39	96	-	57	51	27.6	15.9
1-40	30	-	10.5	13.5	13.8	8
1-41	-	77.5	27.5	35	35.4	20.4
1-42	-	-	31	10.3	-	-
1-45	-	-	-	-	-	-
1-47	82.5	66	-	49.5	11.7	6.7
1-48	46	-	-	15.3	-	-
1-49	4	15.5	81	33.5	41.5	24
1-50	88.5	-	72.5	53.7	11.3	6.5
1-52	91.5	-	86.5	59.3	3.5	2
1-53	93.5	63.5	86	81	15.6	9

Seedling	Year I	Year II	Year III	Mean	Std dev	Std error
1-54	-	-	80.5	26.8	-	-
1-55	-	-	-	-	-	-
1-57	-	-	-	-	-	-
1-58	-	-	74	24.7	-	-
1-59	-	46.5	-	15.5	-	-
1-60	-	-	-	-	-	-
1-61	67	-	-	22.3	-	-
1-62	-	-	-	-	-	-
1-63	44	-	90.5	44.8	32.9	19
1-64	94.5	87.5	85	89	4.9	2.8
1-65	-	-	-	-	-	-
1-67	2.5	34.5	13	16.7	16.3	9.4
1-68	-	-	83.5	27.8	-	-
1-69	-	-	-	-	-	-
1-70	11	-	87.5	32.3	54.1	31.2
1-71	-	-	-	-	-	-
1-72	-	-	42	14	-	-
1-73	6	-	69.5	25.2	44.9	25.9
1-74	-	-	-	-	-	-
1-75	16.5	-	25	13.8	6	3.5
1-76	-	-	-	-	-	-
1-77	-	-	-	-	-	-
1-78	8.5	-	82	30.2	52	30
1-79	-	9	36.5	15.2	19.4	11.2
1-80	-	-	-	-	-	-
1-81	8	42	12	20.7	18.6	10.7
1-83	-	88.5	19	35.8	49.1	28.4
1-84	-	-	-	-	-	-
1-85	1	-	-	0.33	-	-
1-86	1.5	-	-	0.5	-	-
1-87	-	-	-	-	-	-
1-88	-	-	-	-	-	-
1-89	-	-	-	-	-	-
1-92	46	-	32	26	9.9	5.7
1-93	-	13.5	39	17.5	18	10.4
1-94	2	-	-	0.7	-	-
1-95	-	-	-	-	-	-
1-96	-	46	-	15.3	-	-
1-97	-	53	86	46.3	23.3	13.5
1-98	94.5	59.5	14.5	56.2	40.1	23.2
1-100	-	-	-	-	-	-
1-101	95	-	44	46.3	36.1	20.8

Table 3.I continued

Seedling	Year I	Year II	Year III	Mean	Std dev	Std error
1-102	-	-	15.5	5.2	-	-
1-103	11.5	27	-	12.8	11	6.3
1-104	-	-	-	-	-	-
1-105	-	-	-	-	-	-
1-106	-	-	-	-	-	-
1-107	-	-	19	6.3	-	-
1-109	-	46	83	43	-	-
1-110	-	-	-	-	-	-
1-111	-	-	-	-	-	-
1-112	-	-	-	-	-	-
1-113	-	47	80.5	42.5	23.7	13.7
1-114	-	-	-	-	-	-
1-115	-	-	-	-	-	-
1-118	91	-	28	39.7	44.5	25.7
1-120	-	41	50.5	30.5	-	-
1-121	-	-	-	7.7	-	-
1-122	-	23	-	7.7	-	-
1-123	-	-	-	-	-	-
1-124	77.5	-	43.5	40.3	-	-
1-125	-	79	84.5	54.5	3.9	2.2
1-126	-	-	-	-	-	-
1-127	-	94.5	39.5	44.7	38.9	22.5
1-128	-	20.5	18	12.8	1.8	1
1-129	-	-	-	-	-	-
1-130	-	4	86	30	58	33.5
1-131	-	-	-	-	-	-
1-132	53.5	46.5	67	55.7	10.4	6
1-134	80.5	-	83	54.5	1.8	1
1-135	75.5	86.5	83.5	81.8	5.7	3.3
1-136	4	1.5	87	30.8	48.7	28.1
1-137	3	-	-	1	-	-
1-138	-	-	-	-	-	-
1-139	45	-	24.5	23.2	14.5	8.4
1-142	-	-	-	-	-	-
1-143	-	-	-	-	-	-
2-41	2	4.5	10	5.5	4.1	2.4
2-43	-	-	-	-	-	-
2-44	94	66	-	53.3	19.8	11.4
2-45	48.5	-	85	44.5	25.8	14.9
2-46	48	87	43.5	59.5	23.9	13.8
2-47	-	-	-	-	-	-
2-48	47.5	-	66.5	38	13.4	7.8
2-49	-	-	-	-	-	-
2-50	8.5	-	11.5	6.7	2.1	1.2
2-51	-	-	-	-	-	-
2-52	9.5	-	-	3.2	-	-
2-53	-	-	-	-	-	-

Seedling	Year I	Year II	Year III	Mean	Std dev	Std error
2-54	1	-	-	0.3	-	-
2-55	12.5	91	13.5	39	45	26
2-56	85	-	85	56.7	0	0
2-57	98	46	87.5	77.2	27.5	15.9
2-58	22.75	12.5	-	11.8	7.2	4.2
2-59	11.5	-	76.5	29.3	46	26.5
2-60	-	-	85.5	28.5	-	-
2-61	1	-	36	12.3	24.7	14.3
2-62	-	-	-	-	-	-
2-63	-	-	-	-	-	-
2-64	90.5	1	17	36.2	47.7	27.6
2-65	97	75	44	72	26.6	15.4
2-66	2	3.5	-	1.8	1.1	0.6
2-67	-	-	-	-	-	-
2-68	2	10.5	85.5	32.7	46	26.5
2-69	-	-	-	-	-	-
2-70	-	-	15.5	5.2	-	-
2-71	15	49	43.5	35.8	18.3	10.5
2-72	-	-	-	-	-	-
2-73	-	-	-	-	-	-
2-74	-	-	-	-	-	-
2-75	91	90	15.5	65.5	0.7	0.4
2-76	-	-	-	-	-	-
2-77	-	8.5	42	16.8	23.7	13.7
2-78	-	-	-	-	-	-
2-79	-	-	-	-	-	-
2-80	-	-	-	-	-	-
2-81	-	27	39.5	22.2	8.8	5.1
2-82	29	-	21.5	16.8	5.3	3.1
2-83	48	46	-	31.3	1.4	0.8
2-84	-	-	-	-	-	-
2-85	-	-	-	-	-	-
2-86	8	-	39.5	15.8	-	-
2-87	70	-	-	23.3	-	-
2-88	-	73	86.5	53.3	9.5	5.5
2-89	42.5	13.5	5.5	20.5	19.5	11.2
2-90	-	-	-	-	-	-
2-91	-	-	-	-	-	-
2-93	93	-	-	31	-	-
2-94	2.5	30.5	44	25.7	21.2	12.2
2-95	96	89	71	85.3	12.9	7.4
2-96	30.5	-	-	10.2	-	-
2-97	-	-	-	-	-	-
2-98	85	-	84	56.3	0.7	0.4
2-99	-	-	69	23	-	-
2-100	61	-	45.5	35.5	11	6.3
2-101	-	-	11	3.7	-	-

Table 3.1 continued

Seedling	Year I	Year II	Year III	Mean	Std dev	Std error
2-102	-	-	42	14	-	-
2-104	-	-	45	15	-	-
2-107	48.5	-	29.5	26	13.4	7.8
2-108	95	96	-	63.7	0.7	0.4
2-109	-	-	43.5	14.5	-	-
2-110	-	82	34.5	38.8	33.6	19.4
2-111	-	-	43	14.3	-	-
2-112	-	-	86	28.7	-	-
2-113	-	-	70	23.3	-	-
2-114	-	-	43	14.3	-	-
2-115	-	-	-	-	-	-
2-116	94	45	39.5	59.5	30	17.3
2-117	-	-	-	-	-	-
2-118	88.5	93	83.5	88.3	4.8	2.7
2-119	93	-	-	31	-	-
2-120	25.5	37.5	87.5	50.2	32.9	19
2-121	-	70	21	30.3	34.6	20
2-122	2	3.5	27	10.8	14	8.1
2-124	79	74	-	51	3.5	2
2-125	28.5	37.5	-	22	6.4	3.7
2-126	75.5	-	35	36.8	28.6	16.5
2-127	-	46.5	-	15.5	-	-
2-128	92.5	72	85	83.2	10.4	6
2-129	-	94.5	44.5	46.3	35.4	20.4
2-130	-	-	-	-	-	-
2-131	96.5	95	87.5	93	4.8	2.8
2-132	47.5	46	-	31.2	1.1	0.6
2-133	-	-	-	-	-	-
2-134	-	70	-	23.3	-	-
2-136	-	7.5	81	29.5	52	30
2-137	-	16	-	5.3	-	-
2-138	-	-	-	-	-	-
2-139	-	95	17.5	37.5	54.8	31.6
2-141	-	-	-	-	-	-
2-142	-	91	-	30.3	-	-
2-150	-	-	-	-	-	-
3-23	-	-	-	-	-	-
3-41	97	88	41	75.3	6.4	3.7
3-44	47	-	85	44	-	-
3-45	-	79	-	26.3	-	-
3-46	4.5	-	-	1.5	-	-
3-49	2	12	40.5	18.2	20	11.5
3-50	75	69.5	-	48.2	3.9	2.2
3-51	2	-	45	15.7	30.4	17.6
3-52	-	-	-	-	-	-
3-53	44.5	85.5	-	43.3	29	16.7
3-54	17	-	44.5	20.5	19.4	11.2
3-55	-	-	-	-	-	-
3-56	96	94.5	46	78.8	28.4	16.4
3-57	9	7	9.5	8.5	1.3	0.8
3-58	78.5	88.5	81.5	82.8	5.1	3
3-59	-	-	-	-	-	-
3-60	44	-	6.5	16.8	26.5	15.3
3-61	-	-	-	-	-	-
3-62	-	-	-	-	-	-
3-63	-	-	-	-	-	-
3-64	47.5	-	77.5	41.7	21.2	12.2
3-65	-	-	-	-	-	-
3-66	-	-	68.5	22.8	-	-
3-67	-	-	-	-	-	-
3-68	2	-	86	29.3	59.4	34.3
3-69	-	-	-	-	-	-
3-70	-	-	-	-	-	-
3-71	-	95	28	41	47.4	27.4
Average	**45.7**	**50.7**	**52.4**	**49.6**	**22.7**	**13.1**

The means of the phenotypic data collected over a period of three years, for each seedling and trait were further computed and summarised (Table 3.2). The data was then graphically represented in order to monitor the trend of events (Figure 3.2). In terms of stripness, acidity, russeting and size, an increase in the proportion of these fruit quality traits was observed between year I and year III. This increase was greater between year I and year II than between year II and year III. For stripness and size, this variation in proportion was small. For stripness, the proportion rose from 45.7% in year I to 50.7% in year II to 52.4% in year III, and for size from 42.2% in year I to 44.8% in year II to 45.2% in year III.

In terms of juiciness and firmness, a decrease in proportion of these fruit quality traits between year I and year III was detected. For juiciness, this decrease was greater between year I and year II than between year II and year III, declining from 63.1% in year I to 53.8% in year II to 49.1% in year III. On the contrary, for firmness, the reduction was greater between year II and year III than between year I and year II, declining from 69.0% in year I to 64.6% in year II to 56.2% in year III.

In terms of colour, the proportion increased between year I (62.5%) and year II (68.8%) and decreased between year II (68.8%) and year III (64.6%). In terms of form and crispness, the proportion was similar among the three years of investigation.

Table 3.2: Recapitulative table of the means of phenotypic data for nine fruit quality traits.

The means of phenotypic data of the nine fruit quality traits collected over a period of three years from mature apples from the 'Golden Delicious' x 'Dietrich' population and the standard deviation and standard error of the mean phenotypic data for each trait are represented.

Trait	Mean phenotypic data (%)			Standard deviation of the mean phenotypic data (%)	Standard error of the mean phenotypic data (%)
	Year I	Year II	Year III		
Stripness	45.7	50.7	52.4	22.7	13.1
Colour	62.5	68.8	64.6	13.1	7.6
Crispness	63.2	61.6	61.6	13.4	7.7
Juiciness	63.1	53.8	49.1	12.1	7.0
Acid	33.5	42.0	42.0	10.8	6.2
Size	42.2	44.8	45.2	7.6	4.4
Form	48.8	50.0	48.3	10.8	6.2
Russeting	56.0	67.3	70.3	13.8	7.9
Firmness	69.2	64.6	56.3	11.9	6.9

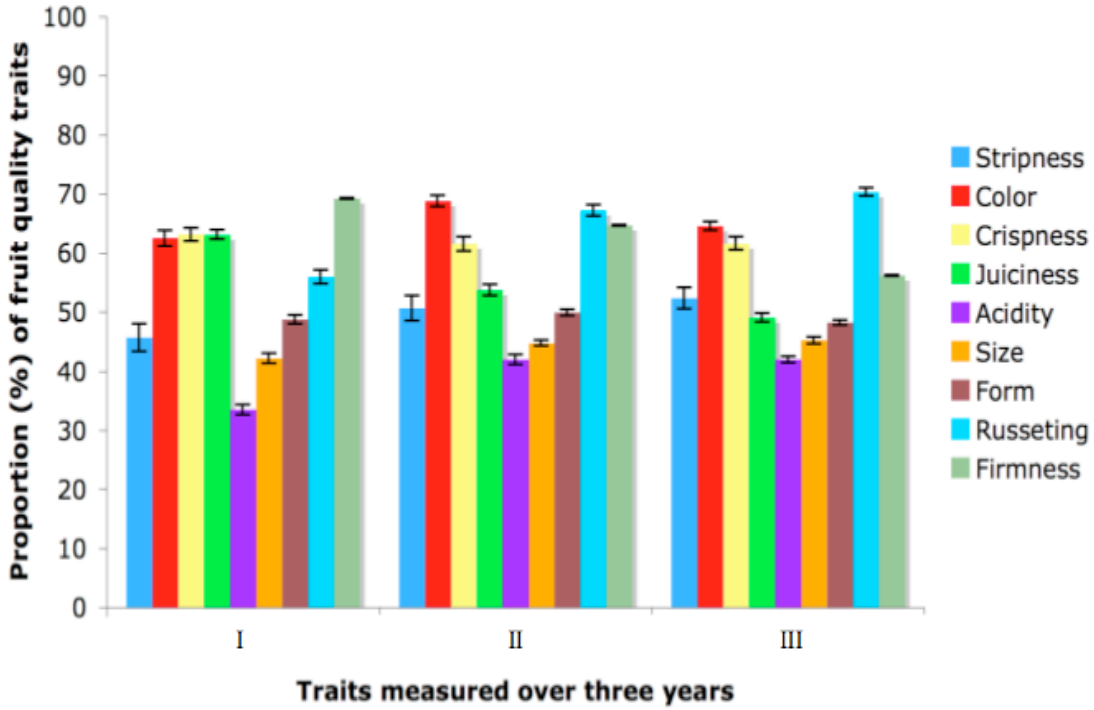


Figure 3.2: Influence of the environment and the genetic factors on the nine fruit quality traits per year.

Variations in the proportion of the fruit quality traits investigated using mature apple from the 'Golden Delicious' x 'Dietrich' population. Each bar represents the average data of the 248 individuals of the population for the quality trait, as indicated, for each year of investigation. Interval bars represent the standard error.

3.2.3 Investigation on the effect of the genetic factors on quality traits

In order to assess the genotypic influence on the fruit quality traits, the variations between individuals, as depicted by the standard error bars, were compared every year of investigation (Figure 3.2). In terms of firmness and russeting, the variations between individuals were statistically significant ($p<0.0001$, ANOVA; HSD=0.23 and 1.01, respectively). These variations were significant between year I and year II, and between year II and III, as revealed by the *t*-test ($p<0.0001$).

In terms of juiciness and acidity, the variations between individuals were significant over the three years of investigations ($p<0.0001$, ANOVA; HSD=1.85). These variations were significant only between year I and year II, and year I and year III, as revealed by the *t*-test ($p<0.0001$) and HSD test (1.85). In terms of size, form, colour, stripness and crispness, the standard error bars were overlapping among the three years of investigation, thus showing that the variations between individuals were not significant ($p>0.05$).

3.2.4 Investigation on the effect of both the environment and the genetic factors on quality traits

In order to assess the influence of both the environmental and the genetic factors on the fruit quality traits, the variations between individuals and the combined three years of phenotypic data were compared (Figure 3.3). In terms of comparison among the proportion of the traits, a wide variation was observed. The traits colour, crispness, russeting and firmness had the highest proportion, while acidity and size had the lowest.

In terms of variations within traits, for stripness, the variation in the proportion was the highest among the traits investigated, as depicted by the standard error bar, with a deviation of +/-13%. In regards to size, the variation in the proportion was the lowest, with a deviation of approximately +/-4%. For colour, crispness and russeting, the deviation in the proportion was approximately +/-8%. For juiciness and firmness, the deviation in the proportion was approximately +/-7%. In regards to acidity and form, the deviation in the proportion was approximately +/-6%.

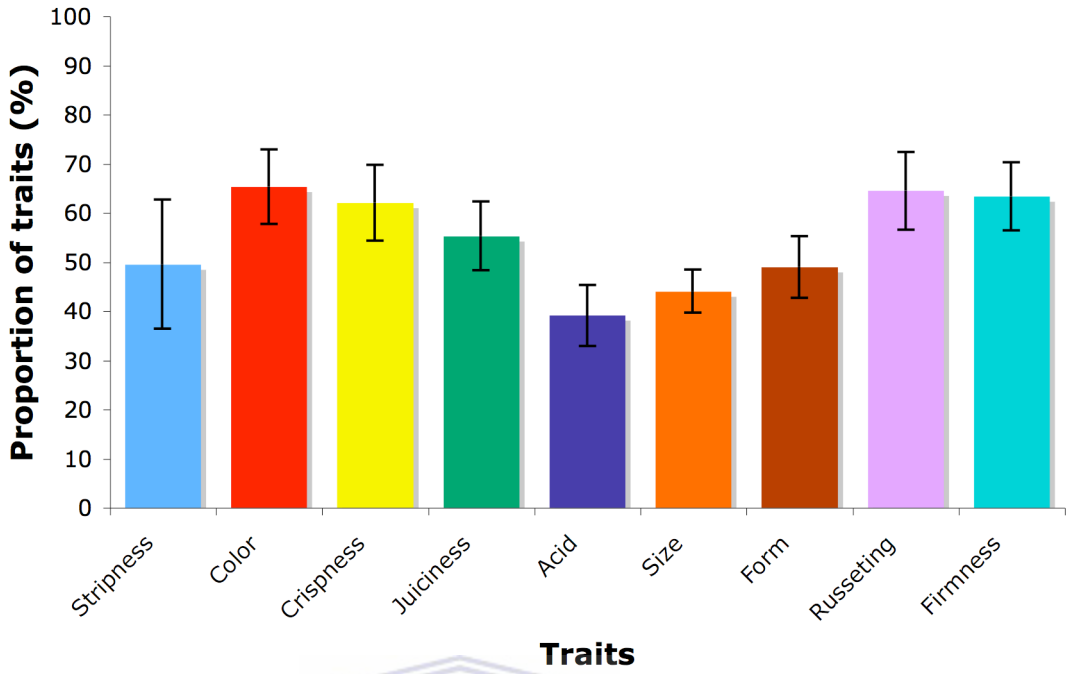


Figure 3.3: Influence of both the environment and the genetic factors on the nine fruit quality traits.

Variations in the proportion of the fruit quality traits investigated using mature apples from the 'Golden Delicious' x 'Dietrich' population. Each bar represents the average data of the 248 individuals of the population for the quality trait, as indicated, for the three years of investigation. Interval bars represent the standard error.

3.3 DISCUSSION

In this chapter, the representation of the population was first assessed. Then, the phenotypic data were analysed to assess the variations between individuals and the influence of the genetic, the environment and the combination of both factors on the fruit quality traits. In this regard, the variation of the phenotypic data was statistically analyzed using either the ANOVA or the student *t*-test, where appropriate.

In terms of the phenotypic data representation within the population, some individuals did bear fruits only one year over the three years of investigation (Table 3.1, Figure 3.1). On the other hand, a few individuals did bear fruits at least twice. Overall, 71.6% of the population was represented. A good representation of the population is necessary to yield enough segregating individuals. Generally, a minimum of 100 individuals is essential for detecting linkage among marker loci and quantitative traits (Tanksley, 1993). In addition, a large population is crucial to maintain polymorphism at any given locus in natural populations, thus providing allelic variation, critical in case of linkage tests to detect polygenes (Tanksley, 1993). The larger a segregating population is, the more likely QTLs with minor effects will reach statistical significance threshold. Thus, small populations can be used to detect QTLs with major effects (Kenis and Keulemans, 2008). In the current study, more than 100 individuals fruited in year II, year III and over the three years of investigation, but only 84 individuals were represented in year I. Therefore in this study, both major and minor QTLs can be detected, with the exception of year I that can be used to detect QTLs with major effects only.

In terms of the influence of the environment, variations in the proportion for the nine quality traits among the three years of investigation were observed (Figure 3.2). While quality traits tend to be inherited, abiotic factors like temperature or humidity during early fruit development can also have an effect on quality, which fluctuates from one year to another. This in turn has an impact on the phenotype (Janick, 1996). In terms of stripness, acidity, russeting, juiciness, firmness and size, greater variations in the proportion were observed between year I and II than between year II and III. On the contrary, for form, colour and crispness, almost no influence was detected over the three years. The results suggest that stripness, acidity, russeting, juiciness, firmness and size were more influenced by the environment than form, colour and crispness. Since more variations were observed between year I and II than between year II and III, this observation suggests that the climatic conditions were different in year I than in year II and III, thus inducing greater variations in the quality traits. This was validated by a critical analysis of the climatic conditions from fruit development to harvest. In year I (2005), higher rainfall was recorded over the four first months of the year than in year II and year III. In addition, greater precipitation (84.34 mm) occurred in year I at harvest than in years II and III (26.15 mm and 38.35 mm, respectively; www.tutiempo.net).

In terms of the influence of the genetic factors, overlapping standard error bars indicate that variations among data are statistically insignificant; while in the case of non-overlapping standard error bars, only statistical analyses can conclude that variations among data are significant (Motulsky, 1995). In terms of influence of the genetic factors on the quality traits, for acidity, juiciness, russeting and firmness, the variations of proportion among the three years were statistically significant (p<0.05, ANOVA; HSD: ranging from 0.23 to 1.85). For acidity and juiciness, these variations were

significant only between year I and year II, and year I and year III ($p<0.05$, Student *t*-test). For stripness, colour, crispness, size and form, the variations among the three years of investigations were not significant as the error bars were overlapping (Figure 3.2). This was confirmed by *p* values greater than 0.05 during ANOVA, HSD test and Student *t*-test analyses.

The quality traits, sweetness and acidity, which both have an effect on juiciness, have been shown to be independently inherited in apples and pears (Visser *et al.*, 1968). The balance between sugar content and acidity play an important role in apple fruit taste (Dolenc and Stampar, 1998). The apple acid taste positively correlates with the acidity, but it is negatively associated with sugar content. Thus, low sugar and high acid content result in a sour taste, while low acid and high sugar contents result in a bland taste. In this study, a high juiciness proportion (63.1%) was measured, while a low acidity (33.5%) was observed for year I. In year II and III, the juiciness proportion decreased, and this was statistically significant ($p<0.0001$), on the contrary, the acidity significantly increased ($p<0.0001$). This negative correlation between juiciness and acidity suggested that the low acidity proportion measured in year I may correlate to higher sugar content and thus sweeter taste in apple fruit juice.

Another negative correlation was observed between firmness and russeting. In this case, the higher the firmness, the lower the russeting on apples, and conversely. A similar relation between the traits was revealed in a study on tomato characterizing the effects of electrical conductivity and mineral nutrition on fruit radial (Hao and Papadopoulos, 2002). This study revealed that a high electrical conductivity was associated with low

russeting as well as high fruit firmness, when supplemented with both magnesium and calcium.

In terms of the influence of the combined effects of the environment and the genetic factors, a low correlation among the nine quality traits was observed (Figure 3.3). Low correlations among traits suggest independent inheritance (Hancock, 2004). However, variations among individuals for each trait were observed, as pointed out by standard error bars. These variations are expected for phenotypic traits that are quantitatively controlled by a number of loci.

A continuous phenotypic variation was observed among the individuals and three years of investigation for the nine quality traits under study (Figures 3.2 and 3.3). This typically represents a quantitative trait under polygenic influence in the absence of selection pressure (Tanksley, 1993). In addition, the persistent variations among individuals and years, which indicate the influence of the environment and genetic factors, suggest that the traits are controlled by joint action of many genes. These genes may be positioned on the genome using genetic markers.

CHAPTER 4

GENETIC MAP CONSTRUCTION

4.1 INTRODUCTION

Genetic mapping is a relatively new tool, which was introduced to assist traditional breeding methods through the identification of QTL (Bradshaw and Stettler, 1995) and their integration into MAS programs (Bus *et al.*, 2000; Missiaggia *et al.*, 2005). The technique involves the construction of genetic linkage maps, which can be performed easily using populations derived from pure lines (Jansen, 2005). The construction of genetic linkage maps is an essential prerequisite for the localization and positioning of genes influencing desirable traits. For this purpose, a mapping population, derived from a cross of two cultivars, is used to test markers and identify the alleles segregating between the parents to the seedlings. The various allele combinations can then be used to construct a map using a program like *JoinMap*® (Van Ooijen and Voorrips, 2001). The size of the population and the distribution of the phenotypic traits are crucial factors to consider when choosing a suitable mapping population.

During map construction, the linkage phase (the grand-parental origin) of the alleles forms the basis of the linkage analysis. Therefore, the first step in map construction is the determination of inheritance vectors (Van Ooijen and Voorrips, 2001). The *JoinMap*® software, through the determination and assignment of inheritance vectors to markers, allows minimising the number of recombinations among adjacent markers related to outbreeding.

In this chapter, a segregation analysis was performed by testing a subset of the available published (Maliepaard *et al*., 1998; Liebhard *et al*., 2002; Vinatzer *et al*., 2004; Silfverberg-Dilworth *et al*., 2006) and newly developed SSR markers. Using the results of this analysis, genetic linkage maps were constructed using a F1 mapping population derived from the crossing of 'Golden Delicious' and 'Dietrich' apple cultivars and compared with the reference map from Maliepaard *et al.* (1998).

4.2 RESULTS

4.2.1 Evaluation of genomic DNA quality

Genomic DNA (gDNA) of apple leaves was extracted from both parents of the 'Golden Delicious' x 'Dietrich' population as well as every individual of the population. In order to control the gDNA quality prior to downstream analyses, the DNA was resolved by agarose gel electrophoresis on a 0.8% agarose gel and visualized by ethidium bromide staining (Figure 4.1). The DNA was observed as a single intact band, thus showing that good quality DNA has been extracted. The DNA was then quantified using a nanodrop, diluted to a final concentration of 5 ng.μl^{-1} and used as a template for megaplex PCR amplifications (section 2.5.3.2).

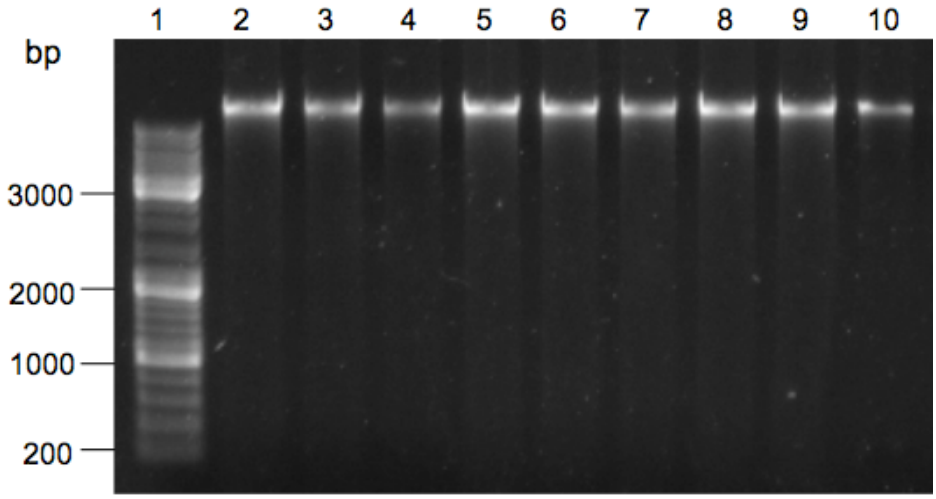


Figure 4.1: Evaluation of the quality of genomic DNA from the 'Golden Delicious' x 'Dietrich' mapping population.

Genomic DNA was separated by 0.8% (w/v) ethidium bromide stained agarose gel. Lane 1: DNA molecular weight marker, lane 2: 'Golden Delicious', lane 3: 'Dietrich', lane 4: individual 1-59, lane 5: individual 1-92, lane 6: individual 1-93, lane 7: individual 1-111, lane 8: individual 1-131, lane 9: individual 2-90 and lane 10: individual 2-108. The individuals were randomly selected from different batches that DNA was extracted.

4.2.2 Detection of PCR amplification products

Following quality control, the gDNA was amplified by megaplex PCR, whereby a set of at least 10 primers were tested in one reaction. A total of 202 published and 238 predicted (newly developed) primers were tested to amplify fragments of gDNA from the two parents of 'Golden Delicious' x 'Dietrich' population as well as the 246 individuals of the population (Table 2.1). Each of the 26 megaplexes carried out in this study consisted of 17 primers that simultaneously amplified DNA during PCR. Only primers that were giving rise to non-overlapping same fluorescently labelled amplified fragments were used together for megaplex PCR amplification (Table 2.1).

Following megaplex PCR, PCR products were electrophoresed on a 6% polyacrylamide gel and polymorphism identified (Figure 4.2). Differing amplicon patterns were observed among parents and individuals. However, the segregation analysis could not be performed after 6% polyacrylamide gels because the resolution of the PCR products was limited and the fluorescent labels were not visible. Therefore, the PCR products were analysed with the ABI 3130xl Genetic Analyser to determine specific sizes.

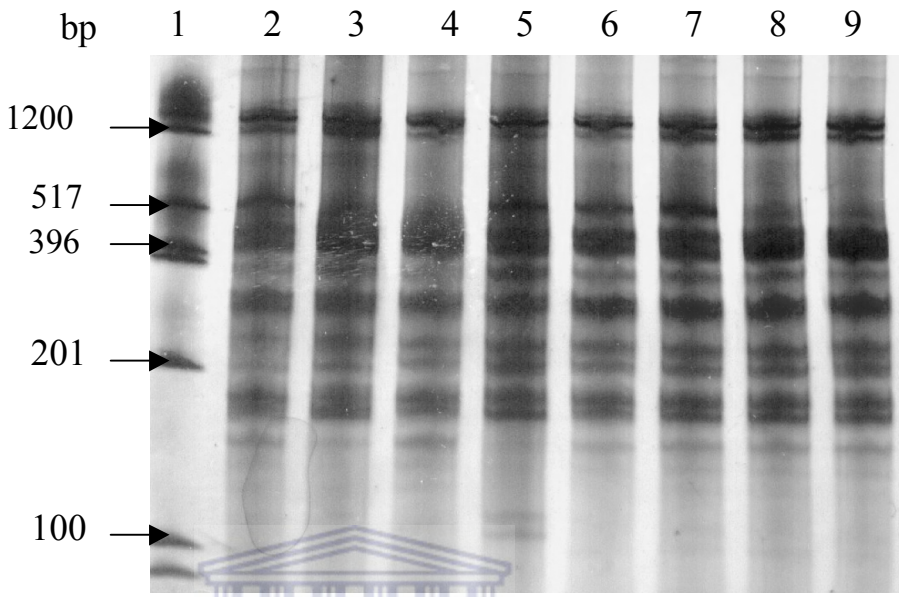


Figure 4.2: Detection of gDNA from the 'Golden Delicious' x 'Dietrich' population after amplification by megaplex PCR.

Following megaplex PCR amplication of the DNA, the efficiency of the amplification was resolved by 6% polyacrylamide gel and stained by silver stain. Lane 1: DNA molecular weight marker pTZ/HinfI, lane 2: 'Golden Delicious', lane 3: 'Dietrich', lane 4: individual 1-92, lane 5: individual 1-93, lane 6: individual 1-111, lane 7: individual 1-131, lane 8: individual 2-90 and lane 9: individual 2-108.

4.2.3 Allele sizing and segregation analysis

Following detection of the efficiency of the megaplex PCR by polyacrylamide gel electrophoresis, DNA fragments were then analysed by capillary electrophoresis on the Genetic Analyser. The results were observed as electropherograms, where each peak represented a PCR product size. The sizes of the fragments were measured against the GeneScan™ LIZ500 size standard using the *GeneMapper*® V4 software. Figure 4.3 shows a typical electropherogram for individual 2-108 after capillary electrophoresis following megaplex PCR of megaplex 17 with the 20 markers fluorescently dyed. Figure 4.4 shows a typical electropherogram for one individual after capillary electrophoresis following megaplex PCR for five markers fluorescently dyed with Pet.

The fragments from the 248 individuals were then compared against the two parents of the 'Golden Delicious' x 'Dietrich' population for allele segregation analysis and genotyping. Individual seedlings were genotyped for each specific locus using the *JoinMap*® V4 coding system (van Ooijen, 2006), as illustrated in Tables 4.1 and 4.2. The threshold was manually adjusted to 20% of base peak intensity to eliminate false positive alleles.

Table 4.1: *JoinMap*® codification for the different forms of segregation (van Ooijen, 2006).

The table lists the possible number of alleles that can be inherited by the progeny, their corresponding codes and different forms of segregation.

Form of segregation	Allele(s) from parent 1	Corresponding code(s)	Allele(s) from parent 2	Corresponding code(s)
ab x cd	1, 2	a, b	3, 4	c, d
ef x eg	1, 2	e, f	1, 3	e, g
nn x np	1	N	1, 2	n, p
nn x np	1	N	2, 3	n, p
lm x ll	1, 2	l, m	1	l
lm x ll	1, 2	l, m	3	l
hk x hk	1, 2	h, k	1, 2	h, k

Table 4.2: Segregation codes and their respective possible genotypes for a cross-pollination (CP) population, adapted from van Ooijen (2006).

Code/ Segregation type	Description	Possible genotypes
ab x cd	Locus heterozygous in both parents, four alleles	ac, ad, bc, bd
ef x eg	Locus heterozygous in both parents, three alleles	ee, ef, eg, fg
hk x hk	Locus heterozygous in both parents, two alleles	hh, hk, kk, h-, k-
lm x ll	Locus heterozygous in the first parent	II, Im, --
nn x np	Locus heterozygous in the second parent	nn, np, --

'x' represent a cross between two parents

a to p represents a distinct allele; '-', '--' represent unknown alleles

h- and k- are dominant genotypes: h- means the genotype can either be hh or hk, and k- either kk or hk.

Figure 4.5 shows a typical allele segregation for marker A421 (Accession number: CO865954). Marker A421 showed existence of heterozygosity between the two parents, with the locus being heterozygous for 'Dietrich' (second parent). Golden Delicious' was homozygous having alleles 'nn', while 'Dietrich' was heterozygous having alleles 'np'. Individual 1 was heterozygous having alleles 'np', inheriting the 'n' allele from 'Golden Delicious' parent and the 'p' allele from 'Dietrich' parent. On the contrary, individual 2 was homozygous having alleles 'nn', inheriting the 'n' allele from both parents. The segregation type was <nnxnp>, with three possible genotypes (Table 2.3).

A total of 440 markers were tested on all the 248 individuals of the 'Golden Delicious' x 'Dietrich' population as well as the two parents. Only 329 of these markers amplified. Following scoring and genotyping of all the alleles, the data were exported to *JoinMap*® for the map construction. Approximately 13% of the amplified markers were homozygous for the two parents. However, 3.1% of the markers that amplified did not amplified on at least 25% of the individuals and were excluded for further analysis with the *JoinMap*® software. Therefore, a total of 289 markers were analysed in *JoinMap*®.

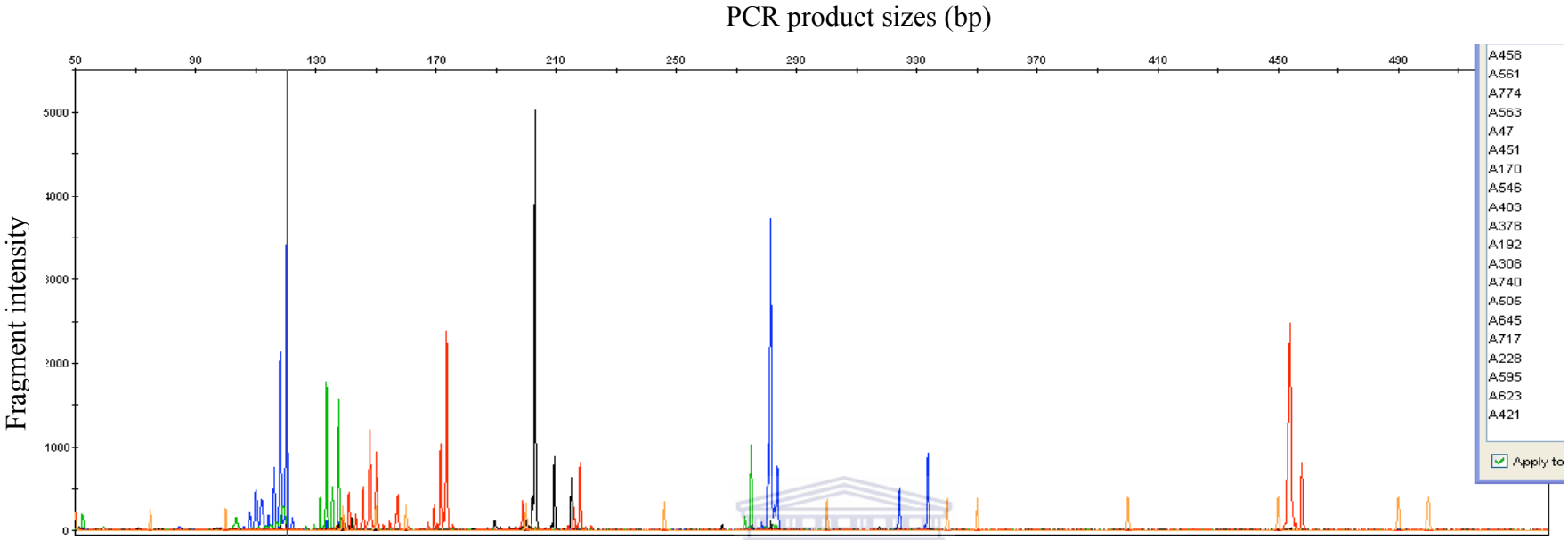


Figure 4.3: Detection of allele inheritance for individual 2-108 from the 'Golden Delicious' x 'Dietrich' population by capillary electrophoresis after megaplex PCR.

Genomic DNA from the individual 2-108 from 'Golden Delicious' x 'Dietrich' population was amplified by megaplex PCR with 20 fluorescently labelled primers in megaplex 17, simultaneously, and then analysed using ABI 3130xl Genetic Analyzer. The markers A458, A451, A192, A505 and A595 were labelled with 6-FAM (blue), the markers A561, A47, A378, A645 and A623 were labelled with Vic (green), the markers A774, A546, A308, A740 and A228 were labelled with Ned (black) and the markers A563, A170, A403, A717 and A421 were labelled with Pet (red). The GeneScan™ LIZ500 size standard is represented by the orange peaks. The electropherogram represents the fragment size and intensity for each peak.

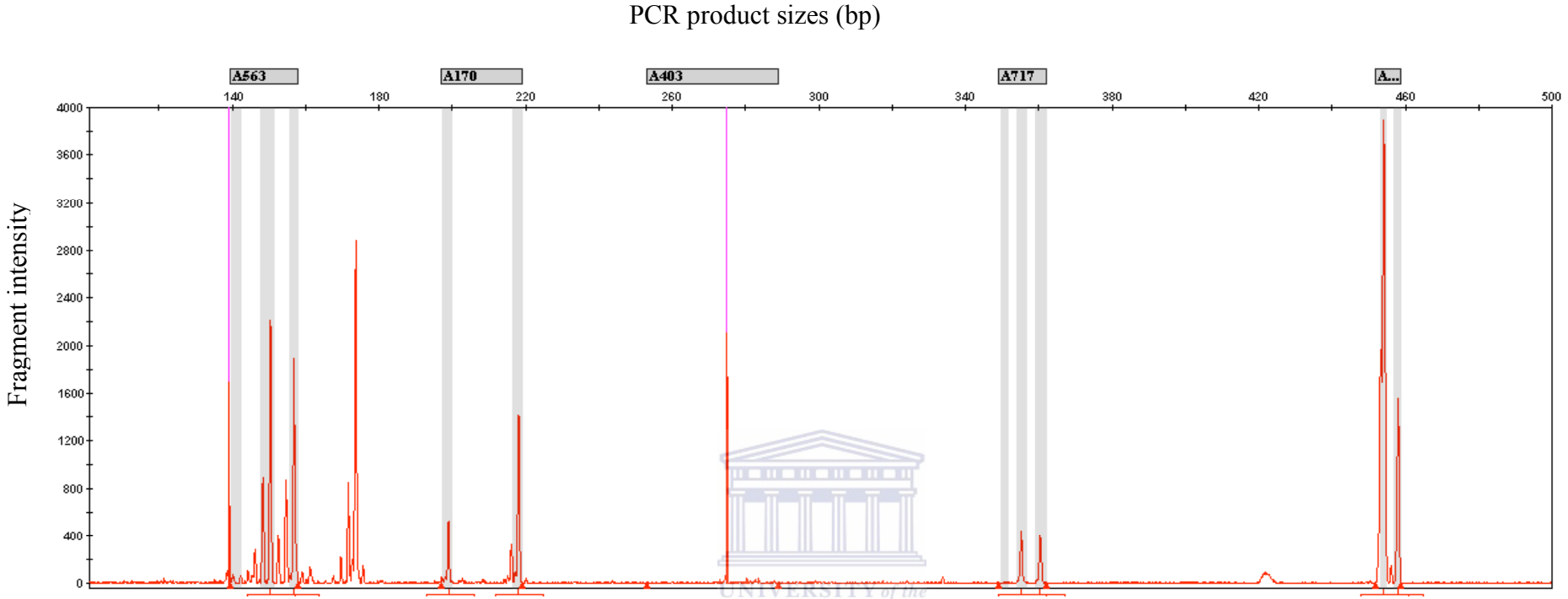


Figure 4.4: Detection of allele inheritance for individual 2-108 from the 'Golden Delicious' x 'Dietrich' population by capillary electrophoresis after megaplex PCR of megaplex 17 for Pet fluorescently labeled markers.

Genomic DNA from the individual 2-108 from 'Golden Delicious' x 'Dietrich' population was amplified by megaplex PCR with 20 fluorescently labelled primers in megaplex 17, simultaneously, and then analysed using ABI 3130xl Genetic Analyzer. The markers A563, A170, A403, A717 and A421 were labelled with Pet (red). The electropherogram represents the fragment size and intensity for each peak. The grey lines represent alleles inherited from both parents and the grey boxes (top) indicate the BIN range for each marker.

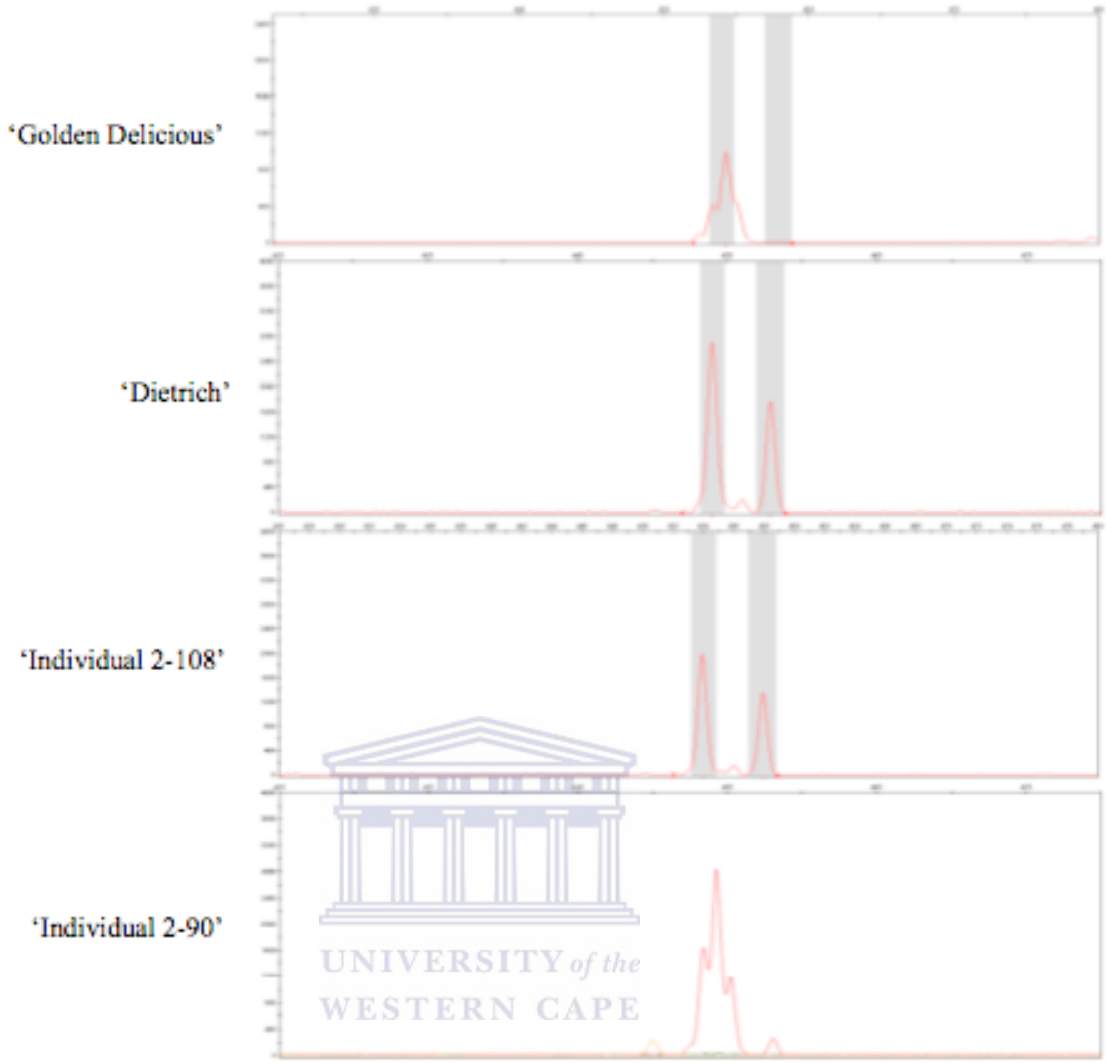


Figure 4.5: Detection of allele inheritance for individuals 2- 108 and 2-90 and from the 'Golden Delicious' x 'Dietrich' population by capillary electrophoresis after megaplex PCR for Pet fluorescently labeled marker A421.

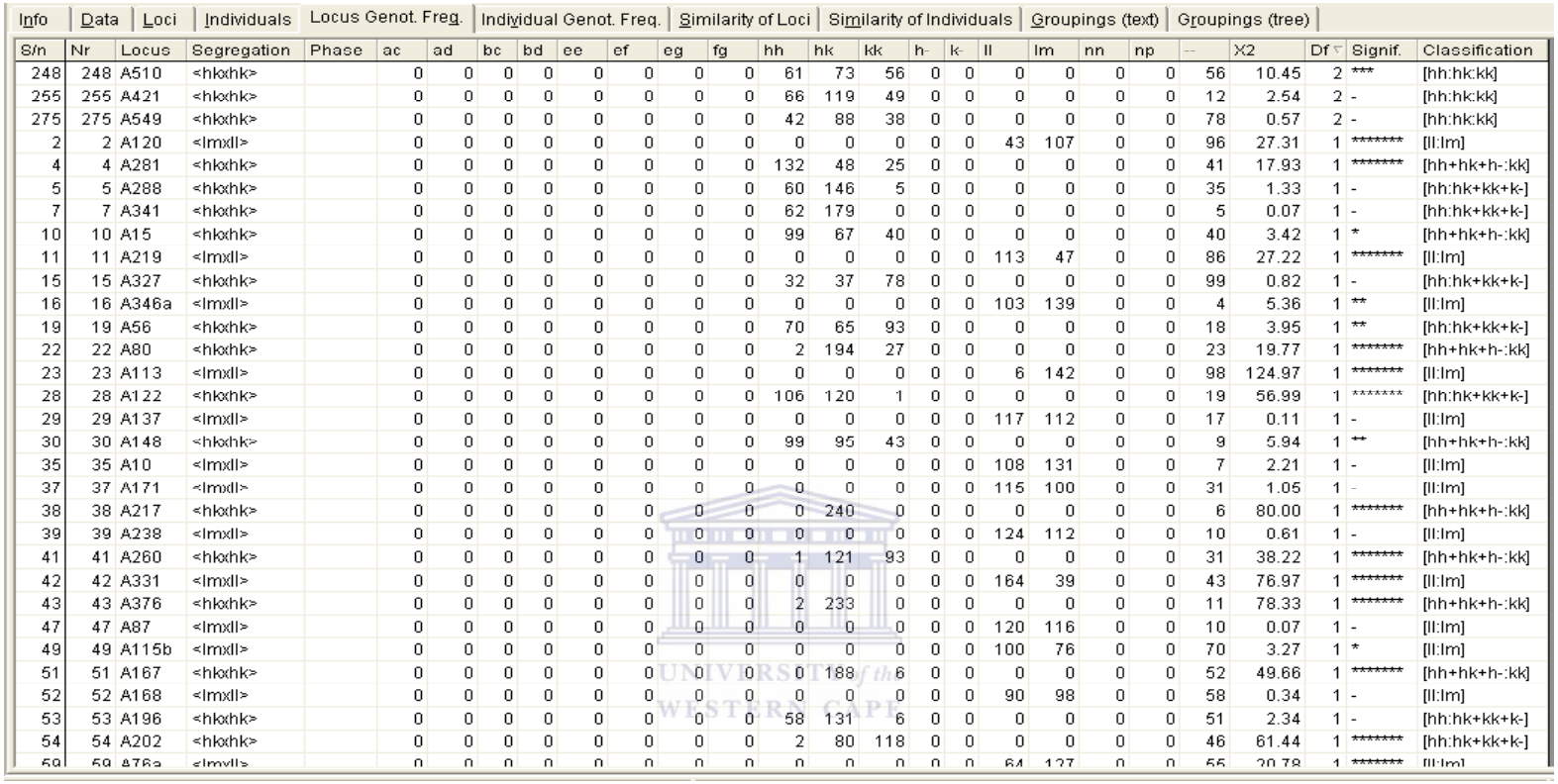
Genomic DNA from the individuals 2-108 and 2-90 from 'Golden Delicious' x 'Dietrich' population were amplified using megaplex PCR with megaplex 17, and analysed using the ABI 3130xl Genetic Analyzer. The marker A421 was labelled with Pet (red). The peaks represent alleles inherited from both parents. Segregation analysis showed that 'Golden Delicious' is homozygous having alleles 'nn', while 'Dietrich' is heterozygous having alleles 'np'. Individual 2-108 is heterozygous having alleles 'np', inheriting the 'n' allele from 'Golden Delicious' parent and the 'p' allele from 'Dietrich'. Individual 2-90 is homozygous having alleles 'nn', inheriting the 'n' allele from both parents.

4.3 CONSTRUCTION OF THE GENETIC LINKAGE MAP

Following analysis of allele segregation using the *GeneMapper*® V4 software, genotypic codes were assigned for each segregation type prior to exporting the data to *JoinMap*®. The *JoinMap*® V4 software was used to compute locus genotypic frequency, individual genotypic frequency and groupings for map construction (van Ooijen and Voorrips, 2006).

4.3.1 Analysis of the locus genotypic frequency data

The locus genotypic frequency allows for the testing of segregation distortion among individuals. From this analysis, the chi-squared (χ^2) value and significance of loci segregation were determined. The χ^2-test indicates the 'goodness of fit' based on the independence of segregation (Figure 4.6).



Info | Data | Loci | Individuals | Locus Genot. Freq. | Individual Genot. Freq. | Similarity of Loci | Similarity of Individuals | Groupings (text) | Groupings (tree)

S/n	Nr	Locus	Segregation	Phase	ac	ad	bc	bd	ee	ef	eg	fg	hh	hk	kk	h-	k-	ll	lm	nn	np	--	X2	Df	Signif.	Classification
248	248	A510	<hkxhk>		0	0	0	0	0	0	0	0	61	73	56	0	0	0	0	0	0	56	10.45	2	***	[hh:hk:kk]
255	255	A421	<hkxhk>		0	0	0	0	0	0	0	0	66	119	49	0	0	0	0	0	0	12	2.54	2	-	[hh:hk:kk]
275	275	A549	<hkxhk>		0	0	0	0	0	0	0	0	42	88	38	0	0	0	0	0	0	78	0.57	2	-	[hh:hk:kk]
2	2	A120	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	43	107	0	0	96	27.31	1	*******	[ll:lm]
4	4	A281	<hkxhk>		0	0	0	0	0	0	0	0	132	48	25	0	0	0	0	0	0	41	17.93	1	*******	[hh+hk+h-:kk]
5	5	A288	<hkxhk>		0	0	0	0	0	0	0	0	60	146	5	0	0	0	0	0	0	35	1.33	1	-	[hh:hk+kk+k-]
7	7	A341	<hkxhk>		0	0	0	0	0	0	0	0	62	179	0	0	0	0	0	0	0	5	0.07	1	-	[hh:hk+kk+k-]
10	10	A15	<hkxhk>		0	0	0	0	0	0	0	0	99	67	40	0	0	0	0	0	0	40	3.42	1	*	[hh+hk+h-:kk]
11	11	A219	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	113	47	0	0	86	27.22	1	*******	[ll:lm]
15	15	A327	<hkxhk>		0	0	0	0	0	0	0	0	32	37	78	0	0	0	0	0	0	99	0.82	1	-	[hh:hk+kk+k-]
16	16	A346a	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	103	139	0	0	4	5.36	1	**	[ll:lm]
19	19	A56	<hkxhk>		0	0	0	0	0	0	0	0	70	65	93	0	0	0	0	0	0	18	3.95	1	**	[hh:hk+kk+k-]
22	22	A80	<hkxhk>		0	0	0	0	0	0	0	0	2	194	27	0	0	0	0	0	0	23	19.77	1	*******	[hh+hk+h-:kk]
23	23	A113	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	6	142	0	0	98	124.97	1	*******	[ll:lm]
28	28	A122	<hkxhk>		0	0	0	0	0	0	0	0	106	120	1	0	0	0	0	0	0	19	56.99	1	*******	[hh:hk+kk+k-]
29	29	A137	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	117	112	0	0	17	0.11	1	-	[ll:lm]
30	30	A148	<hkxhk>		0	0	0	0	0	0	0	0	99	95	43	0	0	0	0	0	0	9	5.94	1	**	[hh+hk+h-:kk]
35	35	A10	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	108	131	0	0	7	2.21	1	-	[ll:lm]
37	37	A171	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	115	100	0	0	31	1.05	1	-	[ll:lm]
38	38	A217	<hkxhk>		0	0	0	0	0	0	0	0	0	240	0	0	0	0	0	0	0	6	80.00	1	*******	[hh+hk+h-:kk]
39	39	A238	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	124	112	0	0	10	0.61	1	-	[ll:lm]
41	41	A260	<hkxhk>		0	0	0	0	0	0	0	0	1	121	93	0	0	0	0	0	0	31	38.22	1	*******	[hh+hk+h-:kk]
42	42	A331	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	164	39	0	0	43	76.97	1	*******	[ll:lm]
43	43	A376	<hkxhk>		0	0	0	0	0	0	0	0	2	233	0	0	0	0	0	0	0	11	78.33	1	*******	[hh+hk+h-:kk]
47	47	A87	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	120	116	0	0	10	0.07	1	-	[ll:lm]
49	49	A115b	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	100	76	0	0	70	3.27	1	*	[ll:lm]
51	51	A167	<hkxhk>		0	0	0	0	0	0	0	0	0	188	6	0	0	0	0	0	0	52	49.66	1	*******	[hh+hk+h-:kk]
52	52	A168	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	90	98	0	0	58	0.34	1	-	[ll:lm]
53	53	A196	<hkxhk>		0	0	0	0	0	0	0	0	0	58	131	6	0	0	0	0	0	51	2.34	1	-	[hh:hk+kk+k-]
54	54	A202	<hkxhk>		0	0	0	0	0	0	0	0	2	80	118	0	0	0	0	0	0	46	61.44	1	*******	[hh:hk+kk+k-]
59	59	A76a	<lmxll>		0	0	0	0	0	0	0	0	0	0	0	0	0	64	127	0	0	55	20.78	1	*******	[ll:lm]

Figure 4.6: Locus genotypic frequency according to loci segregation.

The locus genotype frequency according to the segregation types <hkxhk> and <lmxll> for 30 loci was calculated to determine the goodness of fit (χ^2 value) based on the segregation independence among the 246 individuals of the 'Golden Delicious' x Dietrich' population. The table includes the locus name, segregation type, ratio of individuals for each genotype, χ^2 values, degress of freedom (Df), significance levels and the possible genotype classification.

Since the genotypic expected segregation ratio between the locus allele pairs is 1:1, markers with different ratio displayed segregation distortion. Out of the 329 SSR markers that amplified, 135 markers showed segregation distortion among the tested individuals. For example, marker A281, with a segregation type of <hkxhk>, was expected to have a genotypic ratio of 1:2:1, corresponding to the 'hh', 'hk' and 'kk' allele pairs, respectively. A 5:2:1 ratio however was observed, resulting in a high χ^2 value (Figure 4.6).

A total of 131 and 126 markers, segregating on the 'Golden Delicious' and 'Dietrich' parents, respectively, were highly significant with χ^2 value lower than 10.0 (Table 4.3). These markers showed low distortion and their genotypic ratios approximated the expected ratio for each segregation type. Also, a total of 16 and 18 markers from 'Golden Delicious' and 'Dietrich', respectively, were relatively significant with χ^2 value between 10.1 and 20.0. The segregation of these markers was distorted and the observed genotypic ratios deviated from the expected ratios for each segregation type. In addition, 109 and 116 markers from 'Golden Delicious' and 'Dietrich', respectively, showed low significance level and a χ^2 value greater than 20.1. These markers displayed high segregation distortion and their genotypic ratios deviated from the expected ratios for each segregation type (Table 4.3).

A comparison between locus genotypic frequency and groupings (tree) function results (*JoinMap®*) led to the exclusion of 102 markers for the genetic linkage map construction (Appendix I). Most of the markers excluded were characterised by a χ^2 value greater than 20.1, as well as an <hkxhk> segregation type. Besides, some markers

with χ^2 values below 5.0 were also excluded because of insufficient linkage with other marker groupings. Possible explanations for markers not showing linkage with any other marker or group of markers are: a) the marker had high χ^2 values and insufficient linkage was therefore determined, b) the marker has been published and is known to be situated on the end of the linkage group and /or far away from other segregating markers on the same linkage group, or c) the marker is the only segregating published marker scored on the specific linkage group. Published markers expected to have shown linkage with certain linkage groups were also found in the set of unlinked markers as well as newly developed and published but unmapped markers.

Table 4.3: Grouping of markers according to their chi-square value.

Markers from both the 'Golden Delicious' and 'Dietrich' parents were grouped according to their chi-square (χ^2) value ranges and significance level. The asteriks (*) represents the level of significance.

χ^2 range	Significance level	Markers from 'Golden Delicious'	Markers from 'Dietrich'
0.0-5.0	*	100	99
5.1-10.0	**	31	27
10.1-15.0	***	10	7
15.1-20.0	****	6	11
20.1-∞	*****, ******	109	116

4.3.2 Creation of groups of linked markers

Groups of linked markers were created using the groupings (tree) function and calculation of the map. A total of 32, 38 and 43 linked marker groups were constructed and exported to the *MapChart™* software to construct map charts of the 'Golden Delicious', 'Dietrich' and integrated maps, respectively (Table 4.4). Out of these groups, 28 were positioned to linkage groups (LG) on the 'Golden Delicious', 33 on the 'Dietrich' and 35 on the integrated maps. On the other hand, four, three and eight groups could not be positioned to any linkage group on the 'Golden Delicious', 'Dietrich' and integrated maps, respectively. These groups fail to be positioned because they mostly contained predicted or newly developed markers. The list of markers and their accession number is listed in Table 4.5.

Table 4.4: Groups of markers positioned on the 'Golden Delicious', 'Dietrich' and integrated linkage maps.

Following marker grouping, the positioning of the groups onto the linkage group was attempted using the *MapChart*™ software.

Parent	Number of groups	Number of groups with marker positions	Number of groups without marker positions
'Golden Delicious'	32	28	4
'Dietrich'	38	33	5
Integrated map	43	35	8

Table 4.5: Markers from the marker groups that could not be positioned on the linkage maps.

Marker/primer number	Accession number
A105	CH01g121
A118	CH02d10a
A177	CH05h05
A178	CH05h12
A192	CN491993
A259	CN904905
A341	CO723148
A37	CH02c061
A414	CN489062
A425	CO756781
A451	AF429983
A512	CN944528
A555	Hi02d04
A579	Hi07b06
A603	EB114458
A686	EB106592
A715	SAA715

4.3.3 Construction of the 'Golden Delicious', 'Dietrich' and integrated linkage maps

Three maps were constructed for the 'Golden Delicious' x 'Dietrich' population using a total of 315 loci genotyped over 246 individuals. These include two parental and one integrated maps (combination of the two; Figure 4.7). These 315 loci were amplified from 286 primer pairs. Each linkage map consisted of 17 linkage groups (Table 4.7) that correspond to the number of chromosomes in the apple haploid genome. The map construction was carried out using the *JoinMap®* and *MapChart™* software. The 'Fiesta' x 'Discovery' genetic linkage map from Maliepaard *et al*. (1998) was used as a reference to assign the markers to their respective linkage groups. The positioning of several marker groups to a single linkage group was performed through alignment to the reference markers according to Silfverberg-Dilworth *et al*. (2006).

The linkage maps span a total of 1,437.8 cM and 1,491.5 cM for 'Golden Delicious' and 'Dietrich', respectively. The parental maps could be aligned, as they had 105 fully informative SSR markers in common. The marker distance and order on the linkage groups was similar to published apple maps also constructed using SSR markers, except for few discrepancies: (1) a second locus was amplified for CH04e05 and Hi02c07, and was mapped on LG1 and LG16, respectively; (2) CH05g08 was mapped on LG11 of the 'Golden Delicious' map instead of LG1 on the reference map (Silfverberg-Dilworth *et al*., 2006); (3) CH01f03b was mapped on LG16 of the 'Golden Delicious' map, instead of LG9 on the reference map (Silfverberg-Dilworth *et al*., 2006). On the contrary to the 'Malling 9' x 'Robusta 5' map (Celton *et al*., 2009), the (1) CH05d11 was mapped above CH04g04 and (2) CH05e06 was mapped below CH05f06 in the 'Golden Delicious' map.

In addition, a total of 33 SSR markers developed through this work were positioned on the ‘Golden Delicious’ and Dietrich’ linkage maps, respectively (Table 4.6). A total of 124 SSR markers were mapped for ‘Golden Delicious’, while 116 SSR markers were mapped for ‘Dietrich’ (Table 4.7). The integrated map consisted of 147 SSR markers.

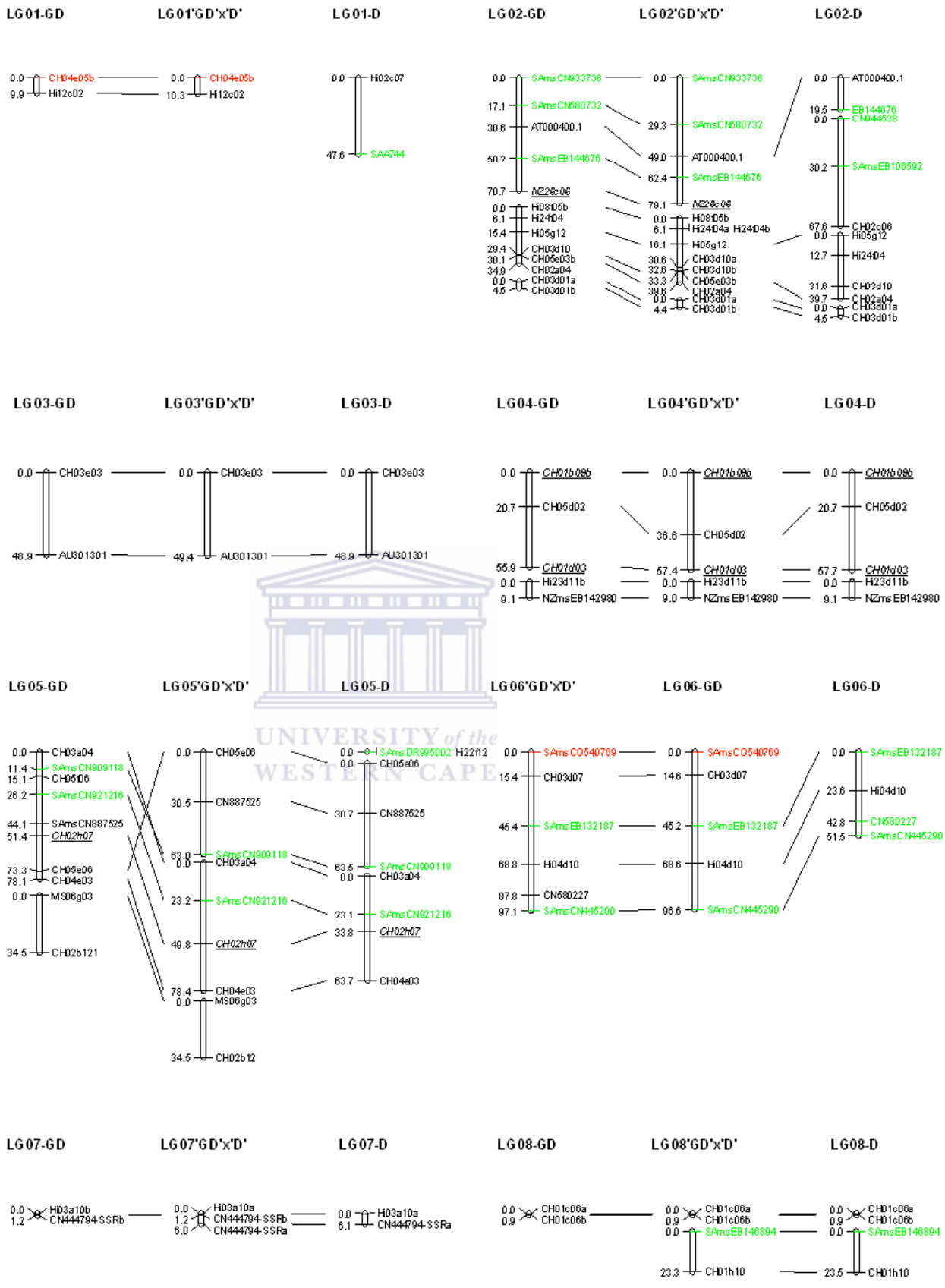
Table 4.6: Newly developed markers from the marker groups that were positioned on at least one of the linkage maps.

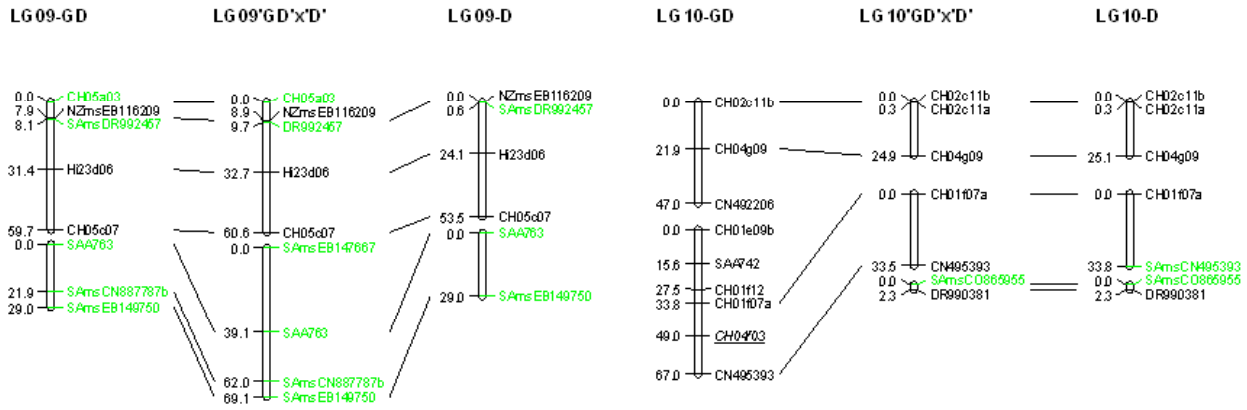
Linkage group (LG)	Marker common number	Marker accession number	Genetic maps 'GD'	'D'
	A254	CN933736	X	
	A220	CN580732	X	
LG2	A602	EB144676	X	X
	A512	CN944528		X
	A686	EB106592		X
	A288	CN909118	X	X
LG5	A505	DR995002		X
	A687	EB132187	X	X
LG6	A307	CN445290	X	X
	A318	CN580227		X
LG8	A647	EB146894	X	X
	A390	CH05a03	X	
	A498	DR992457	X	X
LG9	A638	EB147667	X	
	A241	CN887787	X	
	A592	EB149750	X	X

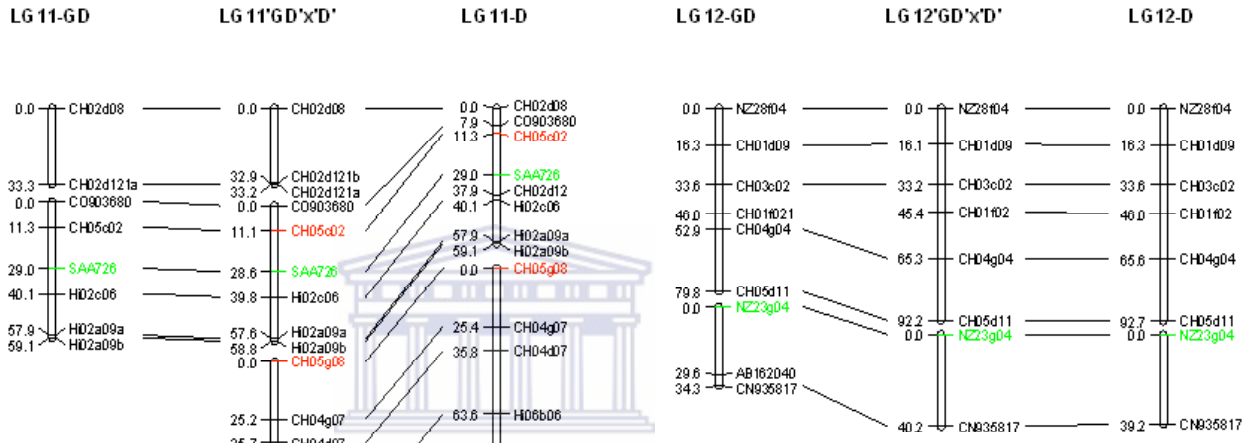
Table 4.6 continued

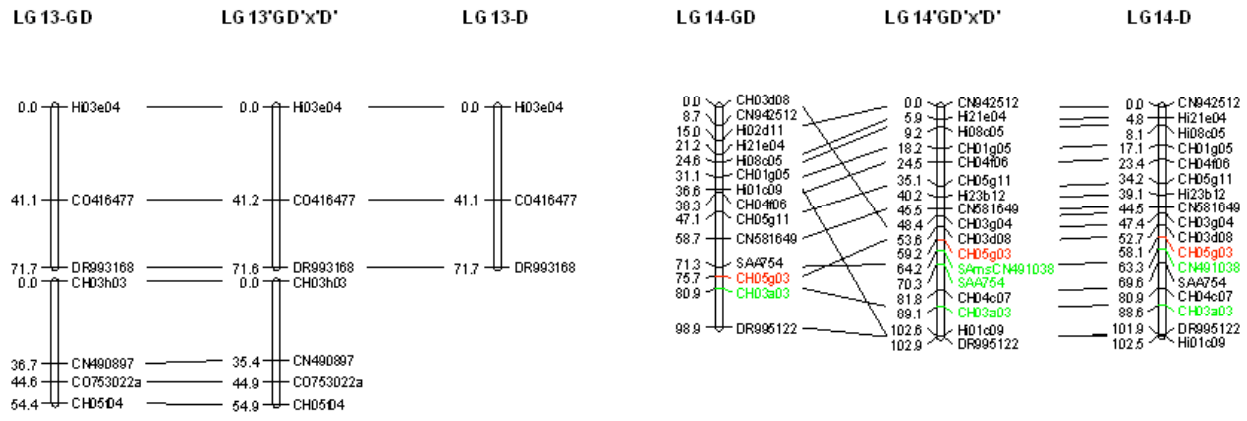
Linkage group (LG)	Marker common number	Marker accession number	Genetic maps 'GD'	Genetic maps 'D'
LG10	A262	CO865955	X	X
	A395	CN495393	X	X
	A726	*SAA726	X	X
LG11	A756	*SAA756	X	X
	A716	*SAA716	X	
LG12	A14	NZ23g04	X	X
LG14	A397	CN491038	X	X
	A148	CH03a03	X	X
LG15	A416	CO168103	X	X
	A661	EB126773	X	X
	A118	CH02d10b		X
LG16	A744	SAA744	X	X
	A728	SAA728		X
LG17	A422	CV627191	X	X
	A736	SAA736	X	X
	A461a	DT000945a		X
	A461b	DT000945b	X	X

*Newly developed markers derived from contigs were renamed as SAnumber.









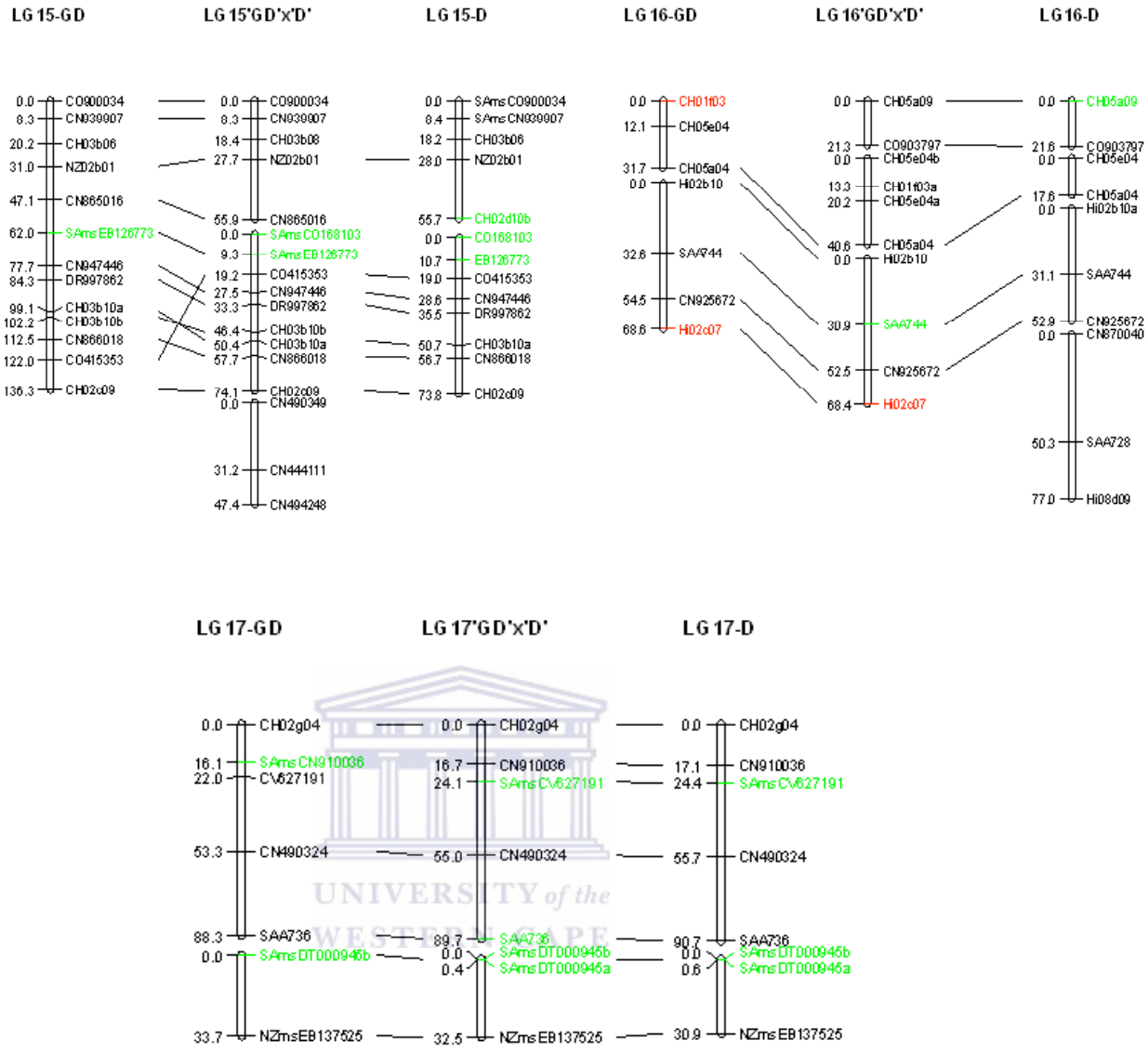


Figure 4.7: 'Golden Delicious' ('GD'), 'Dietrich' ('D') and the integrated genetic linkage maps

The numbering of the linkage groups was according to Maliepaard *et al*. (1998). The marker groups were positioned according to Silfverberg-Dilworth *et al*. (2006). Newly developed and unpublished SSR markers are indicated in green. Published markers whose position is not in accordance with the reference map from Silfverberg-Dilworth *et al*. (2006) are shown in red. Markers underlined and italized were absent in the reference map from Silfverberg-Dilworth *et al*. (2006), but incorporated in maps from Igarashi *et al*. (2008), Celton *et al*. (2009) and van Dyk *et al*. (In press). The lines between linkage groups indicate markers common in both linkage maps.

Table 4.7: Detailed analysis of the number of SSR markers, length and marker density for each linkage group of the'Golden Delicious' and 'Dietrich' genetic maps.

Linkage group (LG)	'Golden Delicious'			'Dietrich'		
	SSR markers	LG length (cM)	Marker density (markers/cM)	SSR markers	LG length (cM)	Marker density (markers/cM)
1	2	9.9	0.20	2	47.6	0.04
2	13	110.1	0.13	11	111.8	0.10
3	2	48.9	0.04	2	48.9	0.04
4	5	65.0	0.08	5	66.8	0.07
5	10	78.1	0.13	9	127.2	0.07
6	5	96.6	0.06	4	51.5	0.08
7	2	1.2	1.7	2	6.1	0.33
8	2	2.2	0.16	4	24.4	0.16
9	8	88.7	0.09	6	82.5	0.07
10	9	114	0.08	7	61.2	0.11
11	8	92.4	0.09	13	146.5	0.09
12	9	114.1	0.08	8	131.9	0.06
13	7	126.1	0.05	3	71.9	0.04
14	14	98.9	0.14	17	102.5	0.17
15	13	136.3	0.10	10	149.5	0.07
16	7	100.1	0.07	10	169.1	0.06
17	7	122.0	0.06	8	121.6	0.07
Total	124	1,437.8		116	1,491.5	

4.3.4 Graphical genotypes analysis

Following the construction of the genetic linkage maps, graphical genotyping of individuals of the 'Golden Delicious' x 'Dietrich' population was carried out using the Graphical Genotyper® (GGT) software (Van Berloo, 1999). The Graphical Genotyper® program displays chromosomal genotypes of individuals based on marker segregation data. Graphical genotypes of a subset of individuals from the 248 individuals of the 'Golden Delicious' x 'Dietrich' population were compared using the 'Linkage group view mode' function (Figure 4.8). In the 'Linkage group view mode', segregation of alleles and association of loci on LG14 of these individuals was observed. Some genomic regions reflect partial linkage disequilibrium, as no recombination was evident. For example, on LG14, individuals (progeny) 1, 3 and 9 reflect linkage disequilibrium among markers H01C09, CH03C03, CH04C07 and SAA754. However, some individuals, including 13, 15, 19, 25, 33, 45, 53 and 61, displayed complete linkage disequilibrium on LG14 (Figure 4.8).

In addition, the graphical genotypes of the 17 linkage groups for each individual were graphically represented using the 'Individual view mode' function (Figure 4.9). A graphical genotype representation of the 17 linkage groups for the individual 1-10 displayed non-random association among some alleles and loci as this was observed only in some of the linkage groups. For example, linkage groups 1, 2 and 10 were heterozygous for every locus. Besides, a high recombination crossover between loci was observed on LG15 (Figure 4.9).

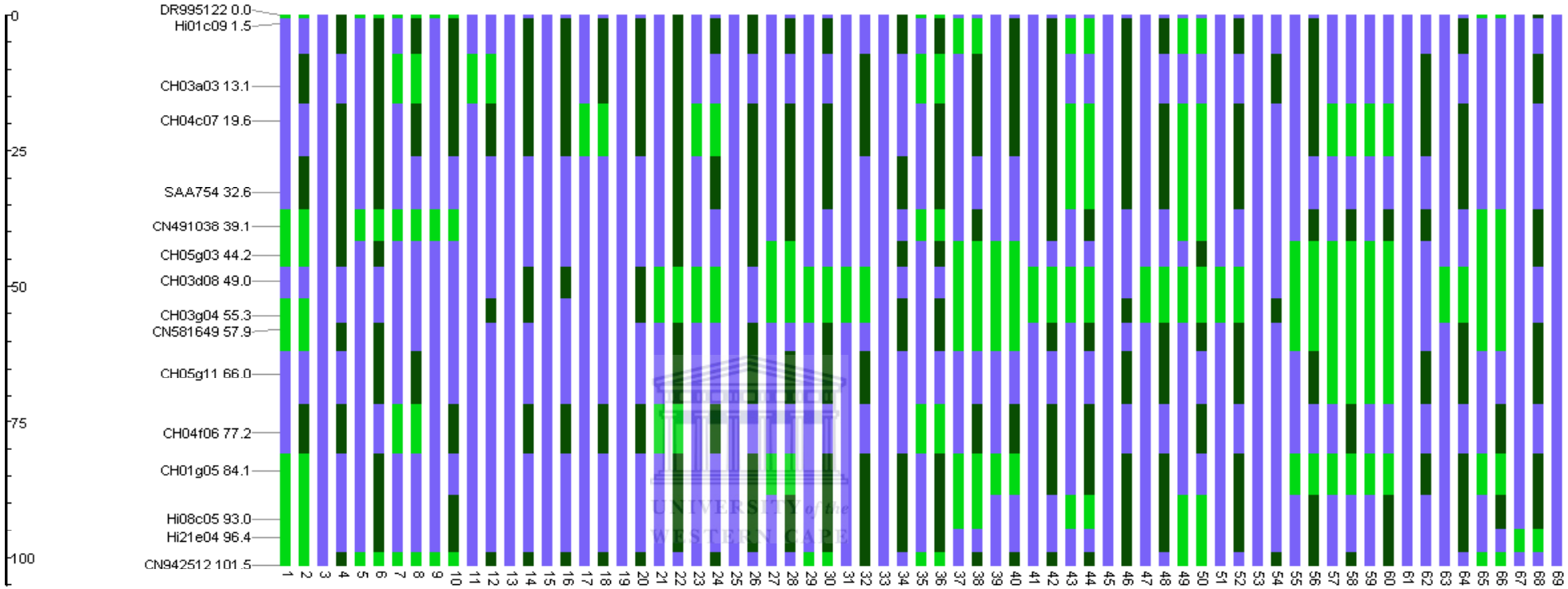
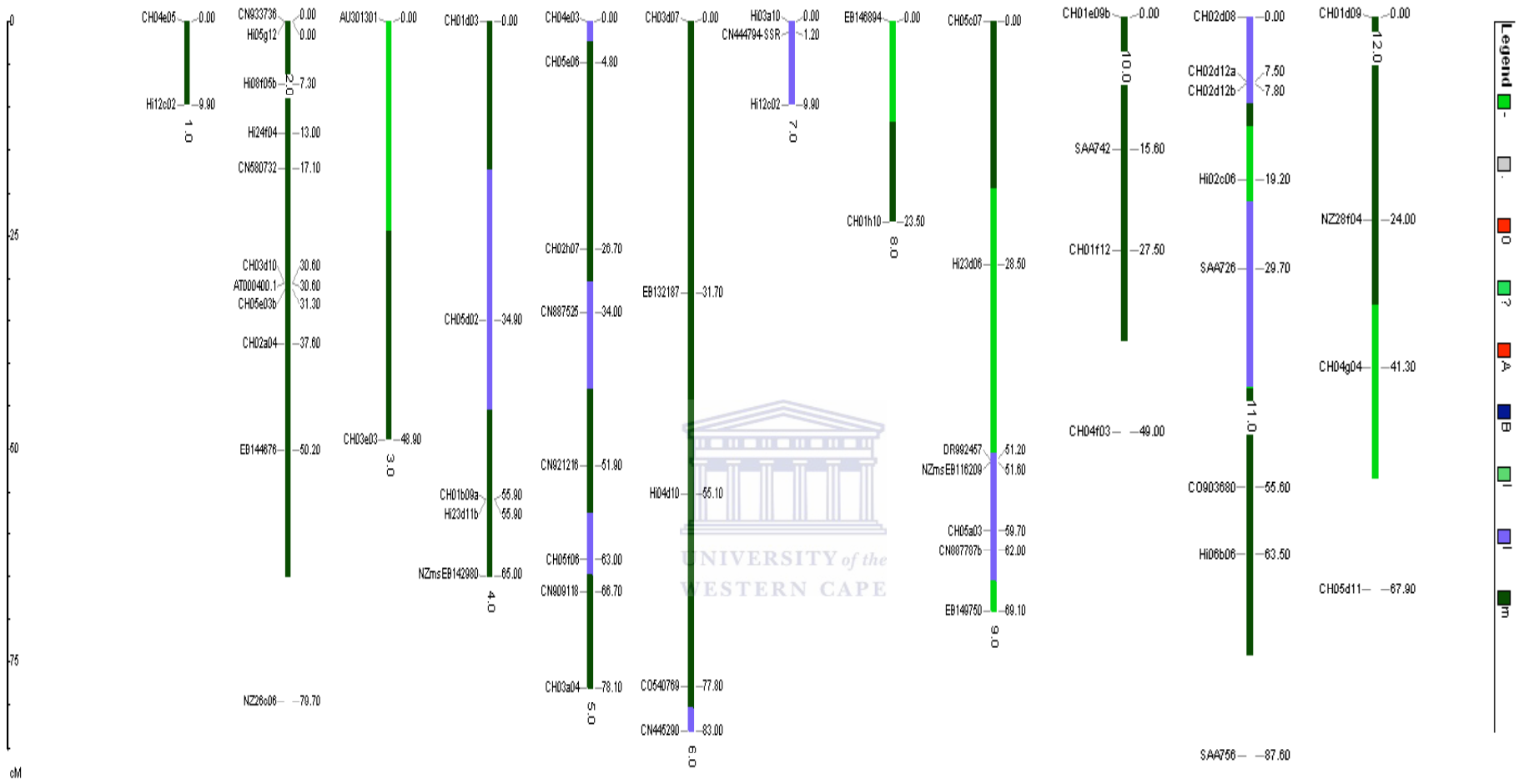


Figure 4.8: Graphical Genotyper representation for markers positioned onto LG14

LG14 is graphically represented using the 'Linkage group view mode' from Graphical Genotyper for individuals 1 to 69 of the 'Golden Delicious' x 'Dietrich' population. The numbers represent the individuals. For each panel, each column represents the chromosomal organisation for a single individual. Individual markers reflect the half-recombinational distance between adjacent markers.



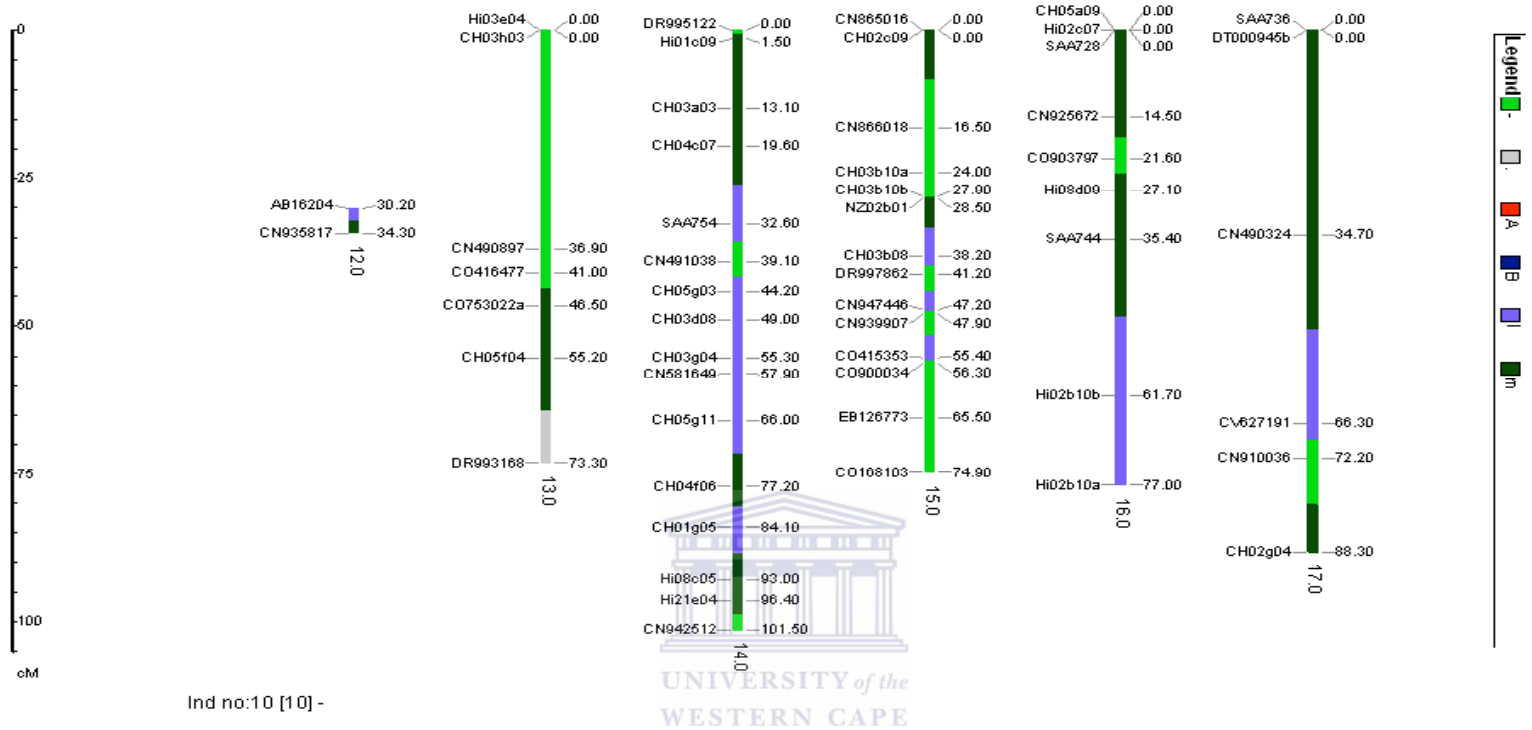


Figure 4.9: Graphical Genotyper representation for markers positioned on 17 linkage groups for individual 1-10.

Individual 1-10 of the 'Golden Delicious' x 'Dietrich' population is graphically represented using the 'Individual view mode' from Graphical Genotyper for the 17 linkage groups. The allele genotypes are colour-coded: 'lm' (deep green) is represented by 'm', 'll' (purple) is represented by 'l' and missing data are represented by '–' (light green). Individual markers in each linkage group reflect the half-recombinational distance between adjacent markers or between an adjacent marker and the end of the linkage group.

4.4 DISCUSSION

The aim of this chapter was to construct genetic linkage maps for the 'Golden Delicious' and 'Dietrich' parents using a population derived from a cross between the two cultivars. Genetic linkage map construction was achieved by directed genotyping utilizing highly polymorphic SSR markers. Directed genotyping is a target-directed, cost- and time- effective approach for the genome-wide genotyping of new crosses, cultivars and breeding lines (Silfverdberg-Dilworth *et al*., 2006). Therefore, in this chapter, a segregation analysis was performed using both published SSR markers (Liebhard *et al*., 2002; Vinatzer *et al.*, 2004; Rose and Falush, 2005; Silfverdberg-Dilworth *et al*., 2006) and those developed in our laboratory, to construct the genetic linkage maps using the 'Golden Delicious' x 'Dietrich' mapping population. Since SSR markers are transferable among genetic linkage maps within a given species and can be used as orthologus markers between species (Yamamoto *et al*., 2001; Silfverdberg - Dilworth *et al*., 2006; Celton *et al*., 2009), these were used in this study to allow a comparative analysis between the newly constructed and previously published maps. This study reports the first development of genetic linkage maps using megaplex PCR for the construction of genetic linkage maps of 'Golden Delicious' and 'Dietrich' apple cultivars.

4.4.1 Analysis of megaplex PCR amplification

Following evaluation of the genomic DNA quality (Figure 4.1), megaplex PCR amplifications were perfomed using 26 megaplex SSR primer mixes (Table 2.1). The implementation of the megaplex PCR relied on the fluorescently labelled markers as well as the size variation of the PCR products. The SSR markers were obtained from

different sources: genomic libraries, publicly available EST databases, literature and others were developed from other *Maloideae* species (Table 2.1). The developed SSRs were screened over a restricted number of apple cultivars (10 in this case; data not shown) to allow estimation of the level of polymorphism (Silfverdberg-Dilworth *et al*., 2006). Both the newly developed and published markers were then megaplexed. However, during the development of these megaplexes, variable annealing temperatures (Tm) were tested to determine the optimum Tm (58°C) for high throughput megaplex PCRs. High Tm was used to promote specificity, while limiting nonspecific amplification of additional non-SSR PCR products that could hamper both scoring and multiplexing (Silfverdberg-Dilworth *et al*., 2006).

4.4.2 Determination of allele sizes

The size of the amplified PCR products for every individual were assessed using capillary electrophoresis using the ABI 3130xl Genetic Analyser and analysed using the *GeneMapper®* software. This process was used to determine the absolute fragment sizes with an accuracy of +/-1 base. However, differences in size estimation between repetitions may occur even though relative size differences among amplicons of tested cultivars are constant (Liebhard *et al*., 2002). These statements were confirmed in the study of Silfverdberg-Dilworth *et al*. (2006) when 2-nt repeat SSRs were efficiently detected with 62% of them being polymorphic.

4.4.3 Analysis of segregating alleles

In this study, 'Golden Delicious' was used as parent 1 and 'Dietrich' as parent 2 (section 2.3.1). Therefore, alleles with the segregation type <lmxll> were only mapped

on ‘Golden Delicious’, while those with the segregation type <nnxnp> were mapped on ‘Dietrich’. The alleles ‘m’ and ‘p’ in this case were the polymorphic alleles that showed segregation. Thus, these segregation types, <lmxll> and <nnxnp>, together with <abxcd> and <hkxhk> were considered informative.

Out of the 440 primers tested on the ‘Golden Delicious’ x ‘Dietrich’ population, 329 (75%) were successfully amplified and tested for polymorphism. Therefore, 111 (25%) did not amplify and thus could not be used for the map construction. A total of 43 SSR primers, representing 13% of the 329 successfully amplified SSR primers showed homozygosity. While, nine SSR primers (equivalent to 3.1%) of those showing heterozygosity did not amplify on at least 30% of the mapping population. The χ^2 values of these markers was greater than 20 and thus, they were not included during the construction of genetic linkage maps. However, their effect on the map was tested, and was shown to induce distortions of the linkage groups.

A total of nine null alleles, markers that could not to be detected, were observed during the determination of allele sizes, possibly because of preferential amplification of smaller alleles or non specific amplification of rogues. In this study, no specific rogues were detected. The failure of primer annealing to the expected DNA fragment may also be due to nucleotide sequence divergence caused by point mutations. However, since null alleles are more frequent in megaplex reactions, a Tm of 58°C was employed to limit false null alleles. This is as a result of competition between primers for *Taq* polymerase and dNTPs, resulting in failure of amplification of larger fragments (Eckert and Kunkel, 1991; Markoulatos *et al.*, 2002).

The presence of null alleles may be detrimental in studies involving parentage testing or cultivar identification (Scalfi *et al*., 2003). On the contrary, in segregating mapping populations, the presence of a null allele does not interfere with the analysis and scoring of markers, some of these markers may be used in the construction of a genetic linkage map (Maharaj, 2007). For the construction of the linkage maps from the 'Golden Delicious' x 'Dietrich' population, some null allele markers were included.

4.4.4 Development of the genetic linkage maps for the apple cultivars 'Golden Delicious' and 'Dietrich'

4.4.4.1 Analysis of the locus genotypic frequency data

Following determination of allele sizes and exportation of data to *JoinMap®,* the locus genotypic frequency, which allows testing of segregation distortion among individuals at each locus, was performed using the χ^2-test on a locus by locus basis, also called *single locus test*, as recommended by Ruiz and Asins (2003). The χ^2-test is a convenient large sample approximation of the exact binomial test with *p*-value of 0.5 (Nixon, 2006). The χ^2-test on a locus by locus basis was used in this study to assess the significance of allele frequency variation from the expected 1:1 ratio. In comparison with the *overall test,* a mathematical test derived by incorporating information from all the loci of a chromosome simultaneously, taking into consideration the fact that loci are genetically linked (i.e. individual loci are not independent), the *single locus test* was detected as more powerful with greater probability of detecting selection (Nixon, 2006). In this regard, different levels of significance in loci segregation were detected (Table 4.3), which are crucial for determining the extent of segregation distortion among individuals at each locus (Zou *et al.*, 2007).

Following the analysis of locus genotypic frequency, all the χ^2 values greater than 20.1, correlating with a low *p* value, corresponded to distorted loci segregation resulting from systematic scoring errors on the individual loci. On the contrary, the χ^2 values lower than 20.0 were highly significant, as indicated by low distortion significance. Therefore, the χ^2 values give an indication of the most affected loci by selection and the direction in which selection is acting (Ruiz and Asins, 2003). The high χ^2 values obtained as a result of segregation distortion led to insufficient linkage among most markers.

4.4.4.2 Groupings analysis and creation of groups

In order to establish groups of linked markers, groupings (tree) were calculated using *JoinMap®*. A total of 32, 38 and 43 groups were generated for the construction of the 'Golden Delicious', 'Dietrich' and integrated maps, respectively (Table 4.4). However, 72 markers failed to link with at least one other marker during the generation of groups. Out of these markers, 34 had the segregation type <hkxhk>, 24 had the segregation type <nnxnp>, 11 had the segregation type <efxeg>, seven had the segregation type <lmxll> and six had the segregation type <abxcd> (Appendix I). Most of the markers with the <hkxhk> segregation type displayed high χ^2 values, even after modifying the genotypic ratio from 1:2:1 (hh:hk:kk) to 1:3 (hh:hk,kk) or 3:1 (hh,hk:kk). The genotypic ratios were modified to confirm the alternative allele distribution in the population, which may reduce the χ^2 values if the alternative genotypic ratio approximates the number of individuals with each genotype. The segregation distortion was then reduced and the probability of a marker to associate with other markers increased. Therefore, the high χ^2 values, as previously mentioned during the analysis of the locus genotypic frequency data, resulted from high allele segregation distortion in the population, and caused the observed genotypic ratio to diverge from the expected genotypic ratio (Zou *et al.*,

2007). It is important to note that some markers with a low χ^2 value, for example marker A603 (χ^2=0.00), could not link to any group. This could also be a result of insufficient linkage with other markers.

In addition, a total of 30 markers were also excluded during the establishment of groups because of unexpected double crossover events. Some of these markers had marker distance of more than 50 cM between adjacent markers. Thus, the marker distance was above the recommended maximum distance of 50 cM between adjacent markers (Celton, personal communication). Besides, some markers could not link to any of the established groups because they were situated at the periphery of the linkage groups.

4.4.4.3 Construction of genetic linkage maps

In order to construct genetic linkage maps, the groups established from groupings (tree) using *JoinMap®* were exported to *MapChart©*. A total of 44, 40 and 45 groups generated were linked for the construction of the 'Golden Delicious', 'Dietrich' and integrated linkage maps, respectively. Out of these groups, 40, 37 and 37 were positioned to form the 17 linkage groups of the 'Golden Delicious', 'Dietrich' and integrated maps, respectively. On the other hand, four, three and eight groups could not be positioned to any linkage group on the 'Golden Delicious', 'Dietrich' and integrated maps, respectively. These groups fail to be positioned because they contained almost exclusively predicted or newly developed markers (Table 4.5), whose location on linkage group is therefore unknown. Markers in these groups also had insufficient linkage to published markers with known linkage groups and thus could not be connected to any other group.

This study reported for the first time the construction of genetic linkage maps using megaplexed SSR markers. The maps contained 167 loci and span 1,437.8 cM and 1,491.5 cM for 'Golden Delicious' and 'Dietrich', respectively. The lengths were comparable to the maps constructed from the 'Fiesta' x 'Discovery' population (Silfverberg-Dilworth *et al.*, 2006), which span 1,145.3 and 1,417.3 for 'Fiesta' and 'Discovery', respectively. The mapping of SSR markers in the present study in common with other published maps for the 'Fiesta' x 'Discovery' (Silfverberg-Dilworth *et al.,* 2006), 'Ralls Janet' x 'Delicious' (Igarashi *et al.*, 2008) and 'Malling 9' x 'Robusta 5' (Celton *et al*, 2009) mapping populations, allowed the alignment of the maps, and thus comparisons. The coverage of the linkage groups was comparable to the reference map of 'Fiesta' x 'Discovery', with the exception of LG1, LG3, LG7 and LG8 for the 'Golden Delicious' map and LG1, LG3, LG6, LG7, LG8, LG10 and LG13 for the 'Dietrich' map (Figure 4.7). In order to fill the gaps in these linkage groups, SSR markers that were monomorphic in this study, but polymorphic in previous studies, the markers may be sequenced to identify SNPs and then mapped to the desired location.

In this study, a total of 167 SSR markers were mapped on the 'Golden Delicious' and 'Dietrich' maps. Out of these, 33 were developed in this work (Table 4.6). In addition, a total of 124 and 116 markers were positioned on 'Golden Delicious' and 'Dietrich' maps, respectively (Table 4.7). Previous studies also reported the use of SSR markers for the constuction of genetic maps. The maps of 'Fiesta' x 'Discovery' (Silfverberg-Dilworth *et al*., 2006) and 'Malling 9' x 'Robusta 5' (Celton *et al*., 2009) contained 300 and 224 SSR markers, respectively. Only 62 SSR markers were positioned for 'Discovery' x 'TN10-8' (Calenge *et al*., 2004) and 20 for 'Telamon' x 'Braeburn' (Kenis and Keulemans, 2005).

To date, only few SSR markers have been positioned on LG1 for every linkage map constructed in apple (Igarashi *et al*., 2008; Kenis and Keulemans, 2008; van Dyk, 2008; Celton *et al*., 2009; Pattocchi *et al*., 2009). In accordance with these previously reported maps, about 10 SSR markers were linked onto LG1. However, in this study, a newly developed SSR marker, SAA744 was, reported mapped on LG1. LG8 of 'Golden Delicious' was the smallest group of all (0.9 cM). However, most of the linkage groups, with the exception of LG1, LG3, LG6, LG7, LG8, LG14 and LG15 for the 'Golden Delicious' map and LG1, LG3, LG6, LG7, LG13 and LG14 for the 'Dietrich' map, were formed by at least two groups not linked together.

4.4.4.4 Analysis of linkage groups

Each of the three genetic linkage maps was composed of 17 linkage groups, representing the 17 chromosomes in apple. Only two markers, each, were positioned in LG1 and LG3, while LG14 had at least 14 markers and was the linkage group with the largest number of markers (Table 4.7).

Only four markers were mapped on LG1 for both parental maps. Most of the markers selected from the reference maps (Silfverberg-Dilworth *et al*., 2006; Igarashi *et al*., 2008; Celton *et al*., 2009) could not be positioned in this study. The marker location at the end of the linkage group or an excessive distance separating the markers may have caused the markers to fail positioning. This may also explain the insufficient linkage observed among markers on this linkage group. One newly developed marker, the SAA744, was positioned on LG1, but for the 'Dietrich' parental map only. Besides, a marker discrepancy was observed for marker CH04e05. This marker was mapped on LG7 of both 'Fiesta' and 'Discovery' maps (Silfverberg-Dilworth *et al*., 2006), while in

the current study, it mapped on LG1 of 'Golden Delicious'; CH05g08 previously mapped on LG1 of both 'Fiesta' and 'Discovery' maps (Silfverberg-Dilworth *et al*., 2006), while in current study, it mapped on LG11 of both 'Golden Delicious' and 'Dietrich' maps. The mapping of a SSR marker to two different linkage groups may indicate homeologous pairs of chromosomes in the apple genome (Maliepaard *et al*., 1998; Liebhard *et al*., 2002; Celton *et al*., 2009).

A total of 16 markers were positioned on LG2 and were enough to densify the linkage group. This assumes that the level of recombination among the markers was significant enough to form a single group. However, the linkage group was truncated into two separate groups. The presence of these two groups, resulted in a linkage group (average of the two parental maps) 43.4 cM and 37 cM longer than for the 'Fiesta' x 'Discovery' (Silfverdberg-Dilworth *et al.,* 2006) and 'Malling 9' x 'Robusta 5' (Celton *et al*., 2009) genetic maps, respectively. These important length differences resulted in an increased marker density for the maps of 'Golden Delicious' and 'Dietrich'. However, the total number of markers was comparable with the two reference maps from Silfverdberg-Dilworth *et al.* (2006) and Celton *et al*. (2009). In addition, the newly designed CN933738, CN580732, CN944528, EB144676 and EB106592 markers significantly linked with markers positioned on LG2.

Only two markers, which were 48.9 cM apart, were positioned on LG3 for both parental maps. Most of the markers from the reference maps (Silfverberg-Dilworth *et al*., 2006; Igarashi *et al*., 2008; Celton *et al*., 2009) could not be positioned on this linkage group in the present study. This could be caused by their location being at the end of the linkage group or an excessive distance separating them. This may also explain the

insufficient linkage observed between the two markers mapped on this linkage group. A total of 10 and 11 markers were mapped on 'Malling 9' and 'Robusta 5', respectively, giving rise to marker densities of 0.09 and 0.16 markers/cM (Celton *et al*., 2009). In comparison, marker densities of only 0.04 markers/cM were obtained for both 'Golden Delicious' and 'Dietrich', suggesting that more markers were required to further saturate the group.

A total of five markers positioned on LG4 for both parental maps, which were common between the parental maps. Differences were observed only for the distance separating adjacent markers. Some markers belonging to this linkage group could not associate as a result of insufficient linkage to the linked markers, suggesting important segregation distortion (χ^2 values greater than 20.1) of the loci. Previously published but unmapped CH01b09a and CH01b09b markers (Liebhard *et al*., 2002) were positioned on the 'Golden Delicious' and 'Dietrich' maps, respectively. However, the low marker density and the overall linkage group length was comparable with the reference maps from Silfverdberg-Dilworth *et al*. (2006).

A total of 12 markers were positioned on LG5. The linkage groups for both 'Golden Delicious' and 'Dietrich' were truncated, giving a total distance of 112.6 cM and 127.2 cM, respectively (Figure 4.7). The linkage group from 'Dietrich' was the longest and even longer than in 'Fiesta' x 'Discovery' (Silfverdberg-Dilworth *et al.,* 2006) and 'Malling 9' x 'Robusta 5' (Celton *et al*., 2009). The longest LG5 in these previous studies was 110.5 cM detected in the 'Discovery' map. The positionment of additional markers like SNPs, AFLPs or DArTs, a sequence-independent, high-throughput and cost-effective whole-genome fingerprinting novel method to discover and score genetic

polymorphic markers, may provide sufficient linkage to separate groups and thus allow their fusion into one single segment. This could also assist in reducing the length of the linkage group. In comparison to the reference maps from 'Fiesta' x 'Discovery' (Silfverdberg-Dilworth *et al.,* 2006), the position of CH05e06 and CH03e04 markers was inverted (Figure 4.7), thus causing the crossing of lines joining common markers between the maps of 'Golden Delicious' and 'Dietrich'. The previously published but unmapped CN909118, CN921218 and DR995002 markers (Liebhard *et al*., 2002) were positioned on LG5 in the current study.

A total of five and four markers linked well with no truncations on the 'Golden Delicious' and 'Dietrich' maps of LG6, respectively. While the length of LG6 was comparable with the reference maps, the marker density was lower, with an average 0.07 markers/cM for 'Golden Delicious' x 'Dietrich' instead of 0.18 markers/cM for 'Malling 9' x 'Robusta 5' (Celton *et al*., 2009). In addition, four published and unmapped markers, namely EB132187, CN445290, CN580227 and CO540769, were linked to LG6 in this study.

Only two markers were mapped on LG7 for both parental maps because most of the markers selected from reference maps (Liebhard *et al*., 2002; Silfverberg-Dilworth *et al*., 2006; Igarashi *et al*., 2008; Celton *et al*., 2009) could not be positioned. This might be due to their location at the end of the linkage group or separation by a large distance, thus reducing the probability of association. The two CN444794 and Hi03a10 markers, positioned in the upper part of the linkage group, were the only two markers common in the maps from both parents. In accordance with the reference maps, LG7 was the linkage group with the second highest marker density after LG8 of 'Golden Delicious'.

Only four markers were mapped on LG8, two on 'Golden Delicious' map and four on 'Dietrich' map, spanning a distance of 0.9 cM and 24.4 cM, respectively. Most markers tested, both newly developed and published, could not link to this linkage group, possibly because of the marker position on the linkage group or a large distance separating them. However, the marker density observed was on average 1.18 markers/cM, which was higher than the average marker density of 0.2 markers/cM observed from the linkage group in the 'Malling 9' and 'Robusta 5' genetic maps, respectively (Celton *et al*., 2009). It is thus important to note that even though the marker densities was higher in this study, the LG8 from the 'Malling 9' and 'Robusta 5' maps were composed of 17 and 11 markers, respectively, most of them being SSRs. This shows that the marker density is influenced by both the number of markers and the length of the linkage group.

A total of 10 markers were mapped on LG9, forming two truncated groups for both parental maps. Although the number of markers positioned was important enough to form a single framework linkage group, the level of association among markers was insufficient. The linkage groups of 'Golden Delicious' and 'Dietrich' were 88.7 cM and 82.5 cM in length, respectively. These maps are comparable to the 'Fiesta' x 'Discovery' (Silfverberg-Dilworth *et al*., 2006) and 'Malling 9' x 'Robusta 5' (Celton *et al*., 2009) reference maps, with the longest LG9 being 72.5 cM in the 'Discovery' map. In addition, the previously published but unmapped CH05a03, DR992457, EB147667, EB149750 and CN87787b markers, as well as the newly developed SAA763 marker were successfully positioned on LG9 in the 'Golden Delicious' and 'Dietrich' genetic linkage maps.

In the ‘Golden Delicious’ map, a total of nine markers were positioned on LG10, but these were separated into two groups. In the ‘Dietrich’ map, only seven markers were positioned, but divided to three separate groups. The occurrence of groups can be caused by an insufficient level of association among the markers. In both the ‘Golden Delicious’ and ‘Dietrich’ maps, LG10 was longer than in the ‘Fiesta’ x ‘Discovery’ (Silfverberg-Dilworth *et al*., 2006) and ‘Malling 9’ x ‘Robusta 5’ (Celton *et al*., 2009) reference maps, the longest LG10 spanning over 97.5 cM in ‘Discovery’. However, the addition of extra markers, like SNPs, AFLPs or DArTs, may provide sufficient linkage to separate groups, and thus allow LG10 to be represented as a single group. This could also assist in reducing the length of the linkage group. When comparing to the reference maps, the SSR markers positioned and the marker density of LG10 were comparable. The previously published but unmapped CH01e09b, CO865955 and CN495393 markers were successfully positioned to LG10.

A total of 17 markers were mapped on LG11. In the ‘Golden Delicious’ map, LG11 was composed of eight markers and divide into two separate groups, while in the ‘Dietrich’ map, LG11 was created from 13 markers and formed as two separate groups. In regards to marker density, a high marker density of 0.29 markers/cM was observed in the ‘Robusta 5’ map (Celton *et al*., 2009), while 0.09 markers/cM was observed in the ‘Golden Delicious’ map. This significant difference is probably influenced by the length of the linkage group, since the length of LG11 for the ‘Golden Delicious’ x ‘Dietrich’ maps was on average 25.5 cM and 36.3 cM greater than for ‘Fiesta’ x ‘Discovery’ and ‘Malling 9’ x ‘Robusta 5’ maps, respectively. However, addition of more markers to saturate LG11 can allow increasing the linkage among individual

groups, and in turn may reduce the linkage group length. In addition, the two newly developed SAA726 and SAA756 markers were successfully positioned to LG11.

A total of nine markers were mapped on LG12. In the 'Golden Delicious' maps, nine markers were positioned on two separate groups, while in the 'Dietrich' map, eight markers were located, forming two separate groups. Concerning marker density, a high marker density was observed in 'Robusta 5', 0.18 markers/cM (Celton *et al*., 2009), while only 0.08 markers/cM were observed in the 'Golden Delicious' map. This important difference can be explained by the length variations of the linkage group and the number of markers. The length of LG12 from the 'Golden Delicious' x 'Dietrich' were on average 36.9 cM and 49.0 cM longer than the linkage group of 'Fiesta' x 'Discovery' and 'Malling 9' x 'Robusta 5' maps, respectively. Besides, the addition of extra markers to saturate LG12 would increase linkage among the groups, which in turn may allow linking them together and thus reducing the linkage group length.

A total of seven markers were mapped on LG13. In 'Golden Delicious', LG13 consisted of seven markers forming two separate groups, while in 'Dietrich', LG13 consisted of three markers linked as one group. In regards to marker density, in a study from Celton *et al.* (2009), a high marker density of LG13 was 0.19 markers/cM observed on 'Robusta 5' map, while a high marker density of 0.04 markers/cM was reached in 'Golden Delicious'. This significant difference can be explained by the low number of markers mapped on the linkage group in comparison with the length of the linkage group. For 'Robusta 5', 10 markers were positioned on a 52.2 cM long linkage group, while for 'Golden Delicious', the linkage group consisted of 7 markers only and a length of 128.0 cM. This lower ratio of marker to length explained the low marker

density in this group, although the number of markers positioned on LG13 on both the maps from Celton *et al*. (2009) and the present study is comparable.

In terms of LG14, 14 markers were mapped for the 'Golden Delicious' map and 17 markers for the 'Dietrich' map. This linkage group was the densest in this study. The CH05g03 marker, which was positioned on LG17 in the maps by Silfverberg-Dilworth *et al*. (2006), was mapped on LG14 in this study. Although this marker was not positioned on LG14 in the reference map by Silfverberg-Dilworth *et al*. (2006), its positioned was conserved since its elimination caused an increase in the linkage group size from about 102 cM to 146 cM and as well, this marker did not link to LG17. The marker therefore played an important role on the size of LG14 and the linkage between markers. Two previously published and unmapped markers, CN491038 and CH03a03, were successfully positioned on this linkage group during the construction of the genetic linkage maps. The overall length of the linkage groups was 98.9 cM and 102.5 cM for 'Golden Delicious' and 'Dietrich' maps, respectively. The average distance between markers in this linkage group was 6.2 cM in this study, providing good framework linkage group coverage for QTL analysis. In terms of marker density, highest marker density from the 'Malling 9' x 'Robusta 5' population was 0.18 markers/cM on the 'Robusta 5' map, while in the 'Golden Delicious' x 'Dietrich' population, 0.17 markers/cM was observed on the 'Dietrich' map. Thus, these marker densities were comparable in both studies.

A total of 15 markers were mapped on LG15. Although a similar marker density was observed, the linkage group of 'Golden Delicious' x 'Dietrich' was on average 18.25 cM longer than the 'Fiesta' x 'Discovery' reference map (Silfverberg-Dilworth *et al*.,

2006). This might be due to the occurrence of truncated groups on 'Dietrich' map. The positionment of extra markers would allow saturating LG15, and thus possibly link the two groups. This could result in the reduction of the linkage group length. The previously published and unmapped CO168103, CH02d10b and EB126773 markers (Liebhard *et al.,* 2002) were successfully positioned in this linkage group. In addition, 29% of the markers in LG15 were common to the linkage group from the reference map by Silfverberg-Dilworth *et al*. (2006).

A total of 13 markers were mapped on LG16, structuring two and four truncated groups in 'Golden Delicious' and 'Dietrich', respectively. This linkage group was the longest of the maps, suggesting a high distortion in the map. The CH01f03b and Hi02c07 markers were positioned on LG16 in this study, while they were mapped on LG9 and LG1, respectively, in the reference map by Silfverberg-Dilworth *et al*. (2006). In regards to marker density, the highest marker density of 0.21 markers/cM was observed for 'Robusta 5' map (Celton *et al*., 2009), while only 0.07 markers/cM was observed in both the 'Golden Delicious' and 'Dietrich' maps. The variations in linkage group length and number of markers might have caused the divergence in marker density observed between the two studies. The length of LG16 from the 'Golden Delicious' and 'Dietrich' maps was on average 61.5 cM and 77.5 cM longer than for the 'Fiesta' x 'Discovery' and 'Malling 9' x 'Robusta 5' populations, respectively. In order to increase linkage among group, achieve a single group and reduce the length of LG16, more markers need to be mapped to saturate the linkage group. The CH05a09 marker that has been previously published but unmapped (Liebhard *et al.,* 2002) was successfully positioned in LG16. In addition, two newly developed markers, SAA744 and SAA728, were positioned in LG16 for both parental maps. In the 'Dietrich' map

only, the SAA744 marker was also positioned in LG1, suggesting that the marker may be a multilocus marker, indicating that it amplifyed on more that one locus (Patocchi *et al.*, 2009).

In terms of LG17, nine markers were mapped forming two separate groups for both parental maps. A total of 88% of the markers were common in both parental maps, but varying distances were observed between adjacent markers. In terms of marker density and number of SSR markers mapped, the linkage group from the 'Golden Delicious' and 'Dietrich' maps were comparable to the 'Fiesta' x 'Discovery' (Silfverdberg-Dilworth *et al*., 2006) and 'Malling 9' x 'Robusta 5' (Celton *et al*., 2009) maps. However, LG17 was longer by 21.8 cM than the LG17 from 'Fiesta' x 'Discovery'.

4.4.5 Graphical genotyping analysis

The graphical genotyping analysis was used in the selection of informative markers, identification of the most useful individuals in a segregating population and detection of errors in the data (Young and Tanksley, 1989; Severson and Kassner, 1995). In addition, recombination crossovers in the mapping population were also identified. This graphical genotyping analysis allowed the rapid and concise assessment of data for large numbers of markers for individuals within the 'Golden Delicious' x Dietrich' population. Severson and Kassner (1995) also reported use of graphical genotyping to compare the entire genome of mosquito populations exhibiting contrasting phenotypes inoder to detect putative associations among genome segments.

In the present study, the graphical genotyping analysis was performed to display chromosomal genotypes by observing the segregation of the SSR marker data on the 17

linkage groups over the entire set of individuals of the ‘Golden Delicious’ x ‘Dietrich’ population. The analysis allowed the visualization of some genetic anomalies in the segregation of markers for several individuals. In the linkage group view mode, some genomic regions reflected partial and complete linkage disequilibrium, as no recombination was evident. In addition, the study showed that some linkage groups like LG1, LG2 and LG10 were heterozygous on every locus, while others like LG15, displayed high recombination crossover among loci.

CHAPTER 5

IDENTIFICATION OF QUANTITATIVE TRAIT LOCI

5.1 INTRODUCTION

Efficient breeding programs and breeding itself together with the selection of high quality apples requires knowledge and understanding of the underlying genetic principle. The availability of high quality genetic linkage maps constructed with molecular markers such as SSRs, enables the study of the whole genome structure and localization of genes of interest (Mohan *et al.*, 1997; Gupta *et al.*, 2003; Liebhard *et al.*, 2003a; Gardiner *et al.*, 2007). In addition, this allows the detection and analysis of major genes and quantitative trait loci (QTL) contributing to quality traits of a genotype (Tanksley *et al.*, 1989; Tanksley, 1993; Liebhard *et al.*, 2003a; Scalfi *et al.*, 2004). The genetic linkage maps and QTL data are essential for breeders who opt to use marker-assisted selection (MAS) because they facilitate the selection of genotypically superior individuals at an early age. Prior to selection of individuals with the required traits, molecular markers showing genetic linkage with the traits of interest are identified. This identification process, which uses genetic linkage maps and phenotypic trait assessment data is termed QTL mapping (section 1.6).

Recently, a more reliable and informative technique was introduced which allows for selection to be based on genotyping of individuals using molecular markers like SSRs (section 1.4.1) (King *et al.*, 2000; Liebhard *et al.*, 2003a; Kenis and Keulemans, 2005; Silfverberg-Dilworth *et al.*, 2006; Sargent *et al.*, 2007; Igarashi *et al.*, 2008; Celton *et al.*, 2009). These co-dominant SSR markers, enable the alignment of genetic linkage maps obtained from different cultivars, and closely related species, resulting in accurate

chromosomal position comparison between QTLs identified in different mapping populations. According to Silfverberg-Dilworth *et al*. (2006), markers with an average distance of one marker per 15 cM and less improve the resolution of the map. This may not necessarily contribute significantly to the number of QTLs identified, but it will facilitate a more accurate or precise identification of the genomic regions that contain a QTL and may even lead to specific gene identification (Vision *et al.*, 2000; Vogl and Xu, 2000).

In this study an attempt was made towards identifying genomic regions containing QTLs controlling fruit quality traits, namely firmness, crispness, juiciness, acidity, form, colour, stripness, russeting and size. The aim was to identify putative QTLs for these quality traits on genetic linkage maps of 'Golden Delicious' and 'Dietrich'. Following identification of QTLs, closely linked SSR markers can be identified, which may be implemented in MAS for breeding cultivars that produce fruits of high quality with consumer acceptable characteristics.

5.2 RESULTS

5.2.1 QTL mapping overview

The QTL analysis of the ANOVA estimates of the traits of interest for each genotype was performed using the *MapQTL®* V5 software. The 'Golden Delicious' and 'Dietrich' maps as well as the integrated genetic linkage maps (Chapter 4), and phenotypic data collected over three years (Chapter 3) were used for QTL analysis. These data was used to detect the parental chromosome carrying the effective allele, the positions and effects of the QTL using the Krustal-Wallis (KW) single locus analysis (nonparametric

mapping). In addition, the study used the maximum likelihood based interval mapping (IM) and the multiple QTL mapping (MQM) approaches.

The IM was carried out following which markers whose values were very close to the significance threshold level were used as cofactors for MQM to pull out all possible QTLs. Putative QTLs were positively identified when the likelihood of the odds (LOD) value of a marker was above the significance threshold line. For example, using A142b (Accession CH02b01b) as a cofactor, one QTL was detected on LG15, while no QTL was in LG13 (Figure 5.1). In addition, the effect of the QTL on the control of the trait of interest was shown by the blue line, explaining phenotypic variance of approximately 50%. However, it is important to note that significance thresholds differ from one year of phenotypic investigation to another and from one linkage group to another. Therefore, in this study, genome wide significance thresholds were used as standard threshold and these were generally higher than linkage group thresholds.

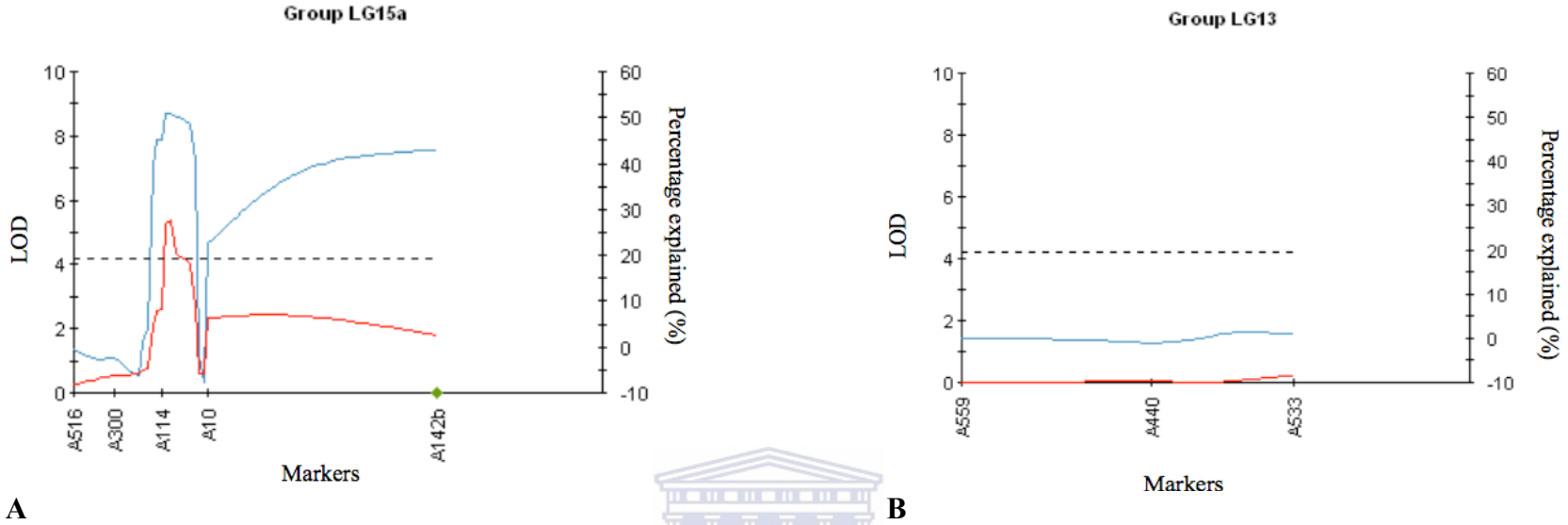


Figure 5.1: Detection of quantitative trait loci on linkage groups 13 and 15 using the likelihood of the odds test.

(A) LG15a shows the positive detection of a potential QTL, the LOD value of the peak of the red line being above 4.2, while (B) LG13 illustrates the absence of QTL, the LOD value of the peak of the red line being below 4.2. The LOD value is represented by the red line. The phenotypic percentage variance explained is represented by the blue line. The position of the cofactor used during MQM mapping is represented by the green sign at marker A142b (Accession CH02b01b, Table 2.1). The dashed line represents the threshold value equivalent to the genome wide value (p=0.95), which was set at 4.2 for QTL identification.

5.2.2 Identification of QTLs associated with fruit quality traits

Following the IM, MQM and KW analyses, a total of 72 putative QTLs controlling the nine fruit quality traits under investigation were identified. Out of these, four QTLs were associated with colour, four with crispness, five with stripness, five with form, six with juiciness, eight with acidity, eight with firmness, nine with size and 23 with russeting. Detailed results are described in sections 5.2.2.1-5.2.2.9. This study reports the first attempt to detect QTLs associated with russeting, stripness and form in apples.

5.2.2.1 Firmness

The QTL analysis conducted on the 'Golden Delicious', 'Dietrich' and 'Golden Delicious' x 'Dietrich' integrated genetic linkage maps identified 13, four and 29 possible QTLs, respectively (data not shown). After a comparative analysis between the IM and MQM data, as well as the nonparametric mapping data, the number of possible genomic regions associated with juiciness was reduced to eight, which were considered as putative QTLs for firmness (Table 5.1).

Out of the eight putative QTLs, six were identified on 'Golden Delicious' and were located on LG2, LG4, LG10, LG12 and LG16. Two QTLs were positioned on LG2, at 13.0 cM and 17.1 cM and were clearly distinguished from each other as they were separated by approximately 4 cM. The identified QTLs for 'Golden Delicious' were more effective in year II than in the other years. Two QTLs were identified on 'Dietrich' and were positioned on LG6 and LG15. These were detected in at least one year of phenotypic assessment, but none of them was identified in year I. All the QTLs detected were above the LOD significant threshold value of 4.2 (Section 5.2.1). In addition, these putative QTLs obtained high phenotypic variance explained values, with

the exception of the QTL detected on LG6, which explained for 15.6% phenotypic variance only.

Table 5.1: Putative QTLs associated with firmness detected in the segregating population of 'Golden Delicious' x 'Dietrich' using Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the locus at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection.

LG	Map	Position of Locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
2	GD	17.1	CN580732	I	4.3	69.0
2	GD	13.0	Hi24f04	II	5.7	34.6
4	GD	0.0	CH01d03	II	9.8	79.4
6	D	9.7	CO540769	III, M	4.8	15.6
10	GD	33.8	CH01f07a	II	5.6	39.1
12	GD	41.3	CH04g04	II	4.7	24.2
15	D	75.7	CH02d10b	II, III	6.7	81.5
16	GD	21.6	CO903797	II	6.2	76.0

5.2.2.2 Juiciness

A total of 14 possible QTLs were identified for the integrated genetic linkage map. The individual genetic linkage maps for 'Golden Delicious' and 'Dietrich' showed six and 11 possible QTLs, respectively. After a comparative analysis of IM and MQM data data as well as nonparametric mapping data, six genomic regions associated with juiciness were considered as putative QTLs (Table 5.2).

In terms of 'Golden Delicious', three of the putative QTLs were identified on LG4, LG9 and LG15. These were detected in every year of phenotypic investigation, including the integrated phenotypic data (M). While, for 'Dietrich', three putative QTLs were detected on LG4, LG15 and LG16, in year I and M, exclusively. In addition, the lowest phenotypic variance explained was 34.1% for a QTL detected on the LG9 of 'Golden Delicious'.

Table 5.2: Putative QTLs associated with juiciness detected in the segregating population of 'Golden Delicious' x 'Dietrich' using Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the locus at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection.

LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
4	GD	0.0	CH01d03	II, III	5.4	57.8
4	D	0.0	CH05e04b	M	4.4	39.9
9	GD	39.1	SAA763	M	4.8	34.1
15	D	47.4	CN444111	I, M	4.4	34.4
15	GD	46.5	CN494248	I, M	6.1	60.9
16	D	76.9	HI08d09	M	5.9	40.6

5.2.2.3 Fruit colour

A total of five possible QTLs were detected for the integrated genetic linkage map, while six and two possible QTLs were identified for the individual, 'Golden Delicious' and 'Dietrich' genetic linkage maps, respectively. Following a comparative analysis of IM and MQM data as well as nonparametric mapping data, four QTLs were considered as putative QTLs controlling colour (Table 5.3). Out of these, three putative QTLs were identified on LG9 and LG15 of 'Golden Delicious' genetic linkage map. Two of the QTLs were located on LG9 at positions 39.1 cM and 62.0 cM. The identified QTLs for 'Golden Delicious' were effective in year II and M only.

In terms of 'Dietrich', one was detected on LG15. This QTL was detected in year I only, an indication that the identified QTL could be unstable. All the QTLs detected were above the LOD significant threshold value of 4.5 (section 5.2.1) and were major QTLs as shown by their high phenotypic variance values.

Table 5.3: Putative QTLs associated with colour detected in the segregating population of 'Golden Delicious' x 'Dietrich' using Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the lous at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection.

LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
9	GD	39.1	SAA763	II	5.6	70.6
9	GD	62.0	CN887787b	II	5.3	70.3
15	GD	30.0	CN444111	M	7.6	72.5
15	D	47.4	CN494248	I	5.2	69.6

5.2.2.4 Fruit size

QTL analysis on the 'Golden Delicious' x 'Dietrich' integrated genetic linkage map identified 23 possible QTLs, while on individual genetic linkage maps of 'Golden Delicious' and 'Dietrich' three and 12 possible QTLs were identified, respectively. However, following a comparative analysis of IM and MQM data as well as nonparametric mapping data, the number of genomic regions associated with fruit size was reduced to nine (Table 5.4).

Out of the nine putative QTLs, three were identified for the 'Golden Delicious' and were located on LG5, LG12 and LG15, while six QTLs were identified for the 'Dietrich' on LG3, LG9, LG15, LG16 and LG17. The QTLs for 'Golden Delicious' showed to be effective in year II and III only. Two QTLs identified on 'Dietrich' and positioned on the same chromosome, LG15, could be clearly distinguished from each other as they were about 73 cM apart. Indeed, the first QTL was positioned at 0.0 cM while the second was at 73.7 cM. The QTLs controlling size were detected in years II and III, as well as M. They explained for a major phenotypic variance, with the exception of the QTL detected on LG17 of 'Dietrich' that explained for 14.4% of phenotypic variance only. All the QTLs detected were above the LOD significant threshold value of 3.8.

Table 5.4: Putative QTLs associated with fruit size detected in the segregating population of 'Golden Delicious' x 'Dietrich' using Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the locus at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection.

LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
3	D	48.9	CH03e03	II	4.6	71.9
5	GD	34.0	CN887525	II	4.6	40.8
9	D	0.0	EB149750	M	4.4	34.3
12	GD	16.9	CH01f02	II	4.3	34.3
15	GD	47.2	CN947446	III	4.6	21.1
15	D	73.7	CH02c09	III	4.1	19.0
15	D	0.0	CN494248	M	5.0	38.8
16	D	50.3	SAA728	M	6.6	38.2
17	D	35.0	CN490324	M	4.1	14.4

5.2.2.5 Stripness

A total of nine possible QTLs were identified for the integrated genetic linkage map. The individual genetic linkage maps for 'Golden Delicious' and 'Dietrich' showed nine and three possible QTLs, respectively. However, after a comparative analysis of IM and MQM data as well as nonparametric mapping data, only five QTLs were considered as putative QTLs controlling stripness (Table 5.5).

Out of the five putative QTLs, two were identified for the 'Golden Delicious' and were positioned on LG3 and LG15, while three were identified for the 'Dietrich' and were located on LG 9, LG15 and LG16. The QTLs were effective in years I and III, as well as M. In addition, they showed a high proportion of phenotypic variance explained, with the lowest being detected on LG16 of 'Dietrich', explaining for 39.1% of phenotypic variance.

Table 5.5: Putative QTLs associated with stripness detected in the segregating population of 'Golden Delicious' x 'Dietrich' using Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the locus at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection

LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
3	GD	48.8	CH03e03	III	19.6	91.6
9	D	29.0	SAA763	M	6.9	44.3
15	GD	46.5	CN494248	I	22.1	95.8
15	D	15.9	CN444111	I	24.1	96.2
16	D	0.0	CN870040	M	6.9	39.1

5.2.2.6 Crispness

The QTL analysis conducted on the 'Golden Delicious' x 'Dietrich' integrated genetic linkage map identified 16 possible QTLs. In addition, the QTL analysis performed on the individual maps for 'Golden Delicious' and 'Dietrich' identified one and five possible QTLs, respectively. However, following a comparative analysis of IM, MQM and nonparametric mapping data, the number of genomic regions associated with fruit crispness was reduced to four, which were considered as putative QTLs (Table 5.6).

Only a single putative QTLs was identified for 'Golden Delicious' and was located on LG9, while three QTLs were identified for 'Dietrich' and were positioned on LG2, LG4 and LG16. These QTLs were detected in year I and II, as well as M. All the QTLs detected were above the LOD significant threshold value, which was set at 4.2 (section 5.2.1). In addition, the QTLs showed high phenotypic variance explained with the lowest being 32.8% for a QTL detected on LG2.

Table 5.6: Putative QTLs associated with crispness detected in the segregating population of 'Golden Delicious' x 'Dietrich' using Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the locus at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection.

LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
2	D	67.6	CN944528	M	4.3	32.8
4	D	0.0	CH01b09b	M	4.9	38.4
9	GD	0.0	EB147667	II	7.9	75.8
16	D	50.3	SAA728	I, M	4.9	51.6

5.2.2.7 Acidity

A total of 14 possible QTLs were identified for the integrated genetic linkage map. The individual genetic linkage maps for 'Golden Delicious' and 'Dietrich' showed 18 and 12 possible QTLs, respectively. Overall, only eight putative QTLs were finally considered after a comparative analysis of IM, MQM and nonparametric mapping data (Table 5.7).

Out of the eight putative QTLs, four were identified on LG2, LG10, LG13 and of 'Golden Delicious', while four QTLs were identified on LG9, LG15 and LG16 of 'Dietrich'. The QTLs were more effective over the three years of phenotypic assessment. The QTL detected on LG15 was the most stable among all the detected QTLs associated with acidity, as it was identified over all the three years of investigation. All the QTLs detected were above the LOD significant threshold values of 4.2 (section 5.2.1). In addition, these displayed high proportion of phenotypic variance explained, with the exception of QTL detected on LG2 of 'Golden Delicious', which explained only 7.3% of phenotypic variance.

Table 5.7: Putative QTLs associated with acidity detected in the segregating population of 'Golden Delicious' x 'Dietrich' using Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the locus at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection.

LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
2	GD	13.0	HI24f04	II	6.9	7.3
9	D	29.4	HI23d06	M	6.9	73.0
10	GD	33.8	CH01f07a	II	6.5	86.6
13	GD	36.9	CN490897	I, II	5.2	68.7
14	GD	49.0	CH03d08	II	5.3	78.1
15	D	75.7	CH02d10b	I, II, III	5.9	46.6
16	D	50.3	SAA728	II	5.5	73.3

5.2.2.8 Form

The QTL analysis conducted on the 'Golden Delicious', 'Dietrich' and integrated genetic linkage maps identified five, seven and eight possible QTLs, respectively. After a comparative analysis of IM and MQM data as well as nonparametric mapping data, the number of genomic regions associated with fruit form was reduced to five, which were considered as putative QTLs (Table 5.8).

Out of the five putative QTLs, a single QTL was identified on LG9 of 'Golden Delicious' map, while four QTLs were identified positioned on LG5, LG11, LG15 and LG16 of 'Dietrich' map. These were detected only in one year of phenotypic assessment. In addition, these QTLs displayed high phenotypic variance explained, with the lowest being 21.5%, detected on the LG5 of 'Dietrich'. All the QTLs detected were above the LOD significant threshold value of 4.0.

Table 5.8: Putative QTLs associated with form detected in the segregating population of 'Golden Delicious' x 'Dietrich' using Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the locus at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection.

LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
5	D	30.7	CN887525	III	4.6	21.5
9	GD	51.6	NZmsEB116209	M	4.2	61.5
11	D	37.9	CH02d12	II	4.9	66.9
15	D	75.7	CH02d10b	M	4.4	77.6
16	D	50.3	SAA728	I	5.0	56.7

5.2.2.9 Russeting

A total of 34 possible QTLs were detected on the integrated genetic linkage map. The individual genetic linkage maps for 'Golden Delicious' and 'Dietrich' showed 28 possible QTLs on each parent. After a comparative analysis of IM, MQM and nonparametric mapping data, the number of genomic regions associated with fruit russeting was reduced to 23, which were then considered as putative QTLs for fruit russeting (Table 5.9).

Out of the 23 putative QTLs, ten were identified on LG4, LG6, LG9, LG10, LG12, LG13, LG14 and LG15 of 'Golden Delicious'. Even though two QTLs each were positioned on LG9 and LG15, they could be clearly distinguished from each other as they were far apart from each other i.e. the markers were positioned at 39.1 cM and 51.4 cM on LG9. For LG15, the markers were mapped at 0.0 cM and 55.4 cM. In terms of 'Dietrich', 13 QTLs were identified, which were positioned on LG2, LG4, LG5, LG9, LG10, LG11, LG15, LG16 and LG17. Two QTLs were identified on LG9, three on LG15 and two on LG16. These QTLs were detected in at least one year of phenotypic assessment. Two QTLs located on LG14 from the 'Golden Delicious' map and LG9 from the 'Dietrich' map were detected on the three years of investigation. The QTLs for 'Golden Delicious' showed to be more effective in all the three years of phenotypic assessment, since the 10 QTLs were detected in the three years of investigation. Further, these putative QTLs showed high proportion of phenotypic variance, with the exception of those detected on LG2, LG4, LG10, LG15 and LG17, which had values lower than 10%.

Table 5.9: Putative QTLs associated with russeting detected in the segregating population of 'Golden Delicious' x 'Dietrich' with Interval Mapping and Multiple-QTL Mapping.

For each QTL, the parental genetic linkage map, the linkage group, the position on the linkage group and the name of the locus at this position, the year of detection, the LOD score and the percentage of phenotypic variance explained are indicated. The letter M represents the mean phenotypic data calculated over the three years of phenotypic data collection.

LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
2	D	0.0	CH02c061	I	4.5	7.6
4	GD	34.9	CH05d02	II	6.0	26.1
4	D	0.0	CH01b09b	II	5.1	8.5
5	D	30.7	CN887525	III	4.9	77.8
6	GD	31.7	EB132187	I	4.6	13.1
9	GD	39.1	SAA763	II, III	6.8	27.7
9	GD	51.2	DR992457	I, III	5.7	22.9
9	D	29.0	SAA763	I, II, III	7.9	80.1
9	D	29.4	HI23d06	M	4.2	13.4
10	GD	33.8	CH01F07a	II	5.6	49.0

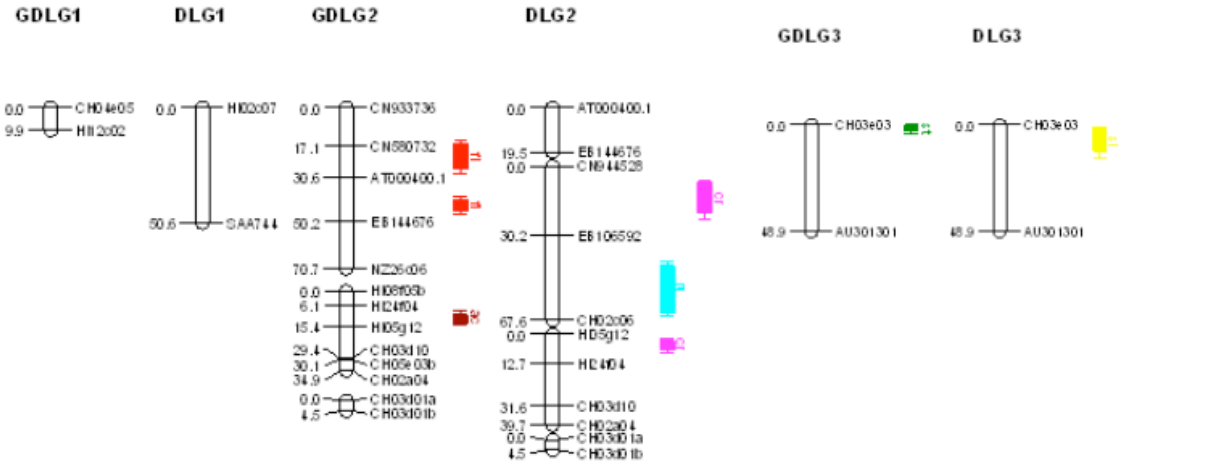
Table 5.9 continued

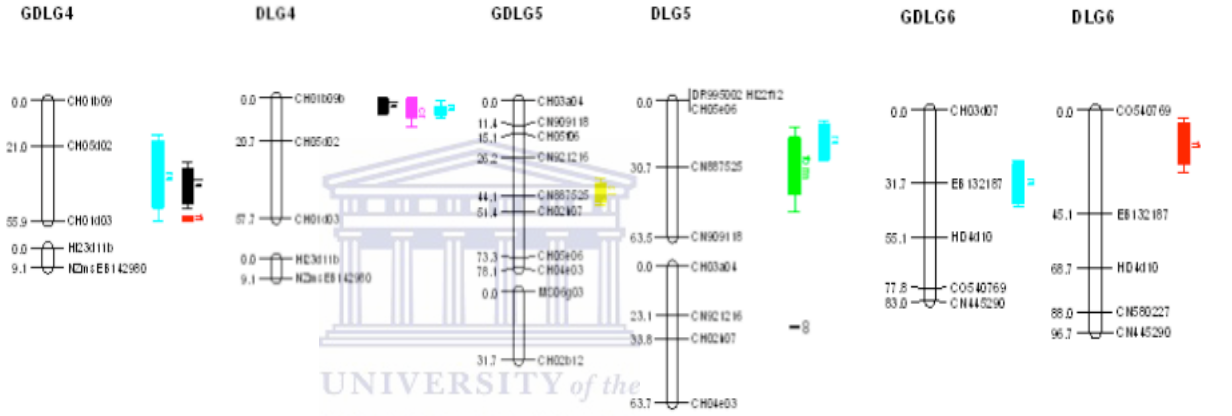
LG	Map	Position of the locus (cM)	Locus	Year	LOD	Phenotypic variance explained (%)
10	D	33.8	CH01f07a	I	5.1	8.2
11	D	40.1	Hi02c06	I, III	5.7	79.6
12	GD	67.9	CH05d11	II	5.0	26.9
13	GD	36.9	CN490897	II, III	5.0	25.5
14	GD	39.1	CN491038	I, II, III	4.9	26.5
15	GD	0.0	NZ02b01	II, III	5.0	39.1
15	GD	55.4	CH05d024	II	5.0	14.8
15	D	15.9	CN444111	I	4.5	89.2
15	D	19.0	CH05d024	I, II	6.6	3.6
15	D	75.7	CH02d10b	I, II	5.8	8.8
16	D	50.3	SAA728	II, M	4.9	37.3
16	D	76.9	HI08d09	II	4.9	76.6
17	D	35.0	CN490324	II	6.0	4.2

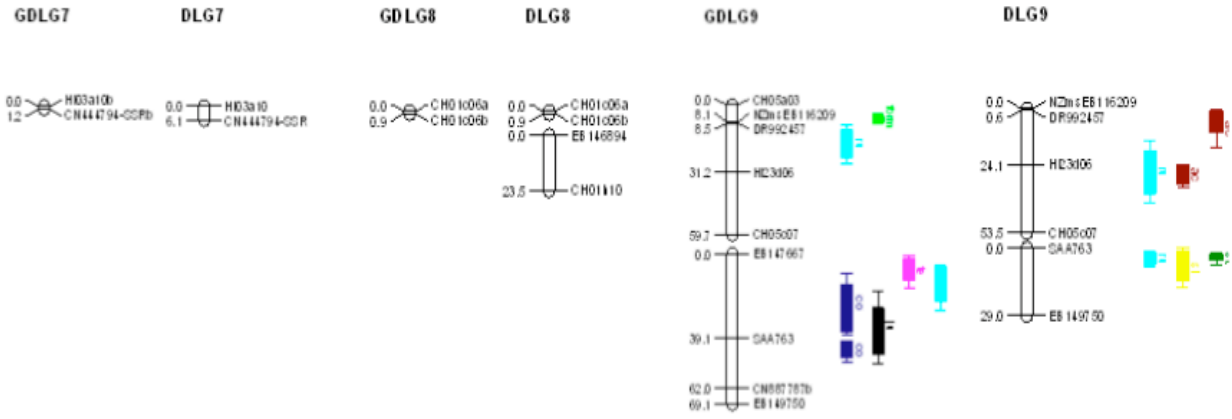
5.2.3 Graphical representation of detected QTLs

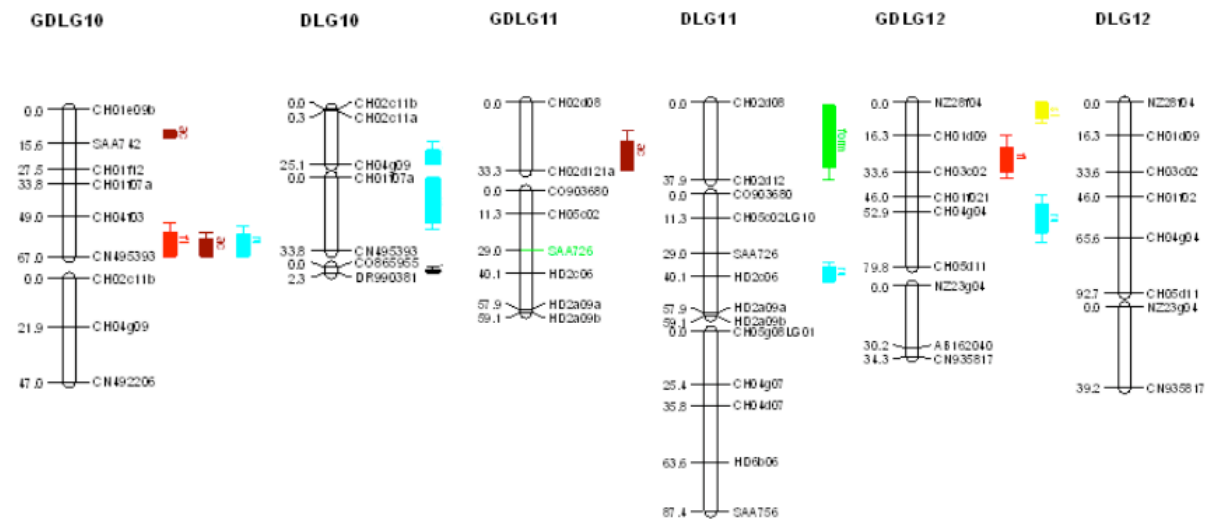
The QTLs identifed for fruit quality were positioned in either the 'Golden Delicious' or 'Dietrich' map (Figure 5.2). The detected QTLs were graphically represented as bars next to the respective linkage group, with a 5% confidence interval. Most of the QTLs were highly associated with LG15 and LG16. In addition, the diagrammatic representations in Figure 5.2 showed that most of the QTLs identified on LG16 were highly overlapping.

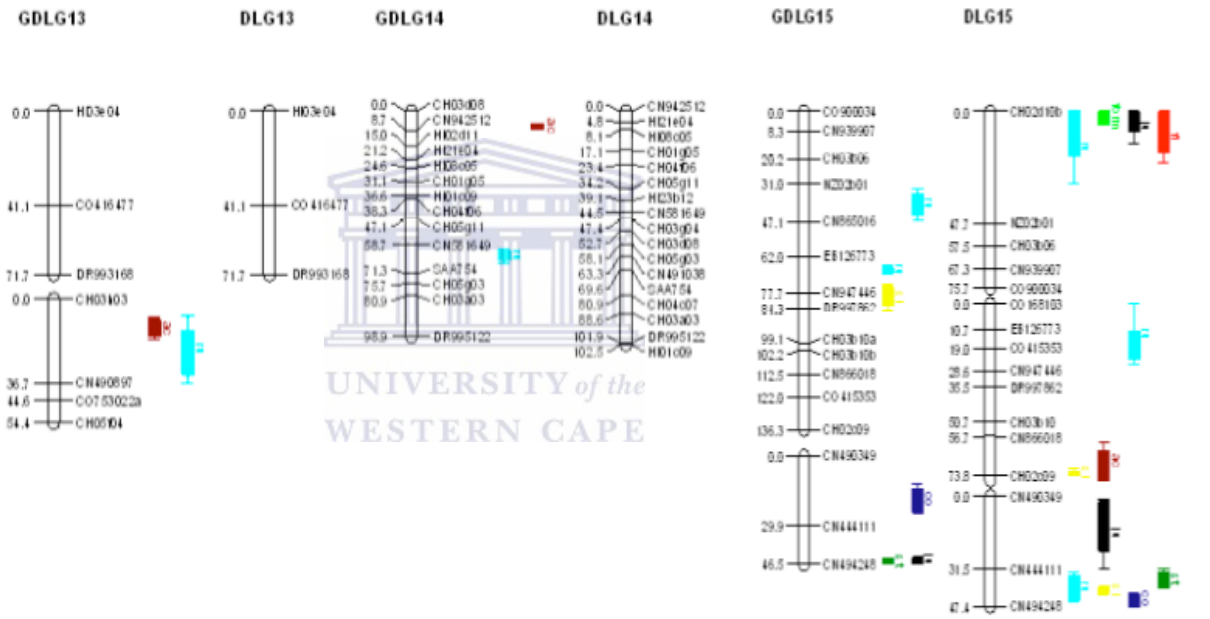
Out of the 72 QTLs detected on the nine fruit quality traits under investigation, 33% of these QTLs were identified for russeting and this represented the maximum number of QTLs for a single trait. On the contrary, the minimum number of QTLs detected for a single trait was observed for colour and crispness, both of which represented 6% each of the total number of QTLs identified in this study. The other quality traits were distributed as follows: 13% for size, 11% for firmness, 9% for acidity, 8% for juiciness, and 7% for stripness and form. A total of 33 QTLs, representing 46%, were detected in the cultivar 'Golden Delicious', while 39 QTLs, representing 54%, were detected in the cultivar 'Dietrich'.











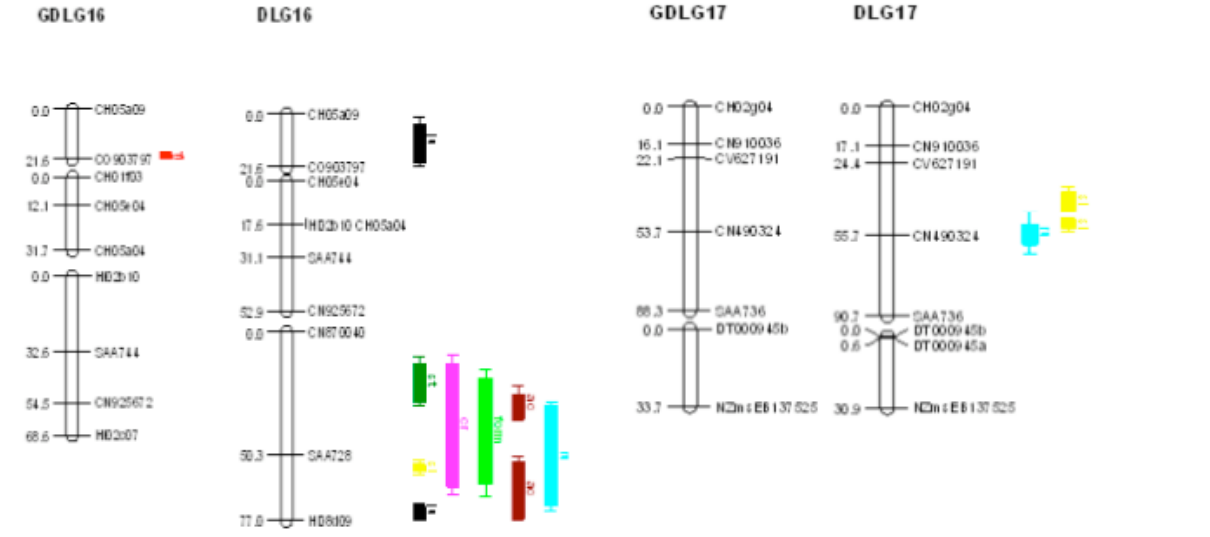


Figure 5.2: Positions of QTLs for nine fruit quality traits on the genetic linkage maps of 'Golden Delicious' ('GD') and 'Dietrich' ('D').

QTLs associated with firmness, crispness, juiciness, acidity, form, colour, stripness, russeting and size were detected by IM and MQM. Only linkage groups with detected QTLs are represented. QTLs are represented by means of bars showing 5% confidence interval bordering important QTL. The linkage groups were numbered according to Maliepaard *et al.* (1998) and the linkage maps were aligned to the reference maps by Silfverberg-Dilworth *et al*. (2006). QTLs for firmness (fi) are represented in red, juiciness (ju) in black, colour (co) in navy blue, size (si) in yellow, stripness (st) in deep green, crispness (cr) in purple, acidity (ac) in brown, form in light green and russeting (ru) in blue.

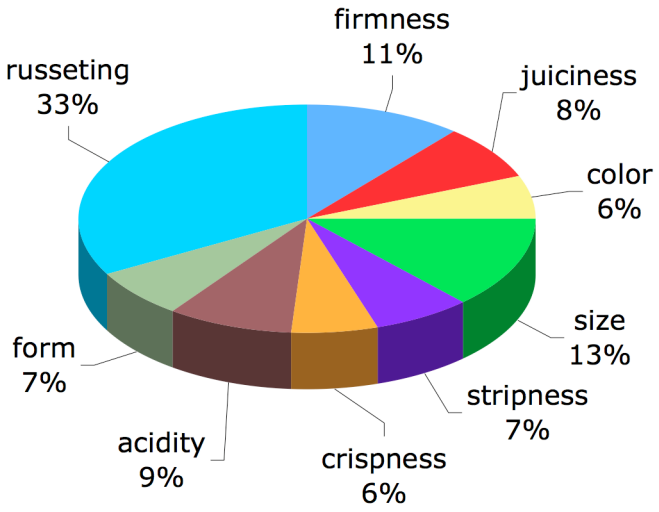


Figure 5.3: Proportion of QTLs identified as controlling each of the fruit quality traits under investigation in the 'Golden Delicious' x 'Dietrich' mapping population.

5.2.4 Nonparametric mapping

Folowing IM, the nonparametric mapping was carried out. In order to perform nonparametric mapping, the Kruskal-Wallis test (KW) was used as the nonparametric mapping algorithm to identify markers that could be associated with fruit quality traits. This analysis detected 46 significant QTLs, as defined by a *p* value below 0.05. In addition, three QTLs with $p>0.1$ were detected (Table 5.10). Markers associated with these QTLs detected by IM and nonparametric mapping were used as cofactors for MQM, thus allowing detecting extra 26 QTLs, which are indicated with no asterick in Table 5.10.

Table 5.10: Recapitulative table of QTLs associated with fruit quality traits, detected in the 'Golden Delicious' ('GD') x 'Dietrich' ('D') mapping population.

The QTLs are listed per linkage group and per year. For each QTL, the LOD score was above the significance threshold for each respective trait. Also for each trait detected with the Kruskal-Wallis test, the significance level (indicated by the asterisks: *p<0.1, **p<0.05, ***p<0.01, ****p<0.005, *****p<0.001, ******p<0.0005, *******p<0.0001) is given. The percentage of variance explained by each QTL and the parental genetic linkage map where the QTLs were detected are indicated.

Trait	Yr	LG2	LG3	LG4	LG5	LG6	LG7	LG8	LG9	LG10	LG11	LG12	LG13	LG14	LG15	LG16	LG17
Firmness	I	69% GD															
	II	[a]35% ** GD		79% GD						39% GD		[a]24% ** GD			[a]82% ** D	[a] 76% ** GD	
	III					16% D 16% * D									82% D		
Juiciness	I									[a]53% *** D					[a]34% ** D 61% GD		
	II			58% GD													

Table 5.10 continued

Trait	Yr	LG2	LG3	LG4	LG5	LG6	LG7	LG8	LG9	LG10	LG11	LG12	LG13	LG14	LG15	LG16	LG17
Juiciness	III			58% GD													
	M			39% D					[a]34% ** GD						[a]34% ** D	[a]41% ** D	
															[a]61% ** GD		
Colour	I														[a]70% ** D		
	II								70% 71% GD								
	M														[a]73% ** GD		
Size	II		72% D		[a]41% ** GD							[a]34% ** GD					
	III														[a]21% ** GD		
															19% D		
	M								[a]34% *** D						[a]39% ** D	[a]38% *** D	14% D

Table 5.10 continued

Trait	Yr	LG2	LG3	LG4	LG5	LG6	LG7	LG8	LG9	LG10	LG11	LG12	LG13	LG14	LG15	LG16	LG17
Stripness	I														96% * GD 96% D		
	III		92% * GD														
	M								44% D							39% D	
Crispness	I															[a]52% ** D	
	M	[a]33% ** D		38% D					[a]76% ** GD							[a]52% ** D	
Acidity	I												[a] 69% ** GD		[a]47% ** D		
	II	[a] 7.3% **** GD								87% GD			87% GD	78% GD	47% D	[a]73% *** D	
	III														65.6 % D		
	M								[a] 73% ***** D								

Table 5.10 continued

Trait	Yr	LG2	LG3	LG4	LG5	LG6	LG7	LG8	LG9	LG10	LG11	LG12	LG13	LG14	LG15	LG16	LG17
Form	I															[a]57% *** D	
	II										67% D						
	III				[a]22% **** D												
	M								[a]62% ** GD						78% D		
Russeting	I	[a] 7.6% ** D				13% GD			[a]23% **** GD	5.6% D					[a]89% *** D		
									[a]80% **** D	[a] 8.2% *** D					8.8% D		
	II			26% GD					[a]28% ** GD	[a]49% ** GD		27% GD	26% GD		15% GD	37% D	[a] 4.2% ** D
				[a] 8.5% *** D					86% D						[a] 3.6% *** D	77% D	

Table 5.10 continued

Trait	Yr	LG2	LG3	LG4	LG5	LG6	LG7	LG8	LG9	LG10	LG11	LG12	LG13	LG14	LG15	LG16	LG17
Russeting	III				[a]78% **** D				[a]25% ** GD		80% D		26% GD	[a]27% ** GD	[a]39% ** GD		
									[a]80% **** D								
	M								13% D							37% D	

[a]-QTLs evaluated based on the amount of phenotypic variation explained by the associated trait, as well as the level of significance of association with the trait of interest during the KW analysis. The markers associated with these QTLs can be used as candidate markers for MAS.

5.2.5 Identification of potential candidate markers

Although 72 putative QTLs were identified in this study, a critical selection of major QTLs and identification of potential candidate markers for use in MAS was necessary. In this regard, the location of suitable candidate markers was performed. A total of 26 SSR markers was associated with the QTLs identified controlling the nine traits studied. Out of these, 18 SSR markers were considered suitable for discriminating individual traits in MAS (Table 5.11). Nine SSR markers were identified as unique to russeting and these are CH02c061 positioned on LG2, DR992457 on LG9, CH01f07a on LG10, CH04g09 on LG10, CN491038 on LG14, CN444111 on LG15, CH05d024 on LG15, NZ02b01 on LG15 and CN490324 on LG17. Three SSR markers were identified for discriminating fruit size, namely EB149750 located on LG9, CH01f02 on LG12 and CN947446 on LG15. Two SSR markers were identified for discriminating fruit firmness, namely CH04g04 positioned on LG12 and CO903797 on LG16. Finally, Hi08d09, which was mapped on LG16, EB147667 on LG9, CN490897 on LG13 and NZmsEB116209 on LG9 were identified as unique to juiciness, crispness, acidity and form, respectively.

Table 5.11: Possible candidate markers associated with QTLs detected for fruit quality traits.

Trait	Marker	Linkage group
Firmness	Hi24f04	2
	CH04g04	12
	CH02d10b	15
	CO903797	16
Juiciness	SAA763	9
	CN444111, CN494248	15
	Hi08d09	16
Colour	CN444111, CN494248	15
Size	CN887525	5
	EB149750	9
	CH01f02	12
	CN947446, CN494248	15
	SAA728	16
Crispness	CN944528	2
	EB147667	9
	SAA728	16
Acidity	Hi24f04	2
	Hi23d06	9
	CN490897	13

Table 5.11 continued

Trait	Marker	Linkage group
Acidity	CH02d10b	15
	SAA728	16
Form	CN887525	5
	NZmsEB116209	9
Russeting	CH02c061	2
	CN887525	5
	SAA763, DR992457, Hi23d06	9
	CH01f07a, CH04g09	10
	CN491038	14
	CN444111, CH05d024, NZ02b01	15
	CN490324	17

5.3 DISCUSSION

The aim of this chapter was to detect QTLs associated with nine apple fruit quality traits following the evaluation of phenotypic data collected over a period of three years (2005-2007) and the construction of genetic linkage maps using a population derived from the apple cultivars 'Golden Delicious' and 'Dietrich'. All the traits under investigation, which could be compared with the parental phenotypes, showed transgressive segregation (Liebhard *et al.*, 2003a). Transgressive segregation is the accumulation in certain progeny of complementary alleles at multiple loci inherited from the two parents (Tanksley, 1993). Since apple is an outbred, and a self-incompatible species, heterozygous rather than homozygous loci are expected. Thus, a combination of superior parental alleles is likely to result in a phenotype exceeding the parental value (Liebhard *et al.*, 2003a).

A direct comparison between the progeny individuals and the parental trees was not always possible as a result of complex nature of fruit quality characters studied. However, a dissection of the complex traits into their responsible factors, as conducted in this study, is also possible without information about the parental phenotypes, since the contribution of the parents can be assigned by means of the genetic linkage maps.

The availability of genetic linkage maps covering entire genomes for QTL mapping overcame the problems associated with point analysis/single approach, where a population is analysed based on a one marker at a time basis (Tanksley, 1993; Van Ooijen, 2004). Therefore, QTL mapping using linked markers makes it possible to compensate for recombination between markers and detected QTLs. This increased the probability of statistically detecting the QTL and also providing an unbiased estimate of

the QTL effect on the character (Van Ooijen, 2004; Yin *et al.*, 2005; Costa *et al.*, 2008; Weebadde *et al.*, 2008).

In this study, individual QTLs were described as either 'major' or 'minor', according to the proportion of phenotypic variation explained by a single QTL (based on the R^2 value). Major QTLs accounted for a relatively large amount of phenotypic variation (>20%), while minor QTLs will usually account for less than 20% of phenotypic variation. In addition, major QTLs may refer to QTLs that are stable (detected in at least one year of phenotypic assessment) across environments, while minor QTLs may be environmentally sensitive, especially when they are associated with disease resistance (Tanksley, 1993; Lindhout, 2002; Pilet-Nayel *et al.*, 2002; Li *et al.*, 2003; Collard *et al.*, 2005).

The identification of putative QTLs is important for identification of potential candidate markers for use in MAS, which could lead to production of cultivars with desired traits. This techinique is time and cost effective, reducing the cost involved in maintaining trees that will only show their 'undesirable' characteristics after years of costly field maintenance. Fruit quality traits are among the most promising characters that can be considered for early marker assisted selection.

5.3.1 Phenotypic data analysis

The identification of QTLs, especially those with minor effect, is influenced by the number of individuals assessed for a given phenotypic trait and the marker density of the genetic linkage maps (Tanksley, 1993). In this study, although the population was large enough for a good quality analysis, the phenotypic data was dependent upon the

number of individuals bearing fruits, which varied from one year to another. According to Figure 3.1, 44%, 34% and 50% of the trees did bear fruits on year I, II and III, respectively, only a portion of the population was represented each of the years of investigation for the phenotypic assessment of the fruits. The lowest population representation was thus in year II. However, a total of 72% of trees did bear fruits at least in one of the three years of phenotypic investigation, giving a good overall representation of the population required to detect QTLs with minor effect.

In addition, the detection of QTLs using the full data set or the three years of phenotypic analysis as well as a mapping population with shared ancestry allowed improving the precision of QTL detection (Tanksley, 1993). However, in this study the majority of the QTLs were detected in one year of phenotypic data, giving rise to unstable QTLs. Only a few QTLs were detected in at least two years of the phenotypic analysis.

5.3.2 Genetic linkage map analysis

The genetic linkage maps of 'Golden Delicious' and 'Dietrich' were constructed using SSR markers, and their overall density was 0.8 markers/cM. This marker density was considerably lower than, of the 'Telamon' and 'Braeburn' genetic linkage maps, which comprised of 4.0 markers/cM (Kenis and Keulemans, 2005, 2008). Whitkus *et al.* (1992) observed that missing marker during map construction and erroneous marker scores can contribute to low marker density. This can give rise to separate individul groups in a linkage group as observed in this study (chapter 4), thus more markers are needed to reduce this effect.

Therefore, in this study, QTLs with major effect were detected in linkage groups with low marker density. For example, in LG3, only two markers were mapped distanced by 48.9 cM, and thus obtaining a very low marker density of 0.04 markers/cM. On this linkage group, two QTLs were detected, one controlling stripness and the other controlling size. Since high-density maps allow identifying, mapping and resolving quantitative traits into discrete genetic factors showing QTLs as Mendelian factors (Mohan *et al*., 1997), it is likely that more QTLs can be detected if more markers were positioned on LG3 and thus a greater marker density. However, the detection of QTLs does not depend only upon the quality of the genetic linkage map but also upon the quality of phenotypic data. In addition, the detection of a QTL is only possible when the QTL segregates in the population (Paterson *et al*., 1991a; Tanksley, 1993).

5.3.3 QTL analysis

To explore the quantitative nature of the nine fruit quality traits investigated in this study, the mean value for each individual per year and mean value over the three years of investigation (phenotypic data), together with the genetic linkage maps (genotypic data) were analysed using the *MapQTL®* software. In order to analyse the phenotypic and genotypic data, the IM, MQM and nonparametric mapping *MapQTL®* algorithms were used. A putative QTL was identified through comparing its presence on both the integrated map and either of the parent maps, in order to exclude false positive QTLs. During QTL mapping, more potential QTLs were identified on the integrated genetic linkage map than on the individual parental genetic linkage maps, and this was observed for all the fruit quality traits studied.

5.3.3.1 Evaluation of the identified QTLs according to the linkage groups

A total of six QTLs were detected on LG2, three on the 'Golden Delicious' map and three on the 'Dietrich' map. The QTLs exceeded the LOD significance threshold (equivalent to the genome wide (GW) for each year) in year I, II and III for russeting, acidity and firmness, respectively (Tables 5.9, 5.7 and 5.1), and years I and M for crispness (Table 5.6). None of these QTLs were detected in every year of investigation, possibly because the proportion of trees bearing fruits fluctuated over the years and the phenotypic data were based on a subset of individuals and not the entire population (Figure 3.1). The phenotypic variance explained ranged from 7.3% to 69% (Table 5.10). Variation in percentage phenotypic variance indicates that QTLs for firmness and crispness had a major influence on the variability contributed by the interactions of the genotype and the environment on the population while QTLs for acidity and russeting had a minor influence.

In terms of LG3, two QTLs were detected on this linkage group, one in each of the parental maps (Figure 5.2), controlling stripness and size. The QTL for stripness was detected using phenotypic data from year III, while the one for size was detected using the data from year II. The markers positioned on this linkage group were 48.9 cM apart. Since IM has been shown to be inefficient for detecting QTLs when the marker loci are >35 cM distant to each other (Van Ooijen, 1999, 2004), this suggests that some putative QTLs may have failed to be detected because of the distance separating markers on LG3. In addittion, a good marker density per linkage group is necessary to represent all regions of the linkage group and therefore to accurately identify QTLs (Lander and Botstein, 1989).

A total of six QTLs were identified on LG4. Three of these were detected on the 'Golden Delicious' map and shared common genetic linkage group region (i.e. were overlapping) between markers CH05d02 and CH01d03. In addition, three QTLs were detected on the 'Dietrich' map, and were mapped on the same location of the linkage group, between markers CH01b09b and CH05d02 (Figure 5.2). One QTL was detected for firmness and one for crispness using the phenotypic data from year II and the mean phenotypic data, respectively. These QTLs were however unstable, as they were detected only in one year of phenotypic assessment. In addition, QTLs associated with juiciness (two) and russeting (two) were detected using the phenotypic data from two years, suggesting that they were stable QTLs. The proportion of phenotypic variance explained by the QTLs detected ranged from 8.5% to 79% (Table 5.10). This variation indicates that the QTLs identified as controlling firmness, crispness and juiciness had a major influence on the variability contributed by the interactions of the genotype and the environment on the population, while QTLs for russeting (phenotypic variance explained <20%) had a minor influence.

In regard to LG5, four QTLs were identified, one on the 'Golden Delicious' map and three on the 'Dietrich' map (Figure 5.2). Stable QTLs detected on LG5 control for russeting, size and form. These were detected using the phenotypic data from years I, II and III, while an unstable QTL was detected controlling colour, using the phenotypic data only from year II. However, the proportion of phenotypic variance explained by the QTLs detected ranged from 22% to 78% (Table 5.10). This variation indicates that the QTLs identified as controlling colour, size, form and russeting had a major influence on the variability contributed by the interactions of the genotype and the environment on the population.

A total of two QTLs were identified on LG6, one associated with russeting on the 'Golden Delicious' map and one associated with firmness on the 'Dietrich' map (Figure 5.2). These QTLs were detected in at least one year of phenotypic assessment, suggesting that these could be stable QTLs. However, both QTLs had a proportion of phenotypic variance explained less than 20%, suggesting that these QTLs could possibly have a minor influence on the population genotype.

On LG9, 13 QTLs were identified, seven in the 'Golden Delicious' map and six in the 'Dietrich' map (Figure 5.2). One QTL controlling russeting on 'Dietrich' was detected in two years of phenotypic assessment, suggesting that this QTL could be stable, while the other four QTLs controlling russeting and the QTLs controlling juiciness, size, stripness, crispness, acidity, colour and form positioned on LG9, were detected only in one year of phenotypic assessment, suggesting that these QTLs could be unstable. However, the proportion of phenotypic variance explained ranged from 13% to 80% (Table 5.10). This variation indicates that the QTLs identified had a major influence on the population genotype with the exception of one QTL controlling russeting. Besides, a QTL controlling colour in LG9, was observed as being unstable (detected only in one year of phenotypic assessment) but with major influence on the variability contributed by the interactions of the genotype and the environment on the population as a result of high proportion of phenotypic variance explained (71%, Table 5.10).

On LG10, a total of three QTLs were identified, only on the genetic linkage map of 'Dietrich' (Figure 5.2). These QTLs were detected only in year I of phenotypic assessment, suggesting that they could be unstable QTLs. However, two QTLs dectected as controlling russeting had low proportion of phenotypic variance explained

(<20%), suggesting that they could have minor influence on the population genotype, while the phenotypic variance explained for juiciness was 53%, indicating that the QTL could have a major influence on the population genotype.

In terms of LG11, two QTLs were detected on the genetic linkage map of 'Dietrich' (Figure 5.2). The two QTLs, one controlling form and one controlling russeting, were detected only in one year of phenotypic assessment, suggesting that these QTLs could be unstable. However, these QTLs had high proportions of phenotypic variance explained, suggesting that even though unstable, they could have a major influence on the population genotype.

A total of three QTLs were identified on LG12 of the parental genetic linkage map of 'Golden Delicious' (Figure 5.2). The QTLs associated with russeting, firmness and size were detected only in one year of phenotypic assessment, suggesting that they could be unstable. However, though unstable, these QTLs had high proportions of phenotypic variance explained, suggesting that they could have a major influence on the population genotype.

In regard to LG13, two QTLs were identified on the parental genetic linkage map of 'Golden Delicious'. The QTLs associated with acidity and russeting, were detected in at least one year of phenotypic assessment, suggesting that they could be stable. In addition, these QTLs had high proportions of phenotypic variance explained, suggesting that they could have a major influence on the population genotype.

In terms of LG14, two QTLs were identified on the parental genetic linkage map of 'Golden Delicious' (Figure 5.2). One QTL was detected as controlling acidity using the phenotypic data from years I and II, suggesting that this could be a stable QTL, while one QTL was detected as controlling russeting using phenotypic data only from year II, suggesting that this could be an unstable QTL. However, both QTLs had high proportions of phenotypic variance explained, suggesting that they could have a major influence on the population genotype.

A total of 17 QTLs were identified on LG15, six on the 'Golden Delicious' map and 11 on the 'Dietrich' map (Figure 5.2). The QTLs associated with russeting, firmness, juiciness, acidity and size were detected in at least one year of phenotypic assessment, suggesting that these QTLs could be stable. The QTLs associated with colour, form and stripness were detected only in one year of phenotypic assessment, suggesting that these QTLs could be unstable. The proportion of phenotypic variation explained ranged from 3.6% to 96% (Table 5.10). This variation indicates that the QTLs identified as controlling russeting, firmness, juiciness, acidity, size colour, form and stripness, with the exception of one QTL associated with size and three QTLs associated with russeting coud have a major influence on the population genotype. This linkage group had the most number of QTLs detected in this study, signifying the important role it plays in controlling fruit quality traits. It represented 24% of the detected QTLs in all the linkage groups.

On LG16, nine QTLs were identified, one on the 'Golden Delicious' map and eight on the 'Dietrich' map (Figure 5.2). The QTLs associated with crispness and russeting were detected in at least one year of phenotypic assessment, suggesting that these could be

stable QTLs. The QTLs associated with form, firmness, acidity, juiciness, stripness and size were detected only in one year of phenotypic assessment, suggesting that these could be unstable QTLs. However, the proportion of phenotypic variation explained ranged from 37% to 76% (Table 5.10), indicating that all the QTLs detected in this linkage group could be major QTLs. Besides, seven of the QTLs identified on the 'Dietrich' map were detected between markers, CN870040 and Hi08d09 (Figure 5.3) suggesting that the QTLs could be located in a gene cluster and/or sharing common nucleotide sequences but could be expressing different proteins following post-transcriptional and post-translational modifications.

A total of two QTLs were identified on LG17 of the genetic linkage map of 'Dietrich' (Figure 5.2). The QTLs associated with russeting and size were detected only in one year of phenotypic assessment, suggesting that they could be unstable QTLs. In addition, QTLs had low proportions of phenotypic variance explained, suggesting that they could have a minor influence on the variability contributed by the interactions of the genotype and the environment on the population.

In regard to LG1, LG7 and LG8, no QTL was detected in these three linkage groups. Only few markers were mapped on these linkage groups, suggesting that this may have restricted the detection of QTLs. Previous studies reported the detection of QTLs associated with firmness (Costa *et al*., 2008) and fruit acidity (Liebhard *et al*., 2003) on LG1 and LG8, respectively. In addition, the *Md-expansin 7* gene, which is associated with fruit softening, was mapped on LG1 in apple using the 'Prima' x 'Fiesta' population and in pear using the 'Passe Crassane' x 'Harrow Sweet' population. This

gene was located close to a major apple QTL for fruit firmness (Costa *et al*., 2008). However, in this study, no QTL could be detected in this region of LG1.

5.3.3.2 Evaluation of the identified QTLs according to fruit quality traits

5.3.3.2.1 Firmness

Most important agronomic traits are quantitatively inherited, being controlled by several genes. Studies carried out on apple-segregating populations have identified important QTLs for firmness on the linkage groups LG1, LG6, LG8, LG10, LG11, LG12, LG14, LG15 and LG16 (King *et al.*, 2000; Maliepaard *et al.*, 2001; Seymour *et al.*, 2002; Liebhard *et al.*, 2003b). In addition, novel insights into fruit-firmness control in apple has recently been reported via a candidate-gene approach investigating the effects of *Md-ACS1* and *Md-ACO1*, two genes involved in ethylene production during ripening, on fruit softening (Oraguzie *et al.*, 2004; Costa *et al.*, 2005; Costa *et al.*, 2008). *Md-ACO1* was mapped on LG10 within the 5% confidence interval border at the same position of a major QTL for fruit firmness, thus, genetically linking ethylene production and apple softening (Costa *et al.,* 2008).

In terms of firmness, in this study, a total of eight QTLs were detected, which were located on LG2, LG6, LG10, LG12, LG15 and LG16 in the 'Golden Delicious' x 'Dietrich' population. The linkage groups associated with QTLs for firmness that are common with those previously detected include LG6, LG10, LG12, LG15 and LG16. In this study, an additional QTL was detected on LG2. However, QTLs for firmness were not detected on LG1, LG8, LG11 and LG14, although previous studies reported the detection of QTLs on these linkage groups. In regards to LG1 and LG8, no QTL for firmness was detected possibly because of insufficient markers positioned on these

linkage group. Concerning LG11 and LG14, although a high marker density was observed, no QTL could be detected. The variations in QTL detection between studies suggest that the QTLs might be cultivar dependent, since Liebhard *et al.* (2003) used 'Fiesta' and 'Discovery' and King *et al.* (2000) used 'Fiesta' and 'Prima' instead of 'Golden Delicious' and 'Dietrich' in the present study. In addition, since the populations were cultivated at different locations and thus were exposed to different environment, the result variations suggest that the QTLs controlling firmness might be influenced by the environmental conditions. However, the results of this thesis can be compared with the previous mentioned studies that identified QTLs for firmness since the genetic linkage maps were constructed in part using co-dominant and transferable SSR markers (King *et al.,* 2000; Liebhard *et al*., 2003; Kenis and Keulemans, 2005, 2008). Overall, the number of QTLs identified controlling firmness represented 11% of the total QTLs detected and associated with fruit quality traits (Figure 5.3).

5.3.3.2.2 Juiciness

A previous genetic study carried out using a 'Fiesta' x 'Prima' population identified QTLs controlling juiciness on LG1, LG12 and LG16 (King *et al*., 2000). Since the 'Fiesta' x 'Prima' genetic linkage map was constructed in part using co-dominant and transferable SSR markers, like for the construction of the 'Golden Delicious' x 'Dietrich' map, the results could be compared (King *et al.,* 2000; Liebhard *et al*., 2003; Kenis and Keulemans, 2005, 2008, Gardiner *et al.,* 2007). In the current study, QTLs for juiciness were identified on LG4, LG9, LG15 and LG16 (Table 5.2). Linkage groups associated with QTLs for juiciness detected in the 'Golden Delicious' x 'Dietrich' population in common with those previously detected include LG16. In this study, additional QTLs for juiciness were identified on LG4, LG9 and LG15. However,

no QTL for juiciness was detected on LG1 and LG12. In regards to LG12, although a high marker density was observed, no QTL was detected in the 'Golden Delicious' x 'Dietrich' population. The variations in QTL detection between these studies suggest that the QTLs might be cultivar dependent, since King *et al*. (2000) used cultivars 'Fiesta' and 'Prima' instead of 'Golden Delicious' and 'Dietrich' like in the present study. In addition, since the populations were cultivated at different locations and thus were exposed to different environments, the result variations suggest that the QTLs controlling firmness might be influenced by the environmental conditions. Overall, the number of QTLs identified for juiciness represented 8% of the total QTLs detected and associated with fruit quality traits (Figure 5.3).

5.3.3.2.3 Colour

Studies carried out on 'Ralls Janet' x 'Delicious' (Igarashi *et al.*, 2008), 'Sciros' x '91.136 B6-77' and 'Geneva' x 'Braeburn' (Chagne *et al*., 2007) populations have identified important QTLs for colour on LG9. The marker, BC226 for fruit skin colour was detected tightly linked to the *Rf* locus (<2 cM; Cheng *et al.*, 1996). It showed a heterozygous combination of alleles in 'Delicious' and mapped on LG9, while in the population of 'Sciros' x '91.136 B6-77' and 'Geneva' x 'Braeburn', it was monomorphic. However, in 'Sciros' x '91.136 B6-77' and 'Geneva' x 'Braeburn' populations, the *Rni* locus, a major genetic determinant of the red foliage and red colour in the core of apple fruit was identified (Chagné *et al*., 2007). In this study, the *MdMYB10* gene, a transcriptional factor, co-segregated with the *Rni* locus and was mapped on LG9 of the apple genome. Besides, the *Rf and Rni* loci were found to located at the bottom of LG9 suggesting that both loci could be located in a gene cluster or even correspond to alleles of the same gene. Interestingly, fruit skin colour and leaf

colour collocate in *Prunus* (Dirlewanger *et al*., 2004), which suggests that there may be a region of synteny between the middle of *Prunus* LG6 and the bottom of *Malus* LG9 (Chagné *et al*., 2007).

In this study, a similar combination of heterozygous alleles was detected in the study of 'Golden Delicious' x Dietrich' population. Besides, detecting an additional QTL on LG15, a QTL controlling colour was also detected in LG9 as reported from the previous mentioned studies. The QTLs detected on LG9 were identified only on the 'Golden Delicious' map. However, by comparison with the previous studies, it would be expected to detect a QTL for red colouration on LG9 of the 'Dietrich' map. Nevertheless, in this study, four QTLs were identified as being associated with fruit colour. Two of the QTLs were positioned on LG9 and two on LG15 (Table 5.3). Three of these QTLs were detected on the 'Golden Delicious' map, while one was detected on 'Dietrich' map. This data suggests that the gene(s) coding for the green colour of fruits from 'Golden Delicious' could be located on LG9 and LG15, while the gene(s) coding for the red colour of fruits from 'Dietrich' could be located on LG15. Further assessment of the position of a QTL associated with red colour on LG9 can be carried out by candidate gene mapping using the *Rni* and *Rf* loci. The number of QTLs identified for colour represents 6% of the total QTLs detected and associated with fruit quality traits (Figure 5.3).

5.3.3.2.4 Size

Previous studies have detected QTLs associated with size on LG2, LG5, LG8, LG9, LG10 and LG17 (Seglias and Gessler, 1997; Liebhard *et al.*, 2003a; Kenis *et al.,* 2006). However, in the present study, QTLs controlling size were detected on LG3, LG5, LG9,

LG12, LG15, LG16 and LG17 (Table 5.4). Of these QTLs, four were newly detected QTLs located on LG3, LG12, LG15 and LG16. The QTLs controlling size, which were common between the previous studies and the study on the 'Golden Delicious' x 'Dietrich' included LG5, LG9 and LG17. By comparison with Kenis *et al*. (2006), Liebhard *et al*. (2003a) and Seglias and Gessler (1997) studies, QTLs associated with size were not detected on LG8. In addition, in this study, a total of 13 QTLs were identified for size representing 13% of the total number of QTLs detected to be associated with fruit quality traits (Figure 5.3).

5.3.3.2.5 Stripness

This study reports the first identification of QTLs associated with stripness in apples. Using the 'Golden Delicious' x 'Dietrich' population, five QTLs were associated with stripness and were positioned on LG3, LG9, LG15 and LG16 (Table 5.5). This study demonstrates the correlation between colour and stripness as a QTL controlling stripness on LG15 was detected close to the QTL controlling colour. This suggests that both QTLs could be located in a gene cluster or even correspond to alleles of the same gene. Besides, several QTLs positioned on different linkage groups, were detected controlling stripness compared to colour, indicating that stripness could be controlled by several genes and thus is a complex trait.

Apple cultivars present a wide range of colour variation in fruit skin ranging from green like 'Granny Smith', through to partially coloured apples like 'Royal Gala' (striped), through to dark red like 'Red Delicious' and 'Dietrich'. Interestingly, colour varies not only in the fruit skin, but also in other plant parts like flower, foliage and fruit flesh (Chagné *et al*., 2007). The degree of stripeness in apples has been shown to be directly

related to the amount of anthocyanin in the fruits of 'Gala' apple strains (Iglesias *et al*., 2008). In this study, the dark red, highly striped fruits contained high anthocaynin levels compared to the light red less striped fruits. This suggest that the regulation of the anthocyanin levels determines the fruit colour and stripness patterns observed on fruits. In addition, the red colour has been associated with *MdMYB10* gene, a transcriptional factor associated with the regulation of anthocyanin biosynthesis, which co-segregate with *Rni* locus, a determinant of red flesh colour (Chagné *et al*., 2007).

5.3.3.2.6 Crispness

Studies carried out on 'Fiesta' x 'Prima' population have identified important QTLs controlling crispness on LG1, LG5, LG10, LG12, LG13 and LG16 (King *et al*., 2000). The results of this study can be compared with the detection of QTLs controlling crispness in the 'Golden Delicious' x 'Dietrich' population because the genetic linkage maps were constructed in part using co-dominant and transferable SSR markers. In the study of the 'Golden Delicious' x 'Dietrich' population, QTLs controlling crispness were identified on LG2, LG4, LG9 and LG16 (Table 5.6). This means that, in this study, additional QTLs controlling crispness were identified on LG2, LG4 and LG9.

The linkage groups associated with QTLs of crispness detected in the 'Golden Delicious' x 'Dietrich' population that were common with those detected in the 'Fiesta' x 'Prima' population included LG16. This implies that QTLs controlling crispness in the 'Golden Delicious' x 'Dietrich' population were not detected on LG1, LG5, LG10, LG12 and LG13. This could be because of low marker density on LG1. Concerning LG5, LG10, LG12 and LG13, although a high marker density was observed, no QTL could be detected. The variations in QTL detection between studies suggest that the

QTLs might be cultivar dependent, since King *et al*. (2000) used cultivars 'Fiesta' and 'Prima' instead of 'Golden Delicious' and 'Dietrich' in the present study. In addition, since the populations were cultivated at different locations and thus were exposed to different environment, the result variations suggest that the QTLs controlling firmness might be influenced by the environmental conditions.

5.3.3.2.7 Acidity

In apple, the predominant factor of variation in flavor is the balance between sugars and acids. Malic acid, being the main substrate for respiration in apples, also represents the principal acid in apple fruit (Hulme and Rhodes, 1971). Studies carried out on the 'Fiesta' x 'Prima' population have identified the exact position of the *Ma* gene for malic acid, which was mapped on LG16 (Maliepaard *et al*., 1998). Further, important QTLs for acidity were detected on linkage groups LG8 and LG16, using both the 'Fiesta' x 'Prima' and 'Fiesta' x 'Discovery' populations (King *et al.*, 2000; Liebhard *et al.*, 2003b; Kenis and Keulemans, 2005). The results of these studies can be compared with the current study of QTLs for acidity detected in the 'Golden Delicious' x 'Dietrich' population because the genetic linkage maps were constructed in part using co-dominant and transferable SSR markers.

In this study, QTLs controlling acidity were identified on LG2, LG9, LG10, LG13, LG14, LG15 and LG16 (Table 5.7). However, additional QTLs controlling acidity were identified in LG2, LG9, LG10, LG13, LG14 and LG15. Linkage groups associated with QTLs controlling acidity detected on the 'Golden Delicious' x 'Dietrich' population that were common with those detected on the 'Fiesta' x 'Prima' and 'Fiesta' x 'Discovery' population included LG16. Concerning LG8, a low marker density was

observed and could have been the reason for failure to detect a QTL in the linkage group.

5.3.3.2.8 Form

Studies performed on peach 'Ferjalou Jalousia' x 'Fantasia' population detected a QTL controlling form on LG6 (Dirlewanger *et al.*, 2004). Although LG6 had no homologous counterparts in the 'Prima' x 'Fiesta' apple population map, results from this study are comparable with Dirlewanger *et al.* (2004) results because peach and apple belong to the Rosaceae family and 'Ferjalou Jalousia' x 'Fantasia' *Pruns* map have homologous counterparts in the 'Prima' x 'Fiesta' apple map. This study, however, reports the first detection of QTLs associated with form in apples. In this study, QTLs associated with form were identified on LG5, LG9, LG11, LG15 and LG16 (Table 5.8). There were no common QTLs detected between the apple and the peach cultivars. In addition, no QTLs were detected on LG6 in the 'Golden Delicious' x 'Dietrich' apple population, as was detected in the peach. This could mean that the QTL controlling form collocated in peach, suggesting that there may be a region of synteny between the peach LG6 and either LG5, LG9, LG11, LG15 or LG16 of the apple. The complexity of form in apples is higher than in peach possibly because the apple genome has 17 chromosomes, 9 chromosomes more than in peach, increasing the chances of gene duplication and thus more genomic regions controlling a trait.

5.3.3.2.9 Russeting

This study reports first the identification of QTLs associated with russeting in apples. Using the 'Golden Delicious' x 'Dietrich' population, 23 putative QTLs associated with russeting were detected. These were positioned on LG2, LG4, LG6, LG9, LG10, LG11,

LG12, LG13, LG14, LG15, LG16 and LG17 (Table 5.9). Russeting was the trait with the highest number of QTLs detected, representing 33% of the identified QTLs in this study (Figure 5.3). The QTLs were distributed over 12 linkage groups. In addition, these results could explain the frequent occurrence of russeting on apple fruits considering that russeting is a result of many causes like humidity, fungal and bacterial infection, and insect bites. This could be because the phenotypic data measured only russeting (in general) but not the cause, thus the large number of QTLs could account for a number of traits like disease resistance. Despite all this, the large number of QTLs detected controlling russeting suggests that many genes control russeting, indicating that it is indeed a complex trait.

5.3.3.3 Potential candidate markers

A total of 26 SSR markers were identified as possible candidate markers for the nine fruit quality traits. Out of these, a total of 18 SSR markers were considered suitable for discriminating individual traits (Table 5.11), suggesting that for exclusive detection of a trait, common markers were not discriminatory and thus not useful. Therefore, only exclusive markers could be selected as potential candidates for use in MAS. Nine SSR markers were identified unique to russeting, three for discriminating size, two for discriminating firmness, and one each for discriminating juiciness, crispness, acidity and form. Candidate markers for discriminating colour and stripness could not be detected even though QTLs were identified for these two traits. However, mapping of candidate genes for colour, showing heterologous segregation among the individuals in the 'Golden Delicious' x 'Dietrich' population, like *MdMYB10* gene, shown to co-segregate with *Rni* locus (Chagné *et al*., 2007) could provide even better candidate markers for colour or even other traits.

In summary, a total of 72 putative QTLs were detected using the 'Golden Delicious' x 'Dietrich' mapping population. Eight putative QTLs were associated with firmness, six with juiciness, four with colour, nine with size, five with stripness, four with crispness, seven with acidity, five with form and 24 with russeting. QTL mapping analysis revealed that fruit quality traits were not completely independent of each other. For example, even though no direct correlation could be revealed among the phenotypic traits, a number of QTLs for different traits were detected overlapping on some linkage groups like LG 15. In addition, a total of 26 SSR markers were identified as possible candidate markers for the nine fruit quality traits. Out of these, 18 SSR markers were considered suitable for discrimination individual traits, however, candidate markers for discriminating colour and stripness could not be detected even though the QTLs were identified for these two traits.

CHAPTER 6

ESTABLISHMENT OF THE 'GOLDEN DELICIOUS' X 'DIETRICH' APPLE FRUIT PULP 2D-PAGE PROTEOME MAP AND PROTEIN IDENTIFICATION

6.1 INTRODUCTION

The fruit is an active plant tissue where several metabolic processes occur like carbohydrate metabolism, defense of tissues against invading pathogens or signal transduction pathways (Carrari and Fernie, 2006; Nosarzewski and Douglas, 2007; Wang *et al.*, 2009). The apple fruit development depends primarily on carbohydrate metabolism and import. Several forms of sugar like sucrose or sorbitol are translocated to the fruit and oxidised for their utilisation, a necessary stage during fruit development (Nosarzewski and Douglas, 2007; Wang *et al.*, 2009). In this study, mature apple fruits from individuals of the 'Golden Deliciuos' x 'Dietrich' population were harvested (section 2.3.1). The apple fruit pulp tissue was used for proteomic analysis.

Proteomics, defined as the large-scale study of protein expression, interaction, structure and functions, allow the global analysis of the subset of genes expressed in a tissue, cell or sub-cellular compartments of various tissues, at specific physiological states (Blackstock and Weir, 1999; Park, 2004; Guarino *et al.*, 2007). The use of proteomic approaches in the area of fruit and vegetable physiology has been increasing over recent years, with reports on tomato (*Solanum lycopersicum*, Rocco *et al.*, 2006), strawberry (*Fragaria ananassa,* Hjerno *et al.*, 2006), pear (Barraclough *et al.,* 2004; Pedreschi *et al.*, 2007; 2008) and apple (Barraclough *et al*., 2004; Guarino *et al*., 2007; Cao *et al*.,

2008). In apple, the studies reported the proteomic analyses of pseudocarp tissue and flower bud tissue using 'Annurca' and 'Fuji' cultivars, respectively (Guarino *et al*., 2007; Cao *et al*., 2008). Thus, proteomics can be used to qualitatively and quantitatively characterise regulatory events occuring during fruit development (Zhao *et al.*, 2008). The characterisation of the fruit proteome can also be used to link the fruit quality proteomic data to the genotype and phenotype.

Gel-based proteomic research mostly relies on the use of two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) to generate high-resolution proteome reference maps under various physiological conditions (Zuo and Speicher, 2000; Barraclough *et al*., 2004; Deytieux *et al.*, 2007; Barel and Ginzberg, 2008).

The aim of this chapter was to establish 2D-PAGE proteome reference maps using total soluble proteins (TSP) isolated from mature apple fruit pulp and to analyse the entire subset of proteins visualized by CBB staining using MALDI-TOF MS for their identification. The positively identified proteins were then characterized using the current knowledge in the literature, sequence databases and bioinformatic tools.

Mature apple fruits from individuals of the 'Golden Deliciuos' x 'Dietrich' population were harvested and subjected to firmness test using a penetrometer. Fruits whose firmness was below 7 kg cm^{-2} were categorised as low firmness, while fruits whose firmness was greater than 7 kg cm^{-2} were categorised as high firmness. The fruits were snap frozen in liquid nitrogen prior to protein extraction. Proteins were extracted from four biological samplesfrom individual apples for each category. The total soluble proteins were pooled together after protein quantification using the Bradford assay.

This sample was then used as a representative of the entire total soluble protein expressed in the apple fruit pulp tissue for proteomic analysis.

6.2 OPTIMISATION OF PROTEIN EXTRACTION AND PROTEIN SEPARATION BY 1D-PAGE

Total soluble proteins were extracted from mature apple fruit pulp using either the TCA/acetone or phenol precipitation (sections 2.6.1.1 and 2.6.1.2, respectively) methods. Following separation by 1D-PAGE, a high background limiting the visualisation of the proteins was observed (Figures 6.1A and 6.2A), a sign of the presence of interfering substances. Therefore, the two methods were optimised (sections 6.2.1 and 6.2.2).

6.2.1 Protein extraction by TCA/acetone precipitation

In order to limit the presence of interfering substances in the TSP, 0.06% (w/v) sodium sulphite was added to TCA/acetone for protein precipitation (section 2.6.1.1). Prior to optimisation, the protein concentration of the pulp extracts was on average 0.8 $\mu g.\mu l^{-1}$ for each sample. In addition, the proteins were not clearly visualised by 1D-PAGE because of the presence of a high background (Figure 6.1A). Following the addition of sodium sulphite the TSP concentration were on average 5.5 $\mu g.\mu l^{-1}$. The protein band pattern of the TSP extracts was more clearly visualised by 1D-PAGE (Figure 6.1B).

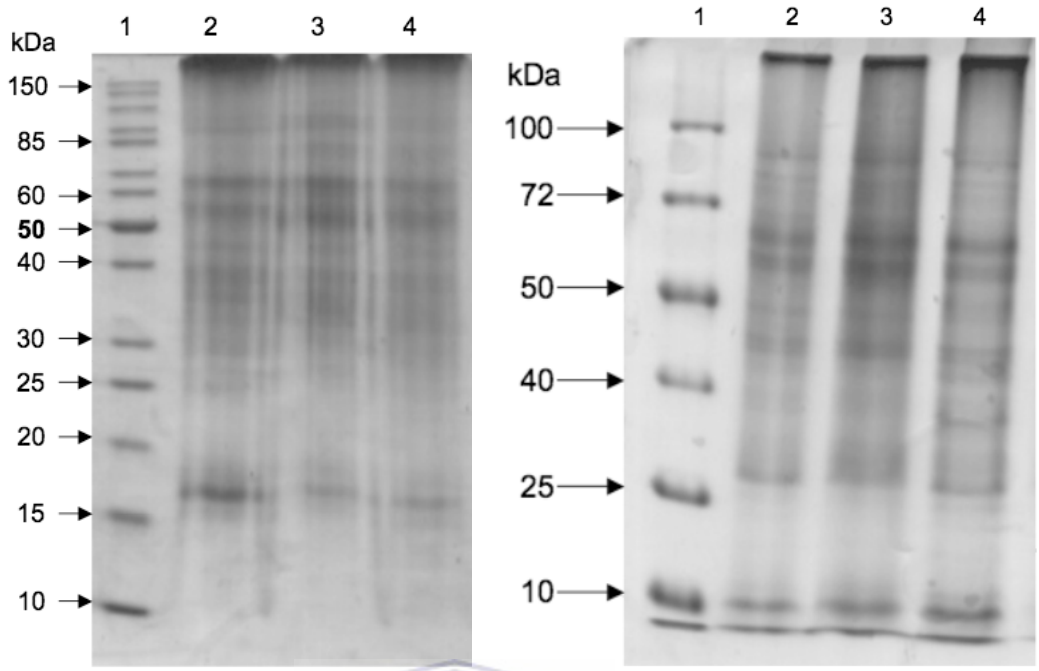


Figure 6.1 1D PAGE profiles of total soluble protein from mature apple fruit pulp after protein extraction by TCA/acetone precipitation (A) before optimisation and (B) after optimisation.

Approximately 20 µg of total soluble proteins from mature apple fruit pulp protein samples were separated by 12% SDS-PAGE and stained with Coomassie brilliant blue. Lane 1: molecular weight marker and lane 2, 3 and 4: proteins extracted from three biological replicates of mature apple fruit pulp samples.

6.2.2 Protein extraction by phenol precipitation

In order to limit the presence of interfering substances in the TSP, proteins were extracted by phenol precipitation with the addition of 50 mM PVPP to the TCA (section 2.6.1.2). The protein concentration was on average 25.5 $\mu g.\mu l^{-1}$ for each sample, both before and after optimisation. Before optimisation, the proteins were not clearly visualised by 1D-PAGE because of the presence of a high background, and this was particularly prominent for the low abundant proteins (Figure 6.2A). In addition, the amount of protein loaded affected the clear observation of the protein bands, as observed when comparing the loading of 20 µg and 30 µg (lane 3 and lane 2, respectively; Figure 6.2A). However, after optimisation, protein visualization was improved and the intensity of the background was reduced (Figure 6.2B). On the contrary to the 1D-PAGE with samples extracted prior to optimization, the increase in protein loading from 20 to 30 µg per lane did not increase the background nor limit protein band visualization, but only increased the abundance of the protein bands.

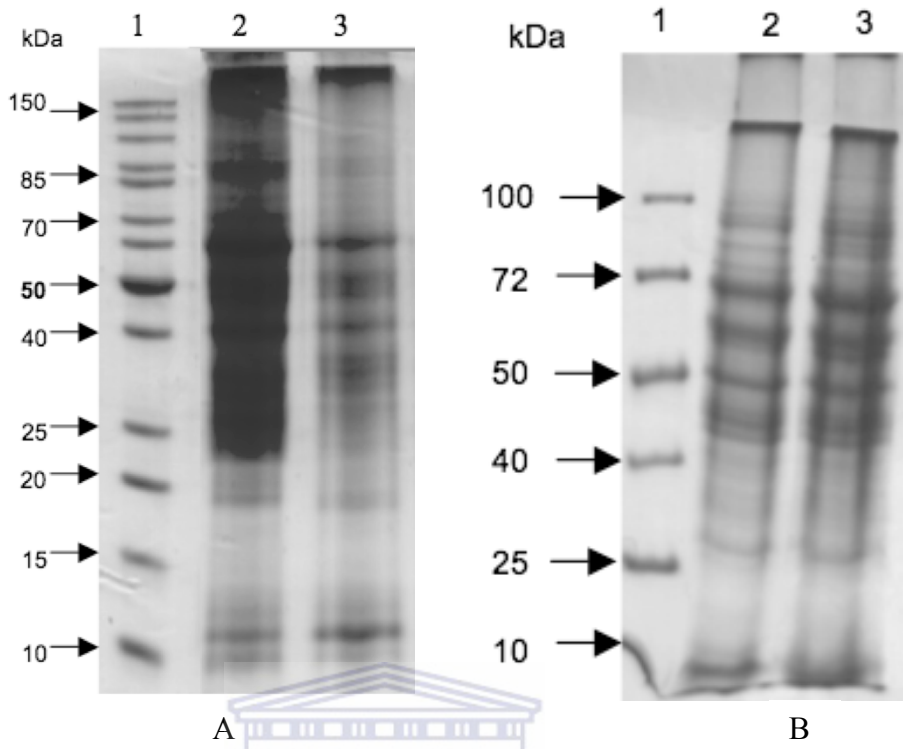


Figure 6.2: 1D PAGE profiles of total soluble proteins from mature apple fruit pulp after protein extraction by phenol precipitation (A) before optimisation and (B) after optimisation.

Total soluble proteins was separated by 12% SDS-PAGE and stained with Coomassie brilliant blue. Lane 1: molecular weight marker. Lane 2: 30 µg and 20 µg of apple fruit pulp tissue protein on panel A and B respectively. Lane 3: 20 µg and 30 µg of apple fruit pulp tissue protein on panel A and B respectively.

6.3 ESTABLISHMENT OF 2D-PAGE PROTEOME MAP

Following TSP extraction and 1D-PAGE separation, protein extracts from mature apple fruit pulp were resolved by 2D-PAGE according to both their p*I* and molecular mass. Two-dimensional proteome maps were established by resolving proteins extracted by either TCA/acetone precipitation (section 2.6.1.1) or phenol precipitation (section 2.6.1.2).

In terms of the 2D-PAGE proteome maps established using extracts from TCA/acetone precipitation, about 60 protein spots were visualised (Figure 6.3A). Most of the proteins resolved were in the molecular range of 20 to 50 kDa and p*I* range of 5 to 7.

In regards to 2D-PAGE proteome maps established using extracts from phenol precipitation, a higher number of protein spots were observed, thus at least 230 protein spots were resolved (Figure 6.3B). The majority of the proteins separated in the molecular weight range of 10 to 100 kDa and p*I* range of 4 to 7. The protein spots from the two proteome maps were labelled for MALDI-TOF MS analysis (Figure 6.4).

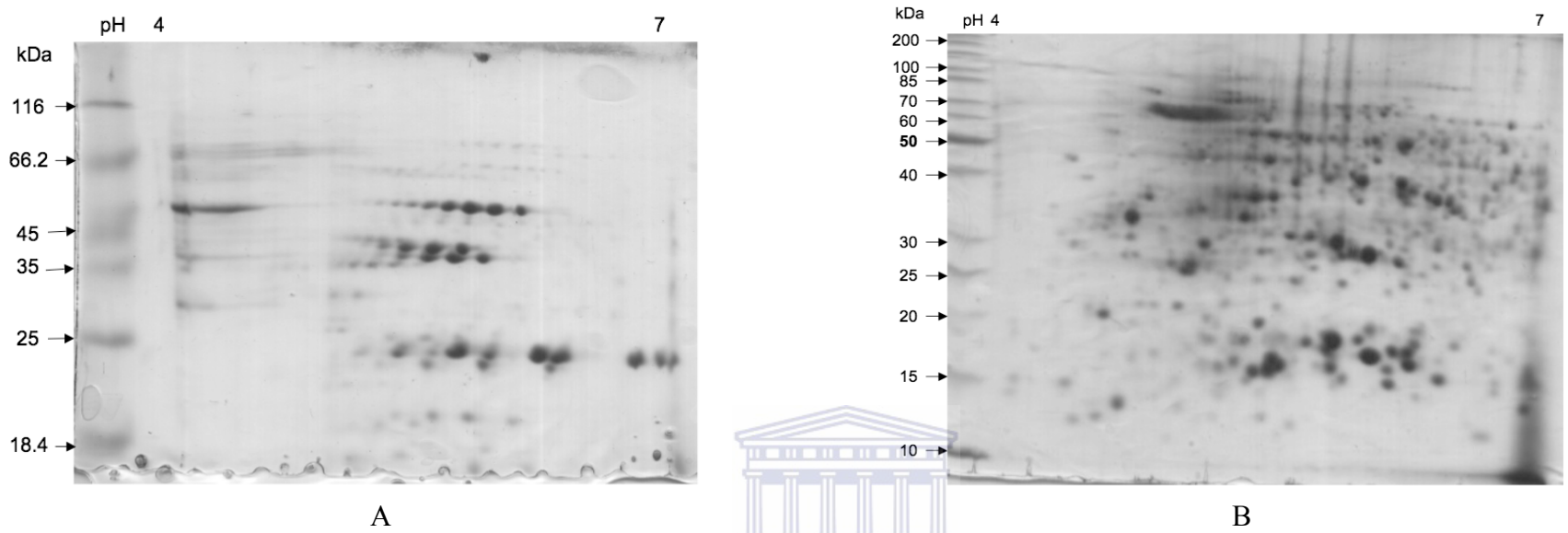


Figure 6.3: 2D-PAGE proteome maps of total soluble proteins from mature apple fruit pulp.

Proteins (200 µg) extracted by (A) TCA/acetone precipitation and resolubilised in extraction buffer and (B) phenol precipitation and resolubilised in IEF rehydration buffer were separated by 2D-PAGE using linear 7 cm IPG strips, pH range 4-7 in the first dimension and 12% SDS-PAGE in the second dimension. Proteins were visualized by Coomassie brilliant blue. The molecular masses of the protein marker are indicated in kDa on the left of each gel. The pH of the IPG strips is indicated by (4) and (7) on the top of each gel.

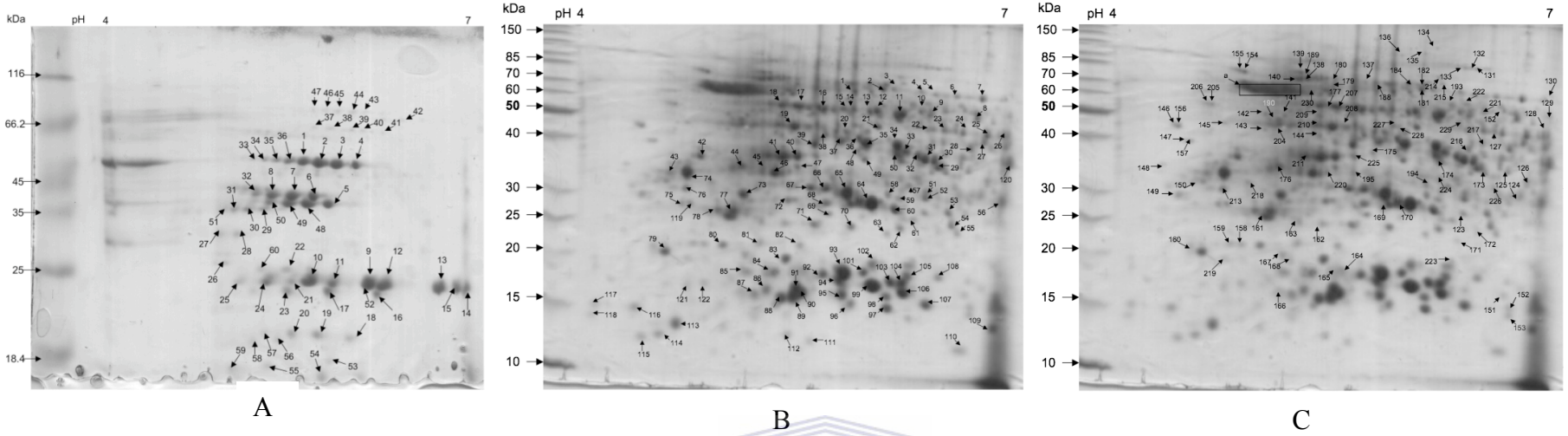


Figure 6.4: 2D-PAGE proteome maps of total soluble proteins from mature apple fruit pulp with protein spots labelled.

Proteins (200 µg) extracted by (A) TCA/acetone precipitation and resolubilised in extraction buffer and (B and C) phenol precipitation and resolubilised in IEF rehydration buffer were separated by 2D-PAGE using linear 7 cm IPG strips, pH range 4-7 in the first dimension and 12% SDS-PAGE in the second dimension. Proteins were visualized by Coomassie brilliant blue staining. The molecular masses of the protein marker are indicated in kDa on the left of each gel. The pH of the IPG strips is indicated by (4) and (7) on the top of each gel.

6.4 IDENTIFICATION OF PROTEINS BY MALDI-TOF MS AND VALIDATION AGAINST *MALUS* EST FROM NCBI

Following separation of proteins by 2D-PAGE, 60 protein spots from protein extracted by TCA/acetone precipitation (Figure 6.4A) and 230 protein spots from protein extracted by phenol precipitation (Figure 6.4B and C), were excised for analysis by MALDI-TOF MS. The excised protein spots were digested with trypsin (section 2.6.7) and the resulting peptide fragments were analysed by MALDI-TOF MS for identification by peptide mass fingerprinting (PMF). The MASCOT search engine, using the MSDB and NCBI databases, was exploited for matching the spectra data (section 2.6.8).

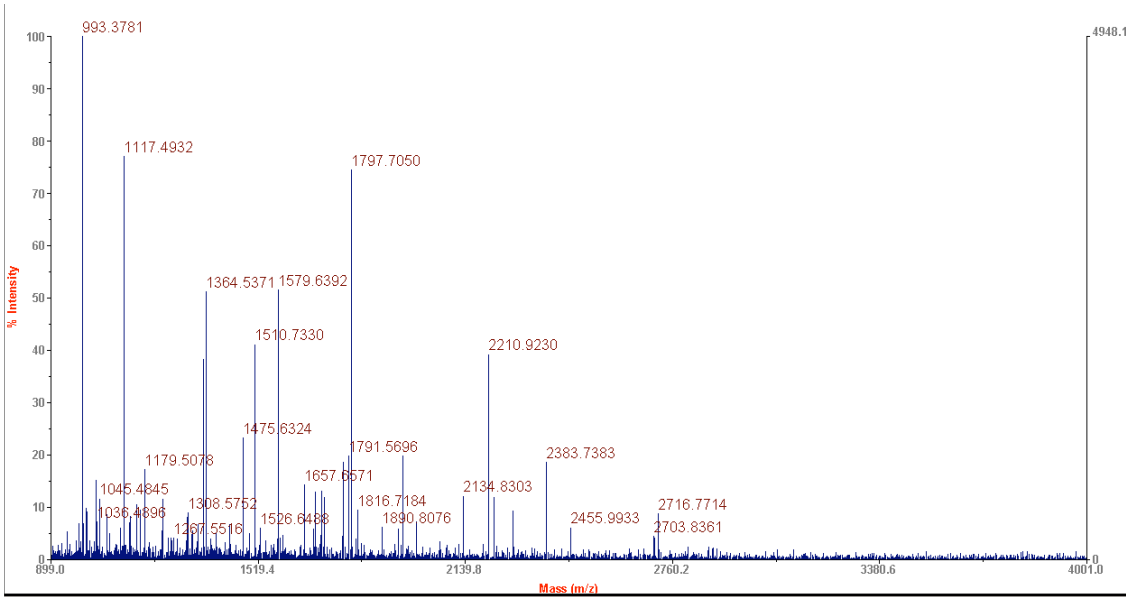
Among the 290 spots analysed by MALDI-TOF MS, 135 (48%) spots were positively identified (Table 6.1). In addition, 115 (40%) spots produced good spectra, as defined by abundant peak detection versus low background noise, but were not positively identified. Finally, 40 (12%) did not produce spectra, thus no peptide peak was detected on the PMF. Positively identified proteins are defined by MOWSE score (probability based score) higher than the threshold value of 64 (when using the MSDB and NCBI database) at $p<0.05$. The p value represents the absolute probability that observed match is a random event (Table 6.2). A typical MALDI-TOF MS spectrum obtained after the digestion of a protein, spot 25, is shown in Figure 6.5A. This protein (spot 25; Figure 6.4B) was identified as a isocitrate dehydrogenase (gi|75247585) from *Cucumis sativus* and its amino acid sequence and matched peptides is shown in Figure 6.5B. The identified proteins were classified into their functional categorises according to Bevan *et al*. (1998), Ndimba *et al*. (2005) and the oilseed proteomic database (http://oilseedproteomics.missouri.edu/, Table 6.2).

All the identified proteins were further analysed by blasting against the publicly available *Malus* EST database from NCBI. This was performed to confirm the validity of the protein identified using MASCOT search engine. Both the complete amino acid sequences of identified proteins and the matched peptide sequences from the PMF as revealed by MASCOT were blasted against the *Malus* EST database (Figure 6.5C). Queries performed using the 15 protein sequences that were identified from *Malus* sp. though successful, showed at least 96% identical, thus not all were 100% identical to the *Malus* ESTs sequences from the NCBI (Table 6.2). A typical cross species protein validation for isocitrate dehydrogenase (gi|75247585; spot 25) from *Cucumis sativus* is shown on Figure 6.5C. A total of 111 (82%) of the 135 identified proteins showed similarities with the existing data of *Malus* EST from the NCBI and a total of 24 proteins showed no significant similarities (Table 6.2).

Table 6.1: Proportion of spots positively identified, spots not positively identified but producing good spectrum and spots producing no spectrum.

Protein spot identification	Protein extracted by TCA/acetone (%)	Protein extracted by phenol (%)	Overall proportion (%)
Positive identification	48	47	48
Spectrum but no positive identification	43	39	40
No spectrum	9	14	12

(A)



(B)

```
  1 MAFQKIKVAN PIVEMDGDEM TRVIWESIKN KLIFPFLELD IKYFDLGLPH
 51 RDATDDKVTI ESAEATLKYN VAIKCATITP DEARVKEFGL KQMWRSPNGT
101 IRNILNGTVF REPILCKNVP RLVPGWTKPI CIGRHAFGDQ YRATDTVIRG
151 PGKLKLVFEG QETQEIEVFN FTGAGGVALA MYNTDESIRS FAEASMATAY
201 EKKWPLYLST KNTILKKYDG RFKDIFQEVY ESQWKSKFEA AGIWYEHRLI
251 DDMVAYALKS EGGYVWACKN YDGDVQSDFL AQGFGSLGLM TSVLVCPDGK
301 TIEAEAAHGT VTRHFRVHQK GGETSTNSIA SIFAWSRGLA HRAKLDDNAS
351 LLEFTEKLEL AYIDTVESGK MTKDLALILH GSKLSRDQYL NTEEFIDAVA
401 EELKSRLLKA
```

(C)

```
Query 1    MAFQKIKVANPIVEMDGDEMTRVIWESIKNKLIFPFLELDIKYFDLGLPHRDATDDKVTI 60
           MAFQKIKVANPIVEMDGDEMTRV W+SIK+KLI PF+ELDIKYFDLGLPHRDATDDKVTI
Sbjct 22   MAFQKIKVANPIVEMDGDEMTRVFWKSIKDKLILPFVELDIKYFDLGLPHRDATDDKVTI  82

Query 61   ESAEATLKYNVAIKCATITPDEARVKEFGLKQMWRSPNGTIRNILNGTVFREPILCKNVP 120
           ESAEATLKYNVAIKCATITPDEAR+KEF-LK-MWRSPNGTIRNILNGTVFREPILCKN+P
Sbjct 83   ESAEATLKYNVAIKCATITPDEARMKEFSLKSMWRSPNGTIRNILNGTVFREPILCKNIP 143

Query 121 RLVPGWTKPICIGRHAFGDQYRATDTVIRGPGKLKLVF--EGQ-ETQEIEVFNFTGAGGV 177
           RL+PGWTKPICIGRHAFGDQYRATD-VI+GPGKLKLVF--EG+-E--E++V++FTG--GV
Sbjct 144 RLIPGWTKPICIGRHAFGDQYRATDAVIKGPGKLKLVFVPEGKDETQLDEVYDFTGAGGV 204

Query 178 ALAMYNTDESIRSFAEASMATAYEKKWPLYLSTKNTILKKYDGRFKDIFQEVYESQWKSK 236
           ALAMYNTDESIR+FAEASM-TA+EKKWPLYLSTKNTILKKYDGRFKDIFQEVYE+-WKSK
Sbjct 205 ALAMYNTDESIRAFAEASMTTAFEKKWPLYLSTKNTILKKYDGRFKDIFQEVYEANWKSK 265

Query 237 FEAAGIWYEHRLIDDMVAYALKSEGGYVWACKNYDGDVQSDFLAQGFGSLGLMTSVLVCP 297
           FEAAGIWYEHRLIDDMVAYALKSEGGYVWACKNYDGDVQSD-LAQGFGSLGLMTSVLVCP
Sbjct 266 FEAAGIWYEHRLIDDMVAYALKSEGGYVWACKNYDGDVQSDMLAQGFGSLGLMTSVLVCP 312

Query 298 DGKTIEAEAAHGTVTRHFRVHQKGGETSTNSIASIFAWSRGLAHRAKLDDNASLLEFTEK 358
           DGKTIEAEAAHGTVTRH+RVHQKGGETSTNSIASIFAW+RGLAHRAKLDDN--LLEFT+K
Sbjct 313 DGKTIEAEAAHGTVTRHYRVHQKGGETSTNSIASIFAWTRGLAHRAKLDDNPRLLEFOEK 373

Query 359 LELAYIDTVESGKMTKDLALILHGSKLSRDQYLNTEEFIDA 398
           LE-A-I-TVESGKMTKDLALILHGSKL+R+--LN--+FIDA
Sbjct 374 LEEACIGTVESGKMTKDLALILHGSKLARNH*LNH*KFIDA 499
```

Figure 6.5: Spectrum of the digested peptides obtained by MALDI-TOF MS for spot 25 from the 2D-PAGE established with protein extracted by phenol precipitation.

Spot 25 was manually excised, trypsin digested and analysed by MALDI-TOF MS for the acquisition of (A) the PMF of the fragmented peptides. The MASCOT tool was then used to interrogate the NCBI and MSDB databases using the mass of every peptide to identify the protein. The protein was positively identified as isocitrate dehydrogenase (gi|75247585) from *Cucumis sativus*. The amino acid sequence of this protein is shown in (B) and matched peptides are highlighted in bold red. (C) represents the validation of the fragmented peptide sequences and/or the complete protein sequence when BLAST searched against the *Malus* EST database from the NCBI to confirm the validity (percentage identity) of the protein identified using the MASCOT search engine. Isocitrate dehydrogenase from *Cucumis sativus* (Query) showed 91% identity to isocitrate dehydrogenase (gb|cv656444.1 and gb|629517.1) from *Malus x domestica* (Subject) cDNA clones Mdst6014g215 and Mdfrt3101i225, respectively.

.

Table 6.2: Recapitulative table of positively identified proteins from apple fruit pulp using MALDI-TOF MS.

The proteins were identified by MALDI-TOF MS and classified into functional classes. The table lists the spot number, corresponding to the numbers in figure 6.4, protein matched, species, accession number, MOWSE score (number greater than 64 being significant at $p<0.05$), the theoretical molecular mass (kDa)/pI, the observed mass (kDa)/pI as seen on the gel, the queries matched and the sequence coverage (%). Protein validation shows the matching percentage of identified proteins against the *Malus sp.* blasted using the NCBI Blastn and the matched queries from the PMF. Spots numbers 1.0 to 15.0 correspond to the spots numbered 1 to 15 on 2D PAGE from Figure 6.4a.

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Mascot Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
	Energy								
	a. Citric Acid cycle/Gluconeogenesis								
25	Isocitrate dehydrogenase	*Cucumis sativus*	gi|75247585	70	46/6.00	40/6.9	12	25	91/11
30	Cytosolic malate dehydrogenase	*Malus x domestica*	gi|78216493	89	36/6.01	35/6.6	8	38	98/7
31	NAD-dependant malate dehydrogenase	*Prunus persica*	gi|75164483	72	35/6.60	35/6.5	6	37	96/4
50	Cytosolic malate dehydrogenase	*Malus x domestica*	gi|78216493	80	36/6.01	36/6.3	8	36	98/6
153	Probable malate dehydrogenase	*Pisum sativum*	gi|7488830	97	42/7.62	15/6.7	10	34	92/6
22	Methylthioalkylamalate synthase	*Arabidopsis lyrata*	gi|122179082	74	55/7.20	40/6.3	8	22	65/4

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
	b. Glycolysis and other carbohydrate metabolism associated proteins								
11	UTP-glucose-1-phosphate uridylyltransferase	*Pyrus pyrifolia*	gi|6136112	118	52/5.99	50/6.2	9	46	97/7
12	Enolase	*Ricinus communis*	gi|1169534	86	48/5.56	48/6.3	9	36	91/6
26	NADP-dependent sorbitol - 6-phosphate dehydrogenase (fragment)	*Prunus emarginata*	gi|75156649	76	28/9.16	40/7.0	10	41	75/6
29	NADP-dependent sorbitol-6-phosphate dehydrogenase (fragment)	*Prunus emarginata*	gi|75156652	68	28/8.78	32/6.6	7	32	75/5
167	NADP-dependent sorbitol-6-phosphate dehydrogenase (fragment)	*Prunus emarginata*	gi|75156649	79	28/9.16	20/5.2	9	44	75/6
202	NADP-dependent sorbitol-6-phosphate dehydrogenase (fragment)	*Prunus emarginata*	gi|75156649	71	28/9.16	62/5.0	8	37	75/6
36.0	R1 (fragment)	*Solanum tuberosum*	gi|75111367	70	60/6.75	55/6.6	14	32	92/5
38.0	Isopentenyl-diphosphate ∂-isomerase II	*Camptotheca acuminata*	gi|6225532	68	35/6.23	66/5.9	9	35	90/5

Table 6.2 continued

Spot [a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
	c. Photosynthesis								
2	Ribulose-1,5-bisphosphate carboxylase (RuBisCO) large subunit	*Azima tetracantha*	gi|75280255	68	53/6.22	65/6.2	7	20	95/5
9	RuBisCO-large subunit (E.C 4.1.1.39) (fragment)	*Keteleeria davidinia*	gi|75314656	76	49/6.78	50/6.7	13	20	93/3
44[b]	RuBisCO-large subunit	*Cuscuta sandwichiana*	gi|122246017	141	53/6.74	34/5.2	41	32	90/7
143	RuBisCO large subunit	*Racinaea ropalocarpa*	gi|122246017	76	53/6.23	40/5.0	10	25	94/8
150	RuBisCO (fragment)	*Larix laricina*	gi|75275869	71	9/5.00	30/4.6	5	70	74/4
20	Ferredoxin-NADP (H) oxidoreductase	*Triticum aestivum*	gi|20302471	72	39/8.29	40/6.0	12	37	85/4
	d. Lipid metabolism								
68	Lipoxygenase	*Fragaria x ananassa*	gi|75295984	70	100/6.34	26/5.8	10	16	76/4
104	Phospholipase D alpha1	*Arachis hypogaea*	gi|122200188	71	91/6.25	16/6.3	7	18	78/4
	Ethylene biosynthesis								
21	S-adenosyl-L-methionine synthetase	*Medicago sativa* subsp. falcata	gi|139478060	123	43/5.77	43/5.8	12	37	96/7

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
41	1-aminocyclopropane-1-carboxylate oxidase (ACO)	*Malus x domestica*	gi|74325224	103	35/5.24	35/5.4	8	33	95/5
141[b]	Ethylene receptor (fragment)	*Actinidia deliciosa*	gi|75128277	140	25/8.12	50/5.6	29	40	80/8
197	Ethylene response sensor protein (ethylene receptor)	*Rumex palustris*	gi|75098246	75	70/6.76	60/5.2	9	38	68/4
211	ACO	*Pyrus communis*	gi|75289373	86	35/5.38	35/5.4	9	10	92/6
	Defense/Detoxifying enzymes								
2.0	Alcohol dehydrogenase	*Garden petunia*	DEPJA1	73	42/6.19	53/5.8	8	21	85/5
11.0	Peroxidase 32	*Thellungiella halophila*	gi|144227396	67	39/7.93	23/5.9	5	28	71/-
33	Putative Quinone oxidoreductase	*Fragaria x ananassa*	gi|15808674	74	36/6.62	36/6.4	8	24	85/6
38	[a]IMP dehydrogenase /GMP reductase	*Medicago truncatula*	gi|122192808	83	141/6.51	40/5.7	6	16	55/2
72	Glutathione-disulphide reductase	*Humulus lupulus*	gi|154292532	78	19/4.94	26/5.5	5	46	71/4
98	Superoxide dismutase 2	*Malus xiaojinensis*	gi|122238163	69	15/5.60	15/6.2	5	43	100/5
105	NADH-plastoquinone oxidoreductase subunit 1	*Piper cenocladum*	gi|115605072	79	21/5.94	16/6.4	6	34	94/4
91	MLA7	*Hordeum vulgare*	gi|75324922	70	108/5.95	15/5.6	9	12	37/1

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
80	Major allegern Mal d 1.03G	*Malus domestica*	gi|75316774	68	18/5.23	20/5.0	6	45	100/6
94	Major allergen mal d 1.02 (AP15)	*Malus x domestica*	gi|42558971	67	17/5.63	16/6.0	6	37	100/6
19.0	Major allergen mal d 1	*Malus x domestica*	gi|42558971	80	18/5.62	19/5.7	9	54	100/9
20.0	Full-Major allergen mal d 1	*Malus x domestica*	gi|42558971	77	18/5.62	19/5.5	5	33	100/5
27.0	Abietadiene synthase	*Abies grandis*	gi|15080737	64	99/5.53	32/5.0	8	13	91/5
28.0	Abietadiene synthase	*Abies grandis*	gi|15080737	69	99/5.53	32/5.2	9	16	91/6
103	Ribonuclease-like PR-10a	*Malus x domestica*	gi|75306007	90	18/5.63	16/6.3	6	41	100/6
123	Dehydroascorbate reductase	*Malus x domestica*	gi|110083901	91	24/6.18	25/6.3	9	18	100/9
193	NBS-containing resistance-like protein	*Prunus serrulata*	gi|111141158	72	17/6.14	50/6.3	7	52	57/2
216	NBS-LRR disease resistance protein homologue (fragment)	*Hordeum vulgare*	gi|7532860	76	107/7.39	35/6.5	19	20	35/3
215	Glutathione peroxidase	*Medicago truncatula*	gi|122244162	76	19/4.86	65/6.2	6	28	73/3
16.0	Temperature-induced lipocalin-1	*Zea mays*	gi|122239267	65	23/5.36	23/6.1	6	34	47/3
24.0	Cysteine protease inhibitor 3 (fragment)	*Solanum tuberosum*	gi|20137687	64	16/8.85	25/5.3	8	48	75/5

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
	Proton transporting ATPases								
8	H^+-transporting two-sector ATPase, α-chain	*Glycine max*	gi|418790	98	55/6.23	45/6.8	11	32	73/5
16	F1-ATP synthase β-subunit (fragment)	*Sorghum bicolor*	gi|75277582	79	49/5.25	50/5.6	13	41	96/6
17	H^+-transporting two-sector ATPase, β-chain	*Hevea brasiliensis*	gi|82027	80	60/5.95	52/5.6	8	21	99/5
18	H^+-transporting two-sector ATPase, β-1 chain	*Nicotiana plumbaginifolia*	gi|82133	91	60/5.95	50/5.5	16	40	97/6
177[b]	F1-ATP synthase, β-subunit	*Sorghum bicolor*	gi|4388533	239	60/5.60	50/5.5	27	77	96/6
47	Putative oligomycin sensitivity conferring protein	*Silene diclinis*	gi|57282891	78	24/9.64	35/5.6	7	39	68/3
226	ABC transporter, trans-membrane region, type 1	*Medicago truncatula*	gi|122190188	69	51/8.50	30/6.6	9	19	69/3
77	Iron binding protein	*Pyrus pyrifolia*	gi|89276793	74	29/5.44	25/5.2	10	42	95/7
	Stress responsive proteins								
154	Heat shock protein	*Ricinus communis*	gi|223534226	83	75/5.35	75/5.0	14	34	93/6
155	Heat shock protein	*Ricinus communis*	gi|223534226	88	75/5.35	75/4.9	10	24	93/5

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
189	Heat shock protein	*Malus x domestica*	gi|6969976	119	71/5.17	70/5.35	25	35	99/20
191	Cell autonomous heat shock cognate protein 70	*Cucurbita maxima*	gi|75299362	88	71/5.17	70/5.45	17	31	96/6
195	Dehydrin	*Malus x domestica*	gi|110238587	74	24/6.61	31/5.6	5	34	96/5
	Transcription and translation								
12.0	TAF15b fragment	*Arabidopsis thaliana*	gi|75223840	66	39/7.93	23/6.4	8	15	83/3
14.0	TAF15b	*Arabidopsis thaliana*	gi|39545916	72	39/7.93	23/7.0	7	19	83/3
13.0	DEAD/DEAH box helicase	*Medicago truncatula*	gi|122181528	77	42/9.12	23/6.8	6	18	-/-
10	38kDa ribosome –associated protein	*Chlamydomonas reinhardtii*	gi|75130662	70	45/9.45	50/6.5	10	37	76/3
17.0	38kDa ribosome – associated protein	*Chlamydomonas reinhardtii*	gi|75130662	68	45/9.45	24/5.8	11	32	76/3
39.0	Probable translation initiation factor	*Barrel medic*	gi|11358946	76	42/9.41	66/5.9	13	39	94/5
62	Maturase K	*Amyema glabrum*	gi|188526390	75	53/9.77	22/6.4	12	28	-/-
63	RNA binding protein	*Zea mays*	gi|195619486	80	54/8.74	22/6.2	8	24	47/-
83	Bromodomain-containing RNA binding protein 1	*Nicotiana benthamiana*	gi|75106815	69	67/7.03	18/5.6	8	15	85/-

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
89	MYB transcription factor MYB142	*Glycine max*	gi|110931736	74	11/7.85	15/5.7	8	74	75/4
96	Putative glycine-rich RNA-binding protein	*Prunus avium*	gi|75325210	68	17/7.82	50/6.25	6	40	92/4
173	Myb-like DNA-binding domain, putative (fragment)	*Medicago truncatula*	gi|122225714	70	41/7.11	30/6.5	9	23	73/-
212	MYB transcription factor 1	*Ostreococcus tauri*	gi|75320693	74	30/9.66	30/4.5	12	40	37/-
229	Small ribosomal protein 4 (fragment)	*Sphagnum cyclophyllum*	gi|7529896	78	22/10.33	40/6.4	10	44	50/-
	Cell growth/division: DNA synthesis/replication								
71	Cell division inhibitor MinD	*Chlamydomonas reinhardtii*	gi|75244812	75	38/9.18	24/5.75	9	18	58/2
113	Cell division inhibitor MinD	*Chlamydomonas reinhardtii*	gi|75244812	73	38/9.18	18/5.0	8	31	58/2
127	Phragmoplastin 12	*Glycine max*	gi|2129825	79	68/8.01	40/6.55	12	27	89/6
228	Probable NAD ADP-ribosyltransferase	*Zea mays*	gi|7489816	85	73/8.78	39/5.95	19	20	75/4
130	Synptobrevin-related protein (fragment)	*Pyrus pyrifolia*	gi|75232217	70	22/8.60	54/7.0	7	29	97/7

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
133	Embryogenic potential marker Dc3	*Daucus carota*	gi|482697	70	17/6.77	75/6.4	8	36	30/-
176	Transposase	*Chlamydomonas reinhardtii*	gi|159464567	93	100/8.91	32/5.1	1	21	-/-
	Signal transduction								
78	GTP-binding signal recognition particle SRP54, G-domain	*Medicago truncatula*	gi|122226059	76	61/9.34	25/5.1	15	38	89/5
84	Putative ABA-binding protein (fragment)	*Arabidopsis lyrata subsp. petraea*	gi|75159326	77	16/9.32	17/5.4	9	56	46/-
156	Probable GTP-binding regulatory protein	*Nicotiana tabacum*	gi|7489093	73	44/6.09	45/4.5	8	28	81/7
161	S-receptor kinase (fragment)	*Arabidopsis lyrata*	gi|75167865	66	36/4.82	25/5.1	6	26	34/-
163	Small-GTP-binding protein	*Lycopersicon esculentum*	gi|75282221	74	22/5.34	24/4.3	8	35	90/5
200	Phytochrome kinase substrate putative	*Ricinus communis*	gi|223550684	82	48/6.86	61/5.0	12	31	41/-
206	Lissencephaly type-1-like homology motif	*Medicago truncatula*	gi|122182451	78	22/6.54	52/4.7	9	23	54/-

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
	Amino acid and purine biosynthesis related enzymes								
36	Glutamine synthetase	*Medicago truncatula*	gi|75276958	67	39/5.38	38/6.1	6	21	92/3
39	β-cyanoalanine synthase 1	*Malus x domestica*	gi|93359259	78	41/7.60	38/5.7	7	19	99/7
138	Aspartate transaminase (AAT) 5 precursor	*Glycine max*	gi|485495	66	51/7.16	70/5.4	8	18	82/4
	Cytoskeleton-related protein								
4	Myosin class II heavy chain (ISS)	*Ostreococcus tauri*	gi|116051804	90	467/6.17	56/6.5	40	11	-/-
86	Myosin class II heavy chain (ISS)	*Ostreococcus tauri*	gi|116057804	85	467/6.17	15/5.4	42	10	-/-
19	Actin isoform B	*Mimosa pudica*	gi|11276971	130	42/5.31	40/5.6	12	47	87/12
141[b]	Putative actin 1 fragment	*Cholispora bungeana*	gi|75286002	140	25/8.12	50/5.6	29	40	95/15
177[b]	Actin-58	*Picea rubens*	gi|231496	239	50/5.6	50/5.5	27	77	97/10
204	Actin	*Morus alba*	gi|110612122	110	37/5.47	42/5.1	18	67	95/17
208	Actin	*Isatis tinctoria*	gi|58013197	93	42/5.31	42/5.5	11	48	95/11
209	Actin	*Gossypium hirsutum*	gi|32186900	151	42/5.23	45/5.4	17	58	95/15
34.0	Actin	*Linum usitatissimum*	gi|74888566	91	42/5.68	55/5.4	11	49	94/10

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
46a	Annexin, putative	*Ricinus communis*	gi|223535020	80	37/8.86	32/5.4	11	32	77/5
144	Dynein heavy chain isoform (DHCIb) (fragment)	*Chlamydomonas reinhardtii*	gi|7484372	81	135/5.92	40/5.4	19	18	-/-
203	Dynein heavy chain isoform (DHCIb) (fragment)	*Chlamydomonas reinhardtii*	gi|7484372	82	135/5.92	60/5.0	14	15	-/-
26.0	Dynein heavy chain isoform Ib.	*Chlamydomonas reinhardtii*	gi|75337416	70	48/6.13	26/5.0	38	15	-/-
	Unknown/unclassified proteins								
1.0	Hypothetical protein orf222	*Nicotiana tabacum*	gi|75103092	67	25/6.54	54/5.7	7	36	-/-
3.0	Hypothetical protein	*Medicago truncatula*	gi|122185315	64	57/8.72	52/6.0	6	11	-/-
28	Hypothetical protein	*Vitis vinifera*	gi|147843754	90	42/6.29	38/6.8	9	28	91/4
44[b]	Hypothetical protein	*Vitis vinifera*	gi|147841490	141	143/7.12	34/5.2	41	32	41/4
82	Hypothetical protein NitaMp024	*Nicotiana tabacum*	gi|57013897	82	16/9.60	20/5.6	10	65	-/-
5.0	AY007207 NID	*Taxus wallichiana*	gi|9965484	65	6/5.3	39/5.85	7	11	-/-
13	AB013353 NID	*Pyrus pyrifolia*	gi|3107931	115	52/5.99	50/6.1	11	43	97/11
79	AB004825 NID	*Solanum tuberosum*	gi|2225883	79	17/5.46	20/5.0	8	57	94/7
4.0	Predicted protein	*Physcomitrella patens*	gi|168016639	71	39/9.73	52/6.2	8	25	-/-

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
15.0	Predicted protein	*Ostreococcus lucimarinus* CCE9901	gi|145351181	73	27/5.74	23/6.9	6	30	-/-
1	Predicted protein	*Physcomitrella patens*	gi|16800824	80	64/9.17	60/6.0	10	20	-/-
5	Predicted protein	*Physcomitrella patens subsp. Patens*	gi|168046260	80	80/8.32	58/6.6	19	25	-/-
23	Predicted protein	*Physcomitrella patens*	gi|168052225	80	135/7.24	41/6.7	27	26	-/-
27	Predicted protein	*Physcomitrella patens*	gi|168011079	88	33/6.07	39/6.9	1	53	45/-
120	Predicted protein	*Physcomitrella patens*	gi|168031400	79	72/8.82	32/7.0	15	22	-/-
152[b]	Predicted protein	*Ostreococcus lucimarinus*	gi|145341236	162	53/7.08	14/6.6	18	35	33/-
		+							
		Physcomitrella patens	gi|168016055		35/9.95				41/-
170	Predicted protein	*Physcomitrella patens*	gi|168063095	77	28/7.10	25/6.0	7	33	-/-
187	Predicted protein	*Populus trichocarpa*	gi|224133614	122	90/5.07	95/5.7	2	33	87/2
188	Predicted protein	*Populus trichocarpa*	gi|224100127	74	59/4.95	65/5.7	31	46	73/9
192	Predicted protein	*Ostreococcus lucimarius* CCE9901	gi|145341236	95	53/7.08	70/5.5	1	31	-/-
32	Unknown	*Zea mays*	gi|194708668	75	47/8.44	35/6.5	10	27	-/-
46b[b]	Unknown	*Populus trichocarpa*	gi|118481067	71	50/8.28		33	26	77/6
			+			32/5.4			
	Hypothetical protein	*Vitis vinifera*	gi|147783587	70	34/9.07		35	32	41/4

Table 6.2 continued

Spot[a]	Protein name and Functional category	Species	Accession number[c]	Mowse score	Theoretical mass (kDa)/pI	Observed mass (kDa)/pI	Queries Matched	Sequence coverage (%)	Identity to *Malus* ESTs (%)/ Queries Matched
181	Unnamed protein product	*Ostreococcus tauri*	gi|116055990	76	51/10.1	60/6.0	14	33	-/-
3	MRGH21	*Cucumis melo*	gi|51477389	76	116/6.48	60/6.3	17	20	-/-
46c	Sialyltransferase-like protein	*Malus x domestica*	gi|70663504	66	44/9.49	35/5.5	9	27	98/6
139	ER-binding protein	*Malus pumila*	gi|58639078	92	74/5.14	75/5.3	18	27	99/16
190	Sequence 87 from patent WO02064764 (fragment)	*Fragaria x ananassa*	gi|25173114	68	64/5.83	70/5.4	8	18	71/3

[a] Spot number corresponding to annotation in figure 6.4

[b] Spot that contains more than one protein

[c] Accession numbers used were according to UniProt and NCBI

- No significant similarity or matched peptides to the *Malus* ESTs data

TAF15b - TATA Binding Protein-Associated Factor 15 binding

Hypothetical means that this protein is, hitherto, uncharacterised (probably named from gene sequencing projects) therefore novel and thus one can bioinformatically predict its function, conduct some assays and establish its function for the first time.

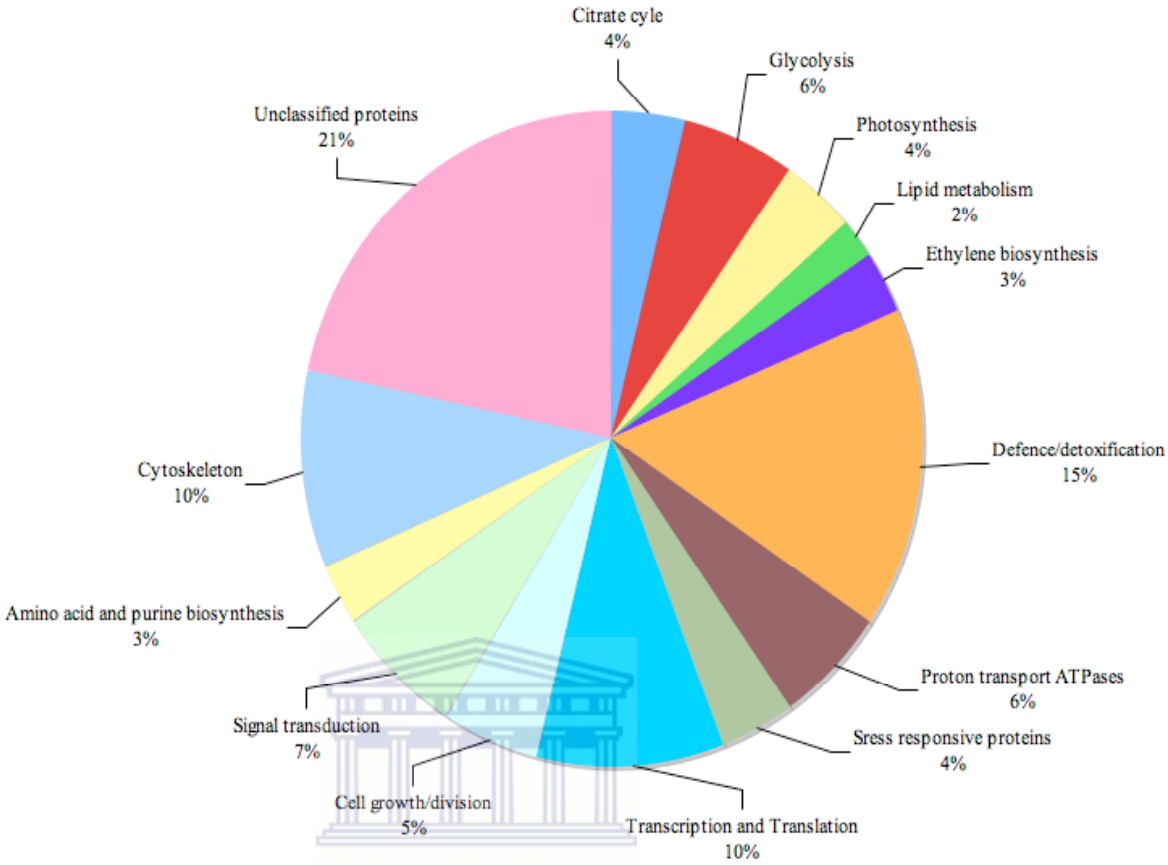


Figure 6.6: Functional classification of fruit pulp proteins identified by MALDI-TOF MS.

The proportion of each functional class represented in the mature apple fruit pulp proteome is represented diagrammatically.

6.5 DISCUSSION

Protein accumulation in a cell or tissue provides an overall response of an organism to endogenous and exogenous cues (Beveridge *et al.*, 2007). Therefore, by measuring the end product of gene expression using proteomic analysis, post-transcriptional, translational and post-translational modifications as well as turnover dynamics of proteins are considered (Park, 2004; Guarino *et al.*, 2007; Ndimba and Thomas, 2008; Ngara *et al.*, 2008). The identification and quantification of proteins expressed in a cell, tissue or entire living organism at a particular time is one approach towards the establishment of protein databases, functions and discovery of biomarkers. Protein characterization particularly determining their function by either direct analysis or measurement of protein expression using the proteomics technologies may lead to the better understanding of gene function (Abbott *et al.*, 1999; Gygi *et al*., 1999). Although several proteomic studies have been carried out using several other plant species like *Arabidopsis thaliana* (Ndimba *et al.,* 2005; Oh *et al*., 2005) and rice (*Oryza sativa*; Rakwal and Agrawal, 2003; Parker *et al*., 2006), plant fruit proteome have been poorly characterised to date. Therefore, this chapter reveals the successful development of preliminary proteome maps of apple fruit pulp and the identification of proteins extracted from the apple fruit pulp.

6.5.1 Optimisation of protein extraction

Total soluble proteins were extracted from mature apple fruit pulp by either TCA/acetone precipitation or phenol precipitation and were resuspended in extraction and IEF rehydration buffers, respectively, containing both urea and thiourea. Urea is a chaotropic substance and serves as a solubilising agent. Chaotrope agents act by

interfering with stabilising intra-molecular interactions mediated by non-covalent forces like hydrogen bonds, dipole-dipole interactions and hydrophobic interactions (Wang *et al.*, 2003). In this respect, it disrupts the three dimensional structure of proteins, which eventually unfold. The addition of thiourea in the buffer, as an additional chaotrope agent, also assists towards protein re-solubilisation while limiting protease activities (Wang *et al.*, 2003). Since the urea-based buffer used to resuspend proteins restricted their proteolytic degradation, interfering substances rather than protein degradation probably caused the smears observed on 1D-PAGE. Further, keeping the samples on ice during sample preparation also prevented proteolysis.

High background was observed after separation of proteins extracted using both extraction methods by 1D-PAGE as shown on Figures 6.1A and 6.2A. This background staining is a sign of interfering substances like phenolic compounds and/or carbohydrates, which are present in abundant quantities in apple fruits (Carpentier *et al.*, 2008). These compounds can cause unspecific interactions with proteins during extraction, and thus reduce the sample concentration (Carpentier *et al.*, 2005; Wang *et al.*, 2006; Carpentier *et al.*, 2008). The two extraction techniques were thus optimised to reduce the background that interfered with protein visualisation after 1D-PAGE.

6.5.1.1 TCA/acetone precipitation

In order to limit the presence of interfering substances in the TSP, 0.06% (w/v) sodium sulphite was added to TCA/acetone during protein precipitation. Sodium sulphite, a reducing agent, was added to inhibit the oxidation of polyphenols, which when oxidized interact with proteins, thus reducing their resolution during 1D- and/or 2D-PAGE (Ljanabi *et al.*, 1999; Singh *et al.*, 2002). After optimisation, protein concentration of

the extracts increased on average by seven fold. In addition, protein visualization after 1D-PAGE was improved and the high background reduced when using sodium sulphite (Figure 6.1B).

Interfering substances, like sugars and phenolic compounds have been shown to induce deviations during quantification by Bradford assay (Banik *et al.*, 2009). Thus, in order to validate whether the variations in protein concentration observed after optimization were caused by higher protein content rather than the presence of interfering substances, proteins were visualized after separation by 1D-PAGE. Following optimization, protein bands were clearly visualized and the background was reduced, suggesting that interfering substances were reduced.

6.5.1.2 Phenol precipitation

The agent PVPP plays a crucial role in trapping interfering compounds by complexing with carbohydrates, alkaloids and phenolic compounds, inhibiting their oxidation and thus preventing their unspecific interactions with proteins (Görg, 2000; Jamet *et al.*, 2008a,b; Wang *et al.*, 2008b). Additionally, PVPP has also been shown to improve protein stability during extraction (Carpentier *et al.*, 2008). The phenolic compounds and/or carbohydrates are in abundant quantities in apple fruits and cause unspecific interactions with proteins during extraction, reducing the sample concentration (Carpentier *et al.*, 2005; Wang *et al.*, 2006; Carpentier *et al.*, 2008). Therefore, in order to limit the presence of interfering substances in the TSP, 50 mM PVPP was added to TCA during the protein precipitation step. While similar protein concentrations were measured before and after optimisation, the resolution of the proteins improved after 1D-PAGE (Figure 6.2). In additon, the background staining was reduced. As a result,

the protein band pattern was more clearly visualised. After optimization, increase protein loading onto the 1D-PAGE only resulted in increased protein content, without the occurrence of high background shadowing band visualisation. This confirms the efficient removal of interfering substances that were present in the extracts prior to optimization.

6.5.1.3 Comparison between the TCA/acetone and phenol precipitation protein extraction methods

In terms of protein recovery efficiency, after optimization, the sample concentrations were approximately five times greater using the optimized phenol precipitation method than with the TCA/acetone precipitation method. These two methods have previously been used for protein extraction from banana (*Musa spp*) meristems, meristems of *in vitro* apple plantlets and leaves of *in vitro* potato (*Solanum tuberosum* L.) plantlets (Carpentier *et al.*, 2005; Saravanan and Rose, 2004). This is in correlation with the results reported by Saravanan and Rose (2004) who showed that protein concentration in samples extracted by phenol precipitation was higher than in samples extracted by TCA/acetone precipitation. Interestingly, Saravanan and Rose (2004) revealed that using phenol precipitation, more glycoproteins were extracted than when using TCA/acetone precipitation (Saravanan and Rose, 2004). This suggests that more glycoproteins might have been extracted from apple fruit pulp using phenol precipitation.

In terms of protein band pattern after optimization, similar protein band patterns were observed by 1D-PAGE separation (Figures 6.1B and 6.2B). A low abundance of high molecular weight proteins (>100 kDa) was also observed in both gels. In terms of

protein visualization, the phenol precipitation method resulted in a more efficient removal of interfering substances, generating high quality 1D-PAGE with reduced background (Figure 6.2). This was mainly because when using the phenol precipitation method, the upper phenol rich phase (collected) contains cytosolic and membrane proteins, lipids and pigments, while the aqueous lower phase (discarded) contains carbohydrates, nucleic acids and insoluble debris. The protein extracts in the phenol phase were then precipitated out by ammonium acetate in acetone (salting effect; Carpentier *et al.*, 2005).

6.5.2 Establishment of 2D-PAGE proteome maps

Following protein separation by 1D-PAGE, proteome maps of apple fruit pulp were established by resolving TSP on 2D-PAGE using protein samples from four biological replicates extracted by either TCA/acetone or phenol precipitation (section 6.3). Whilst the 1D-PAGE technique separates proteins according to their molecular weight only, proteins are resolved according to their pI and molecular weight during 2D-PAGE, thus allowing precise comparison between samples.

The apple pulp TSP were initially resolved using linear 7 cm IPG strips, pH range 3-10 for the isoelectric focusing step and the protein separation according to their molecular weights (data not shown). The majority of proteins were detected clustered in the pH 4-7 region, thus a narrower pH range was used to increase protein resolution and separation. The proteome reference maps were therefore developed using linear 7 cm IPG strips, pH range 4-7, exclusively (Figure 6.3).

Both vertical and horizontal streaking were observed on the 2D-PAGE after CBB

staining and were minimal after optimisation (Figure 6.3). Vertical streaking can be a result of insufficient equilibration, excess of DTT in the IEF rehydration buffer or insufficient SDS in the running buffer for SDS-PAGE (Wang *et al*., 2003; Saravanan and Rose, 2004) and/or impurities in the samples or solutions. To limit vertical streaking, the equilibration steps were extended from 15 to 20 min, 8 mM DTT was added to the IEF rehydration buffer and 0.1% (v/v) SDS was included to the running buffer. Horizontal streaking may be caused by impurities in the agarose overlay or equilibration solutions, the presence of interfering substances in the samples, or under focusing of the protein (Wang *et al*., 2003; Saravanan and Rose, 2004). In this regard, horizontal streaking was limited using freshly prepared equilibration solutions, agarose overlay and running buffer for every gel. The focusing period during IEF was also extended from 4 to 5 h. In addition, horizontal streaking may also occur in case of insoluble or too acidic samples (Wang *et al*., 2006). To limit these problems, 0.02% (v/v) carrier ampholytes (pH 3-10), a buffer that acts as cyanate scavenger and thus improves protein stability over a pH range of 3-10 by minimizing protein aggregation due to charge-charge interactions, was used (Rahimpour *et al*., 2007).

In terms of protein focusing and resolution, the phenol precipitation method resulted in a more efficient removal of interfering substances, generating higher quality 2D-PAGE with low background and vertical streaking (Figure 6.4B and C). Studies by Carpentier *et al.* (2005) and Saravanan and Rose (2004) discussed earlier have shown a similar trend with the results obtained in this study.

After successive experiments comparing efficiency of extraction buffer and IEF buffer (section 2.2) on focusing and protein resolution on 2D-PAGE, a better resolution of the

protein was obtained on 2D-PAGE when protein was resolubilised in extraction buffer and IEF extraction buffer using TCA/acetone and phenol precipitation, respectively.

In terms of proteome profile differences between the two extraction methods, the majority of protein spot molecular weight ranged between 18 kDa and 66.2 kDa (Figure 6.3A), and between 10 kDa and 120 kDa (Figure 6.3B) for the TCA/acetone and phenol precipitation, respectively. In addition most of the spots were concentrated in the pH ranges 5 to 7 and 4 to 7 for the TCA/acetone and phenol precipitation, respectively. The two proteome maps had very different spot patterns. The narrow pH gradients used increased resolution and facilitated comparative studies between the two protocols implemented in this study.

In terms of number of spots resolved, an average of 60 and 230 protein spots were visualized after separation by 2D-PAGE of TSP extracted using TCA/acetone and phenol precipitation, respectively (Figure 6.4). More protein spots were visualized after extraction using phenol precipitation compared to TCA/acetone precipitation, although equivalent amount of proteins were loaded on 2D-PAGE, according to the Bradford results. Phenol is known to acts as a dissociating agent by decreasing molecular interaction between proteins and other plant secondary compounds like organic acids and phenolic compounds (Carpentier *et al*., 2005). Therefore, the variation in protein spot number between extraction methods might be attributed to a greater capacity of the phenol to dissociate protein complexes.

6.5.3 Identification and validation of the total soluble proteins

Following the establishment of 2D-PAGE proteome maps of apple fruit pulp TSP from

combined protein samples extracted by either TCA/acetone or phenol precipitation from the seedlings of the cross between the cultivars ‘Golden Delicious’ and ‘Dietrich’ (Figure 6.4), all the CBB stained protein spots were excised from the gels, trypsin digested and analysed by MALDI-TOF MS.

The peptide mass fingerprints (PMF) obtained were used to interrogate the NCBI and MSDB databases using the MASCOT search engine for identification of the proteins (Table 6.2). The majority of the proteins did digest well and a PMF was obtained. However, a PMF could not be obtained for some of the protein spots, possibly because of low peptide concentration (Ochs, 1983). Since the apple genome has not yet been fully sequenced yet, searches were performed against ‘Viridiplantae’ and ‘other green plants’ (as classified by MASCOT) using the NCBI and MSDB databases. The results were thus highly dependent on the similarity of the apple proteins to proteins from other green plants.

Identifications with MOWSE scores higher than 85 were automatically considered as positive assignments because of their high significance. All other assignments with MOWSE score greater than 64 were considered positive if more than 10% of the protein sequence was covered. The sequence coverage represents the ratio of peptides from the protein spot of interest to match to the sequence of protein in the database (Perkins *et al.*, 1999; Damodarana *et al.*, 2007). If more than one protein satisfied the criteria, the entry with the highest MOWSE score was assigned.

In terms of comparison between the expected and theoretical masses, discrepancies were observed for several proteins. In the case of some proteins, the theoretical mass

was greater than the observed mass on 2D-PAGE. For example, in the case of phospholipase D alpha 1 (spot 104; Figure 6.4B; Table 6.2), the theoretical mass is 91 kDa, while the observed mass was approximately 16 kDa. These variations may be a result of protein degradation or denaturation, causing the fragmentation of the protein during extraction and/or sample preparation. In addition, the spot analysed may represent a subunit of the full-length protein that unfolded during resuspension in the denaturing urea/thiourea-based buffer. Conversely, the theoretical mass of some proteins was lower than that observed on 2D-PAGE. For example, in the case of RuBisCO large subunit (spot 2; Figure 6.4B; Table 6.2), the observed mass (65 kDa) was greater than the theoretical mass (53 kDa). These mass increments may be signatures of post-translation modifications, like phosphorylation or glycosylation (Thomas, 2008). In accordance with this mass shift, RuBisCO has been shown to become phosphorylated (Agarwal *et al*., 1993). In some cases, the protein sequence in the database is only a fragment of the full-length sequence, and thus may explain the mass variations for some protein spots like RuBisCO (spot 150; Figure 6.4C; Table 6.2).

A total of 60 and 230 CBB stained and well-resolved spots (Figure 6.4) were selected from 2D proteome maps established with TSP extracted using the TCA/acetone and phenol precipitation methods, respectively, for their identification by MS. The MALDI-TOF MS analysis resulted in the positive identification of 135 spots (i.e. with MOWSE score greater than 64), out of the 290 proteins originally trypsinised. This represents 48% success rate of protein identification. From these positively identified proteins, 11 and 106 spots were excised from the 2D-PAGE proteome map obtained using proteins extracted using the TCA/acetone and phenol precipitation methods, respectively. The

identified proteins were then classified according to their functions (Table 6.2).

Several spots were identified as the same protein, for example spots 30 and 50 detected as cytosolic malate dehydrogenase, and spots 26, 167 and 202 were all identified as fragments of NADP-dependent sorbitol-6-phosphate dehydrogenase. Some spots, listed at the end of Table 6.2, were not positively identified although good-quality spectrum, as defined by abundant peptide peaks and low noise, were obtained. Thus, 52% (155) of the spots were not positively identified. Out of these, 40% (115) of the spots had good spectra, while 12% (40) did not have good quality spectrum, i.e. the spectrum obtained had high background *versus* peptide signal, and consequently could not be identified (listed at the end of Table 6.2).

Only 30 protein spots were identified against plants belonging to the Rosaceae family. Out of these, 15 protein spots were identified against plants from the *Malus* genus. This may be as a result of the limited genetic and proteomic data in the Rosaceae family. However, all the identified proteins were further blasted against the publicly available *Malus* EST database from NCBI. This was performed to confirm the validity of the protein identified using MASCOT search engine. Both the complete amino acid sequences of identified proteins and the matched peptide sequences from the PMF as revealed by MASCOT were blasted against the *Malus* EST database. The preliminary queries were performed using the 15 protein sequences that were identified in *Malus* in order to demonstrate validity of the search method. Though successful, the sequences were at least 96% identical, thus not all were 100% identical to the *Malus* ESTs sequences from the NCBI (Table 6.2). This could be due to sequence modifications during post-transcription and post-translation. Of important to note was that all the

matched peptide sequences from PMF matched completely when blasted against the *Malus* EST database. When the blast search was applied to the rest of the identified proteins, 82 % of the proteins showed similarities with the existing data of *Malus* EST from the NCBI (Table 6.2). This therefore demonstrated the utility of *Malus* EST in proteomic data validation. Some proteins showed low similarities and a total of 24 proteins showed no significant similarities. These may be a result of short length ESTs in the database or poor sequence quality of raw EST entries (Mooney and Thellen, 2004). In addition, these results clearly indicate the limited genetic data in the Rosaceae family. The genomes of these plant species, which include apple, pear and peach, are not yet fully sequenced, but sequencing efforts are currently initiated for apple, strawberry and peach (Rees, personal communication). This will help towards improved rate of positive identification for proteins with good-quality spectra.

To date, only one proteomic analysis has been published using apple flesh (pulp) to identify the major soluble components in 'Annurca' cultivar (Guarino *et al*., 2007). Guarino *et al*. (2007) reported the separation of TSP by 2D-PAGE and the positive identification of 44 protein spots, involved in energy production, ripening and stress response. However, no comprehensive analysis of the total proteome from apple fruit pulp has been published to date.

6.5.4 Classification of proteins

Following their identification, proteins were classified into 11 functional categories as established by Bevan *et al*. (1998), Ndimba *et al.* (2005) and the oilseed proteomic database (http://oilseedproteomics.missouri.edu/). The most represented functional class of the positively identified proteins was associated with energy metabolism (16%).

However, the largest group was represented by unclassified proteins (21%), followed by energy metabolism (16%), defense/detoxification related proteins (15%), transcription and translation related proteins (10%), cytoskeleton related proteins (10%), signal transduction (7%), proton transport ATPase (6%), cell growth and division (5%), stress responsive proteins (4%), ethylene biosynthesis (3%)and amino acid and purine biosynthesis (3%). Targeted bioinformatics tools (section 6.5.4.12) were used for the determination of conserved domains to predict the putative functions of all the unclassified proteins, as a first step toward their characterisation.

6.5.4.1 Energy associated proteins

Energy associated proteins were grouped into four sub-categories, namely citric acid cycle associated proteins, glycolysis and other carbohydrate metabolism associated proteins, photosynthesis associated proteins and lipid metabolism associated proteins. This classification was based on the involvement of energy-associated proteins in the transformation of macronutrients, carbohydrates, lipids and proteins to provide energy for driving cellular processes that require energy.

6.5.4.1.1 Citric acid cycle associated proteins

Isocitrate dehydrogenase (IDH; EC 1.1.1.42, spot 25) is an enzyme involved in the third step of the citric acid cycle. It catalyzes, in a two-step process, the oxidative decarboxylation of isocitrate, producing α-ketoglutarate and CO_2 while converting NAD^+ to NADH. Several studies have reported the detection of IDH in fruits (Gallardo *et al*., 1995; Sadka *et al*., 2000b; Katz *et al*., 2007). In young peach fruits, the majority of the protein activity was measured in plant mitochondria. However, as the fruit matures, the activity was also measured in the cell cytosol (Etienne *et al.*, 2002).

Accordingly, in this study, IDH was detected in mature fruit pulp, thus enzyme was probably located in the apple cytosol rather than in mitochondria. In citrus fruit pulp, the IDH has been associated with fruit development (Sadka *et al.*, 2000b). During the late stage of peach ripening, IDH has been shown to act on the metabolism of organic acids (Etienne *et al.*, 2002). Following purification of IDH isoforms from TSP extracts of ripe tomato fruit tissues by ion exchange chromatography, the results of activity assay suggested that IDH might play a role in glutamate accumulation (Gallardo *et al.*, 1995; Baxter *et al.*, 2005). Accordingly, in this study, IDH was detected in mature fruit pulp, suggesting that the enzyme was probably located in the apple cytosol and thus may play a role in organic acids metabolism and/or glutamate accumulation in the mature apple fruit pulp.

Four protein spots were identified as malate dehydrogenase (EC 1.1.1.37; spots 30, 31, 50 and 153). Malate dehydrogenase catalyzes the reversible conversion of malate into oxaloacetate, using NAD^+ as cofactor (Katz *et al*., 2007). The enzyme is also involved in gluconeogenesis, the synthesis of glucose from smaller molecules. Malate dehydrogenase is known to play a role in malic acid biosynthesis during fruit development (Diakou *et al.,* 2000). Malic acid is one of the organic acids, which serve as a respiratory substrate during the ripening process of apples (Katz *et al*., 2007) and peaches (Etienne *et al*., 2001). In apples, it has been shown to accumulate during the early stages of fruit development and thus induces the acidic taste (Katz *et al.*, 2007)

6.5.4.1.2 Glycolysis and other carbohydrate metabolism associated proteins

In this study, two proteins were identified as involved in glycolysis and four in sorbitol metabolism. It is widely reported that fruit development induces changes of the

carbohydrate content in the fruit pulp. In terms of glycolysis, spot 11 was positively identified as UTP-glucose-1-phosphate uridylyltransferase, also known as UDP-glucose pyrophosphorylase (EC 2.6.7.9). This enzyme synthesizes UDP-glucose from glucose-1-phosphate and plays a role in sucrose biosynthesis in fruit cells (Katz *et al*., 2007). In banana, Pua *et al.* (2000) showed that UTP-glucose-1-phosphate uridylyltransferase expression was greater in fruit pulp than in peel during the ripening process. These results emphasize the importance of the enzyme in sugar synthesis in pulp during fruit ripening.

Spot 12 was identified as enolase (EC 4.2.1.11), also called phosphoenolpyruvate, is the metalloenzyme responsible for the catalysis of 2-phosphoglycerate to phosphoenolypyruvate, the only dehydration and penultimate step of glycolysis (Pancholi, 2001). A study reported the detection of an elevated expression of the genes coding for enolase in tomato fruits, thus suggesting the increased need for energy in nonphotosynthetic tissues (Van Der Straeten *et al.*, 1991). The identification of the enolase as well as the other enzymes involved in the glycolysis pathway in citrus fruit suggested that a fraction of the sucrose accumulating in fruits is used for energy production (Katz *et al.*, 2007). Therefore, the detection of enolase in mature apple fruit pulp, a non-photosynthetic tissue, suggests the use of sucrose to respond to the increasing energy needs of the fruits.

In terms of sorbitol metabolism, four protein spots were identified as NADP-dependent sorbitol-6-phosphate dehydrogenase (SPDH; EC 1.1.1.140; spots 26, 29, 167 and 202). The S6PDH is involved in the reversible reduction of glucose to sorbitol (Kanayama *et al.,* 2005). The sugar alcohol, sorbitol, also known as glucitol, is the primary product of

photosynthesis i.e. major photosynthate (Kanamarua *et al*., 2004). Once translocated to the fruit, sorbitol-6-phosphate is catabolised to fructose and glucose by NAD-dependent sorbitol dehydrogenase and sorbitol oxidase, respectively (Yamaki, 1980). In apples, the biosynthesis of sorbitol is an important step of carbohydrate metabolism, since the carbohydrates assimilated in leaves have been shown to be translocated into the vacuoles of fruit cells (Yamaki, 1987). In our study, NADP-dependent SPDH was detected in the fruit tissue and not NAD-dependent sorbitol dehydrogenase as expected, an indication that the protein is also present in the fruit tissue. A Blast search for the similarity between NAD and NADP-dependent sorbitol-6-phosphate dehydrogenase was performed and showed that they are not identical. Therefore, this suggests that sorbitol metabolism is active in fruit pulp cells. However, it is also important to affirm the activity of SPDH in the fruit cells to validate this suggestion.

R1 fragment (spot 36.0) proteins are involved in starch metabolism and more specifically, regulating the degree of phosphorylation of starch (Lorbeth *et al*., 1998). Semicrystalline starch is hydrophobic, and may be difficult for hydrolytic enzymes to act on intact starch granules, thus it can be assumed that phosphorylation changes the surface charge distribution and increases the hydrophilicity and hydration of starch (Yu *et al*., 2001). Thus, the R1 protein reversibly binds to assimilatory starch granules and this reversible targeting is due to a direct protein-carbohydrate interaction without other proteins being acquired (Ritte *et al*., 2000). To support this, an antisense repression of the R1 transcript in potato led to a strong reduction of the amount of phosphate that was covalently bound to a small portion of glucose monomers within amylopectin (Yu *et al*., 2001). The binding of R1 to transitory starch granules during net degradation as described by Ritte *et al*. (2000) strongly suggests that this process is an essential event

within starch breakdown. R1 protein has been detected in higher plants including potato tuber and leaves (Lorbeth *et al*., 1998) and banana fruit (Ritte *et al*., 2000), even though the level of R1 did not correlate with the content of starch derived phosphate. This study, however, reports the detection of R1 protein in apple fruit pulp, which may be implicated in starch metabolism since the apple fruit has up to 10.5% carbohydrate content.

6.5.4.1.3 Photosynthesis

Six proteins involved in photosynthesis were identified, five as ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO; EC 4.1.1.29) large subunit (spots 2, 9, 44, 143 and 150), and one as ferredoxin–NADP(H) oxidoreductase (EC 1.18.1.2; spot 20). RuBisCO, the most abundant protein on Earth, is formed of eight identical large subunits coded by the *rbcL* gene and synthesized in chloroplasts and eight identical small subunits coded by the *rbcS* gene in the nucleus (Kellogg and Juliano, 1997). RuBisCO is the enzyme responsible for the conversion of inorganic carbon, as CO_2, into organic compounds like 3-phosphoglycerate (Kellogg and Juliano, 1997; Griffiths, 2006). While photosynthesis mainly occur in leaves, the process can also occur in green unripe fruits but at rates too low to cover the entire fruit energy requirement (Wheelwright and Logan, 2004). In terms of comparison between the observed and predicted mass, spots 44 and 143 resolved at lower mass than predicted (Figures 6.4; Table 6.2), suggesting that the RuBisCO subunit was degraded either in the cells by proteolysis or during the process of protein extraction. The degradation of the enzyme by proteolysis has already been reported in chloroplasts, which contain various proteases (Kokubun *et al*., 2002).

Ferredoxin-NADP(H) oxidoreductase (EC 1.18.1.2) is a subunit of the chloroplasts cytochrome b_6f complex, with a molecular weight of ~35 kDa (Zhang *et al.*, 2001b). The enzyme participates in a wide variety of photosynthetic reactions, including the reduction of $NADP^+$, cyclic photo-phosphorylation, CO_2 assimilation and light regulation. Ferredoxin-NADP(H) oxidoreductase has been detected in non-photosynthetic tomato fruit tissues (Green *et al.*, 1991). The enzyme is suggested to support ferredoxin-dependent biosynthetic processes, like nitrogen assimilation, that occur in proplastids and amyloplasts of leaves, roots and the red and green pericarp of tomato (Knaff and Hirasawa, 1991).

6.5.4.1.4 Lipid metabolism

Two proteins involved in lipid metabolism were identified in the apple fruit pulp. Spot 104 was detected as phospholipase D alpha 1 (E.C. 3.1.4.4), an enzyme that hydrolyzes membrane lipids to generate phosphatidic acid and a free-head group, during the biosynthesis of fatty acid precursors (Jandus *et al.*, 1997; Wang, 1999; Pinheroa *et al.*, 2003; Wang, 2005). Phospholipase D alpha also act in diverse plant processes including retailoring of membrane phospholipids (Whitaker *et al.*, 2001), programmed cell death, root growth, freezing tolerance and other stress responses (Wang, 2005). The enzyme also plays a role as regulator of reversible protein phosphorylation, a role achieved through regulating the activity of protein kinases and phosphatases (Wang *et al.*, 2008a; Wang *et al.*, 2008c). Expression analyses revealed that the levels of the gene coding for phospholipase D alpha in tomato fruits increased with fruit development and ripening while moderate levels were detected in leaves and low levels in roots and stems (Whitaker *et al.*, 2001). Therefore, detection of phospholipase D alpha in the mature

apple fruit pulp reported in this study correlates positively with its detection in tomato fruit during development and ripening.

Spot 68 was identified as lipoxygenases (LOX; linoleate:oxygen reductase, E.C. 1.13.11.12), nonheme iron-containing enzymes that catalyze the addition of molecular oxygen at either the C-9 or C-13 residue of fatty acids with a 1,4-pentadiene structure (Veronico *et al.,* 2006). Linoleic and linolenic acids are the most abundant fatty acids in the lipid fraction of plant membranes and are the major substrates for LOX (Veronico *et al*., 2006). Royo *et al.* (1999) observed that lipoxygenase plays a role as growth regulator, and thus an important physiological function during plant growth and development. In addition, the enzyme has been shown as implicated in plant defense against pathogen infection and wound stress in potato by acting as an antimicrobial compound and plant-signaling molecule (Royo *et al.*, 1999). Therefore, in this study, lipoxygenase may play a role in fruit development and/or maturation or in defense against pathogen infection.

6.5.4.2 Ethylene biosynthesis related proteins

Four protein spots involved in ethylene biosynthesis were identified in mature apple fruit pulp, namely S-adenosyl-L-methionine (SAM) synthetase (EC 2.5.1.6; spot 21), 1-aminocyclopropane-1-carboxylate oxidase (ACO, EC 1.14.17.4; spots 41 and 211), ethylene receptor (spot 141) and ethylene response sensor protein (spot 197).

On the contrary to the other plant hormones, ethylene is a gaseous hormone (López-Gómez *et al.*, 2008). It is released in climacteric fruits like apples, bananas, avocados and papaya, as they approach maturity to induce their ripening (López-Gómez *et al.*,

2008). In higher plants, ethylene is produced from L-methionine. ATP activates methionine to form SAM, a major methyl group donor involved in the transmethylation of proteins, nucleic acids, polysaccharides and fatty acids through the catalytic activity of SAM synthetase (Peleman *et al.*, 1989; Watkins *et al.*, 2000). In addition, SAM acts as a precursor in the biosynthesis of the hormone ethylene (Yang and Hoffman, 1984), while the decarboxylated form of SAM serves as a propylamine group donor in polyamine synthesis (Whittaker *et al.*, 1995).

Following synthesis of SAM, ethylene is produced in a two-step process. The first step consists in the production of the non-protein amino acid ACC by the enzyme ACC synthase (ACS) and with pyridoxal phosphate acting as a co-factor (Capitani *et al.*, 1999). Formation of ACC is the rate-limiting step in ethylene biosynthesis (Bleecker and Kende, 2000). Signals such as auxin or wounding can induce ethylene synthesis through increasing ACS gene expression (Iglesias-Fernandez and Matilla, 2009).

The second step involved the production of ethylene from ACC, which is catalyzed by ACO. This reaction is oxygen-dependent (Bleecker and Kende, 2000; Catusse *et al.*, 2008). The enzyme ACO is constitutively expressed in most vegetative plant tissues like leaves.

Once produced, ethylene binds to the N-terminal domain of an ethylene receptor protein, a membrane-localized protein. The intracellular portion of the protein possesses a kinase domain, which is activated upon binding of ethylene. Such receptor protein kinases are termed two-component-systems because they are composed of a sensor (ethylene binding site) and a response regulator. Transduction of the ethylene signal is

assumed to occur through a series of phosphorylation events, carried out by a cascade of protein kinases similar to the mitogen activated protein kinase pathway and probably through other poorly understood steps. Finally, a transcription factor is activated, inducing early gene transcriptional activators, which in turn induces late genes. These late-induced genes may encode enzymes, which degrade the middle lamellae or the cell wall during fruit ripening and abscission (Bleecker and Kende, 2000).

The enzymes involved in ethylene biosynthesis together with the ethylene response sensor protein are implicated in fruit ripening by inducing ethylene production during the onset of ripening (Hayama *et al*., 2006). Ethylene in turn induces biochemical and physiological changes in fruits like texture (firmness) or colour development by inducing transcriptional activation (Haji *et al*., 2005; Hayama *et al*., 2006; Catusse *et al.*, 2008). Therefore, detection of ethylene biosynthesis related proteins is an indication that the fruit was undergoing biochemical and physiological changes associated with fruit maturation and ripening. This is in accordance with the state of the fruits (mature) during the time of harvesting.

6.5.4.3 Defense/detoxifying enzymes

Several proteins implicated in cell detoxification were identified in mature apple fruit pulp. The enzymes alcohol dehydrogenase (ADH, EC 1.1.1.1; spot 2.0) and peroxidase 32 (EC 1.11.1.7; spot 11.0) were detected from samples extracted using the TCA/acetone precipitation method. The protein dehydroascorbate reductase (DHAR, EC 1.8.5.1; spot 123), putative quinone oxidoreductase (EC 1.6.5.5; spot 33), inosine 5'-monophosphate (IMP) dehydrogenase (EC 1.1.1.205; spot 38), glutathione-disulphide reductase (EC 1.8.1.7; spot 72), superoxide dismutase (SOD) 2 (EC

1.15.1.1; spot 98), NADH-plastinoquinone oxidoreductase subunit 1 (EC 1.4.3.6; spot 105), powdery mildew resistance locus-containing protein (MLA7) (spot 91), two isoforms of major allergen (Mal) d 1 (1.03G and AP15; spots 80 and 94 respectively), ribonuclease-like PR-10a (spot 103), nucleotide binding site (NBS)-containing resistance-like protein (spot 193), NBS-leucine-rich region (LRR) disease resistance protein homologue (spot 216) and glutathione peroxidase (EC 1.11.1.9; spot 215) were detected from samples extracted using the phenol precipitation method.

The enzyme ADH facilitates the interconversion between alcohols and aldehydes or ketones with the reduction of NAD^+ to NADH. The enzyme has been shown to play significant roles in plant detoxification through the oxidation of alcohols including primary, secondary, cyclic secondary or hemi-acetal to aldehydes or ketones (Salentij *et al*., 2003). In addition, ADH has been shown to be up regulated during ripening in both apple and tomato, suggesting that the enzymes play a crucial role in the formation of colour and flavor compounds during fruit ripening (Janssen *et al.,* 2008). ADH together with phenylalamine ammonia lyase (PAL) have been shown to increase fruit firmness during ripening by lignifying fruit tissues (Chong *et al.*, 2006).

Peroxidase 32 is widely distributed in the plant tissues and plays a role at the end of the lignin-forming process and in the protection of tissues damaged by, or infected with, pathogenic microorganisms (Civello *et al.*, 1995). The enzyme is generally detected in its glycosylated form and associated with membranes, although soluble isoenzymes were observed in banana (Toraskar and Modi, 1984), peach (Tijskensa *et al.*, 1997), strawberry (Civello *et al.*, 1995; Martínez *et al.*, 2001) and apple fruits (Fernández-Trujillo *et al.*, 2003; Valderrama and Clemente, 2004). A study from Valderrama and

Clemente (2004) using 'Gala' and 'Fuji' cultivars revealed that the enzyme was implicated in the quick darkening of freshly cut apple slices, through the oxidation of peroxide in the presence of molecular oxygen.

Dehydroascorbate reductase (DHAR) is an enzyme involved in the reduction of dehydroascorbate to ascorbic acid in a reaction requiring glutathione. Ascorbic acid is a major antioxidant in plants and is involved in ROS detoxification (Asada and Takahashi, 1987). DHAR plays a critical role in regenerating ascorbic acid from its oxidised state and regulating the cellular ascorbic acid redox state, which in turn affects cell responsiveness and tolerance to environmental reactive oxygen species (ROS) like superoxide, singlet oxygen, ozone and hydrogen peroxide (Asada and Takahashi, 1987). Using tobacco, a study by Chen and Gallie (2006) revealed that the effect of DHAR expression on leaf aging inversely correlated with the level of lipid peroxidation, suggesting that the enzyme contributes to plant growth by maintaining photosynthetic functioning through efficient ascorbic acid recycling that in turn limits ROS-mediated damage and slows leaf aging and probably fruit aging.

The MLA proteins belong to the coiled-coil, nucleotide-binding site, leucine-rich repeat (CC-NBS-LRR) class of genes implicated in specific recognition between host and pathogen (Zhou *et al.*, 2001; Jones, 2001). The CC-NBS-LRR class is the most prevalent among cloned plant resistance genes and is known to function in the recognition of bacterial, fungal, viral and nematode pathogens (Baker *et al.*, 1997; van der Biezen and Jones, 1998). In barley (*Hordeum vulgare*), approximately 30 alleles of the barley *Mla* locus specify resistance to the causal agent of powdery mildew disease (Wei *et al.*, 2002) and *Mla1, Mla6* and *Mla12* trigger a rapid and absolute resistance,

while *Mla7* and *Mla12* trigger a delayed and intermediate response (Wise and Ellingboe, 1983; Jorgensen, 1994; Wei *et al.*, 1999; Shen *et al.*, 2003; Halterman and Wise, 2004). In delayed resistance, termination of fungal growth occurs after the formation of haustorium and secondary hypha, leading to the death of the infected as well as surrounding cells (Boyd *et al.*, 1995; Freialdenhoven *et al.*, 1994; Kruger *et al.*, 2003).

Leucine-rich repeat (LRR) and nucleotide binding site (NBS) domains, detected in most resistance (R) genes, play a primary role in the detection of pathogens and the initiation of specific plant defenses. Although they shared high similarity among them, the NBS-LRR resistance proteins differ primarily at the N-terminus and can either have a Toll/interleukin-1 receptor (TIR) or a coiled-coil (CC) domain (Burch-Smith and Dinesh-Kumar, 2007). They are thus categorized according to their N-terminal either in the TIR-NBS-LRRs or the CC-NBS-LRRs class. They confer resistance to a wide range of organisms like, bacteria, fungi, viruses or nematodes (Whitkus *et al.*, 1992; Lawrence *et al.*, 1995; Van Der Biezen and Jones, 1998; van der Vossen *et al.*, 2000). The activation of these plant R-proteins following microbial avirulence gene product invasion ultimately results in localized cell death (Heath, 2000; Dangl and Jones, 2001).

The putative TIR-NBS type R protein 4 contains both TIR and NBS domains (Swiderski *et al*., 2009). The TIR-NBS protein family, along with TIR-X, represent the two major classes of disease resistance proteins in plants. The prevalence of TIR-encoding sequences in diverse plant genomes suggest that this is an important protein motif. However, these protein families do not posses LRR domains, unlike most R-genes (Gessler *et al*., 2006).

Major allergen (Mal) d1, a non-specific lipid transfer protein, is a homolog of Bet v 1, the major birch pollen allergen and causes a class-II food allergy (Breiteneder and Ebner, 2000). Class-II food allergy is initiated by protein contact with the immune system through inhalation of plant products like pollen from several tree species and grasses (Gao *et al*., 2005c). Mal d1 possess pathogenesis-related properties like PR-10a, another pathogen-related protein and is induced in response to biotic stimulus to act as a defense responsive protein (Sancho *et al*., 2006). Mal d1 is mainly expressed in apple fruit skin (Sancho *et al.*, 2006), but low levels of the protein are also detected in apple fruit pulp (Hovmalm *et al.*, 2006). In accordance with Hovmalm *et al.* (2006), the results in this study confirmed the presence of Mal d 1 in apple fruit pulp. The MS results also revealed that at least two isoforms of Mal d 1 are present in apple fruit pulp, Mal d 1.02 and Mal d 1.03, with Mal d 1.02 being more abundant than Mal d 1.03. These findings are in correlation with analyses of transcript expressions of the Mal d 1 isoforms that revealed Mal d 1.02 being 10 times more expressed than Mal d 1.01 and Mal d 1.03 (Marzban *et al.*, 2005). However, the Mal d 1.01 form was not identified in the present study, although its transcript has been detected in apple fruit pulp previously.

Quinone oxidoreductase and NADH-plastinoquinone oxidoreductase are enzymes induced during ripening that are negatively regulated by auxin. It catalyses the formation of 4-hydroxy-2, 5-dimethyl-3(2H)-furanone (HDMF), a key flavor compound discovered in strawberry, through the hydrogenation of the α, β-unsaturated carbonyl compounds and the highly reactive precursor 4-hydroxy-5-methyl-2-methylene-3(2H)-furanone (HMMF; Raab *et al.*, 2006). The detection of HMMF has also been reported in tomato and pineapple (*Ananas comosus*) fruits by using HPLC-

ESI-MS. This suggested that a similar HDMF biosynthetic pathway occurs in all plant species producing fruits (Raab *et al.*, 2006; Klein *et al.*, 2007). In addition, antioxidant activities, whose activity has been proposed to lower the risk of cancer and other diseases in human gut, were observed in vegetables and fruits as a result of quinone-oxidoreductase activity (Janzowski *et al.*, 2004). A study from Yoder (2001) revealed that the highly active quinone-oxidoreductase acts on xenognostic molecules, like quinones for the defense of plants against parasites or for chemical signalling among plants. These xenognostic molecules are a subset of allelopathic molecules released by a plant, which can modify the growth and development of a second plant.

Superoxide dismutase plays a role in protecting plants against oxidative stresses. Gupta *et al*. (1993) reported that the increase in abundance of superoxide dismutase and ascorbate peroxidase, an active peroxide-scavenging enzyme, induced a greater oxidative stress protection (Gupta *et al.*, 1993). An increase in superoxide dismutase activity also correlated with the 'senescent breakdown', a physiological disorder observed in apple fruits (Du and Bramlage, 1995).

6.5.4.4 Proton transporting ATPases

Eight protein spots were identified as H^+-transporting ATP synthases (H^+-ATPases), which are required for the maintenance of ion homeostasis in cells (Zhu, 2003). These protein spots include three H^+-transporting two-sector ATPase (EC 3.6.3.14; spots 8, 17 and 18), two F1-ATP synthase (EC 3.6.1.34; spots 16 and 177), one putative oligomycin sensitivity conferring protein (spot 47), one ATP-binding cassette (ABC) transporter, transmembrane region type 1 (spot 226) and one iron binding protein (spot 77).

The H^+-ATPase enzymes are a class of enzymes that catalyze the synthesis of ATP through the phosphorylation of ADP, releasing energy. The enzymes then harness this released energy to establish an electrochemical potential gradient for H^+ across the tonoplast. The gradient is utilized to drive H^+/Ca^+, H^+/Na^+, H^+/sucrose antiports present in the tonoplast (Yokoi *et al.*, 2002). The H^+-ATPases are integral membrane proteins mainly located in the tonoplast of plant cells. The enzyme is composed of the F0 (membrane associated segment) and F1 (soluble segment) domains (Penefsky and Cross, 1991; Saviani *et al.*, 1998). Both the F0 and F1 domains are composed of α- and β-subunits, but with different conformations. The F0 domain is anchored in the membrane on the C-terminal of the oligomycin sensitive conferring protein (essential for the reconstitution of oligomycin-sensitive H^+-ATPase) and is responsible for proton translocation. On the other hand, the F_1 domain, composed of five subunits with a stoichiometry of $\alpha_3\beta_3\gamma\delta\varepsilon$, contains the catalytic centre responsible for ATP synthesis that is detected at the N-terminal of the ATPase complex. The synthesis of ATP occurs at the β-subunit, but the α-subunit is essential for β-subunit activity (Saviani *et al.*, 1998). The detection of the anion-sensitive H^+-transporting two-sector ATPase, α- and β-subunits in the fruit pulp may suggest its implicated function in transportation activities associated with fruit development and maturation.

The ABC transporter transmembrane region type 1 protein belongs to the ABC superfamily, which uses the hydrolysis of ATP to provide energy to diverse biological systems. The ABC superfamily comprises of transmembrane proteins that function in the transport of a wide variety of substrates across extra- and intra- cellular membranes, including metabolic products, lipids and sterols. They are involved in the export of a wide variety of substrates ranging from small ions to macromolecules. The export

systems are also involved in the extrusion of noxious substances from the intracellular to the extracellular environment. The ABC transporter gene is also implicated in seed size (Orsi and Tanksley, 2009), thereby differentiating the domesticated apple from the wild type. Therefore, the identification of the ABC transporter transmembrane region type 1 protein confirms that the apple population used in this study is not a wild type.

Iron binding protein plays an important role in sequestering intracellular iron ion involved in the generation of reactive hydroxyl radicals through a Fenton reaction (Deak *et al*., 1999). During the Fenton reaction, ferritin is formed, which is involved in the protection of plant cells from oxidative damage induced by a wide range of stresses. The iron binding protein is also involved in cellular ion homeostasis and positive regulation of cell proliferation thus promoting positive regulation of growth rate (Deak *et al.*, 1999; Hell and Stephen, 2003).

6.5.4.5 Stress responsive proteins

Five stress responsive proteins were identified in the apple fruit pulp, four were heat shock proteins (Hsp, spots 154 & 159, 189, and 191) and one was dehydrin (spot 195). The family of Hsp proteins are expressed in every part of the plant and generally up regulated in response to stress exposure. These proteins are involved in protein folding by binding to the polypeptide chains of other proteins. In addition, Hsp proteins act as molecular chaperones since their binding to other proteins protects them from denaturing and degrading. It is suggested that Hsp proteins may play a role in environmental stress tolerance, like resistance to chilling or facilitating intracellular movement of vital cellular enzymes across organelle membranes (Ndimba *et al.*, 2005). Besides, they have been shown as involved in photosynthesis and photosystem II, by

acting on the water-oxidising quinone-reducing complex. Therefore, the Hsp may play a role in thermotolerance and promotion of fruit colour changes during maturation as suggested by Neta-Sharira *et al.* (2005).

Dehydrins are a family of plant proteins induced in response to environmental stresses that cause cellular dehydration, like water stress, cold acclimation, salinity or occurring during embryogenesis or the late stages of fruit development (Caruso *et al.*, 2002; Brini *et al.*, 2006; Kosova *et al.*, 2007). Some dehydrins also accumulate in response to the plant hormone abscisic acid (ABA), whose levels have been shown to increase in response to osmotic stress (Chandler and Robertson, 1994). In apple and pear, dehydrin abundance has been detected as fluctuating according to the seasons, reaching maximum accumulation in winter until early spring. This higher level correlated with an increase in plant hardiness (Wisniewski *et al.*, 1999). Therefore, the dehydrins detected in fruit pulp may play a role in maintaining fruit firmness.

6.5.4.6 Transcription and translation related proteins

Eleven proteins were identified as involved in transcription and translation (Table 6.2). Two protein spots were identified as TATA box binding protein (TBP)-associated factors (TAFs) 15b fragments (spots 12.0 and 14.0). These proteins bind to nucleotides and act as co-activators and core promoter recognition factors within the transcriptional factor II D (TFIID) complex. As a component of the polymerase II pre-initiation complex, these proteins play an important role in the initiation of transcription by RNA polymerase II (Albright and Tjian, 2000). In this study, the TAFs were detected by 2D-PAGE coupled with MS analysis resolved with samples extracted by TCA/acetone precipitation only (Figure 6.4a).

Spots 63 and 83 were identified as RNA-binding proteins, which are typically detected in both the cytoplasm and the nucleus. These proteins associate with double stranded or single stranded RNAs through RNA recognition motif and regulate post-transcriptional events, like RNA splicing and translation of RNA.

Spots 89, 173 and 212 were detected as myeloblastosis (MYB) transcriptional factors, a superfamily of transcriptional factors. These MYB transcriptional factors act by controlling gene expression through binding to DNA and are involved in diverse pathways including secondary metabolism like the anthocyanin pathway, signal transduction pathway and disease resistance (Yanhui *et al.*, 2006). They also play a regulatory role during developmental processes and plant defense responses to abiotic stress. In apple fruit, Allana *et al*. (2008) reported that a distinct class of MYB transcriptional factors, MYB transcriptional factor R2R3 was involved in controlling the biosynthesis of anthocyanins, the pigments that confers the red, purple, or blue colouration. Since the sequences coding for the *MYB* genes have been shown to be virtually identical among apple varieties, it is suggested that it is rather the differential expression of these genes that induce colour patterning (Allana *et al.*, 2008). This same study reported the identification of two *MYB* genes controlling the red colouration of fruit skin and flesh. These genes have been mapped as candidate markers for marker-assisted selection (Takos *et al.*, 2006; Ban *et al.*, 2007; Espley *et al.*, 2009). In addition, MYB transcriptional factors are the primary determinant influencing fruit colour variations among cultivated perennial plant varieties as well as annual species like potato, tomato or pepper (Chiou and Yeh, 2008).

DEAD/DEAH box helicase (EC 3.6.1.-; spot 13.0) is an enzyme involved in various

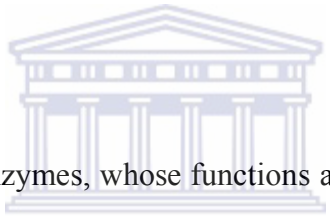
aspects of RNA metabolism, including nuclear transcription, pre-mRNA splicing, ribosome biogenesis, nucleocytoplasmic transport, translation, RNA decay and organellar gene expression (Cordina *et al.*, 2006).

Spot 62 was identified as maturase K (*matK*). Maturases are splicing factors for the plant group II introns from premature RNAs. While they generally contain three domains, the *matK* gene contains only fractions of the reverse-transcriptase (RT) domain, and no evidence of the zinc-finger-like domain (Mohr *et al.*, 1993). However, *matK* displays the domain X (the proposed maturase functional domain) and is believed the only chloroplast gene to contain it (Neuhaus and Link, 1987). *MatK* is suggested to function in the chloroplast as a post-transcription splicing factor (Ems *et al.*, 1995; Jenkins *et al.*, 1997; Vogel *et al.*, 1999; Barthet and Hilu, 2007). To date, only three studies have presented evidence for the existence of a *matK* protein in plants (du Jardin *et al.*, 1994; Liere and Link, 1995; Vogel *et al.*, 1999). These studies also indicated that *matK* can be truncated in some plant species. However, this thesis reports the first detection of *matK* in fruit pulp.

Spot 10 and spot 163 were identified as 38 kDa ribosome associated protein and small ribosomal protein 4, respectively. During translation, these two proteins regulate proteins from the plastid ribosome leading to an efficient initiation of translation (Manuell *et al.*, 2005).

Putative glycine-rich RNA-binding protein (spot 96) belongs to the RNA binding protein family. The protein consists of glycine rich C-terminal in addition to the consensus sequence type RNA binding domain (Naqvi *et al.*, 1998). It plays a role in

nucleic acid binding and has been implicated in the regulation of alternative splicing (Kim *et al.,* 2007). In plants, the expression of RNA binding proteins is regulated by both internal signals, like the circadian clock and the abscisic acid levels, and external signals, like wounding. The putative glycine-rich RNA-binding protein can be induced in response to dehydration stress or mercuric chloride treatment. In maize (*Zea mays*), it has been shown to interact with several proteins as well as RNA molecules at specific moments (Naqvi *et al*., 1998). In addition, a study using ryegrass (*Lolium sp.*) revealed the role of glycine rich RNA binding protein in stress tolerance (Shinozuka *et al*., 2006). The precise role of putative glycine-rich RNA-binding protein in the fruit pulp is however not clear but it may be involved in the regulation of alternative splicing in the pulp cells.



The detection of these enzymes, whose functions are associated with gene transcription (mRNA synthesis) and mRNA translation (protein synthesis) reveals that cells from the apple fruit pulp were actively involved in synthesizing proteins involved in a wide range of metabolic activities.

6.5.4.7 Cell growth/division: DNA synthesis/replication related proteins

Seven proteins were identified from the mature apple fruit pulp as associated with cell growth/division. These include two spots identified as cell division inhibitor MinD (spots 71 and 113), phragmoplastin 12 (spot 127), embryogenic potential marker Dc3 (spot 133), transposase (EC 2.6.7.-; spot 176), NAD ADP-ribosyltransferase (EC 2.4.2.31; spot 228) and synptobrevin-related protein (spot 130).

The cell division inhibitor MinD protein is essential for controlling cell division, a critical step necessary to avoid uncontrolled cell growth (Maple and Moller, 2005). A study by Cordell *et al*. (2001) revealed that MinD, a peripheral membrane ATPase belonging to the ATPase family, binds to MinC, a cell division inhibitor protein, in order to inhibit cell division.

Phragmoplastin 12 is involved in the early membrane fusion that occur during cell wall formation. This step involves the fusion of Golgi apparatus-derived vesicles in the center of the phragmoplast during cytokinesis in plant cells. In addition, it has been shown in soybean (*Glycine max*) that the overexpression of GFP-phragmoplastin acts as a dominant negative, slowing down the completion of cell plate formation, and often resulting in an oblique cell plate (Lukowitz *et al*., 1996; Gu and Verma, 1997).

Embryogenic potential marker, Dc3, is a protein synthesised and regulated during somatic embryogenesis. It has been suggested as a molecular marker for the acquisition of embryogenic potential (Dudits *et al.*, 1995). In a study from Seffens *et al*. (1990), the *Dc3* gene family from carrot (*Daucus carota*) was detected similar to genes from other plant species that are expressed in response to environmental and developmental cues. These results suggested a possible role for the embryogenic potential marker Dc3 during seed dessication and possibly water stress responses in plants (Seffens *et al.*, 1990).

Transposase is an enzyme that catalyses transposon movement from one site of the chromosome to another by a cut and paste mechanism by binding to its ends. This process does not induce change in the chromosome size. It is also involved in a

replicative transposition mechanism, which in this case, causes an increase in chromosome size equal to the length of the replicated transposition (Bundock and Hooykaas, 2005). In apples, a high copy number of the *Spring* transposable elements, in every haploid genome have been observed. As they possess terminal inverted repeat sequences, these transposons are capable of forming palindromic hairpin-like structures, causing sequence variations in duplicated regions, like the promoter regions of the *Mdomt1* and *Mdomt2* genes (Han and Korban, 2007). The *Spring* elements also play a role in the regulation of mRNA stability and the alteration of the expression of the gene coding for ACC synthase in apple fruit (Sunako, 1999; Han and Korban, 2007).

NAD ADP-ribosyltransferase also called poly (ADP-ribose) synthase, catalyses the DNA-dependent covalent attachment of ADP-ribose to nuclear proteins. Environmental and endogenous genotoxic agents, that cause DNA strands to denature, induce the activity of ADP-ribosyltransferase. Thus, the enzyme was shown to play a role in cellular processes like DNA repair, recombination, cell proliferation and death, as well as genomic stability (Bork *et al.*, 1997; Tong *et al.*, 2001).

The synaptobrevin-related protein is an intrinsic membrane protein that associates with small synaptic vesicles, and plays an essential role in the regulation of vesicle transport or membrane trafficking during cytokinesis (Edamatsu and Toyoshima, 2003). The R-helix of the enzyme begins to coil into a groove upon the vesicle docking, bringing the vesicle membrane and the target membrane into close proximity (Verma and Hong, 2008). The coiling releases enough free energy to drive vesicle fusion. In fission yeast (*Saccharomyces cerevisiae*), the heterotetrameric complexes of synaptobrevin have been identified in vesicle trafficking among endo-membrane compartments in the cell

(Verma, 2001). In addition, synaptobrevin and its related proteins constitute the small membrane proteins located at intracellular vesicles, organelle, and plasma membrane (Jahn and Südhof, 1999).

6.5.4.8 Signal transduction related proteins

Seven proteins associated with signal transduction were identified in this study (Table 6.2). These include three GTP-binding proteins (spots 78, 156 and 163), two kinases (spots 161 and 200), one putative abscisic acid (ABA)-binding protein (spot 84) and one lissencephayl type-1-like homology motif (spot 206).

The superfamily of GTP-binding proteins consists of several members, including translational factors, tubulins and signal-transducing GTP-binding proteins. GTP-binding proteins have the ability to bind and subsequently hydrolyse guanine nucleotides. They act as molecular signal transducers whose activity and inactivity status depends on the binding of GTP and GDP, respectively. The GTP-binding proteins are composed of major subfamilies, namely the small GTP-binding proteins (G-protein) family (Terryn *et al.*, 1993; Vernoud *et al.*, 2003). The small GTP-binding proteins play a critical role in the regulation of a wide range of cellular processes, like growth, differentiation, and intracellular transportation. They are also involved in defense signal-transduction pathways (Seo *et al.*, 1995).

Protein kinase (EC 2.6.11.1-EC 2.6.11.20) superfamily regroup the enzymes catalysing the reversible transfer of the γ-phosphate from ATP to amino acid side chains of proteins, a process called phosphorylation, inducing changes in protein activity, folding, or interaction with other proteins (Stone *et al.*, 1998). Protein phosphorylation has been

implicated in response to various signals like light, pathogen invasion, hormones, temperature stress, or nutrient deprivation (Stone *et al.*, 1998; Hardie, 1999). S-receptor kinase, functioning with co-receptors lacking kinase domains, is involved in the signalling of diverse arrays for plant development and defense (Becraft, 2002). In addition, this kinase also plays a role in self-incompatibility and hormone perception signalling (Jakubowska *et al.*, 2007). On the other hand, phytochrome kinase are involved in the modulation of light signalling by sensing red and far-red light. This process is necessary for plants to monitor optimal growth and development (Sineshchekov, 2005).

Putative abscisic acid (ABA)-binding proteins are ABA-receptors mediating ABA signals during fruit development (Zhang *et al.*, 2001a). The activity of these proteins has already been detected in the cytosol of developing apple fruit flesh (Zhang *et al.*, 2001a).

Lissencephaly type-1-like homology motif has been proposed to contribute to the regulation of microtubule dynamics, either by signaling/mediating dimerisation, or by binding directly to the cytoplasmic dynein heavy chain (Emes and Ponting, 2001; Kim *et al.*, 2004).

6.5.4.9 Amino acid and purine biosynthesis related proteins

Three proteins whose functions are related to amino acid and purine biosynthesis were identified in apple fruit pulp. These include glutathione synthetase (EC 6.3.2.3; spot 36), β-cyanoalanine synthase 1 (EC 4.4.1.9; spot 39) and aspartate transaminase (AAT) 5 precursor (EC 2.6.1.57; spot 138).

Glutathione synthetase is an enzyme catalyzing the ATP dependent biosynthesis of glutathione from γ-glutamylcysteine and glycine. The reduced form of glutathione synthetase is maintained by the NADP-dependent glutathione reductase. In the sulphur assimilation pathway, glutathione synthetase acts as a reductant for the reduction of 5′-adenosylsulphate (APS) to sulphite. It is also the major carrier form of reduced sulphur in plants (Gomez *et al.*, 2004; Thomas, 2008).

The enzyme β-cyanoalanine synthase 1, the mitochondrial cysteine synthase, was identified. It plays a crucial role in cyanide metabolism by catalyzing the slow isotopic α-H exchange in cysteine, an end-product amino acid (Yamaguchi *et al.*, 2000; Maruyama *et al.*, 2001).

The enzyme AAT 5 precursor, from the aspartate biosynthetic pathway, is one of the most important enzymes of the amino acid metabolism in living organisms. It catalyses the transfer of an amine group from L-aspartic acid to α-oxoglutarate to form oxaloacetate and glutamate in the presence of the coenzyme pyridoxal-5-phosphate (Wadsworth *et al.*, 1995; Murooka *et al.*, 2002; Wadsworth, 2006).

6.5.4.10 Cytoskeleton related proteins

Two spots were identified as myosin class II heavy chains (spots 4 and 86), six spots as actin (spots 19, 141, 177, 204, 208 and 209), two spots as dynein heavy chain isoform (DHCIb; spots 144 and 203) and one spot as annexin (spot 46). Myosin and dynein had no identity when validated against the *Malus* EST database from NCBI probably because of the limited information on the EST data available for the *Malus* species. This may be rectified when the genome sequence is completed. However, myosin, actin

and dynein form the cytoskeleton motor proteins (Hirokawa and Tekamura, 2003; Mallik and Gross, 2004).

Myosin proteins are molecular/actin motors that use the chemical energy stored in ATP and convert it into a mechanical displacement of kinetic energy. They can bind to and move along actin filaments but mainly operating on the plus-end of the actin filaments. In this way, myosins drive actin-based cell movements (Reddy and Day, 2001; Jiang and Ramachandran, 2004). The highly conserved actin protein shares high similarities among plant species. This highly dynamic and fibrous protein is essential in cellular processes, like cell division, cytoplasmic streaming, cell motility and signal transduction by transmitting internal stress. The folding of actin is stimulus responsive particularly following stress (Hussey *et al.*, 2002; Ketelaar *et al.*, 2007). Dyneins together with kinesins are microtubule motors capable of a sliding movement useful to establish and maintain the structural integrity of the cell during mitosis (Smith, 2002).

In this study, myosin and dynein had no significant sequence similarities to *Malus* EST from NCBI and on the contrary, actin had 100% identity. However, the identification of myosin, actin, dynein and kinesin in the apple fruit pulp proteome suggests their structural role for the fruit integrity. Their abundance is likely to influence the shape/form and firmness of fruits.

Annexins are abundant proteins involved in cell expansion and fruit maturation (Bianco *et al.*, 2009). During fruit ripening, annexins interact with the callose (1,3-β-glucan) synthase, which possess an intrinsic nucleotide phosphodiesterase activity. Annexins also bind to F-actin and assist in the modulation of actin activities. In addition, annexins

have peroxidase activity, a crucial role for detoxification in plants (Basel, 1997; Clark *et al.*, 2001). Gene expression analyses conducted in both strawberry and pepper reported an increment of annexin during fruit development until fruit maturation (Wilkinson *et al.*, 1995; Proust *et al.*, 1996). In strawberry, it has been speculated that the abundance or distribution of an annexin-like protein might influence ion fluxes, membrane cytoskeletal attachments, or other aspects of plasmalemma function that change during fruit maturation and senescence (Bianco *et al*., 2009). In apple, like in strawberry, annexins could be involved in exocytosis of cell wall degrading enzymes, an act to sequester Ca^{2+} ions and phospholidids released from the degrading cell wall matrix, or serve as signal molecules to help regulate the ripening developmental program (Wilkinson *et al.*, 1995; Basel, 1997).

The presence of cytoskeleton related proteins in apple fruit pulp was expected as they form one of the fundamental cellular adaptive strategies for growth and development involving cell size adjustments. In addition, the identification of actin in the apple fruit pulp suggests that the protein participates in signal transduction by transmitting internal stress, probably providing mechanical strength to the cell cortex and spatial organization of the cytoplasm.

6.5.4.11 Unknown/unclassified proteins

Several proteins identified from the apple fruit pulp whose biological functions are unknown were also identified in this study (Table 6.2). In this category, six protein spots, representing approximately 6% of the positively identified protein spots (Table 6.2), were identified as hypothetical proteins. In addition, 13 (approximately 10%) and three (approximately 2%) protein spots were identified as predicted and unknown

proteins, respectively. Finally, seven protein spots, which represented approximately 6%, could not be categorised because their functions have not been clearly defined yet. These included AY007207 NID (spot 5.0), AB013353 NID (spot 13), AB004825 NID (spot 79), MRGH21 (spot 3), sialyltransferase-like protein (spot 46), ER-binding protein (spot 139) and sequence 87 from patent WO02064764. For example, the protein spot 46 was identified as a sialytransferase-like protein, which is involved in sialic acid metabolism (Gebbie *et al.*, 2005). Previously considered as absent in plant cells, Zeleny *et al.* (2006) demonstrated its presence in minute quantities, but their precise function has not been determined to date.

6.5.4.12 Functional annotation of hypothetical and predicted proteins

The functional annotation of hypothetical, predicted and unknown proteins was carried out using proteomics bioinformatics tools. In this regard, several databases were searched for the identification of sequence and domain homologs using the amino acid sequence of the best match obtained with the MASCOT search engine. Several publicly available databases were used for the detection of conserved domains, namely SIMAP (http://webclu.bio.wzw.tum.de/portal/web/simap/seqfinder), Expert protein analysis system (ExPASy; http://au.expasy.org/cgi-bin/prosite/ScanView.cgi), NCBI BLASTp (http://blast.ncbi.nlm.nih.gov/Blast.cgi), Integrated resources of protein domains and functional sites (InterProScan; http://www.ebi.ac.uk/Tools/InterProScan), the Arabidopsis information resources (TAIR; http://www.arabidopsis.org/) and Scansite (http://scansite.mit.edu/cgi-bin/motifscanseq). Common domains were detected and the sequence homologs were also blasted against the *Malus* EST databases from NCBI prior to confirmation of the identified identity as a first step towards protein functional prediction.

Functionalised proteins were categorised based on the annotated protein function. Six functional categories were determined, namely glycolysis and other carbohydrate metabolism related proteins (AB013353 NID; spot 13 and AB004825 NID; 79), defense/detoxification related protein (unknown protein; spot 32), signal transduction related proteins (hypothetical protein; spots 28 and 44), transcription and translation related proteins (predicted protein; spot 5), cell growth and division related proteins (predicted protein; spot 187), and heat shock (stress responsive protein; ER-binding protein, spot 139).

AB013353 NID (spot 13)

The putative conserved domains including casein kinase II phosphorylation site, protein kinase C phosphorylation site and N-glycosylation site were detected in the amino acid sequence of AB013353 NID, in common with super UTP-glucose-1-phosphate uridylyltransferase, also called UDP-glucose pyrophosphorylase. The results of sequence similarity search using NCBI BLASTp, SIMAP and InterProScan databases revealed that the protein shared 100% sequence identity with super UTP-glucose-1-phosphate uridylyltransferase. UTP-glucose-1-phosphate uridylyltransferase plays a role in sucrose biosynthesis in fruit cells, through the reversible conversion of UDP-glucose to glucose-1-phosphate. This enzyme has also been detected as spot 11 (section 6.5.3.1.2). As observed on the 2D-PAGE proteome map established using TSP extracted by phenol precipitation (Figure 6.4B), spots 11, 12 and 13 resolved at a similar molecular mass but varying p*I*. After assigning to spot 13 its the predicted function, the results from the protein characterization revealed that these three spots are all involved in glycolysis and other carbohydrate metabolism.

AB004825 NID (spot 79)

In terms of AB004825 NID, a putative conserved domain, eukaryotic initiation factor 5A hypusine signature, was detected from the amino acid sequence of the protein, in common with the eukaryotic translation factor 5A from potato (gi|2225883). Using the NCBI BLASTp, SIMAP and InterProScan tools, the amino acid sequence was shown to share 100% identity with eukaryotic translation factor 5A. This protein is a highly conserved GTPase activating protein in eukaryotes and is the only protein containing the post-translationally synthesised unusual amino acid hypusine (Park, 2006). It plays a critical role for the association of the large and small subunits of ribosome. A study by Wang *et al*. (2003) suggested that this eukaryotic translation factor 5A protein was involved in programmed cell death associated not only with leaf senescence but also fruit and petal senescence. Further work using *Arabidopsis* have also shown that eukaryotic translation factor 5A-2 can induce programmed cell death in response to pathogen attack by the virulent *Pseudomonas syringae* pv *tomato* DC3000 (Hopkins *et al*., 2008). Therefore, the protein AB004825 NID can be classified as a defense/detoxification related protein.

Hypothetical protein (spots 28 and 44)

The putative conserved domain, phosphoglycerate kinase signature, was recognized in the amino acid sequence of the two hypothetical proteins identified as spots 28 and 44. These domains are also detected in the ADP-specific phosphofructokinase (AT1G79550) from *Arabidopsis thaliana*. A 100% sequence identity was detected using NCBI BLASTp, SIMAP and InterProScan search engines, while 95.76% identity was obtained using TAIR. ADP-specific phosphofructokinase, a member of ribokinases, is a glycolytic enzyme, which has been well characterized in

hyperthermophilic archaea. The form occurring in archaea is highly homologous to hypothetical proteins present in several eukaryotes like *Drosophila melanogaster* (AAF49769), *Caenorhabditis elegans* (T32780) and *Homo sapiens* (AAH06112; Ito *et al*., 2001). The enzyme has a phosphotransferase activity, with its alcohol group acting as an acceptor in the carbohydrate metabolic process (Ronimus *et al*., 2001).

Predicted protein (spot 5)

Putative conserved domain, zinc finger cyscyshiscys (CCHC) type, was detected in the amino acid sequence of the predicted protein identified as spot 5, which was common with the domains observed in peptidyl-prolyl cis-trans isomerase from *Croceibacter atlanticus* (HTCC2559). A 100% sequence identity was detected between the predicted protein and peptidyl-prolyl cis-trans isomerase using NCBI BLASTp and SIMAP. Peptidyl-prolyl cis-trans isomerase, also termed rotamase, is an enzyme that accelerates protein folding by catalysing the cis-trans isomerisation of proline imidic peptide bonds in oligopeptides (Henriksson *et al*., 2004). In peas, peptidylprolyl cis-trans isomerase activity was detected in the cytosol, mitochondria, and chloroplasts.

Unknown protein (spot 32)

In common with the immediate-early fungal elicitor protein CMPG1 from maize (gblACG47370.1), a RING finger domain was detected in the amino acid sequence of the unknown protein identified as spot 32. Using NCBI BLASTp, the sequences of the two proteins were detected as sharing 87% sequence identity, but 11% gaps were also identified. The immediate-early fungal elicitor CMPG1 protein is implicated in the efficient activation of defense mechanisms triggered by multiple resistance genes in plants (Muratani and Tansey, 2003). This protein contribute in triggerring defense

mechanisms by activating either a positive regulator or transcription factors through their ubiquitination activities, which in turn activates plant defense responses (Muratani and Tansey, 2003). The ubiquitination also regulates transcription by modifying histones and, therefore, chromatin. In addition, CMPG1 has been shown to activate plant defense by directing the degradation of negative regulators (Brays *et al.,* 2005). Therefore, this unknown protein can be categorized as a defense related protein, with respect to the annotated function of its homologous protein CMPG1.

Predicted proteins (spots 152 and 192)

Using the bioinformatic tool NCBI BLASTp, ExPASy scan view and SIMAP, the putative conserved domains, FAD-linked oxidase, C- and N- terminal domains, from the amino acid sequence of hypothetical proteins were detected as common with those domains of flavin adenine dinucleotide (FAD) binding protein. The two sequences share 100% sequence identity and 0% gap. FAD binding proteins are a diverse class of proteins without universal sequence conservation that employ their tightly bound FAD moiety to transfer one or two electrons from a reduced substrate to an acceptor (Mattevi *et al*., 1997). Thus the predicted proteins identified as spot 152 and 192 can be predicted as involved in eliciting oxidase activities for plant development.

Predicted protein (spot 187)

The putative conserved domains detected in the amino acid sequence of the predicted protein identified as spot 187 were common with those of cell division cycle protein 48, also termed vasolin-containing protein. The two sequences share 91% sequence identity and 0% gap, as determined with NCBI BLASTp and SIMAP. The cell division cycle protein 48 is among the proteins involved in the regulation of cell division and growth

processes (Yamada *et al*., 2003). Cell division cycle protein 48 is primarily localized in the nucleus and, during cytokinesis, in the phragmoplast, a site where membrane vesicles are targeted in the deposition of new cell wall materials. The protein is highly expressed in proliferating cells of vegetative shoots, roots, floral inflorescences and flowers (Yamada *et al*., 2003). Based on its homology to cell division cycle protein 48, spot 187 can predicted as a cell division associated protein. As its expression is detected in growing cells, the identification of this protein in apple fruit pulp by MALDI-TOF MS suggests that fruit cells also undergo proliferation during its development.

Estrogen receptor -binding protein (spot 139)

Spot 139 was identified as an Estrogen receptor (ER)-binding protein. The amino acid sequence of this protein was detected as sharing 95% sequence identity and 0% gap with heat shock proteins from cucumber (*Cucumis sativus*), using NCBI BLASTp, InterProScan and SIMAP. ER was first detected in lipid extracts from *Solanum glaucophyllum* and tomato organs, is concentrated in the nucleus and interacts with estrogen-binding sites. Previously, ER-binding protein has been predicted to either function as a sterol carrier or a storage molecule (Milanesi and Boland, 2004). However, the bioinformatics analyses carried out in this study suggest that ER-binding protein can be classified as a heat shock protein (stress responsive protein).

CHAPTER 7

IDENTIFICATION OF THE DIFFERENTIALLY REGULATED PROTEINS BETWEEN THE HIGH AND LOW FIRMNESS APPLE FRUIT PULP

7.1 INTRODUCTION

Firmness is an important textural property in apple and one of the key parameters for the determination of fruit maturity and post harvest quality. Recently, a multispectral scattering technique was developed for the nondestructive assessment of fruit firmness and soluble solids content (Lu, 2004; Peng and Lu, 2007).

Apple fruit firmness is influenced by a number of factors. These include preharvest factors like genetics (variations with cultivar), cultural practices such as pruning, and application of fertilisers or growth regulators and postharvest factors have an enormous effect on apple firmness like maturity at harvest, postharvest dips, for example calcium dips, cooling and storage conditions (DeEll *et al*., 2001). In addition, Volz *et al*. (2004) showed that crop load, cell size, cell packing and intracellular air spaces also influence the overall firmness of a fruit. This study demonstrated that firmer fruits had smaller cortical cells; higher cell packing and more intracellular air spaces. Further, Hallet *et al*. (2005) illustrated that pectin plays a crucial role as a structural element in creating and preserving firmness in fruits. In this case, pectin present in the middle lamella acts as an adhesive agent between adjacent cells providing structure and cohesion to the fruit tissue. However, degradation of cell wall and pectin during ripening alters the chemical and structural composition of cells leading to fruit softness (Hallet *et al*., 2005).

To date, the characterization and differentiation of genetic relationships between species and cultivars is mainly based on the use of molecular markers (Wunsch and Hormaza, 2007). However, the information is still inadequate to fully characterize polygenic characters like fruit firmness among apples from the same cultivar or different cultivars, although genetic relationships is sufficient to estimate their genetic variability.

In Chapter 6, 2D-PAGE proteome reference maps were established using TSP isolated from mature apple fruit pulp of the seedlings derived from a cross between the cultivars 'Golden Delicious' and 'Dietrich'. In addition, all the CBB stained proteins were analysed by MALDI-TOF MS for their identification, validated by blasting against the *Malus* EST database from NCBI and then characterized. However, this chapter reports the identification of differentially regulated proteins between the high and low firmness apple fruit pulps harvested from the seedlings derived from a cross between the cultivars 'Golden Delicious' and 'Dietrich'.

Therefore, the aim of this chapter was to further characterise the total apple pulp proteome in order to gain insights into the molecular mechanisms that regulate and/or control fruit firmness. A comparative proteomic analysis of the high and low firmness phenotypes was therefore carried out. In this regards, 2D-PAGE proteome maps were generated using TSP from the high and low firmness phenotypes of apple fruit pulp and comparatively analysed using the PDQuest software. The differentially regulated proteins were then identified by MALDI-TOF MS, validated as described in chapter 6 and then characterised.

7.2 RESULTS

Apple fruits were harvested from the selected twenty fruiting individuals from the 'Golden Delicious' x 'Dietrich' population (section 2.3.1; Table 7.1). The individuals were selected based on the preliminary results obtained by M.K. Soeker (unpublished data) between 2005 and 2007 (section 2.3.3). The harvested fruits were stored for two days at room temperature prior to any analysis. The fruit skin was then peeled off from three regions of each fruit prior to measuring firmness using the penetrometer. The fruits were snap frozen in liquid nitrogen and then stored at -80°C prior to their use for proteomic analyses (section 2.6).

Table 7.1: Classification of the 20 individuals selected for fruit firmness measurements and used for proteomics analyses.

The individuals were classified according to their fruit firmness group and the date of fruit collection. The individual number indicates the tree row (1 or 2) in the orchard and the individual tree number (after the '-' sign).

Individual number	Fruit firmness group	Fruit collection date
1-59	High	10/03/08
2-48	High	10/03/08
2-75	High	10/03/08
2-118	High	10/03/08
2-119	High	10/03/08
2-96	High	26/03/08
1-124	High	26/03/08
2-108	High	03/04/08
2-41	High	16/04/08
2-88	High	03/04/08
2-104	Low	10/03/08
2-57	Low	26/03/08
2-136	Low	03/04/08
2-95	Low	26/03/08
2-69	Low	03/04/08
2-134	Low	03/04/08
2-124	Low	03/04/08
1-27	Low	03/04/08
1-67	Low	03/04/08
1-39	Low	03/04/08

7.2.1 COMPARATIVE ANALYSIS OF PROTEIN EXPRESSION BETWEEN THE HIGH AND LOW FIRMNESS PHENOTYPES OF APPLE FRUIT PULP AFTER SEPARATION BY 1D-PAGE

Fruits with penetrometer results ranging from 0-7 kg cm^{-2} were categorised as low firmness, while fruits ranging from 7-13 kg cm^{-2} were categorised as high firmness. Total soluble proteins were extracted from both high and low firmness mature apple fruit pulp using an optimised phenol precipitation method, as described in section 6.2 instead of the TCA/acetone precipitation. The former method was preferred following optimisation (section 6.2) as more protein spots were obtained, as observed in Figure 6.3, thus increasing the chances of detecting differentially regulated proteins. Three biological replicates for each phenotype were then separated according to their molecular weight by 1D-PAGE (Figure 7.1). Similar protein band patterns were visually observed between replicates and the two phenotypes. Proteins typically resolved in the molecular range 10 to 150 kDa, but only few proteins were visualised below the 15 kDa molecular weight of the marker.

In terms of comparison between the high and low firmness phenotypes, the abundance of some proteins was altered (Figure 7.1). The abundance of two protein bands, resolving at approximately 70 kDa and above the 20 kDa band of the molecular weight marker, increased in every replicate from the high firmness phenotype in comparison with the low firmness phenotype (Figure 7.1; panel a and b). In addition, the abundance of one protein band, visualized just above the 15 kDa band of the molecular marker, was visually detected as decreased in the high firmness phenotype in comparison with the low firmness phenotype (Figure 7.1; panel c). This band was even totally absent in some of the high firmness phenotype.

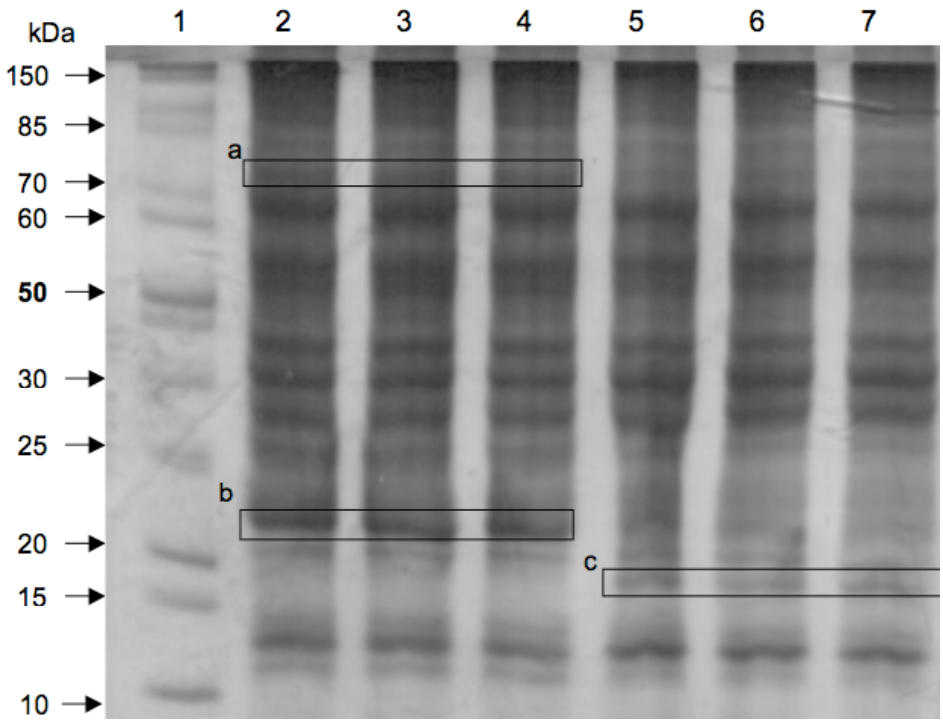
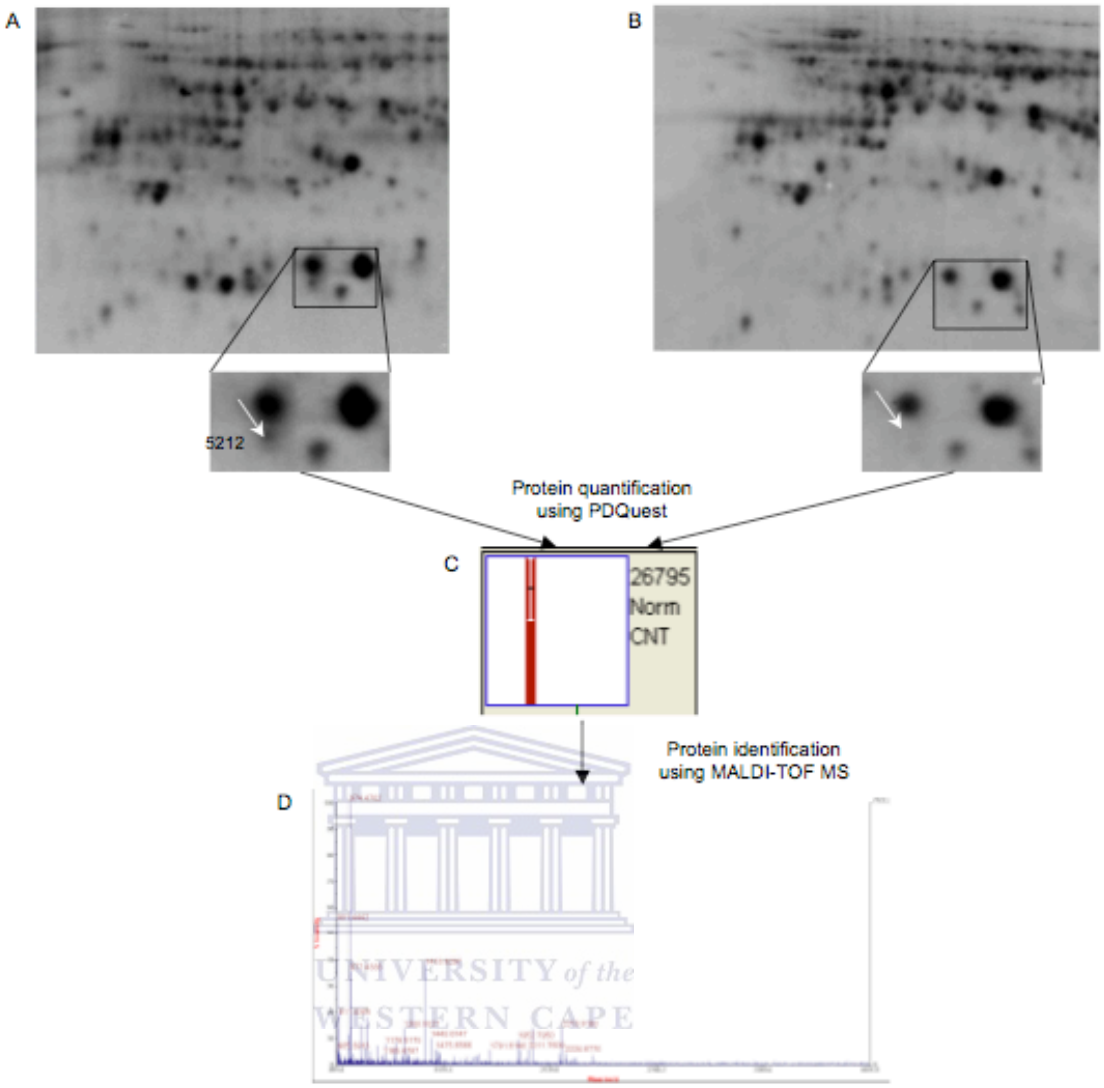


Figure 7.1: Influence of fruit firmness on the 1D-PAGE profiles of total soluble proteins from high and low firmness mature apple fruit pulp.

Total soluble proteins (20 µg) were separated by 12% SDS-PAGE and stained with Coomassie brilliant blue. Lane 1: molecular marker. Lane 2, 3 and 4: protein extracts from the high firmness phenotype of individuals 1-124, 2-96 and 2-41, respectively. Lane 5, 6 and 7: proteins from the low firmness phenotype of individuals 2-124, 1-27 and 1-67, respectively. Panels a, b and c indicate differentially expressed protein bands.

7.2.2 COMPARATIVE ANALYSIS OF PROTEIN EXPRESSION BETWEEN HIGH AND LOW FIRMNESS APPLE FRUIT PULP AFTER SEPARATION BY 2D-PAGE

In order to identify the proteins differentially expressed between the high and low firmness phenotypes (Table 7.1), 2D-PAGE proteome maps were established. Four biological replicates of total soluble proteins for each phenotype from individuals of the 'Golden Delicious' x 'Dietrich' apple population were resolved by 2D-PAGE according to their pI in the first dimension and their molecular weight in the second dimension. The replicated proteome maps were used for the detection of differentially expressed proteins between the high and low firmness phenotypes using the PDQuest software. An illustration of the comparative analysis of high and low firmness phenotyes 2D-PAGE using PDQuest and the identifcation of the differentially regulated proteins is shown on Figure 7.2. The 2D-PAGE proteome maps established were based on the developed and optimised protocol discussed in Chapter 6, using linear 7 cm IPG strips pH range 4-7 only. Using this narrower pH range, proteins resolved in the molecular range 10 to 150 kDa and pH range 4 to 7. An average of 217 and 212 protein spots were visualized after CBB staining in extracts from the high and low firmness phenotypes, respectively (Figure 7.3). However, only differentially regulated spots detected by PDQuest were labelled on the 2D-PAGE.



E

```
  1 GVYTFENEYT SEIPPPRLFK AFVLDADNLI PKIAPQAIKH AEILEGDGGP
 51 GTIKKITFGE GSQYGYVKHK IDSVDEANYS YAYTLIEGDA LTDTIEKVSY
101 ETKLVASGSG SIIKSISHYH TKGDVEIKEE HVKAGKEKAH GLFKLIESYL
151 KGHPDAYN
```

F

```
>gb|GO527722.1|  Mdwdr7511B17.g1 Apple_EST_Mdwdr Malus x domestica cDNA 5' similar
to gb|AAB01362.1| Ap15 gene product gb|AAC26136.1| major
allergen Mal d 1 [Malus x domestica] gb|AAD26545.1|AF124822_1
major allergen mal d 1 [Malus x domestica] gb|AAD26549.1|AF124826_1
major allergen mal d 1 [Malus x domestica] gb|AAD26550.1|AF124,
mRNA sequence.
Length=580

 Score =  295 bits (754),  Expect = 2e-80, Method: Compositional matrix adjust.
 Identities = 158/158 (100%), Positives = 158/158 (100%), Gaps = 0/158 (0%)
 Frame = +1

Query  1    GVYTFENEYTSEIPPPRLFKAFVLDADNLIPKIAPQAIKHAEILEGDGGPGTIKKITFGE  60
            GVYTFENEYTSEIPPPRLFKAFVLDADNLIPKIAPQAIKHAEILEGDGGPGTIKKITFGE
Sbjct  91   GVYTFENEYTSEIPPPRLFKAFVLDADNLIPKIAPQAIKHAEILEGDGGPGTIKKITFGE  270

Query  61   GSQYGYVKHKIDSVDEANYSYAYTLIEGDALTDTIEKVSYETKLVAsgsgsiiksisHYH  120
            GSQYGYVKHKIDSVDEANYSYAYTLIEGDALTDTIEKVSYETKLVASGSGSIIKSISHYH
Sbjct  271  GSQYGYVKHKIDSVDEANYSYAYTLIEGDALTDTIEKVSYETKLVASGSGSIIKSISHYH  450

Query  121  TKGDVEIKEEHVKAGKEKAHGLFKLIESYLKGHPDAYN  158
            TKGDVEIKEEHVKAGKEKAHGLFKLIESYLKGHPDAYN
Sbjct  451  TKGDVEIKEEHVKAGKEKAHGLFKLIESYLKGHPDAYN  564
```

Figure 7.2: Comparative analysis of the high (A) and low (B) apple fruit pulp firmness to and further identification of one differentially regulated protein spot using MALDI-TOF MS.

Total soluble protein samples (200 μg) were extracted from the high and low firmness apple fruit pulp harvested from seedlings of a population derived from the cross between 'Golden Delicious' and 'Dietrich' by phenol precipitation. Protein extracts were resolubilised in IEF rehydration buffer and separated by 2D-PAGE using linear 7 cm IPG strips, pH range 4-7 in the first dimension and 12% SDS-PAGE in the second dimension. Proteins were visualized by CBB. The gel images from four biological replicates were comparatively analysed using the PDQuest software. To minimize experimental variations, spot intensity was normalised. Spots were automatically detected and matched, and manually edited. The normalised quantity of each matched spots was shown by (C), the spot review bars. The red bar corresponds to the normalised intensity of four spots from four replicates in the high (A) firmness 2D-PAGE. In this example, the protein spot was only detected in the high firmness phenotype, as shown by the absence of green bar, which corresponds to the low (B) firmness phenotype. Differentially regulated spots were manually picked, trypsinised and analysed by MALDI-TOF MS for the acquisition of the PMF of the fragmented peptides. (D) represents the spectra of peptides from the digested spot. Using MASCOT to interrogate the NCBI and MSDB databases with the peptide list, the protein spot was positively identified as 'major allergen mal d 1.02' (gi|42558971), from *Malus x domestica*. The amino acid sequence of this protein is shown in (E) and matched peptides are highlighted in bold red. (F) represents the validation step when the PMF of the fragmented peptide sequences and/or the complete protein sequence is blasted against the *Malus* EST database from the NCBI to confirm the validity (percentage identity) of the protein identified using the MASCOT search engine.

.

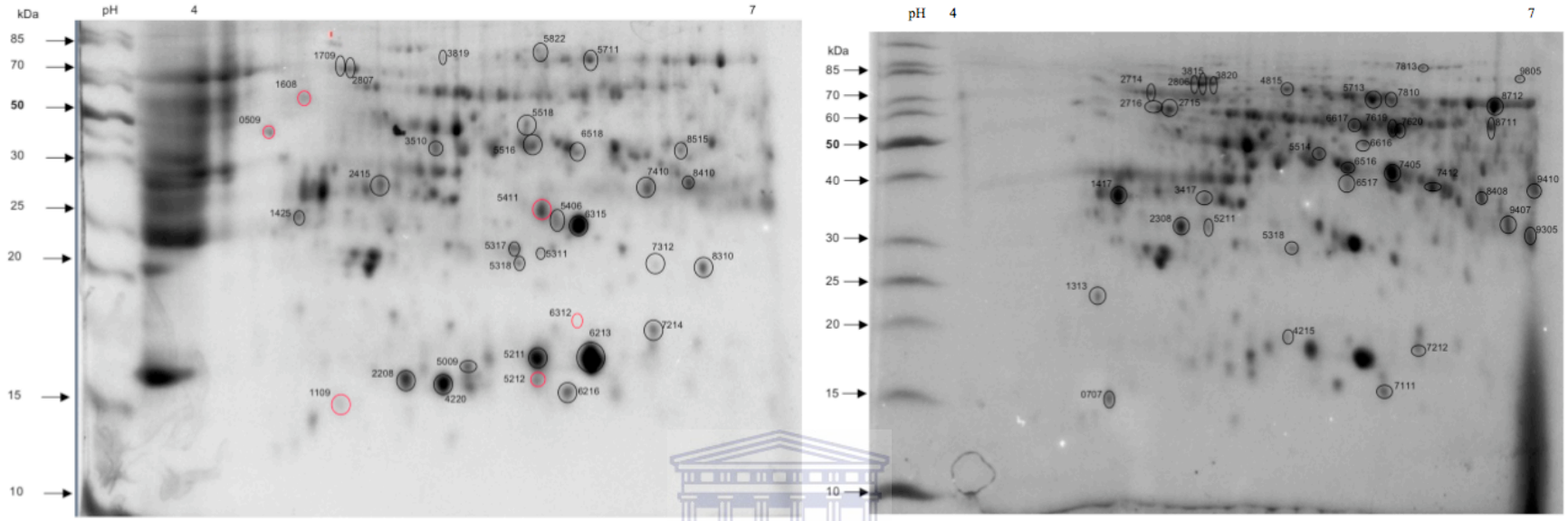


Figure 7.3: Influence of fruit firmness on the 2D-PAGE profiles of total soluble proteins from high (A) and low (B) firmness mature apple fruit pulp.

Total soluble protein samples (200 µg) extracted by phenol precipitation and resolubilised in IEF rehydration buffer were separated by 2D-PAGE using linear 7 cm IPG strips, pH range 4-7 in the first dimension and 12% SDS-PAGE in the second dimension. Proteins were visualized by CBB. Black circles indicate the differentially expressed proteins, as determined by PDQuest analysis (either significantly at $p<0.05$ or quantitatively at 2-fold minimum), between the high and low firmness phenotypes. Red circles indicate spots detected only in one of the two phenotypes. The molecular masses of the protein marker are indicated in kDa on the left of each gel. The pH of the IPG strips is indicated by (4) and (7) on the top of each gel. The spots numbers were generated by the PDQuest software during comparative analysis.

The similarity between the proteome maps was analysed using the scatter plot generated during the PDQuest analysis (Figure 7.4). The scatter plot was used to determine either the up- or down-regulation fold change for each protein spot. A regression line was obtained by plotting on a log scale the normalized abundance of each spot from the high firmness phenotype against the low firmness phenotype. The coefficient of correlation of this regression line was 0.525146.

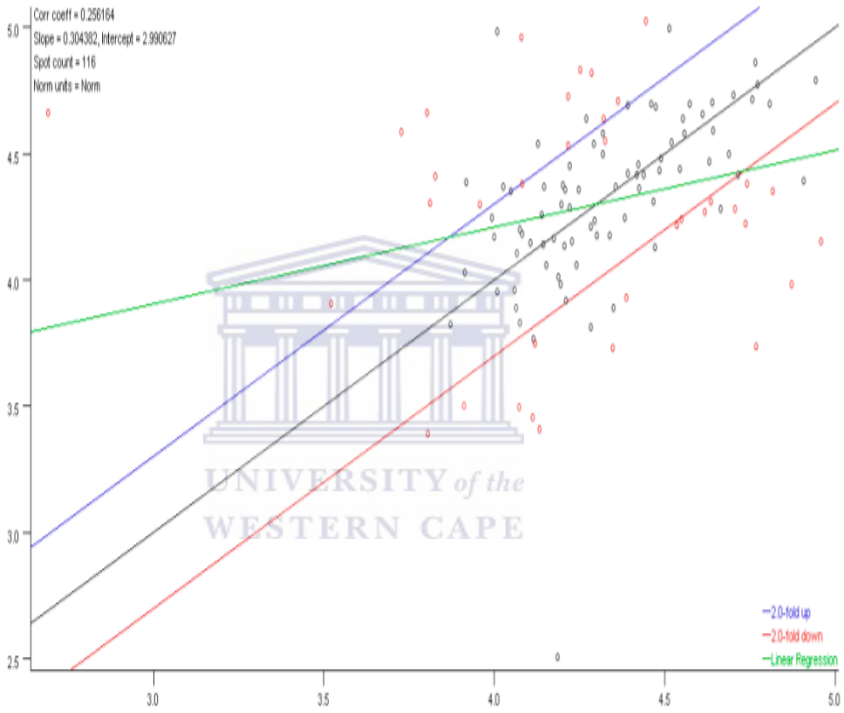


Figure 7.4: Comparison of the abundance of spots between the high and low firmness phenotypes.

The normalized abundance of each spot in the high firmness (X-axis) was plotted on a log scale against its abundance in the low firmness phenotype (Y-axis). The green line represents the regression line. Spots falling around the black centerline are of similar abundance in both phenotypes. The blue and red lines set the limit for the 2-fold up- and down-regulation, respectively, in the high firmness phenotype in comparison to the low firmness phenotype. The coefficient of correlation, slope, intercept, spot counts and normalization unit are indicated on the top left of the graph. Red circles indicate protein spots whose expression was significantly (p<0.05) different between the two phenotypes.

In terms of comparison between the high and low firmness phenotypes, a total of 54 spots were detected as differentially expressed (Figure 7.3). Using the Student *t*-test (T-test), the abundance of 21 and 15 spots was up- and down- regulated, respectively, in the high firmness phenotype in comparison to the low firmness phenotype. Using the quantitative test (Q-test), the abundance of 15 and 16 spots was detected as up- and down- regulated, respectively, by a minimum of 2.0-fold, in the high firmness phenotype. Nine protein spots were detected as up-regulated in the high firmness phenotype using both the T- and Q-test (Figure 7.5A, represented as T∩Q=9). In addition, nine protein spots were detected as down-regulated in the high firmness phenotype using both tests (Figure 7.5B, represented as T∩Q=9).

Taken as a whole both the student *t*-test and the quantitative analysis, the abundance of 49 spots was detected as differentially regulated when comparing the two phenotypes. A total of 27 and 22 spots were up- and down- regulated, respectively, in the high firmness phenotype (Table 7.1). Qualitatively, five protein spots identified as predicted protein, major allergen mal d 1, cell division inhibitor MinD, unidentified protein and lissencephaly type-1-like homology motif (Figure 6.4, spots 65, 94, 113, 146 and 206, respectively), were only visualized in extracts from the high firmness phenotype. As revealed by the T-test, the abundance of the unidentified spot 56 (SSP 9405) was detected as being the least up-regulated spot, with 1.5-fold. The abundance of myosin class II heavy chain (spot 86/SSP 2208) was detected as being the spot with the maximum up regulation of 11.0 fold. The identified and unidentified spots, which were differentially regulated as well as their variation in abundance, expressed in fold, between the high and low firmness phenotypes are presented in Table 7.2.

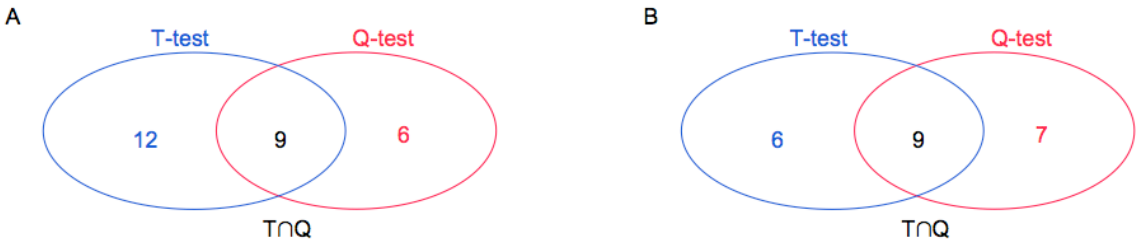


Figure 7.5: Number of differentially expressed protein spots detected between the high and low firmness apple fruit pulp 2D-PAGE proteome maps using PDQuest analysis.

Proteins whose expression was (A) up regulated or (B) down regulated, as revealed by either the Student *t*-test (T-test) or quantitative test (Q-test), in the high firmness phenotype in comparison to the low firmness phenotype are graphically represented by Venn diagrams. Proteins detected by both tests are shown in the intersection (T∩Q).

Table 7.2: Differentially expressed proteins in the high firmness phenotype in comparison with the low firmness phenotype.

Differentially expressed protein spots detected by PDQuest analysis were analysed by MALDI-TOF MS. Positively identified proteins were listed according to their function. Their regulation factor of the protein spot abundance is specified in fold increase/decrease. The protein spot numbers correspond to the labelling indicated in Figure 6.4B and C. The SSP numbers correspond to the spot labelling from the PDQuest analysis. The identified protein sequences were BLAST searched against *Malus* EST database from NCBI to determine the highest scoring match, -- represents no data to validate and NS- stands for no significant similarity. Protein spots whose expression increased or decreased in the high firmness phenotype in comparison to the low firmness phenotype are indicated by (+) and (-), respectively. Protein spots visualised in the high firmness phenotype only are indicated by the sign '*'. 'T' and 'Q' represent proteins that were detected as differentially regulated using the student *t*-test ($p<0.05$) and quantitative analysis (minimum of 2-fold up or down), respectively.

Spot number (SSP)	Protein name	Identity to *Malus* EST (%)	Protein	
			Regulation high firmness phenotype	Factor increase (test)
Citric acid cycle				
31/7412	NAD-dependent malate dehydrogenase	96	-	<3.0 (Q, T)
50/5516	Cytosolic malate dehydrogenase	98	+	~2.5 (Q)
Carbohydrate metabolizing proteins				
29/8410	NAD-dependent sorbitol-6-phosphate dehydrogenase	75	+	~2.0 (T)
202/2715	NADP-dependent sorbitol-6-phosphate dehydrogenase	75	-	~2.0 (Q, T)
Photosynthesis				
20/5514	Ferredoxin-NADP (H) oxidoreductase	85	-	~3.5 (T)
Lipid metabolism				
68/5318	Lipoxygenase	76	-	<2.5 (T)

Table 7.2 continued

Spot number (SSP)	Protein name	Identity to *Malus* EST (%)	Protein	
			Regulation high firmness phenotype	Factor increase (test)
Ethylene biosynthesis				
197/2714	Ethylene response sensor protein (ethylene receptor)	68	-	~2.0 (Q, T)
21/6616	S-adenosyl-L-methionine synthetase	96	-	~9.5 (T)
Defense/detoxifying enzymes				
94/5212	Major allergen mal d 1.02	100	+*	
123/7312	Dehydroascobate reductase	100	+	~5.0 (T)
91/4220	MLA7	37	+	~3.0 (T)
Cell growth/division: DNA synthesis/replication				
176/2415	Transposase	NS	+	>2.0 (Q, T)
113/1109	Cell division inhibitor MinD	58	+*	
Signal transduction				
206/1608	Lissencephaly type-1-like homology motif	54	+*	
20/1709	Phytochrome kinase substrate putative	85	+	<2.0 (Q, T)
Cytoskeleton related proteins				
144/3510	Dynein heavy chain isoform (DHCIb, fragment)	NS	+	>2.0 (Q, T)
86/2208	Myosin class II heavy chain	NS	+	~11.0 (T)
Unclassified/unidentified proteins				
3/7810	MRGH2	NS	-	~2.5 (Q)
13/6617	AB013353 NID	97	-	~4.0 (T)

Table 7.2 continued

Spot number (SSP)	Protein name	Identity to *Malus* EST (%)	Protein	
			Regulation high firmness phenotype	Factor increase (test)
79/1313	AB004825 NID	94	-	~7.0 (Q, T)
28/8408	Hypothetical protein	91	-	<2.0 (Q, T)
82/3314	Hypothetical protein NitaMp024	NS	+	<2.0 (Q, T)
170/5406	Predicted protein	NS	+*	
188/5822	Predicted protein	73	+	~2.5 (Q)
69/5317	No positive hit	--	+	~4.5 (Q)
34/5518	No positive hit	--	+	>2.5 (Q)
198/2807	No positive hit	--	+	>2.0 (Q, T)
178/3819	No positive hit	--	+	~2.0 (T)
146/Q509	No positive hit	--	+*	
35/6518	No positive hit	--	+	<5.0 (Q, T)
57/7410	No positive hit	--	+	2.5 (Q, T)
169/5311	No positive hit	--	+	~4.5 (T)
184/5711	No positive hit	--	+	2.5 (T)
108/7214	No positive hit	--	+	>7.5 (T)
54/8310	No positive hit	--	+	~6.5 (T)
93/5211	No positive hit	--	+	~3.5 (T)
65/5411	No positive hit	--	+	~2.5 (Q, T)
99/6213	No positive hit	--	+	~3.5 (Q)
64/6315	No positive hit	--	+	~4.0 (Q, T)
56/9405	No positive hit	--	+	>1.5 (T)
217/8515	No positive hit	--	+	~2.0 (Q, T)

Table 7.2 continued

Spot number (SSP)	Protein name	Identity to *Malus* EST (%)	Protein	
			Regulation high firmness phenotype	Regulation factor fold (test)
137/4815	No positive hit	--	-	~2.0 (Q)
182/7810	No positive hit	--	-	~2.5 (Q)
97/6216	No positive hit	--	+	3.5 (T)
45/3417	No positive hit	--	-	~3.5 (Q, T)
140/3815	No positive hit	--	-	<2.5 (Q)
107/7111	No positive hit	--	-	~7.0 (Q, T)
35/6516	No positive hit	--	-	~3.0 (T)
49/6517	No positive hit	--	-	~2.0 (Q)
74/1417	No positive hit	--	-	~3.0 (Q)
73/2308	No positive hit	--	-	>7.5 (Q, T)
126/9407	No positive hit	--	-	~7.0 (Q)
NI/9805	No positive hit	--	-	~2.5 (Q, T)
NI/2716	No positive hit	--	-	~2.0 (T)

7.3 DISCUSSION

Fruit firmness is a critical determinant to estimate a fruit shelf life. According to Harker *et al*. (2003), fruits with low firmness have reduced consumer acceptability on the market unlike high firmness fruits. In addition, high firmness fruits have a longer shelf life than low firmness fruits (Harker *et al.*, 2003). Reduction in fruit firmness (fruit softening) has been mainly associated with endogenous ethylene production pathway (Abeles and Biles, 1991; Bleecker and Kende, 2000; Hermann *et al.*, 2007). However, it is not known whether there are other pathways and/or proteins involved in fruit firmness, causing some apples to have a longer shelf life than others. Thus, there is a need to characterise the molecular mechanisms involved in maintaining fruit firmness. A proteomics approach was used to generate 2D-PAGE proteome maps of high and low firmness apple fruit pulp and differentially regulated proteins were then identified by MALDI-TOF MS and characterised.

7.3.1 Comparative analysis of protein expression between the high and low firmness apple fruit pulp after separation by 1D-PAGE

Following protein extraction from high and low firmness apple fruit pulp by phenol precipitation, 1D-PAGE for both phenotypes were successfully established. Comparatively, differentially regulated protein bands were observed on both 1D-PAGE analyses of the phenotypes (Figure 7.1). From the visual observation of the gel, some protein bands were visually observed as differentially regulated. These groups of protein bands were characterized by molecular weights of 15-20 kDa, 20-25 kDa and ~70 kDa. The protein bands observed could either represent a single protein or a cluster

of proteins. However, this was not possible to resolve using 1D-PAGE. Therefore, to clearly resolve the protein extracts, 2D-PAGE was carried out (section 7.3).

7.3.2 Comparative analysis of protein expression between the high and low firmness apple fruit pulp after separation by 2D-PAGE

Following protein separation by 1D-PAGE, proteome maps were established by resolving TSP on 2D-PAGE using apple fruit pulp protein samples from both high and low firmness phenotypes extracted by phenol precipitation. Whilst the 1D-PAGE technique separates proteins according to their molecular weight only, proteins are resolved according to their pI and molecular weight during 2D-PAGE, thus allowing precise comparison between the two phenotypes. In this way individual spots, each representing a single protein, can be well visualised. Also, the spots were round shaped and not oval, a sign that good focusing and second dimension molecular weight separations were achieved. Therefore, an informative comparative analysis between high and low firmness phenotypes was carried out (Figure 7.3).

During comparative analysis using 2D-PAGE of high and low firmness phenotype proteomes by PDQuest, a total of 217 and 212 protein spots were visualized on high and low firmness phenotype proteome maps, respectively (Figure 7.3). PDQuest (qualitative, quantitative and student *t*-test, $p<0.05$) analysis revealed that 54 protein spots were differentially expressed in the high firmness phenotype, when compared to the low firmness phenotype (Figures 7.3 and 7.4; Table 7.2). Only the spots differentially expressed, as revealed by the PDQuest analysis, were characterised in this chapter. Therefore, of the 54 spots that were shown to be differentially regulated, 27 and 22 protein spots were up- and down- regulated, respectively in the high firmness

phenotype, in comparison with the low firmness phenotype. In addition, five spots were expressed in the high firmness phenotype only (Figure 7.3).

In terms of the 2D-PAGE proteome profile similarity between the high and low firmness phenotypes, the coefficient of correlation of the regression line obtained by plotting the normalized abundance of every single spot was 0.525146. The correlation coefficient denotes the strength of a linear relationship between two variables following the assumption of normality (normal distribution). In other words, the closer the coefficient is to 1, the more similar the groups become. Therefore, this low coefficient of correlation revealed that a relatively low similarity in proteome profiles between the two phenotypes was observed, since gel to gel consistency was maintained. A low correlation between proteome profiles suggested that a high number of proteins are influenced by the apple pulp firmness phenotypes (Figure 7.4). This also possibly suggested that an intricate set of molecular mechanisms are influenced by and/or induce the firmness phenotype in apple fruit pulp.

However, the presence of outlier spots detected during regression analysis, for example spot positioned at coordinates (2.55; 4.65) may have influenced the regression line and thus the low coefficient of correlation obtained (Figure 7.4). Therefore, the similarity between the two phenotypes may be higher than reflected by the scatter plot. The scatter plot, which is a graphical representation of the relationship between two variables (high and low firmness phenotypes in this case), allows representing a straight line of best fit. This regression line thus separates non-random variations and permits predicting the fold of variation for a given protein.

7.3.3 Identification and characterization of the differentially expressed proteins between the high and low firmness phenotypes

In order to characterise the differentially expressed proteins between the high and low firmness phenotypes from apple fruit pulp, the spots were identified by MALDI-TOF MS. Out of the 54 protein spots detected as differentially expressed during the PDQuest analysis, 24 were positively identified including two hypothetical and two predicted proteins (Table 7.2). The identified spots matched with those identified in chapter 6 (Table 6.1). The functions of 16 of these proteins were determined but eight remained unclassified, i.e. whose function has not yet been clearly assessed (Table 7.2). In addition, a total of 30 spots remained unidentified although good spectra were obtained. This is probably due to the limited genomic data coverage publicly available for apple and other species of the Rosaceae family.

The positively identified proteins were classified according to their function into eight functional categories, as established by Bevan *et al*. (1998), Ndimba *et al.* (2005) and the oilseed proteomic database (http://oilseedproteomics.missouri.edu/; Table 7.2). These categories include energy associated proteins like citric acid cycle associated proteins, carbohydrate metabolising proteins, photosynthesis and lipid metabolism, ethylene biosynthesis, defense/detoxifying enzymes, cell growth/division, signal transduction, cytoskeleton related proteins, unclassified proteins, and unidentified proteins.

7.3.3.1 Energy associated proteins

7.3.3.1.1 Citric acid cycle-associated proteins

Two proteins whose expression varied between the high and low firmness phenotypes were identified as related to the citric acid cycle-associated proteins. Spot 7412 was identified as NAD-dependent malate dehydrogenase and spot 5516 as cytosolic malate dehydrogenase. The PDQuest analysis showed that the abundance of cytosolic malate dehydrogenase was up-regulated in the high firmness phenotype and NAD-malate dehydrogenase was down-regulated in the high firmness phenotype.

Malate dehydrogenase, in the presence of NAD^+ as a cofactor, catalyses the reversible conversion of malate into oxaloacetate (Katz *et al*., 2007). The enzyme is also involved in gluconeogenesis, the synthesis of glucose from smaller molecules. The NAD+-dependent malate dehydrogenase is present in mitochondria, where it is only active in dark condition (Berkemeyer *et al*., 1998). To allow the movement of oxaloacetate out of the mitochondria, malate dehydrogenase reduces it to malate, and it then traverses the inner mitochondrial membrane. Once in the cytosol, malate is oxidized back to oxaloacetate by the cytosolic malate dehydrogenase for its utilisation in the citric acid cycle (Berkemeyer *et al.*, 1998; Matic *et al.*, 2005). Up-regulation of cytosol malate dehydrogenase may imply an increased citric acid synthesis in the high firmness phenotype. Citric acid is the primary organic acid that contributes to the overall fruit organoleptic quality (Bianco *et al.,* 2009). Normally, citric acid declines gradually during fruit development. However, in this study, the up-regulation of cytosolic malate dehydrogenase may suggest that citric acid does not decline quickly in the high firmness phenotype in comparison to the low firmness phenotype.

The expression of the mitochondrial malate dehydrogenase has been shown to be induced by an increase carbonylation (Backhausen *et al*., 1998), an oxidative protein modification due to the introduction of a carbonyl radical like carbon monoxide into proteins (Moller *et al.,* 2007). Carbonylation of proteins induces a decrease or even a total inhibition of their activity and increases their susceptibility to proteolysis (Pantke *et al.*, 1999; Berlett and Stadtman, 1997), largely because of the unfolding of the targeted protein domains (Grune *et al*., 2004). However, it has been shown in potato that proteins can escape degradation and form high molecular weight aggregates that accumulate with age and promote fruit senescence (Backhausen *et al.*, 1998). Therefore, the higher accumulation of the NAD^+-dependent malate dehydrogenase observed in the low firmness phenotype may be due to an accumulation of carbonylated proteins, not degraded by proteolysis, thus leading to faster/quicker fruit senescence, which relates to the softening of the apple fruit pulp.

7.3.3.1.2 Carbohydrate metabolizing proteins

The two spots 8410 and 2715 were identified as NAD-dependent sorbitol-6-phosphate dehydrogenase and NADP-dependent sorbitol-6-phosphate dehydrogenase, respectively. These two enzymes are involved in the carbohydrate metabolism. NAD-dependent sorbitol-6-phosphate dehydrogenase was up-regulated in the high firmness phenotype and on the contrary, NADP-dependent sorbitol-6-phosphate dehydrogenase was down-regulated in the high firmness phenotype.

These two enzymes are involved in sorbitol metabolism. While the NADP-dependent sorbitol-6-phosphate (S-6-P) dehydrogenase reducess S-6-P from glucose-6-phosphate (Kanayama *et al.*, 2005), the NAD-dependent S-6-P dehydrogenase converts sorbitol in

the leaves to fructose (fruit sugar), which is then translocated to the fruit (Watkins *et al.*, 2000; Kanamarua *et al.*, 2004). In Rosaceae species like apple, sorbitol present in fruits is converted into glucose when translocated into the cells in an apoplastic manner, normally in the absence of NAD/NADP (Ohkawa *et al.*, 2008).

In a study by Kanamarua *et al.* (2004) using transgenic apple plants that overexpressed a NADP-dependent S-6-P dehydrogenase cDNA, the activity of NADP-dependent S-6-P dehydrogenase was detected as positively correlated with the amount of sorbitol. In this study, they revealed that the sucrose content increased markedly when the sorbitol content decreased. In this study, the expression of the NADP-dependent S-6-P dehydrogenase increased in the high firmness phenotype, while NAD-dependent S-6-P dehydrogenase decreased. Taken together, these results suggest that higher content of S-6-P might accumulate in the high firmness phenotype, while the fruit content in fructose and glucose may be lower in the high firmness phenotype in comparison with the low firmness phenotype. However, it is necessary to conduct enzyme activity assay to verify whether the accumulated NADP-dependent S-6-P dehydrogenase is active in the fruit cells.

7.3.3.1.3 Photosynthesis

The abundance of spot 5514 was down-regulated in the high firmness phenotype by approximately 3.5 fold. This spot was identified as ferredoxin-NADP oxidoreductase, an enzyme involved in photosynthesis.

The redox enzyme ferredoxin-$NADP^+$-oxidoreductase catalyses the terminal step of the photosynthetic electron transport chain, namely the reduction of $NADP^+$ by ferredoxin,

in chloroplasts (Shin *et al*., 1963; Shin and Arnon, 1965). Two isoforms of the enzyme, encoded by different genes, one photosynthetic and the other heterotrophic, are present in plants (Matsubara and Hase, 1983). They link to different metabolic pathways and are regulated by different factors (Gummadova *et al.*, 2007). These two enzymes have also been reported in tomato fruit (Green *et al*., 1991). The photosynthetic ferredoxin-NADP(H) oxidoreductase catalyses the final step in the photosynthetic electron transport chain (Zhang and Cramer, 2004), while the heterotrophic ferredoxin-NADP(H) oxidoreductase supplies reducing power in non-photosynthetic tissue to a variety of metabolic processes, including nitrate assimilation (Bowsher *et al.*, 1992). The presence of both FNR and ferredoxin has already been reported in nonphotosynthetic tissues in tomato, suggesting that plant cells depend on ferredoxin and FNR for the assimilation of inorganic nitrogen (Green *et al.*, 1991). The lower abundance of the enzyme in the high firmness phenotype suggests that a small amount of $NADP^+$ will be reduced and thus low reducing power is available for other metabolic processes to occur in the high firmness phenotype. Since the major metabolic process that occurs in fruit is the production of fructose, the reduction in ferredoxin-$NADP^+$-oxidoreductase expression may lead to reduced fructose content in the fruits.

7.3.3.1.4 Lipid metabolism

Spot 5318 was identified as lipoxygenase, a lipid metabolising enzyme. Its expression was down-regulated in the high firmness phenotype in comparison to the low firmness phenotype by a factor of approximately 2.5 fold.

Lipoxygenases form a family of iron-containing enzymes that catalyses the dioxygenation of polyunsaturated fatty acids in lipids containing a cis,cis-1,4-

pentadiene structure (Liavonchanka and Feussner, 2006). In pea (*Pisum sativum*), the lipoxygenase has been shown to be involved in plant growth and development, pest resistance and senescence and responses to wounding (Veronico *et al.*, 2006). In addition, the enzyme predominately accumulates during fruit ripening and has been shown to incorporate oxygen into unsaturated fatty acids producing hydroperoxides, which aid in textural associated changes in tomato fruit ripening (Kausch and Handa, 1995; Ramakrishna *et al.*, 2003; Faurobert *et al.*, 2007). Kausch and Handa (1995) showed that fruit firmness significantly increased after the removal of lipoxygenase. The co-suppression effect was stably inherited in tomato fruits (Faurobert *et al.*, 2007). These findings are in correlation with the PDQuest results, which detected lower expression of lipoxygenase in the high firmness phenotype.

7.3.3.2 Ethylene biosynthesis

The abundance of two ethylene biosynthesis related proteins, ethylene response sensor protein (spot 2714) and S-adenosyl-L-methionine synthetase (spot 6616), was detected as down-regulated in the high firmness phenotype in comparison to the low firmness phenotype. The ethylene response sensor protein and S-adenosyl-L-methionine synthetase were down-regulated by a factor of approximately 2.0 and 9.5 fold, respectively.

In climacteric fruits like apples and tomatoes, the increase in ethylene accelerates the ripening process (Alexander and Grierson, 2002), inducing colour development (Fraser *et al.*, 1994) and cell wall softening (Brummell and Harpster, 2001). Ethylene biosynthesis in higher plants begins with the formation of *S*-adenosyl-L-methionine (SAM) from methionine by the action of the enzyme SAM synthetase. Then, SAM is

catalysed to ACC by ACS. The last step consists in the conversion of ACC to ethylene, which is carried out by ACO (Kende, 1993).

The ethylene response sensor protein, also known as ethylene receptor, is located on the cell membrane and regulates adaptive responses to a broad range of environmental signals (Konishi and Yanagisawa, 2005; Yin *et al.*, 2008). The enzyme also plays a role in growth, development and quality of fruits (Alexander and Grierson, 2002). It has been demonstrated that an up-regulation in the expression of ethylene response sensor protein induces increase in ethylene sensitivity (Ciardi *et al.*, 2000).

The ripening process is characterized by the partial disassembly of the fruit cell wall, which in turn induces fruit softening, and textural changes. The activity of several enzymes, like expansins and other cell wall components, has been shown to be regulated by ethylene during ripening, thus causing structural changes in the cell walls, which lead to fruit softening in the apple cultivar 'McIntosh' (Hrazdina, 1994; Hrazdina *et al.*, 2003) and in tomato (Alexander and Grierson, 2002). In addition, Hrazdina *et* al. (2003) revealed that a reduction in ethylene production resulted in significant delay in softening resulting in improved storability.

The lower expression of SAM in the high firmness phenotype suggests that the production of ethylene is lower than in the low firmness phenotype. In addition, the reduction in the accumulation of the ethylene response sensor protein suggests that the ethylene sensing is reduced in the high firmness phenotype. Taken together, these results may partially explain the difference of fruit firmness between the two phenotypes. However, accurate measurements of ethylene content in fruits from both

phenotypes are necessary to confirm that lower rates of the hormone are produced in the high firmness phenotype.

7.3.3.3 Defense/detoxifying enzymes

The abundance of three proteins involved in plant defense was detected as changing between the two phenotypes. Spot 5212 was only detected in the high firmness phenotype, as revealed by the qualitative analysis and was identified as major allergen Mal d 1 (AP15). Spots 4220 and 7312, identified as a powdery mildew resistance locus-containing protein (MLA7) and dehydroascorbate reductase were up-regulated in the high firmness phenotype.

Mal d 1 is a major allergen protein that is involved in disease resistance and allergic response. It also plays a role in senescence, nucleotide-, cytokinin- and brassinosteroid-binding (Beuning *et al.*, 2004). In 'Royal Gala' apples, the gene coding for Mal d 1 was the one dominantly expressed in ripe fruits among a family of 12 members (Puehringer *et al.*, 2003, Beuning *et al.*, 2004). Studies in various plant species including tomato, strawberry or banana revealed the synthesis of some of the major allergen proteins to be associated with fruit ripening (Bianco *et al*., 2009). In apple, the mRNA transcript levels coding for several proteins have also been detected as up-regulated during the ripening of 'Mondial Gala' apple fruits. In correlation with a study by Beuning *et al.* (2004) that showed higher mRNA coding for Mal d 1 in ripe fruits than in unripe fruits, the transcript coding for Mal d 1 was detected among the proteins induced with fruit ripening (Goulao and Oliveira, 2006). In this study, the detection of Mal d 1 in the pulp of ripe fruits is therefore in agreement with the literature. The results suggest that fruits with high firmness might be more allergenic than fruits with low firmness. In addition,

this is the first study reporting the variation in Mal d 1 expression in relation with fruit pulp firmness.

The up-regulation of MLA7 protein in the high-firmness phenotype may suggest that higher resistance to powdery mildew may be observed in high firmness fruit pulp in comparison to low firmness fruit pulp. However, further work is necessary to determine if the resistance to powdery mildew is a cause or consequence of the firmness state of the fruit pulp.

The up-regulation of dehydroascorbate reductase in the high firmness phenotype may suggest an increased reduction of dehydroascorbate to ascorbic acid in the high firmness fruit pulp in comparison to the low firmness fruit pulp. According to Chen and Gallie (2006) the effect of dehydroascorbate reductase expression on leaf aging inversely correlated with the level of lipid peroxidation, suggesting that the enzyme contributes to plant growth by maintaining photosynthetic functioning through efficient ascorbic acid recycling, which in turn limits ROS-mediated damage and slows leaf aging and probably fruit aging. This may suggest that the up-regulation of dehydroascorbate reductase in the high firmness phenotype slows down fruit aging, thus maintaining its firmness.

7.3.2.4 Cell growth/division: DNA synthesis/replication associated proteins

The abundance of two proteins, whose functions are associated with cell growth/division: DNA synthesis/replication, was shown as differentially regulated between the two phenotypes. Spot 2415, identified as transposase, was up-regulated 2-fold in the high firmness phenotype in comparison to the low firmness phenotype. Spot

1109, identified as cell division inhibitor MinD, was only identified in the high firmness phenotype. Detection of cell inhibitor MinD coincides with the ripening process since the fruits used in this study were mature (ripe). Cell inhibitor MinD is expected to be highly expressed post fruit cell expansion and beginning of fruit ripening (Janssen *et al.,* 2008). Generally, Cell inhibitor MinD would be expected from both phenotypes, but in this study it was only identified in the high firmness. According to Janssen *et al*. (2008), expression of cell division proteins determine fruit shape, cell size texture, thus detection of cell division inhibitor MinD only in the high firmness may imply that it plays a role maintaining fruit firmness.

The transposase enzyme catalyses transposon movement from one site of the chromosome to another either as a cut and paste mechanism by binding to its ends, or as a replicative transposition mechanism (Bundock and Hooykaas, 2005). The interaction of transposable elements and regulatory sequences of genes can lead to alterations in the level of transcription (Weil and Wessler, 1990). Harada *et al*. (2000) suggested that 1-aminocyclopropane-1-carboxylate synthase gene (*Md-ACS-1* and *-2*), whose promoter is mutated by a transposase contributes to the long-term storage capability of some apple cultivars through the reducing ethylene production during ripening (Harada *et al.*, 2000). Thus, the up-regulation of transposase in the high firmness phenotype may imply that some genes influencing the fruit pulp softening may be altered through the insertion of transposable elements resulting in the modification of their functions (Harada *et al.*, 2000).

7.3.2.5 Signal transduction associated proteins

The PDQuest analysis revealed that the abundance of spot 1608 and spot 1709, whose function is associated with signal transduction, varied between the high and low firmness phenotypes. Spot 1608, identified as lissencephaly type-1-like homology motif, was only detected in the high firmness phenotype. Spot 1709, identified as phytochrome kinase substrate putative, was up-regulated by 2 fold in the high firmness phenotype.

Lissencephaly type-1-like homology motif has recently been demonstrated in animals as being part of the proteins involved in microtubule dynamics, cell migration, nucleokinesis and chromosome segregation (Suzuki *et al.*, 2008). In addition, armadillo-repeat-containing protein 8 (ARMc8), a key component of the C-terminal of lissencephaly type-1-like homology motif, has been shown to regulate proteosome-dependent degradation by interacting with α-catenin, a process important in regulating gluconeogenesis (Suzuki *et al.*, 2008). Following a blast search using NCBI on other green plants, less than 56% sequence homology was detected. However, lissencephaly type-1-like homology motify was found to be present in other plant genomes like *Arabidopsis thaliana*. This study reports for the first time the expression of lissencephaly type-1-like homology motif in the Rosaceae family. Taking into account the function of this motif in animals, its detection in the high firmness phenotype of the apple fruit pulp only suggested that it functions for intracellular organization, cell shape, motility and expansion through its role on microtubules.

Phytochrome kinase belongs to a small protein family that interacts with phytochromes (Fankhauser *et al.*, 1999), the molecular light switches that regulate various aspects of

plant growth and development (Kendrich and Kronenberg, 1994; Quail *et al.*, 1995; Fankhauser, 2000; Smith, 2000). The kinase has been shown to be a negative regulator of phytochrome signaling (Fankhauser *et al.*, 1999), by blocking the interaction of phytochrome with its putative signal transducers (Kim *et al.*, 2004). The kinase is also differentially regulated by phosphorylation under red-light conditions *in vivo* (Fankhauser *et al.*, 1999). The up-regulation of the kinase in the high firmness phenotype suggests the kinase role in the fruit development by negatively regulating phytochrome signaling.

7.3.2.6 Cytoskeleton related proteins

Two proteins whose functions are implicated in the cytoskeleton were detected as differentially regulated between the high and low firmness phenotypes. These include dynein heavy chain isoform (DHCIb; spot 3510) and myosin class II heavy chain (spot 2208), which were both up-regulated in the high firmness phenotype. The abundance of DHCIb and myosin class II heavy chain increased by approximately 2.0 and 11.0 fold, respectively. These two proteins, though identified by MASCOT could not be validated against the *Malus* EST database from the NCBI. This could be due to the limited data on the publicly accessible EST databases.

Myosin class II heavy chain, whose expression was the most different between the two phenotypes, is a cytoskeleton motor protein and is one component of the actin motor. Dyneins together with kinesins are microtubule motors capable of a sliding movement useful to establish and maintain the structural integrity of the cell during mitosis (Smith, 2002). The increase expression of these two proteins involved in the cytoskeleton, which in turn influenced the cell structure, together with the identification of the

lissencephaly type-1-like homology motif in the high firmness phenotype only, suggests that they might play an important role in maintaining the cell structure of the high firmness phenotype.

7.3.3.7 Unclassified proteins

Several proteins whose biological functions are unknown were also detected as differentially regulated between the two phenotypes. The expression of spots 3314, 7312 and 5822 up-regulated in the high firmness phenotype in comparison with the low firmness phenotype, by a factor of approximately 2.0, 5.0 and 2.5 fold, respectively. The abundance of spots 7810, 6617, 1313 and 8408 was down-regulated in the high firmness phenotype in comparison with the low firmness phenotype, by a factor of approximately 2.5, 4.0, 7.0 and 2.0 fold, respectively.

The spot positions of unclassified proteins on 2D-PAGE, SSP numbers from PDQuest, percentage identity to *Malus* EST database from NCBI and nature of their changes in expression between the high and low firmness phenotypes are shown in Table 7.1 and Figure 7.3. Four unclassified spots were identified to be up-regulated and four down-regulated in the high firmness phenotype. Two predicted protein (spots 5822 and 5406) and hypothetical protein NitaMp024 (spot 3314) were up-regulated in the high firmness phenotype, while MRGH2 (spot 7810), AB013353 NID (spot 6617), AB004825 NID (spot 1313) and hypothetical protein (spot 8408) were down-regulated in the high firmness phenotype in comparison with the low firmness phenotype.

Following functionalisation using bioinformatics tools, some of the unclassified proteins were annotated predicted functions based on sequence and domain homology

(section 6.5.4.12). AB013353 NID was detected as having common conserved domains with UTP-glucose-1-phosphate uridylyltransferase, which plays a role in sucrose biosynthesis in fruit cells, through the conversion of UDP-glucose to glucose-1-phosphate (section 7.6.3.1.2). Thus, AB013353 NID was classified under carbohydrate metabolizing proteins. In terms of AB004825 NID was detected as having common conserved domains with the eukaryotic initiation factor 5A hypusine signature, which plays a critical role in the assembling of large and small subunits of ribosomes and is involved in programmed cell death. Thus, AB004825 NID was classified as a defense/detoxification protein. In addition, hypothetical protein was detected as having common conserved domains with the phosphoglycerate kinase signature, which is known to have phosphotransferase activity and has been implicated to play a role in carbohydrate metabolic processes (Ronimus *et al*., 2001). The roles of these proteins in regulating firmness have not been characterized to date.

7.3.2.8 Unidentified proteins

The expression of 30 unidentified proteins was detected as influenced by the firmness phenotype. Although good spectra were obtained for these proteins, these proteins remain unidentified probably because of the limited genomic data available in apple and other plant species from the Rosaceae family. Out of the 30 proteins, 17 were up-regulated and 12 were down-regulated in the high firmness phenotype, in comparison with the low firmness phenotype. In addition, one protein was visualized only in the high firmness phenotype as shown on Figure 7.3 and Table 7.2.

In this chapter, 27 and 22 spots were observed up-regulated and down-regulated, respectively, in the high firmness phenotype. In addition, five spots were detected only

in high firmness phenotype. Further, seven functional categories were deduced from the protein identification. These include energy metabolism (citric acid cycle associated proteins, carbohydrate metabolising proteins, photosynthesis and lipid metabolism), ethylene biosynthesis, defense/detoxifying enzymes, cell growth/division, signal transduction, cytoskeleton related proteins and unclassified proteins. The differential expression of these proteins may account for the difference in firmness between individual apples in an individual tree or between individual trees and cultivars. Therefore, this study reports the success in identifying the differentially regulated proteins between the high and low firmness phenotypes of apple fruit pulp tissue.

CHAPTER 8

FINAL DISCUSSION

The overall aim of this thesis was to characterise the molecular mechanisms influencing fruit quality in apples using both genomics and proteomics approaches. In order to achieve this, the study first aimed at constructing genetic linkage maps from the 'Golden Delicious' x 'Dietrich' population for the identification of QTLs related to fruit quality traits. In this regard, genomic DNA was extracted from leaves from the 248 individuals of the population, and used in megaplex PCR, a recently adopted and optimized method. A total of 286 heterozygous SSR markers were used to construct the genetic linkage maps of 'Golden Delicious' and 'Dietrich' using the *JoinMap*® software. Following map construction, QTLs associated with the firmness, juiciness, crispness, form, colour, size, acidity, stripness and russeting fruit quality traits were detected using the phenotypic data collected from mature apple fruits over a period of three years and the *MapQTL*® software.

To gain insight into the molecular mechanisms controlling fruit pulp firmness, the pulp of mature apple fruits was used to generate 2D-PAGE proteome maps. The spots resolved using CBB staining were identified and characterised. To achieve this goal, total soluble proteins (TSP) extracted from pulp of mature apple fruits from the 'Golden Delicious' x 'Dietrich' population were resolved by 2D-PAGE and visualised by CBB staining. TSP were isolated using two optimised methods, either by phenol or TCA/acetone precipitation and resolubilised in urea/thiourea based extraction buffer. Protein spots were then excised out of the gels, trypsin digested, identified by MALDI-TOF MS and characterised using bioinformatics tools. To further determine proteins

involved with fruit firmness, 2D-PAGE proteome maps from both high and low firmness apple fruit pulp were comparatively analysed using the PDQuest software.

8.1 CONSTRUCTION OF GENETIC LINKAGE MAPS

The genetic linkage maps from the 'Golden Delicious' x 'Dietrich' population were successfully constructed by mapping a total of 167 SSR markers. Each map comprised of 17 linkage groups, representing the 17 chromosomes of the apple haploid genome. The individual maps of 'Golden Delicious' and 'Dietrich' contained 142 and 116 markers and spanned 1,437.8 cM and 1,491.5 cM, respectively. However, several of the 17 linkage groups were truncated, thus producing a total of 44, 40 and 45 groups for the 'Golden Delicious', Dietrich' and integrated map, respectively (Table 4.4; Figure 4.7). This suggests that more markers were needed to link the groups and obtain better representation of the linkage groups. Out of the 167 SSR markers positioned on the linkage maps, 33 were newly developed, and distributed over 13 linkage groups.

The linkage maps in this study were constructed only with SSR markers. On the contrary to other markers, like RAPDs, RFLPs or DarTs, SSR markers are co-dominant and transferable across species (Liebhard *et al.,* 2003a; Gardiner *et al*., 2007; Igarashi *et al*., 2008; Celton *et al*., 2009; Pattochi *et al*., 2009), and thus allow comparison with other published maps. The 17 linkage groups obtained in this study, with the exception of LG1, LG3, LG7 and LG8, were comparable with the published maps of 'Fiesta' x 'Discovery' (Maliepaard *et al*., 1998; Silfverberg-Dilworth *et al*., 2006) and 'Malling 9' x 'Robusta 5' (Celton *et al*., 2009). Although many markers previously mapped on LG1, LG3, LG7 and LG8 in previous studies (Maliepaard *et al*., 1998; Liebhard *et al*., 2002; Silfverberg Dilworth *et al*., 2006; Igarashi *et al*., 2008; Celton *et al*., 2009;

Pattochi *et al*., 2009; van Dyk *et al.* 2009) were tested, only a few linked on these linkage groups on the maps of 'Golden Delicious' and 'Dietrich'.

'Golden Delicious' has been classified as a disease resistant cultivar as it is resistant to silver leaf disease (Bus *et al.*, 2000). However, markers like CHVf1, which have been shown to be associated with disease resistance, mapped on LG1 in previous studies (Bus *et al.*, 2008), could not link to either LG1 or other linkage groups in this study. Mapping more markers to 'Golden Delicious' and 'Dietrich' maps could help provide sufficient linkage between markers that are associated with disease resistance to previously identified linkage groups. However, failure of these markers to be positioned on any of the linkage groups could suggest that 'Golden Delicious' x 'Dietrich' population may not be suitable for studies related to disease resistance.

Some homeologous markers were observed on LG1. For instance, CH04e05, previously positioned on LG7 (Silfverberg-Dilworth *et al*., 2006) was mapped on the LG1 of the 'Golden Delicious' map while CH05g08, which was previously mapped on LG1 (Silfverberg-Dilworth *et al*., 2006) was positioned on LG11 in the current study.

In terms of comparison between the 'Golden Delicious' and 'Dietrich' linkage maps, a total of 105 markers were common in the two maps, representing 63% of the markers positioned. In comparison with the genetic maps by Liebhard *et al*. (2002), Silfverberg-Dilworth *et al*. (2006), Igarashi *et al*. (2008) and Celton *et al*. (2009), the maps constructed in this study showed higher allele sharing between the two parents. This suggested that the cultivars 'Golden Delicious' and 'Dietrich' share a common parent along their lineage. However, because of the limited data available on the genetic

background for the cultivar 'Dietrich', further pedigree studies are necessary to confirm this. Nevertheless, this may explain the high distortions within the map like the oversized LG11, LG12 and LG15, when compared to other apple genetic maps.

In conclusion, high similarity with regards to the SSR markers positioned was observed between the genetic linkage maps of 'Golden Delicious' and 'Dietrich', and the published maps of 'Prima' x 'Fiesta' (Maliepaard *et al*., 1998), 'Fiesta' x 'Discovery' (Liebhard *et al*., 2002; Silfverberg-Dilworth *et al*., 2006), 'Ralls Janet' x 'Delicious' (Igarashi *et al.*, 2008) and 'Malling 9' x 'Robusta 5' (Celton *et al*., 2009). Thus, these maps were deemed suitable for the identification of QTLs as an approach to understand the genetic factors influencing fruit quality traits.

8.2 IDENTIFICATION OF QTLs

Following construction of the genetic linkage maps, QTLs associated with nine fruit quality traits in the 'Golden Delicious' x 'Dietrich' population were detected and compared with the QTLs previously detected in the reference maps derived from the 'Prima' x 'Fiesta' (King *et al*., 2000, 2001), 'Fiesta' x 'Discovery' (Liebhard *et al*., 2003), 'Telamon' x 'Braeburn' (Kenis and Keulemans, 2008) and 'Ralls Janet' x 'Delicious' (Igarashi *et al.*, 2008) populations. The comparison is important and valuable for the identification of QTLs of interest towards marker assisted breeding. In addition, this allows estimating the influence of abiotic factors (climatic conditions and agronomic factors) on the alleles associated with the QTLs.

A total of 72 putative QTLs associated with nine fruit quality traits in the 'Golden Delicious' x 'Dietrich' mapping population were detected in this study. Out of these,

eight putative QTLs were associated with firmness, six with juiciness, four with colour, nine with size, five with stripness, four with crispness, eight with acidity, five with form and 23 with russeting. Russeting represented 33% of the QTLs identified and this represented the maximum number of QTLs for a single trait so far documented.

In terms of linkage groups, 17 QTLs associated with eight traits, namely juiciness, russeting, size, form, firmness, acidity, colour and stripness were detected on LG15. This represented the maximum number of QTLs and traits observed for a single linkage group. Out of the 17 QTLs detected, five were associated with russeting and three with size. To date, only QTLs associated with the crop load and fruit weight have been assessed and detected on LG15 (Liebhard *et al*., 2003a). Therefore, this study identified for the first time QTLs associated with these eight traits.

In contrast, the minimum number of QTLs detected on a single linkage group was on LG3, which were associated with size and stripness. Only two markers, that were 48.9 cM apart, were positioned on LG3 resulting in a low marker density. Since IM has been shown to be inefficient for the detection of QTLs when the marker loci are over 35 cM distant to each other (Van Ooijen, 1999, 2004), it is possible that more putative QTLs could have not been detected. Therefore, a good marker density is desirable to ensure the identification of all QTLs associated with the fruit quality traits under investigation on this linkage group.

Previous studies reported the detection of QTLs associated with firmness (Costa *et al*., 2008) and fruit acidity (Liebhard *et al*., 2003) on LG1 and LG8, respectively. However, in this study, no QTLs were detected on LG1, LG7 and LG8 possibly because of the

low marker density of these linkage groups, which might have restricted the detection of QTLs.

In terms of the fruit quality traits, in addition to the previously detected QTLs controlling fruit quality traits, new QTLs were mapped on the linkage groups of the 'Golden Delicious' x 'Dietrich' population. These include LG2 for firmness; LG4, LG9 and LG15 for juiciness; LG15 for colour; LG3, LG12, LG15 and LG16 for size; LG2, LG4 and LG9 for crispness; LG2, LG9, LG10, LG13, LG14 and LG15 for acidity.

In terms of colour, studies carried out on 'Ralls Janet' x 'Delicious' (Igarashi *et al.*, 2008), 'Sciros' x '91.136 B6-77' and 'Geneva' x 'Braeburn' (Chagné *et al*., 2007) populations have identified important QTLs for colour on LG9. In these studies, the *Rf* and *Rni* loci have been associated with fruit skin colour, and red foliage and red colour in the core of apple fruit, respectively. Both loci were mapped at the bottom of LG9 suggesting that they could be located in a gene cluster or even correspond to alleles of the same gene. Interestingly, fruit skin colour and leaf colour collocate in *Prunus* (Dirlewanger *et al*., 2004), which suggests that there may be a region of synteny between the middle of *Prunus* LG6 and the bottom of *Malus* LG9 (Chagné *et al*., 2007). However, in this study, besides detecting an additional QTL on LG15, a QTL controlling colour was also detected in LG9. The QTLs detected on LG9 were identified only on the 'Golden Delicious' map. However, by comparison with the previous studies, it would be expected to detect a QTL controlling red colouration on LG9 of the 'Dietrich' map. This data suggests that the gene(s) coding for the green colour of fruits from 'Golden Delicious' could be located on LG9 and LG15, while the gene (s) coding for the red colour of fruits from 'Dietrich' could be located on LG15. A

further assessment on the position of a QTL associated with red colour on LG9 can be carried out by candidate gene mapping using the *Rni* and *Rf* loci.

Apple cultivars present a wide range of colour variation in fruit skin ranging from green like 'Granny Smith', the partially coloured (striped) apples like 'Royal Gala', through to dark red like 'Red Delicious' and 'Dietrich'. The degree of stripness in apples has been shown to be directly related to the amount of anthocyanin in the fruits of 'Gala' apple strains (Iglesias *et al*., 2008). This suggests that the regulation of the anthocyanin levels determines the fruit colour and stripness patterns observed on fruits. In this study, QTLs controlling colour and stripness detected on LG15 were close to one another, suggesting that both QTLs could be located in a gene cluster or even correspond to alleles of the same gene. In addition, a common linkage group was also detected, LG9, even though the QTLs controlling colour were on the 'Golden Delicious' map, while a QTL controlling stripness was on the 'Dietrich' map. Additionally, more QTLs controlling stripness were detected on LG3 and LG16, indicating that stripness is controlled by several genes and is thus a complex trait.

In terms of form, five QTLs associated with this trait were detected using the 'Golden Delicious' x 'Dietrich' population. This study reported the first identification of QTLs associated with form in a plant species from the Rosaceae family, with the exception a QTL detected on LG6 in peach (Dirlewanger *et al*., 2004). Homologous regions have been detected between apple and peach (Gasic *et al*., 2008) suggesting that this QTL may be collocating in the two species. Therefore, a region of synteny between the peach LG6 and either the apple LG5, LG9, LG11, LG15 or LG16 may exist. The detection of several QTLs controlling form indicated the complexity of this trait in apple in

comparison with peach. Since the apple genome has 17 chromosomes, instead of eight in peach (Dirlewanger *et al*., 2004), the variation in QTL number could also be due to gene duplication and thus the occurrence of more genomic regions controlling a trait in apple.

In terms of russeting, 23 QTLs associated with this trait were detected using the 'Golden Delicious' x 'Dietrich' population. This study reported the first identification of QTLs associated with russeting in the Rosaceae family. The QTLs were positioned on LG2, LG4, LG6, LG9, LG10, LG11, LG12, LG13, LG14, LG15, LG16 and LG17. The several QTLs detected spanning over 12 linkage groups could explain the frequent occurrence of russeting on apple fruits considering that russeting is a consequence of many causes, like humidity, fungal and bacterial infection, and insect bites. However, the large number of QTLs detected controlling russeting suggests that many genes control russeting and is thus a complex trait.

In terms of comparison with previous studies on fruit quality, several QTLs have been detected positioned on LG1 and LG8 (King *et al*., 2000; Liebhard *et al.,* 2003b; Kenis and Keulemans, 2008). In this study, no QTLs could be detected in these linkage groups, possibly because of the low marker density of these linkage groups, which might have restricted the detection of QTLs. Although, no QTL has been reported detected on LG2 in the previous studies, in this study, six QTLs associated with firmness, acidity, crispness and russeting have been reportedly positioned on LG2. This can be a result of the high heterozygosity among cultivars influencing the detection of a locus and more often a QTL on a linkage group (Kenis and keulemans, 2008). In addition, the variations in QTL detection between studies suggest that the QTLs might

be cultivar dependent, since King *et al.* (2000) used cultivars 'Fiesta' and 'Prima' instead of 'Golden Delicious' and 'Dietrich' that was used in the present study. Further, since the populations were cultivated at different locations and thus were exposed to different environments, the result variations suggest that the QTLs controlling firmness might be influenced by the environmental conditions.

Identification of markers for MAS

Following identification of QTLs, markers related to these QTLs were identified. These candidate markers can be used to screen gene inheritance in individuals of a given population as early as the seedling stage, thus greatly hastening the selection of individuals with interesting traits (Mohan *et al.*, 1997; Liebhard *et al.*, 2002; Graham *et al.*, 2009). The greatest number of SSR markers identified for the selection or discrimination of individuals according to their genetic background for a given trait was nine detected for russeting. Three markers were associated with size, two with fruit firmness and one each with juiciness, crispness, acidity and form selection. However, no marker could be identified for colour and stripness, even though QTLs were identified for these two traits.

8.3 ESTABLISHMENT OF 2D-PAGE PROTEOME MAPS FROM APPLE FRUIT PULP AND PROTEIN IDENTIFICATION

This part of the thesis reported the preliminary proteomics analysis of mature apple fruit pulp, harvested from the seedlings of the 'Golden Delicious' x 'Dietrich' population, to characterise its total proteome. This analysis relied on the extraction of TSP, the establishment of high-quality 2D-PAGE proteome maps and identification of proteins by MALDI-TOF MS. Protein extraction was optimised to resolubilise the most proteins

possible and thus obtain a good representation of the pulp proteome, while limiting interfering substances that would reduce the gel quality.

Highly glycosylated proteins and interfering substances like sugars have been shown to hinder protein absorbance during quantification of proteins by Bradford thus causing erroneous estimation in the actual protein content of a given sample (Banick *et al.*, 2009). In addition, glycosylated proteins are usually not detected during CBB staining of gels (Robertson *et al.*, 1997). Sugars and phenolic compounds can interfere with protein focusing during the first dimension of 2D-PAGE (Heazlewood and Millar, 2006). These protein modifications and compounds have probably limited the map quality and limited the detection of the total proteome of apple fruit pulp.

Following extraction, proteins were separated by 1D-PAGE as quality control assessment method. The background staining of these gels was high, and limited the good visualization of the protein bands. Carbohydrates, which account for approximately 12% in mature apple fruit (Vieths *et al*., 1993), may have induced the background staining of the gels. In addition, the background may have been caused by the presence of phenolic compounds, degraded proteins and/or long storage of protein samples (Barraclough *et al.,* 2004). Similar predicaments were observed during the separation of TSP from apple fruit skin (H.V.H. Mathye, personal communication) and the resolution of proteins from the extracellular matrix of *Arabidopsis thaliana* and sorghum (*Sorghum bicolor*) cell suspension cultures (L.A. Thomas, personal communication). These interfering substances were partly removed through the use of phenol, PVPP or sodium sulphite, allowing a better visualization of proteins.

This study reported the establishment of 2D-PAGE proteome maps of TSP extracted using either phenol precipitation or TCA/acetone precipitation. These two extraction methods generated completely different 2D-PAGE proteome profiles. Previous studies also reported variations in proteome profiles during extraction of proteins from banana, apple skin and potato when comparing TCA/acetone method and phenol precipitation method (Carpentier *et al.*, 2005; Wang *et al*., 2006).

The TCA/acetone precipitation method is widely used in proteomics studies as it effectively precipitates proteins (Damerval *et al*., 1986; Barraclough *et al*., 2004), whilst reducing proteolytic activities by inactivating proteases (Wu and Wang, 1984). However, proteins are difficult to redissolve after precipitation with TCA, and a portion is lost remaining in the pellet. Barraclough *et al.* (2004), also observed that protein extracts precipitated using TCA do not completely redissolve in lysis buffer resulting in loss of proteins in low concentration. The difficulties of protein resolubilisation may explain to some extent the proteome profiles disparity between the two extraction methods and between this study and previous studies.

In this study, the combination of 2D-PAGE and MALDI-TOF MS allowed the identification of 111 apple pulp protein species despite the limited genomic data of the Rosaceae family. Out of the 290 spots excised from the 2D-PAGE, trypsinised and analysed by MALDI-TOF, good spectra were obtained from 250 (86%) of the digested spots. A total of 135 (48%) spots were positively identified as 111 distinct proteins. Of these, 111 proteins were validated through blasting against the *Malus* EST database from the NCBI. Spots with good spectra failed to be positively identified probably because of the limited proteomic data in apple and the low sensitivity of the MALDI-

TOF MS (L.A. Thomas, personal communication). However, availability of the spectra for these unidentified proteins is a step towards their identification, which may be possible when the apple genome sequence as well as other plant species from the Rosaceae family will become available. Another alternative would be to use a more sensitive MS such as the Q-TOF or MALDI-TOF-TOF or even through protein sequencing using tandem MS.

Following their identification, proteins were classified into 11 functional categories as established by Ndimba *et al.* (2005), TAIR (AraCyc, http://www.arabidopsis.org/biocyc/index.jsp) and the oilseed proteomic database (http://oilseedproteomics.missouri.edu/). These categories included energy related proteins, ethylene biosynthesis proteins, defense/detoxifying proteins, proton transporting ATPases, heat shock proteins, transcription and translation related proteins, cell growth/division related proteins, signal transduction, amino acid and purine biosynthesis related proteins, cytoskeleton related proteins and unclassified proteins. Apart from unclassified proteins, energy metabolism associated proteins represented the largest category with 17% of the positively identified proteins. All the major metabolic categories of proteins involved in fruit development were represented. Only 30 proteins were identified against plants belonging to the Rosaceae family and of these, 15 proteins were identified against *Malus* and indicating that there is limited genomic data coverage in this species. Once available, the identification of proteins from apple and other plants from the Rosaceae family will be greatly facilitated.

Several of the proteins identified in this study belonged to the energy metabolism and defense/detoxification categories. The energy metabolism proteins previously detected

in fruit pulp (Guarino *et al*., 2007) were identified. In terms of the ethylene biosynthesis associated proteins, four major proteins involved in ethylene biosynthesis, namely SAM synthetase, ACO, ethylene receptor and ethylene response sensor protein (section 6.5.4.2), were detected.

In terms of unclassified proteins, the detection of conserved domains to predict their putative functions was carried out using bioinformatics tools. The bioinformatic analysis of the amino acid sequence of eight of the 28 unclassified proteins allowed their functionalisation and categorisation into six functional categories, glycolysis and other carbohydrate metabolism related, defense/detoxification, signal transduction, transcription and translation, cell growth and division, and heat shock proteins.

8.4 COMPARATIVE ANALYSIS BETWEEN THE HIGH AND LOW FRUIT PULP FIRMNESS PHENOTYPES

This part of the thesis reported the establishment of 2D-PAGE proteome maps of TSP extracted from apple fruit pulp of both the high and low firmness phenotype of the seedlings from a 'Golden Delicious' x 'Dietrich' population. After comparative analysis of the proteome maps between the two phenotypes using the PDQuest software, the differentially regulated proteins were identified by MALDI-TOF MS. The aim was to gain insight into the molecular mechanisms influencing fruit firmness.

In this study, differentially expressed proteins were defined as proteins whose abundance was up- or down- regulated by a minimum factor of 2, quantitatively and the variation in expression was statistically significant at 95% during Student *t*-test analysis when comparing the high and low firmness phenotypes. The quantitative analysis and

student *t*-test revealed that 49 proteins were differentially regulated, with 27 and 22 protein spots being up- or down- regulated, respectively, in the high firmness phenotype in comparison to the low firmness phenotype. Out of these, 24 proteins were positively identified and represented seven functional categories: energy related proteins, ethylene biosynthesis, defense/detoxifying enzymes, cell growth/division, signal transduction, cytoskeleton related proteins and unclassified proteins. In the energy category, proteins involved in the citric acid cycle, carbohydrate metabolism, photosynthesis and lipid metabolism were identified.

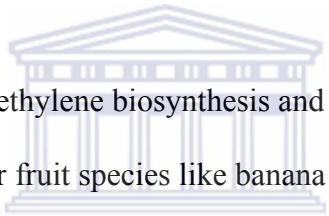
The qualitative analysis showed that five spots were expressed in the high firmness phenotype only. These proteins were positively identified as lissencephaly type-1-like homology motif, cell division inhibitor MinD, major allergen Mal d 1 (AP15), predicted protein. One of the five protein spots could not be positively identified.

In this study, dynein heavy chain isoform and myosin class II heavy chain were the most up-regulated proteins in the high firmness phenotype in comparison to the low firmness phenotype according to the PDQuest analysis.

Previous studies have shown that myosin in complex with actin participates in transmitting internal stress in fruit pulp, provides mechanical strength to cell cortex and is involved in the spatial organization of the cytoplasm (Hussey *et al.*, 2002; Ketelaar *et al.*, 2007). Dynein proteins are known to be essential for the establishment and maintenance of the cell structural integrity during mitosis (Smith, 2002). The up-regulation of these two proteins in the high firmness phenotype could demonstrate their

specific role in the maintenance of cell structure and in the firmness characteristics of apple fruit pulp.

The expression of some proteins was detected as down-regulated in the high firmness phenotype. These include proteins involved in citric acid cycle, and lipid and ethylene biosynthesis. These proteins have been implicated in inducing fruit softening. For example, mitochondrial malate dehydrogenase, which is induced in response to carbonylation (Backhausen *et al.*, 1998), promotes fruit senescence. Also, ethylene biosynthesis associated proteins, which are induced during fruit ripening, have been shown to promote cell wall softening (Brummell and Harpster, 2001).



The correlation between ethylene biosynthesis and fruit ripening has been extensively studied in apple and other fruit species like banana (Jiang *et al.*, 2004), cherry (*Prunus avium*, Gong *et al.*, 2002) and pear (Fu *et al.*, 2007). Hrazdina *et al.* (2003) showed that the activity of hydrolytic enzymes under the control of the ripening hormone ethylene in 'McIntosh' apples induced structural changes in cell walls, thus leading to fruit softening. The same study revealed that a reduced expression of SAM synthetase, an enzyme that catalyze the conversion of methionine to the ethylene precursor SAM, limited apple fruit softening. In this study, SAM synthetase was down-regulated in the high firmness phenotype. This lower expression may have reduced the biosynthesis of ethylene and thus may have a role in apple fruit softening.

The specific role of all the differentially expressed proteins in regulating apple fruit pulp firmness could not be clearly investigated/appreciated. For example, the role of proteins involved in detoxification/defense and the unclassified remains unclear. In

addition, it is important to note that some proteins may be differentially regulated as a consequence of variation in fruit firmness rather than the cause of the firmness status. Further studies involving techniques like gene silencing, for example are necessary to gain insights on the specific roles of these proteins in apple fruit pulp firmness.

8.5 FUTURE WORK

8.5.1 Genomics

In this study, genetic linkage maps of 'Golden Delicious' and 'Dietrich' were constructed and QTLs associated with nine fruit quality traits were identified. However, the identification of QTLs on some linkage groups was not successful because of the low coverage of SSR markers in these linkage groups. In addition, most linkage groups have at least two unlinked groups, which make the total distance spanned by the maps larger than some of the previously published maps. Therefore, in order to link the separate groups and saturate the linkage groups to obtain better genome coverage, more markers like the DArTs could be included to the map. The DArT methodology offers a high multiplexing level, and the assays generate whole-genome fingerprints by scoring the presence *versus* absence of DNA fragments in genomic representations generated from genomic DNA samples through the process of complexity reduction. This technology has successfully been used in rice, barley, wheat (*Triticum aestivum*) and cassava (*Manihot esculenta,* Wenzl *et al.*, 2004) and recently in sorghum (Mace *et al.*, 2008). The DArT fingerprints will be useful for accelerating plant breeding, and for the characterisation and management of genetic diversity in domesticated apple species as well as in their wild relatives. Thus, the saturated genetic linkage maps will enable the detecting of QTLs in short periods of time. Therefore, this allows focusing on the most

crucial factors associated with fruit quality in plant breeding reliability and precision.

8.5.2 Proteomics

The results in this study revealed that the presence of sugars in protein extracts from apple fruit pulp influence the quantification of protein by Bradford assay. Although, the response of BSA was shown not to be affected by pheolic and pectic compounds (Weiss and Bisson, 2001), recently, it has been shown to be influenced by polysaccharides and disaccharides present in the protein extracts and to a lesser extent by monosaccharides (Banick *et al.*, 2009). Therefore, optimisation of Bradford assay for high sugar containing extracts or highly glycosylated proteins can be carried out to improve efficiency in protein quantification assays. In this regard, metabolomics can be employed to detect the sugars present in the apple pulp, and then test their influence on the Bradford assay.

In addition, a large number of proteins were positively identified, including predicted and hypothetical proteins, but these could not be classified into any of the known functional categories. However, with the complete sequencing of the apple fruit transcriptome, further studies can be carried out using bioinformatics tools and databases like ExPASy, InterProScan, TAIR or SIMAP, to characterise the protein sequences using the transcriptome data. In addition, MS/MS or MALDI-TOF/TOF can also be used to obtain more positive protein identities.

Further investigations can be carried out to identify functional roles of the observed firmness regulatory proteins using molecular genetic approaches like gene knockouts and RNA silencing. These approaches will provide the phenotypic roles of the

regulatory proteins identified in this study. Thus, this will allow designing of markers associated with genes coding for firmness regulatory proteins, which can be used for mapping of candidate genes controlling firmness.

8.6 CONCLUDING REMARKS

This study related the first identification of QTLs associated with nine fruit quality traits using the 'Golden Delicious' x 'Dietrich' mapping population. The study also reported the first comprehensive establishment of 2D-PAGE proteome maps of mature apple fruit pulp and their characterisation. In addition, the comparative proteomic analysis between the high and low firmness phenotypes for the identification of the differentially regulated proteins was carried out. Finally, the relation between genetic and proteomic data was established in this study through the identification of regulatory mechanisms related to fruit firmness.

While the genomic study allowed the identification of putative regions on the genome controlling traits of interests, the proteomic study related the identification of proteins whose expression was influenced and/or controlled firmness and thus of the expressed genes encoding for these proteins. These expressed genes may be mapped onto the genetic maps, a process termed candidate gene mapping, and compared against the QTLs detected for firmness. Therefore, this approach bridges data from the proteomic, genomic and phenotypic analyses and provides a link between crop biotechnologies and fruit breeding for the selection of cultivars and/or individuals displaying the desired characteristics.

APPENDICES

Appendix I. List of markers excluded from the genetic linkage map construction during marker order determination.

Markers excluded, their segregation type, χ^2 value, degree of freedom and LOD value are indicated.

Marker *	Segregation type	χ^2 Value	Degree of freedom	LOD
A100	<hkxhk>	3.09	2	3
A116	<nnxnp>	32.12	1	6
A122	<hkxhk>	56.99	1	4
A130	<nnxnp>	10.30	1	3
A131	<efxeg>	18.54	3	2
A135	<nnxnp>	80.88	1	6
A145	<nnxnp>	0.40	1	4
A159	<nnxnp>	40.02	1	3
A161	<nnxnp>	16.25	1	6
A167	<hkxhk>	49.66	1	2
A181	<nnxnp>	19.62	1	4
A195	<nnxnp>	4.97	1	3
A196	<hkxhk>	2.34	1	5
A207	<lmxll>	48.85	1	5
A212	<hkxhk>	33.44	2	4
A217	<hkxhk>	80.00	1	2
A241	<lmxll>	20.78	1	3
A245	<hkxhk>	18.52	2	3
A254	<hkxhk>	74.36	1	4

Appendix I continued

Marker *	Segregation type	χ^2 Value	Degree of freedom	LOD
A265	<abxcd>	51.16	3	2
A266	<efxeg>	75.08	3	6
A301a	<efxeg>	49.61	3	2
A301b	<hkxhk>	18.52	1	2
A305	<hkxhk>	53.33	1	2
A316	<hkxhk>	2.08	1	6
A331	<lmxll>	65.15	1	6
A335	<efxeg>	124.35	3	6
A339	<nnxnp>	3.17	1	2
A34	<abxcd>	180.50	3	4
A343	<hkxhk>	61.36	1	2
A344	<lmxll>	12.20	1	9
A346b	<hkxhk>	22.98	1	2
A372	<lmxll>	11.68	1	6
A376	<hkxhk>	78.33	1	2
A379	<hkxhk>	74.69	1	3
A38	<efxeg>	109.92	3	3
A381	<hkxhk>	72.33	1	2
A398	<efxeg>	72.47	3	6
A40	<nnxnp>	11.07	1	2
A400	<nnxnp>	31.78	1	4
A421	<hkxhk>	2.54	2	3
A425	<efxeg>	85.36	3	9
A466	<nnxnp>	17.56	1	3

Appendix I continued

Marker *	Segregation type	χ^2 Value	Degree of freedom	LOD
A448	<hkxhk>	5.90	2	3
A47	<nnxnp>	5.60	1	5
A490	<nnxnp>	1.28	1	2
A5	<hkxhk>	1.64	2	2
A507	<hkxhk>	0.65	1	3
A508	<efxeg>	10.42	3	3
A510	<hkxhk>	10.45	2	3
A538	<abxcd>	4.04	3	2
A551	<nnxnp>	2.70	1	4
A561a	<abxcd>	8.06	3	3
A561b	<efxeg>	40.49	3	2
A563	<efxeg>	12.59	3	3
A567	<nnxnp>	0.19	1	2
A574	<hkxhk>	9.16	2	2
A583	<hkxhk>	51.20	1	2
A603	<nnxnp>	0.00	1	3
A617	<hkxhk>	6.10	1	3
A638	<abxcd>	16.34	3	6
A664	<hkxhk>	61.33	1	2
A688	<nnxnp>	2.59	1	4
A70	<nnxnp>	8.38	1	3
A718	<lmxll>	0.61	1	3
A73	<nnxnp>	24.78	1	4
A738	<nnxnp>	0.20	1	2

Appendix I continued

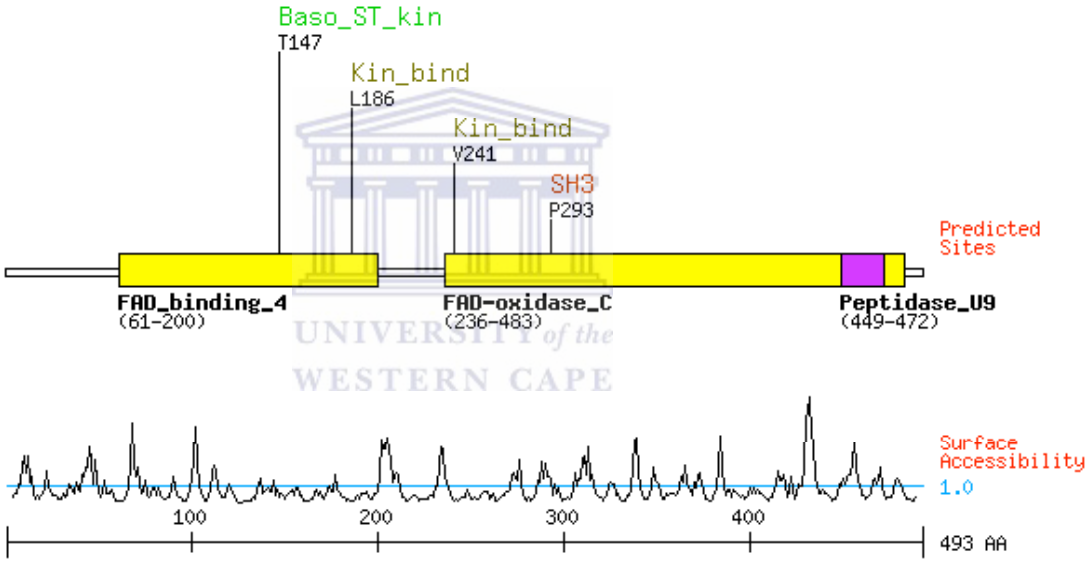
Marker *	Segregation type	χ^2 Value	Degree of freedom	LOD
A742	<nnxnp>	18.89	1	6
A75	<efxeg>	22.32	3	6
A759	<nnxnp>	5.40	1	2
A766	<hkxhk>	48.01	1	4
A768	<lmxll>	36.92	1	3
A769	<abxcd>	13.07	3	3
A774	<hkxhk>	36.00	1	3
A775	<nnxnp>	1.55	1	2
A78	<hkxhk>	11.76	2	4
A780	<hkxhk>	74.36	1	3
A781	<lmxll>	5.30	1	2
A785	<nnxnp>	9.85	1	3
A792	<abxcd>	4.55	3	3
A793	<nnxnp>	2.01	1	3
A80	<hkxhk>	69.10	1	3
A801	<hkxhk>	65.00	1	2
A802	<lmxll>	4.30	1	3
A806	<hkxhk>	48.45	1	3
A813	<hkxhk>	57.88	1	3
A81b	<lmxll>	19.98	1	3
A822	<hkxhk>	69.43	2	2
A826	<hkxhk>	34.42	2	2
A829	<hkxhk>	6.98	2	4
A88	<efxeg>	4.62	3	4

*-Marker accession numbers described in section 2.4.4.2, Table 2.2.

Appendix II. Typical conserved domains search for functional annotation to predicted and hypothetical proteins using proteomic-bioinformatics tools.

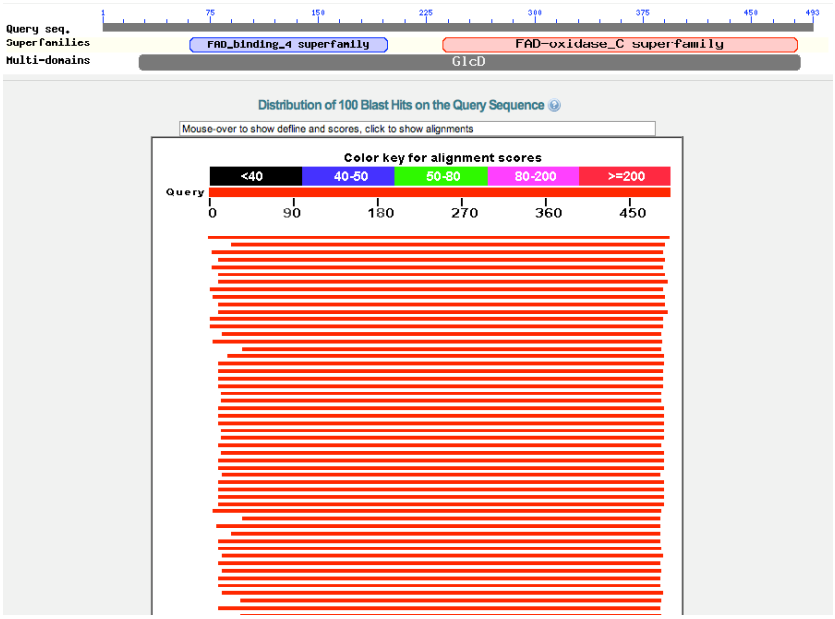
Spot 152/192: 80 for gi|145341236: Predicted protein [*Ostreococcus lucimarinus* CCE9901]

```
     1 MGGAGASDAR ARQWPFSVVR ASDADAFVDA LGGDATRVLT RAEDVKKYSV
    51 DWMGKYVGAS AVVVLPRTTE EVSKVMRHCH ARRIAVVPQG GNTGLVGGGT
   101 PTRDEVVVSL ERMRDIVSID EDAGCAVCEA GVVLEELESA VRARGMTVPL
   151 DLGAKGKCQM GGCVSTNAGG LRLLRYGSLR GSVLGLEVVL PNGDVLDLVR
   201 TLRKDNTGYD LKQLFIGAEG TLGVVTKVAI STPRAPRSVN VALFGLESFA
   251 KCVEMLKLAR GLLGEILSAY EFFDRESLDL VLAQLSGTRD PLPGKPCEFY
   301 VVIETSGSDA KHDTAKLDAF LNVVKSQRIV VDGVVGRDEK HAFALWTLRE
   351 RISVALKYAG AVYKYDLSLP TARMYNLVVV LRDRLRPMFG SRVKVLGYGH
   401 AGDGNLHLNV SCAEYDDAIE RAIEPFVYEY TRDERGSVSA EHGLGVMKAE
   451 EIHYSKDAKA VELMATMKRA LDPFGIMNPY KVLPAAAVGL SKL
```

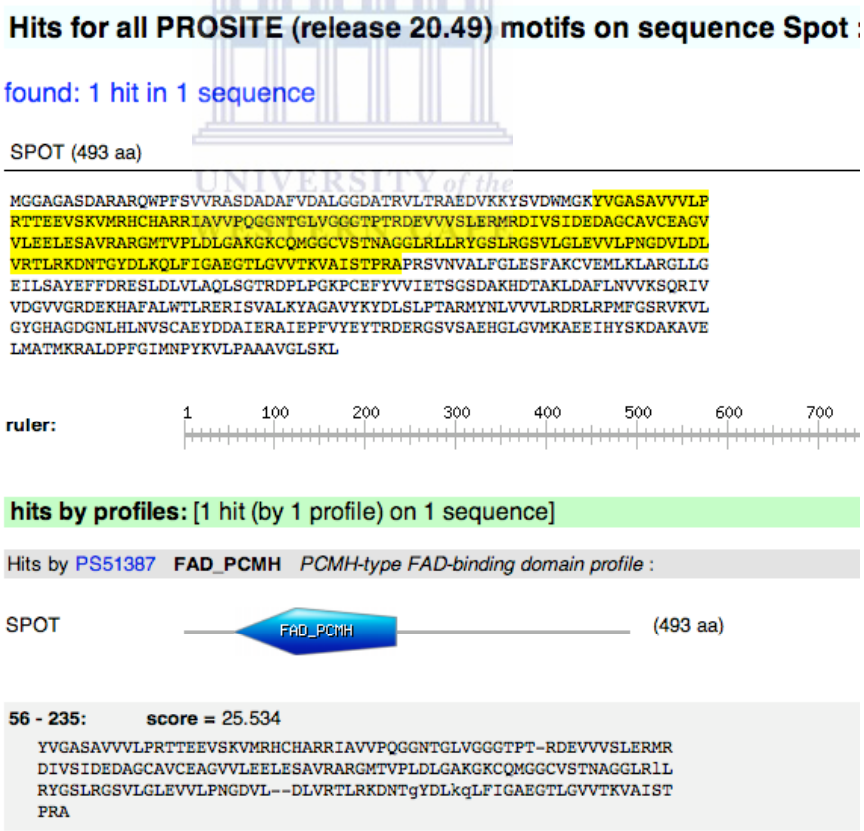


http://scansite.mit.edu/cgi-bin/motifscan_seq

Appendix II continued:



http://blast.ncbi.nlm.nih.gov/Blast.cgi



http://au.expasy.org/cgi-bin/prosite/ScanView.cgi

REFERENCES

Abbott, J. A., Saftner, R. A., Gross, K. C., Vinyard, B. T. and Janick, J. (2004). Consumer evaluation and quality measurement of fresh-cut slices of 'Fuji,' 'Golden Delicious,' 'GoldRush,' and 'Granny Smith' apples. *Postharvest Biology and Technology, 33*, 127-140.

Abbott, J. A., Klein, J. D., Campbell, T. A., Conway, W. S., Sams and C.E (2000). Sensory and firmness measurements of calcium- and heat-treated apples. *Journal of Texture Studies, 31*, 109-121.

Abbott, A. (1999). A post-genomic challenge: learning to read patterns of protein synthesis. *Nature, 402,* 715-720.

Abbott, J. A., Watada, A. E. and Massie, D. R. (1984). Sensory and instrument measurement of apple texture. *Journal of American Society for Horticultural Science, 109*, 221-228.

Abeles, F. B. and Biles, C. L. (1991). Cellulase activity in developing apple fruits. *Scientia Horticulturae, 47*, 77-87.

Adam-Blondon, A., Roux, C., Claux, D., Butterlin, G., Merdinoghi, D. and This, P. (2004). Mapping 245 SSR markers on *Vitis vinifera* genome: a tool for grape genetics. *Theoretical and Applied Genetics, 109*, 1017-1027.

Aebersold, R. and Mann, M. (2003). Mass spectrometry based proteomics. *Nature, 422*, 198-207.

Agarwal, K. K., Saluja, D. and Sachar, R. C. (1993). Phosphorylation of RuBisCO in *Cicer arietium*: non-phosphoprotein nature of RuBisCO in *Nicotiana tabacum*. *Phytochemistry, 34*, 329-335.

Agrawal, G. K. and Thelen, J. J. (2006). Large scale identification and quantification profiling of phosphoproteins expressed during seed filling in oilseed rape. *Molecular and Cell Proteomics, 5,* 2044-2059.

Ahn, S. N., Bollich, C. N., McClung, A. M. and Tanksley, S. D. (1993). RFLP analysis of genomic regions associated with cooked-kernel elongation in rice. *Theoretical and Applied Genetics, 87*, 27-32.

Alavoine, F., Crochon, M. and Bouillon, C. (1990). Practical methods to estimate taste quality of fruit: How to tell it to the consumer. *Acta Horticulturae, 259*, 61-68.

Albright, S. R. and Tjian, R. (2000). TAFs revisited more data reveal new twists and confirm old ideas. *Gene, 242*, 1-13.

Aldwinckle, H. S., Lamb, R. C. and Gustafson, H. L. (1977). Nature and inheritance of resistance to *Gymnosporangium juniperi-virginianae* in apple cultivars. *Phytopathology, 67*, 259-266.

Alexander, L. and Grierson, D. (2002). Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. *Journal of Experimental Botany, 53*, 2039-2055.

Allana, A. C., Hellensa, R. P. and Lainga, W. A. (2008). MYB transcription factors that colour our fruit. *Plant Science, 13*, 99-102.

Alston, F. H. and Briggs, J. B. (1968). Resistance to *Sappaphis devecta* (WIk.) in apple. *Euphytica, 17*, 468-472.

Alston, F. H. and Briggs, J. B. (1970). Inheritance of hypersensitivity to rosy apple aphid *Dysaphis plantaginea* in apple. *Canadian Journal of Genetics and Cytology, 12*, 257-258.

Alston, F. H. and Briggs, J. B. (1977). Resistance genes in apple and biotypes of *Dysaphis devecta*. *Annual Review of Applied Biology, 97*, 75-81.

Ames, G. K., Johnson, D. T. and Rom, R. C. (1984). The effect of European red mite feeding on the quality of 'Miller Sturdeespur' apple. *Journal of American Society for Horticulture Science., 109*, 834-837.

Antofie, A., Lateur, M., Oger, R., Patocchi, A., Durel, C. E. and Van de Weg, W. E. (2007). A new versatile database created for geneticists and breeders to link molecular and phenotypic data in perennial crops: the AppleBreed DataBase. *Bioinformatics, 23*, 882-891.

Apple journal variety descriptions: http://www.applejournal.com.

Asada, K. and Takahashi, M. (1987). Production and scavenging of active oxygen in photosynthesis. In *Photoinhibition*, Kyle, D. J., Osmond, C. B. and Arntzen, C. J. (Eds.). Elsevier, Amsterdam, 227-287.

Azuma, H., García-Franco, J. G., Rico-Gray, V. and Thien, L. B. (2001). Molecular phylogeny of the Magnoliaceae: the biogeography of tropical and temperate disjunctions. *American Journal of Botany, 88*, 2275-2285.

Babojelic, M. S., Ivancic, K., Druzic, J., Kovac, A. and Voca, S. (2007). Chemical and sensory characteristics of three apple cultivars (*Malus x domestica* Borkh.). *Agriculturae Conspectus Scientificus, 72*, 317-322.

Babu, R., Nair, S. K., Kumar, A., Venkatesh, S., Sekhar, J. C., Singh, N. N. *et al.* (2005). Two-generation marker-aided backcrossing for rapid conversion of normal maize lines to quality protein maize (QPM). *Theoretical and Applied Genetics, 111*, 888-897.

Backhausen, J. E., Vetter, S., Baalmann, E., Kitzmann, C. and Scheibe, R. (1998). NAD-dependent malate dehydrogenase and glyceraldehyde 3-phosphate dehydrogenase isoenzymes play an important role in dark metabolism of various plastid types. *Planta, 205*, 359-366.

Baker, B., Zambriski, P., Staskawicz, B. and Dinesh-Kumar, S. P. (1997). Signalling in plant-microbe interactions. *Science, 276,* 726-733.

Baldi, P., Patocchi, A., Zini, E., Toller, C., Velasco, R. and Komjane, M. (2004). Cloning and linkage mapping of resistance gene homologues in apple. *Theoretical and Applied Genetics, 109*, 231-239.

Ban, Y., Honda, C., Hatsuyama, Y., Igarashi, M., Bessho, H. and Moriguchi, T. (2007). Isolation and functional analysis of a *MYB* transcription factor gene that is a key regulator for the development of red colouration in apple skin. *Plant and Cell Physiology, 48*, 958-970.

Banick, S. P., Pal, S., Ghorai, S., Chowdhurry, S. and Khowala, S. (2009). Interference of sugars in the Coomassie Blue G dye binding assay of proteins. *Analytical Biochemistry, 386*, 113-115.

Barbier-Brygoo, H. and Joyard, J. (2004). Focus on plant proteomics. *Plant Physiology and Biochemistry, 42*, 913-917.

Barel, G. and Ginzberg, I. (2008). Potato skin proteome is enriched with plant defense components. *Journal of Experimental Botany, 59*, 3347-3357.

Barraclough, D., Obenland, D., Laing, W. and Carroll, T. (2004). A general method for two-dimensional protein electrophoresis of fruit samples. *Postharvest Biology and Technology, 32*, 175-181.

Barthet, M. M. and Hilu, K. W. (2007). Expression of *mat*K: Functional and evolutionary implications. *American Journal of Botany, 94*, 1402-1412.

Basel, B. (1997). Structures and functions of annexins. *Cellular and Molecular Life Sciences, 53*, 546-553.

Batlle, I. and Alston, F. H. (1996). Genes determing leucine aminopeptidase and mildew resistance from the ornamental apple, 'White Angel'. *Theoretical and Applied Genetics, 93,*179-182.

Battle, L., Alston, F. H. and Evans, K. M. (1995). The use of the isoenzymic marker gene Got-1 in the recognition of incompatibility S alleles in apple. *Theoretical and Applied Genetics, 90*, 303-306.

Baxter, C. J., Sabar, M., Quick, W. P. and Sweetlove, L. J. (2005). Comparison of changes in fruit gene expression in tomato introgression lines provides evidence of genome-wide transcriptional changes and reveals links to mapped QTLs and described traits. *Journal of Experimental Botany, 56*, 1591-1604.

Becker, J., Vos, P., Kuiper, M., Salamini, F. and Heun, M. (1995). Combined mapping of AFLP and RFLP markers in barley. *Molecular Genome and Genetics, 249*, 65-73.

Becraft, P. W. (2002). Receptor kinase signaling in plant development. *Annual Review Cell Development Biology, 18*, 163-192.

Beecher, B., Bowman, J., Martin, J. M., Bettge, A. D., Morris, C. F., Blake, T. K. *et al.* (2002). Hordoindolines are associated with a major endosperm-texture QTL in barley (*Hordeum vulgare*). *Genome, 45*, 584-591.

Belfanti, E., Silfverberg-Dilworth, E., Tartarini, S., Patocchi, A., Barbieri, M., Zhu, J., Vinatzer, B. A., Gianfranceschi, L., Gessler, C. and Sansavini, S. (2004). The HcrVf2 gene from a wild apple confers scab resistance to a transgenic cultivated variety. *Proceedings of the National Academy of Sciences of the United States of America, 101,* 886-890.

Ben-Chaim, A., Paran, I., Grube, R. C., Jahn, M., van Wijk, R. and Peleman, J. (2001). QTL mapping of fruit-related traits in pepper (*Capsicum annuum*). *Theoretical and Applied Genetics, 102*, 1016-1028.

Bénaouf, G. and Parisi, L. (2000). Genetics of host-pathogen relationships between *Venturia inaequalis* races 6 and race 7 and *Malus* species. *Phytopathology, 90*, 236-242.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research, 27*, 573-580.

Beranova-Giorgianni, S. (2003). Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: strengths and limitations. *Trends in Analytical Chemistry, 22*, 273-281.

Berkemeyer, M., Scheibe, R. and Ocheretina, O. (1998). A novel, non-redox-regulated NAD-dependent malate dehydrogenase from chloroplasts of *Arabidopsis thaliana* L. *Journal of Biological Chemistry, 273*, 27927-27933.

Berlett, B. S. and Stadtman, E. R. (1997). Protein oxidation in aging, disease and oxidative stress. *Journal of Biology and Chemistry, 272,* 20313-20316.

Beuning, L., Bowen, J., Persson, H., Barraclough, D., Bulley, S., Macrae, E. (2004). Characterisation of Mal d 1-related genes in *Malus*. *Plant Molecular Biology, 55*, 369-388.

Bevan, M., Bancroft, I., Bent, E., Love, K., Goodman, H., Dean, C. *et al.* (1998). Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature, 391,* 485-488.

Beveridge, C. A., Mathesius, U., Rose, R. J. and Gresshoff, P. M. (2007). Common regulatory themes in meristem development and whole-plant homeostasis. *Plant Biology, 10*, 44-51.

Bhadauria, V., Zhao, W., Wang, L., Zhang, Y., Liu, J., Yang, J. *et al.* (2007). Advances in fungal proteomics. *Microbiological Research, 162*, 193-200.

Bianco, L., Lopez, L., Scalone, A. G., Di Cardi, M., Desiderio, A., Benvenuto, E. *et al.* (2009). Strawberry proteome characterisation and its regulation during fruit ripening and in different genotypes. *Journal of Proteomics, 72,* 586-607.

Blackstock, W. P. and Weir, M. P. (1999). Proteomics: quantitative and physical mapping of cellular proteins. *Biotechnology, 17*, 121-126.

Bleecker, A. B. and Kende, H. (2000). Ethylene: a gaseous signal molecule in plants. *Annual Review of Cell and Developmental Biology, 16*, 1-40.

Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S. F. and Koonin, E. V. (1997). A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *The Federation of American Societies for Experimental Biology Journal, 11*, 68-76.

Botstein, B., White, R. L., Skolnick, M. and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics, 32*, 314-331.

Bouchra, C., El-Otmani, M., Luro, F., Spairi, I. and Tijane, M. (2008). Changes of fruit quality traits in 7 clementine selections during maturation. *Journal of Environmental, Agricultural and Food Chemistry, 7*, 2565-2577.

Bowsher, C. G., Boulton, E. L., Rose, J., Nayagam, S. and Emes, M. J. (1992). Reductant for glutamate synthase is generated by the oxidative pentose phosphate pathway. *The Plant Journal, 2*, 893-898.

Boyd, L. A., Smith, P. H., Foster, E. M. and Brown, J. K. M. (1995). The effects of allelic variation at the *Mla* resistance locus in barley on the development of *Erysiphe graminis* f. sp. *hordei* and host responses. *The Plant Journal, 7,* 959-968.

Bradford, M. M. (1976). A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry, 72*, 248-254.

Bradshaw, B. D. J. and Stettler, R. F. (1995). Molecular genetics of growth and development in *Populus*: Mapping QTLs with large effects on growth, form and phenology traits in a forest tree. *Genetics, 139*, 963-973.

Bramlage, J. W. (1983). Measuring fruit firmness with a penetrometer. *Post Harvest Pomology, 1*, 1-13.

Bray, S., Musisi, H. and Bienz, M. (2005). *Bre*1 is required for Notch signaling and histone modification. *Development of the Cell, 8,* 279-286.

Breiteneder, H. and Ebner, C. (2000). Molecular and biochemical classification of plant-derived food allergens. *Journal of Allergy and Clinical Immunology, 106*, 27-36.

Brini, F., Hanin, M., Lumbreras, V., Irar, S., Page, M. and Masmoudi, K. (2006). Functional characterization of DHN-5, a dehydrin showing a differential phosphorylation pattern in two Tunisian durum wheat (*Triticum durum* Desf.) varieties with marked differences in salt and drought tolerance. *Plant Science, 172*, 20-28.

Brouk, B. (1975) *Plants consumed by man.* Academic press, London, UK, pp. 479.

Brown, A. G. and Harvey, D. M. (1971). Sweetness and acidity in the cultivated apple. *Euphytica, 20*, 68-80.

Brown, G. R., Kadel, E. E., Bassoni, D. L., Kiehne, K. L., Temesgen, B., van Buijtenen, J. P. *et al.* (2001). Anchor reference loci in loblolly pine (*Pinus taeda* L.) for intergrating pine genomics. *Genetics, 159*, 799-809.

Brown, G. S., O' Loughlin, J. and Jotic, P. (1998). A comparison of fruit maturity and quality of four strains of 'Fuji' apples. *Acta Horticulturae, 464*, 491-451.

Brown, S. K. (1995). Genetic Improvement of Apple: The Roles of Plant Breeding and Biotechnology. *New York Fruit Quaterly*, *3*, 2-5.

Brown, T. A. (2006). *Gene cloning and DNA analysis*. *An introduction*. 5th Edition. Wiley-Blackwell Publishing Ltd, United Kingdom, pp. 386.

Brummell, D. A. and Harpster, M. H. (2001). Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Molecular Biology Reporter, 47*, 311-340.

Bundock, P. and Hooykaas, P. (2005). An *Arabidopsis* hAT-like transposase is essential for plant development. *Nature, 436*, 282-284.

Burch-Smith, T. M. and Dinesh-Kumar, S. P. (2007). The functions of plant TIR domains. *Science*, *401*, 46-51.

Burr, B. and Burr, F. A. (1991). Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. *Trends in Genetics, 7*, 55-60.

Bus, V., Gardiner, S., Basset, H., Ranatunga, C. and Rikkerink, E. (2000). Marker assisted selection for pest and disease resistance in the New Zealand apple breeding program. *Acta Horticulturae, 538*, 541-547.

Bus, V. G. M., Chagné, D., Bassett, H. C. M., Bowatte, D., Calenge, F., Celton, J.-M. *et al.* (2008). Genome mapping of three major resistance genes to woolly apple aphid (*Eriosoma lanigerum* Hausm.). *Tree Genetics and Genomics, 4*, 233-236.

Bus, V. G. M., Laurens, F. N. D., Van de Weg, W. E., Rusholme, R. L., Rikkerink, E. H. A., Gardiner, S. E. *et al.* (2005a). The *Vh8* locus of a new gene-for-gene interaction between *Venturia inaequalis* and the wild apple *Malus sieversii* is closely linked to the *Vh2* locus in *Malus pumila* R12740-7A. *Phytopathology, 166*, 1035-1049.

Bus, V. G. M., Rikkerink, E. H. A., Van de Weg, W. E., Rusholme, R. L., Gardiner, S. E., Bassett, H. C. M. *et al.* (2005b). The *Vh2* and *Vh4* scab resistance genes in two differential hosts derived from Russian apple R12740-7A map to the same linkage group of apple. *Molecular Breeding, 15*, 103-116.

Calenge, F., Drouet, D., Denance, C., Van de Weg, W. E., Brisset, M. N., Paulin, J. P. *et al.* (2005). Identification of a major QTL together with several minor additive or epistatic QTLs for resistance to fire blight in apple in two related progenies. *Theoretical and Applied Genetics, 111*, 128-135.

Calenge, F., Faure, A., Goerre, M., Gebhardt, C., Van de Weg, W. E., Parisi, L. *et al.* (2004). Quantitative trait loci (QTL) analysis reveals both broad-spectrum and isolate-specific QTL for scab resistance in an apple progeny chanllenged with eight isolates of *Venturia inaequalis*. *Phytopathology, 94*, 370-379.

Cao, S., Zhang, Q., Zhu, Z., Guo, J., Chen, Y. and Xue, H. (2008). Preliminary proteomics analysis of the total proteins of flower bud induction of apple trees. *Frontiers of Agriculture, 2*, 467-473.

Capitani, G., Hohenester, E., Feng, L., Storici, P., Kirsch, J. F. and Jansonius, J. N. (1999). Structure of 1-aminocyclopropane-1-carboxylate synthase, a key enzyme in the biosynthesis of the plant hormone ethylene. *Journal of Molecular Biology, 294*, 745-756.

Carpentier, S. C., Panis, B., Vertommen, A., Swennen, R., Sergeant, K., Renaut, J. *et al.* (2008). Proteome analysis of non-model plants: a challenging but powerful approach. *Mass Spectrometry Reviews, 27*, 354-377.

Carpentier, S. C., Witters, E., Laukens, K., Deckers, P., Swennen, R. and Panis, B. (2005). Preparation of protein extracts from recalcitrant plant tissues: an evaluation of different methods for two-dimensional gel electrophoresis analysis. *Proteomics, 5*, 2497-2507.

Carrari, F. and Fernie, A. R. (2006). Metabolic regulation underlying tomato fruit development. *Journal of Experimental Botany, 57*, 1883-1897.

Caruso, A., Morabito, D., Delmotte, F., Kahlem, G. and Carpin, S. (2002). Dehydrin induction during drought and osmotic stress in *Populus*. *Plant Physiology and Biochemistry, 40*, 1033-1042.

Celton, J.-M., Tustin, D. S., Chagné, S. E. and Gardiner, S. E. (2009). Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. *Tree Genetics and Genomics, 8*, 556-562.

Cevik, V. and King, G. J. (2000). Molecular genetic analysis of the *Sd*1 aphid resistance locus in *Malus. Acta Horticulturae, 538*, 553-559.

Cevik, V. and King, G. J. (2002). High-resolution genetic analysis of the *Sd*-1 aphid resistance locus in *Malus* spp. *Theoretical and Applied Genetics, 105,* 346-354.

Chagné, D., Carlisle, C. M., Blond, C., Volz, R. K., Whitworth, C. J., Oraguzie, N. C. *et al*. (2007). Mapping a candidate gene (*MdMYB10)* for red flesh and foliage colour in apple. *BMC Genomics, 8,* 212-222.

Chagné, D., Chaumeil, P., Ramboer, A., Collada, C., Guevara, A., Cervera, M. T. *et al.* (2004). Cross-species transferability and mapping of genomic and cDNA SSR's in pines. *Theoretical and Applied Genetics, 109*, 1204-1214.

Chahal, G. S. and Gosal, S. S. (2002). *Principles and Procedures of Plant Breeding: Biotechnology and Conventional Approaches*. Alpha Science, Pangbourne Harrow, United Kingdom, pp. 604.

Chandler, P. M. and Robertson, M. (1994). Gene expression regulated by abscisic acid and its relation to stress tolerance. *Annual Review of Plant Physiology and Plant Molecular Biology, 45*, 113-141.

Chen, Z. and Gallie, D. (2006). Dehydroascorbate reductase affects leaf growth, development and function. *Plant Physiology, 142*, 775-787.

Cheng, F. S., Weeden, N. F. and Brown, S. K. (1996). Identification of co-dominant RAPD markers tightly linked to fruit skin colour in apple. *Theoretical and Applied Genetics, 93*, 222-227.

Chevalier, F., Martin, O., Rofidal, V., Devauchelle, A. D., Barteau, S., Sommerer, N. *et al.* (2004). Proteomic investigation of natural variation between *Arabidopsis* ecotypes. *Proteomics, 4*, 1372-1381.

Chiou, C. and Yeh, K. (2008). Differential expression of MYB gene (OgMYB1) determines colour patterning in floral tissue of *Oncidium Gower Ramsey*. *Plant Molecular Biology, 66*, 379-388.

Cho, R. J., Mindrinos, M., Richards, D. R., Sapolsky, R. J., Anderson, M., Drenkard, E. *et al.* (1999). Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nature Genetics, 23*, 203-207.

Chong, C., Changjie, X., Xian, L., Ferguson, I. and Kunsong, C. (2006). Accumulation of lignin in relation to change in activities of lignification enzymes in loquat fruit flesh after harvest. *Postharvest Biology and Technology, 40*, 163-169.

Churchill, G. A. and Doerge, R. W. (1994). Empirical threshold values for quatitative trait mapping. *Genetics, 138*, 963-971.

Ciardi, J. A., Tieman, D. M., Lund, S. T., Jones, J. B., Stall, R. E. and Klee, H. J. (2000). Response to *Xanthomonas campestris* pv. vesicatoria in tomato involves regulation of ethylene receptor gene expression. *Plant Physiology, 123*, 81–92.

Civello, P. M., Martinez, G. A., Chaves, A. R. and Anon, M. C. (1995). Peroxidase from strawberry fruit (*Fragaria ananassa* Duch.): Partial purification and determination of some properties. *Journal of Agriculture and Food Chemistry, 43*, 2596-2601.

Clare, A. (2005). *Integration of genomic and phenotypic data*. *In*: Azuaje, F. and Dopazo, J. (Eds), Data analysis and visualization in genomics and proteomics, Wiley, London,UK, 83-98.

Clark, G. B., Sessions, A., Eastburn, D. J. and Roux, S. J. (2001). Differential expression of members of the annexin multigene family in *Arabidopsis*. *Plant Physiology, 126*, 1072-1084.

Clauser, K. R., Baker, P. and Burlingame, A. L. (1999). Role of accurate mass measurement (+/–10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Analytical Chemistry, 71*, 2871–2882.

Cliff, M., Sandford, K., Wismer, W. and Hampson, C. (2002). Use of digital images for evaluation of factors responsible for visual preference of apples by consumers. *Hortscience, 37*, 1127-1131.

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

Collins, F. S., Brooks, L. D. and Chakravarti, A. (1998). A DNA polymorphism discovery resource for research on human genetic variation. *Genome Research, 8*, 1229-1231.

Conner, P. J., Brown, S. K. and Weeden, N. F. (1998). Molecular-marker analysis of quantitative traits for growth and development in juvenile apple trees. *Theoretical and Applied Genetics, 96*, 1027-1035.

Coombe, B. G. and McCarthy, M. G. (2000). Dynamics of grape berry growth and physiology of ripening. *Australian Journal of Grape and Wine research, 6*, 131-135.

Copenhaver, G. P., Browne, W. E. and Preuss, D. (1998). Assaying genome-wide recombination and centromere functions with *Arabidopsis* tetrads. *Genetics, 95*, 247-252.

Cordell, S. C., Anderson, R. E. and Löwe, J. (2001). Crystal structure of the bacterial cell division inhibitor MinC. *The European Molecular Biology Organisation Journal, 20*, 2454-2461.

Cordina, O., Banroquesa, J., Tannera, N. K. and Lindera, P. (2006). The DEAD-box protein family of RNA helicases. *Gene, 367*, 17-37.

Costa, F., Stella, S., Van de Weg, W. E., Guerra, W., Cecchinel, M., Dallavia, J. *et al.* (2005). Role of the genes *Md-ACO1* and *Md-ACS1* in ethylene production and shelf life of apple (*Malus x domestica* Borkh). *Euphytica, 141*, 181-190.

Costa, F., Van de Weg, W. E., Stella, S., Dondini, L., Pratesi, D., Musacchi, S. *et al.* (2008). Map position and functional allelic diversity of Md-Exp7, a new putative expansin gene associated with fruit softening in apple (*Malus* × *domestica* Borkh.) and pear (*Pyrus communis*). *Tree Genetics and Genomics, 4*, 575-586.

Crassweller, R. M., Walker, J. and Shewfelt, R. L. (1984). Colour evaluation of seventeen strains of Delicious. *Fruit Variety Journal, 39*, 21-24.

Crowhurst, R. N., Allan, A. C., Atkinson, R. G., Beuning, L. L., Davy, M., Friel, E. *et al.* (2005). *The HortResearch apple EST database: A resource for apple genetics and functional genomics.* Poster presented at the Plant and Animal Genomes XIII Conference, San Diego.

Currie, A. J., Janeshanandam, C., Noiton, D. A., Garrick, D., Shelbourne, C. J. A. and Oraguzie, J. (2000). Quantitative evaluation of apple (*Malus* × *domestica* Borkh.) fruit shape by principal component analysis of Fourier descriptors. *Euphytica, 111*, 221-227.

Daillant-Spinnler, B., MacFie, J. H., Beyts, P. K. and Hedderley, D. (1996). Relationships between perceived sensory properties and major preference directions of 12 varieties of apples from the southern hemisphere. *Food Quality and Preference, 7*, 113-126.

Damerval, C., Devienne, D., Zivy, M. and Thiellement, H. (1986). Technical improvements in 2D electrophoresis increase level of genetic variation detected in wheat seedling proteins. *Electrophoresis, 7,* 52-54.

Damodarana, S., Wood, T. D., Nagarajanc, P. and Rabina, R. A. (2007). Evaluating peptide mass fingerprinting-based protein identification. *Genomics, Proteomics and Bioinformatics, 5*, 152-157.

Dangl, J. L. and Jones, J. D. (2001). Plant pathogens and integrated defense responses to infection. *Nature, 411*, 826-833.

Darvasi, A., Weinreb, A., Minke, V., Weller, J. I. and Soller, M. (1993). Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics, 134*, 943-951.

Dayton, D. F. (1977). Genetic immunity to apple mildew incited by *Podosphaera leucotricha*. *Horticulturae Science, 12*, 225-226.

Dayton, D. F. and Williams, E. B. (1968). Independent genes in *Malus* for resistance to *Venturia inaequalis*. *Proceedings of the American Society of Horticultural Science, 92*, 89-94.

Deak, M., Horvath, G.V., Davletova, S., Torok, K., Sass, L., Vass, I. *et al*. (1999). Plants ectopically expressing the iron-binding protein, ferritin, are tolerant to oxidative damage and pathogens. *National Biotechnology, 17*, 192-196.

DeEll, J. R., Khanizadeh, S., Saad, F. and Feree, D. C. (2001). Factors affecting apple fruit firmness. A review. *Journal of the American Pomology Society, 55,*8-22.

De Salvador, F. R., Fisichella, M. and Fontanari, M. (2006). Correlations between fruit size and fruit quality in apples with high and standard crop load levels. *Journal of Fruit and Ornamental Plant Research, 14*, 113-122.

Deytieux, C., Geny, L., Lapaillerie, D., Claverol, S., Bonneu, M. and Done` che, B. (2007). Proteome analysis of grape skins during ripening. *Journal of Experimental Botany, 58*, 1851-1862.

DFPT (2009). Key deciduous fruit statistics, *Optimal Agricultural Business Systems*, 84.

Diakou, P., Svanella, L., Raymond, P., Gaudillère, J.-P. and Moing, A. (2000). Phosphoenolpyruvate carboxylase during grape berry development: Protein level, enzyme activity and regulation. *Australian Journal of Plant Physiology, 27,* 221-229.

Dirlewanger, E., Cosson, P., Tavaud, M., Aranzana, J., Poizat, C., Zanetto, A. *et al.* (2002). Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). *Theoretical and Applied Genetics, 105*, 127-138.

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

Dobrzanski, B. and Rybzynski, R. (2002). Colour change of apple as a result of storage, shelflife, and bruising. *International Agrophysics, 16*, 268-272.

Doerge, R. W. and Rebaï, A. (1996). Significance thresholds for QTL interval mapping tests. *Heredity, 77*, 459-464.

Doganlar, S., Frary, A., Daunay, M.-C., Lester, R. N. and Tanksley, S. D. (2002). Conservation of gene function in the Solanaceae as revealed by comparative mapping of domestication traits in eggplant. *Genetics, 161*, 1713-1726.

Dolenc, K. and Stampar, K. (1998). Determining the quality of different cherry cultivars using the HPLC method. *Acta Horticulturae, 468*, 705-712.

Drysdale, R. (2001). Phenotypic data in FlyBase. *Briefings in Bioinfomatics, 2*, 68-80.

Du, Z. and Bramlage, W. J. (1995). Superoxide dismutase activities in senescing apple fruit (*Malus x domestica* Borkh.). *Journal of Food Science, 59*, 581-584.

du Jardin, P., Portetelle, D., Harvengt, L., Dumont, M. and Wathelet, B. (1994). Expression of intron-encoded maturase-like polypeptides in potato chloroplasts. *Current Genetics, 25*, 158-163.

Dudits, D., Gyorgyey, J., Bogre, L. and Bako, L. (1995). *In vitro* embryogenesis in plants. *In*: Thorpe, T. A. (Ed.), *Embryogenesis in Plants.* Kluwer Dordrecht, The Netherlands, 267-308.

Dugo, M. L., Satovic, Z., Millan, T., Cubero, J. I., Rubiales, D., Cabrera, A. *et al.* (2005). Genetic mapping of QTLs controlling horticultural traits in diploid roses. *Theoretical and Applied Genetics, 111*, 511-520.

Duke, J. (1983). *Handbook of Energy Crops*. Unpublished.

Durel, C. E., Parisi, L., Laurens, F., Van de Weg, W. E., Liebhard, R. and Jourjon, M. F. (2003). Genetic dissection of partial resistance to race 6 of *Venturia inaequalis* in apple. *Genome, 46*, 224-234.

Durel, C. E., van der Weg, W. E., Venisse, J. S. and Parisi, L. (2000). Localization of a major gene for apple scab resistance on the European gene map of the Prima x Fiesta cross. *OILB/WPRS Bull, 23,* 245-248.

Eccher, T., Hajnajari, H., Di Lella, S. and Elli, A. (2008). Gibberellin content of apple fruit as affected by genetic and environment factors. *Acta Horticulturae, 774,* 221-228.

Eckert, K. A. and Kunkel, T. A. (1991). DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods Application, 1*, 7-24.

Edamatsu, M. and Toyoshima, Y. Y. (2003). Fission yeast synaptobrevin is involved in cytokinesis and cell elongation. *Biochemical and Biophysical Research, 301*, 641-645.

Emes, R. D. and Ponting, C. P. (2001). A new sequence motif linking lissencephaly, Treacher Collins and oral–facial–digital type 1 syndromes, microtubule dynamics and cell migration. *Human Molecular Genetics, 10*, 2813-2820.

Ems, S. C., Morden, C. K., Dixon, K. H., Wolfe, K. H., DePamphilis, C. W. and Palmer, J. D. (1995). Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant *Epifagus virginiana. Plant Molecular Biology, 29,* 721-733.

Erdin, N., Tartarini, S., Broggini, G. A. L., Gennari, F., Sansavini, S., Gessler, C. *et al*. (2006). Mapping of the apple scab resistance gene *Vb. Genome, 49,* 1238-1245.

Espley, R. V., Brendolise, C., Chagné, D., Kutty-Amma, S., Green, S., Volz, R. *et al.* (2009). Multiple repeats of a promoter segment causes transcription factor autoregulation in red apples. *Plant Cell, 21*, 168-183.

Etienne, C., Moing, A., Dirlewanger, E., Raymond, P., Monet, R. and Rothan, C. (2002). Isolation and characterization of six peach cDNAs encoding key proteins in organic acid metabolism and solute accumulation: involvement in regulating peach fruit acidity. *Physiology of Plants, 114*, 259-270.

Evans, K. M. and James, C. M. (2003). Identification of SCAR markers linked to *Pl-w* mildew resistance in apple. *Theoretical and Applied Genetics, 106,* 1178-1183.

Fankhauser, C. (2000). Phytochromes as light-modulated protein kinases. *Seminars in Cell Developmental Biology, 11*, 467-473.

Fankhauser, C., Yeh, K.-C., Lagarias, J. C., Zhang, H., Elich, T. D. and Chory, J. (1999). PKS1, a substrate phosphorylated by phytochrome that modulates light signaling in *Arabidopsis*. *Science, 284*, 1539-1541.

FAO (2008). *Agricultural products: Fresh fruits and vegetables.* 1st Edition, United Nations, Rome, pp. 285.

Faurobert, M., Mihr, C., Bertin, N., Pawlowski, T., Negroni, L., Sommerer, N. *et al.* (2007). Major proteome variations associated with cherry tomato pericarp development and ripening. *Plant Physiology, 143*, 1327-1346.

Fernández-Trujillo, J. P., Nock, J. F., Kupferman, E. M., Brown, S. K. and Watkins, C. B. (2003). Peroxidase activity and superficial scald development in apple fruit. *Journal of Agriculture and Food Chemistry, 51*, 7182-7186.

Fraser, P. D., Truesdale, M. R., Bird, C. R., Schuch, W. and Bramley, P. M. (1994). Carotenoid biosynthesis during tomato fruit-development. *Plant Physiology, 105*, 405-413.

Freialdenhoven, A., Scherag, B., Hollricher, K., and Collinge, D. B. (1994). *Nar-1* and *Nar-2*, two loci required for *Mla12*-specified race-specific resistance to powdery mildew fungi. *The Plant Cell, 6*, 983-994.

Fu, L., Cao, J., Li, Q., Lin, L. and Jiang, W. (2007). Effect of 1-methylcyclopropene on fruit quality and physiological disorders in Yali pear (*Pyrus bretschneideri* Rehd.) during storage. *Food Science and Technology International, 13*, 49-54.

Gallardo, F., Galvez, S., Gadal, P. and Canovas, F. M. (1995). Changes in $NADP^+$-linked isocitrate dehydrogenase during tomato fruit ripening-characterization of the predominant cytosolic enzyme from green and ripe pericarp. *Planta, 196*, 148–154.

Galli, Z., Halasz, G., Kiss, E., Heszky, L. and Dobranszki, J. (2005). Molecular identification of commercial apple cultivars with microsatellite markers. *Horticulturae Science, 40*, 1974-1977.

Gao, Z. S., van de Weg, W. E., Shaart, J. G., Schouten, H. J., Tran, D. H., Kodde, L. P. *et al.* (2005a). Genomic cloning and linkage mapping of the *Mal d 1* (*PR-10*) gene family in apple (*Malus domestica*). *Theoretical and Applied Genetics, 111,* 171-183.

Gao, Z. S., van de Weg, W. E., Shaart, J. G., van Arkel, G., Breiteneder, H., Hoffman-Sommergruber, K. *et al.* (2005b). Genomic characterization and linkage mapping of the apple allergen genes *Mal d 2* (thaumatin-like protein) and *Mal d 4* (profiling). *Theoretical and Applied Genetics, 111,* 1087-1097.

Gao, Z. S., van der Weg, W. E., Schaart, J. G., van der Meer, I. M., Kodde, L., Laimer, M. *et al.* (2005c). Linkage map positions and allelic diversity of two Mal d 3 (non-specific lipid transfer protein) genes in the cultivated apple (*Malus domestica*). *Theoretical and Applied Genetics, 104,* 1856-1859.

Gardiner, S. E., Bus, V. G. M., Rusholme, D., Chagné, D. and Rikkerink, E. H. A. (2007). Apple. Chittaranjan Kole (Ed.), *Genome mapping and molecular breeding in plants. Fruits and nuts*. Springer, Berlin, Germany, 1-53.

Gardiner, S. E., Murdoch, J., Meech, S., Rusholme, R., Bassett, H., Cook, M. *et al.* (2003). Candidate resistance genes from an EST database prove a rich source of markers for major genes conferring resistance to important apple pests and diseases. *Acta Horticulturae, 622,* 141-151.

Garfin, D. E. (2003). Two-dimensional gel electrophoresis: an overview. *Trends in Analytical Chemistry, 22*, 263.

Gebbie, L. K., Burn, J. E., Hocart, C. H. and Williamson, R. E. (2005). Genes encoding ADP-ribosylation factors in *Arabidopsis thaliana* L. Heyn.: genome analysis and antisense suppression. *Journal of Experimental Botany, 56*, 1079-1091.

Gerald, D. G., Birth, G. S., Smittle, D. A. and Leffler, R. G. (1989). Near infrared analysis of soluble solids in intact cantaloupe. *Journal of Food Science, 54*, 393-395.

Gessler, C. and Patocchi, A. (2007). Recombinant DNA technology in apple. *Advanced in Biochemical Engineering/Biotechnology, 107*, 113-132.

Gessler, C., Patocchi, A., Sansavini, S., Tartarini, S. and Gianfranceschi, L. (2006). *Venturia inaequalis* resistance in apple. *Plant Science, 25*, 473-503.

Gianfranceschi, L., Seglias, N., Tarchini, R., Komjanc, M. and Gessler, C. (1998). Simple sequence repeats for the genetic analysis of apple. *Theoretical and Applied Genetics, 96*, 1069-1076.

Gianfranceschi, L. and Soglio, V. (2004). The European project HiDRAS: innovative multidisciplinary approaches to breeding high quality disease resistant apples. *Acta Horticulturae, 663*, 327-330.

Gliha, R., Kravos, A., Lomberger, F., Zupan, M. and Ambrozic, B. (1981). Evaluation of fruit quality of apple cultivars by sensory tests. *Journal of Yugoslavia Pomology, 1*, 595-602.

Gomez, L. D., Noctor, D., Knight, M. R. and Foyer, C. H. (2004). Regulation of calcium signalling and gene expression by glutathione. *Journal of Experimental Botany, 55*, 1851-1859.

Gong, Y., Fan, X. and Mattheis, J. P. (2002). Responses of 'Bing' and 'Rainier' sweet cherries to ethylene and 1-methylcyclopropene. *Journal of the American Society for Horticultural Science, 127*, 831-835.

Görg, A. (2000). Two-dimensional electrophoresis with immobilized pH gradients for proteome analysis. *Electrophoresis, 21*, 1037-1053.

Goulao, L. F. and Oliveira, C. M. (2006). Molecular identification of novel differentially expressed mRNAs up-regulated during ripening of apples. *Plant Science, 172,* 306-318.

Graham, J., Hackett, C. A., Smith, K., Woodhead, M., Hein, I. and McCallum, S. (2009). Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit. *Theoretical and Applied Genetics, 118*, 1143-1156.

Grandillo, S. and Tanksley, S. D. (1999). Identifying the loci responsible for natural variation in fruit size and shape in tomato. *Theoretical and Applied Genetics, 99*, 978-987.

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

Green, L. S., Yee, B. C., Buchanan, B. B., Kamide, K., Sanada, Y. and Wada, K. (1991). Ferredoxin and ferredoxin NADP reductase from photosynthetic and non-photosynthetic tissues of tomato. *Plant Physiology, 96*, 1207-1213.

Greene, D. W. and Autio, W. R. (1990). Evaluation of ripening and fruit quality of Gala and McIntosh apples at harvest and following air storage. *Fruit Variety Journal, 44*, 117-123.

Greene, D. W. and Autio, W. R. (1993). Evaluation of new apple cultivars. *Fruit Notes, 58*, 414.

Grieve, M. (1984). *A modern herbal*, 1st Edition. Leyel, C. F. (Ed.), Dover publications. New York, pp. 919.

Griffiths, H. (2006). Designs on Rubisco. *Nature, 441*, 940-994.

Gu, X. and Verma, D. P. (1997). Dynamics of phragmoplastin in living cells during cell plate formation and uncoupling of cell elongation from the plane of cell division. *The Plant Cell, 9*, 157-169.

Guardiola, J. L. and Garcia-Luis, A. (2000). Increasing fruit in citrus. Thinning and stimulation of fruit growth. *Plant Growth Regulation, 31,* 121-132.

Guarino, C., Arena, S., De Simone, L., D'Ambrosio, C., Santoro, S., Rocco, M. *et al.* (2007). Proteomic analysis of the major soluble components in Annurca apple flesh. *Molecular and Nutritional Food Research, 51*, 255-262.

Guilford, P., Prakash, S., Zhu, J. M., Rikkerink, E., Gardiner, S., Bassett, H. *et al.* (1997). Microsatellites in *Malus x domestica* (apple): abundance, polymorphism and cultivar identification. *Theoretical and Applied Genetics, 94*, 249-254.

Gupta, A. S., Webb, R. P., Holaday, A. S. and Allen, R. D. (1993). Overexpression of superoxide dismutase protects plants from oxidative stress (Induction of

ascorbate peroxidase in superoxide dismutase-overexpressing plants). *Plant Physiology, 103*, 1067-1073.

Gupta, P. K., Rustgi, S., Sharma, S., Singh, R., Kumar, N. and Balyan, H. S. (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Theoretical and Applied Genetics, 270*, 315-323.

Gygax, M., Gianfranceschi, L., Liebhard, R., Kellerhals, M., Gessler, C., and Patocchi, A. (2004). Molecular markers linked to the apple scab resistance gene *Vbj* derived from *Malus baccata jackii*. *Theoretical and Applied Genetics, 109,* 1702-1709.

Gygi, S. P., Rochon, Y., Franza, B. R., Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Molecular Cell Biology, 19,* 1720-1730.

Haji, T., Yaegaki, H. and Yamaguchi, M. (2005). Inheritance and expression of fruit texture melting, non-melting and stony hard in peach. *Scientia Horticulturae, 105*, 241-248.

Halterman, D., and Wise, R. P. (2004). A single amino acid substitution in the sixth leucine-rich repeat of barley MLA6 and MLA13 alleviates dependence on RAR1 for disease resistance signaling. *Plant Journal, 38,* 215-226.

Hallet, C., Sutherland, P. W., Harker, F. R. and MacRae, E. A. (2005). Fruit cell wall, texture and convinience. *Microscopic and Microanalysis, 11,* 1144-1145.

Han, Y. and Korban, S. S. (2007). Spring: a novel family of miniature inverted-repeat transposable elements is associated with genes in apple. *Genomics, 90*, 195-200.

Hanke, V., Kim, W. S. and Geider, K. (2002). Plant transformation for induction of fire blight resistance: transgenic apples expressing apples expressing viral EPS-depolymerase. *Acta Horticulturae, 590*, 393-395.

Hao, X. and Papadopoulos, A. P. (2002). Effects of electrical conductivity and mineral nutrition on fruit radial. *Acta Horticulturae, 633*, 239-244.

Harada, T., Sunako, T., Wakasa, Y., Soejima, J., Satoh, T. and Niizeki, M. (2000). An allele of the 1-aminocyclopropane-1-carboxylate synthase gene (Md-ACS1) accounts for the low level of ethylene production in climacteric fruits of some apple cultivars. *Theoretical and Applied Genetics, 101*, 742-746.

Hardie, D. G. (1999). Plant protein serine/theorine kinases: Classification and Functions. *Annual Review of Plant Physiology, 50*, 97-131.

Harker, F. R., Gunson, F. A. and Jaeger, S. R. (2003). The case for fruit quality: an interpretive review of consumer attitudes, and preferences for apples. *Postharvest Biology and Technology, 28*, 333-347.

Harker, F. R., Maindonald, J., Murray, S. H., Gunson, F. A., Hallett, I. C. and Walker, S. B. (2002). Sensory interpretation of instrumental measurements 1: texture of apple fruit. *Postharvest Biology and Technology, 24*, 225-239.

Harris, S. A., Robinson, J. P. and Juniper, B. E. (2002). Genetic clues to the origin of apple. *Trends in Genetics, 18*, 426-430.

Hayama, H., Tatsuki, M., Ito, A. and Kashimura, Y. (2006). Ethylene and fruit softening in the stony hard mutation in peach. *PostHarvest Biology and Technology, 41*, 16-21.

Heath, M. C. (2000). Hypersensitive response-related death. *Plant Molecular Biology, 44*, 321-334.

Heazlewood, J. L., and Millar, A. H. (2006). Plant proteomics: challenges and resources. *Blackwell Publishing, 28,* 2-3.

Helentjaris, N. J., Slocum, T., Ruggero, M. and Schaefer, B. A. (1987). Restriction fragment length polymorphism analysis of loci associated with insect resistance in tomato. *Crop Science, 27*, 797-803.

Hell, R. and Stephen, U. W. (2003). Iron uptake, trafficking and homeostasis in plants. *Planta, 216,* 541-551.

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

Hemmat, M., Brown, S. K., Aldwinckle, H. S., Mehlenbacher, S. A. and Weeden, N. F. (2003). Identification and mapping of markers for resistance to apple scab from 'Antonovka' and 'Hansen's baccata #2. *Acta Horticulturae, 622,* 153-161.

Hemmat, M., Weeden, N. F., Conner, P. J. and Brown, S. K. (1997). A DNA marker for columnar growth habit in apple contains a simple sequence repeat. *Journal of the American Society for Horticultural Science, 122*, 347-349.

Hemmat, M., Weeden, N. F., Manganaris, A. G. and Lawson, D. M. (1994). Molecular marker linkage map for apple. *Journal of Heredity, 85*, 4-11.

Henriksson, L. M., Johansson, P., Unge, T. and Mowbray, S. L. (2004). X-ray structure of peptidyl-prolyl cis-trans isomerase A from *Mycobacterium tuberculosis*. *European Journal of Biochemistry, 271*, 4107-4113.

Herrera, C. M. (1992). Interspecific variation in shape: allometry, phylogeny and adaptation to dispersal agents. *Ecology, 75*, 1832-1841.

Hirokawa, N. and Tekamura, R. (2003). Biochemical and molecular characterization of diseases linked to motor proteins. *Trends in Biochemical Sciences, 28*, 558-565.

Hjerno, K., Alm, R., Canbäck, B., Matthiesen, R., Trajkovski, K., Björk, L. *et al.* (2006). Down-regulation of the strawberry Bet v 1-homologous allergen in concert with the flavonoid biosynthesis pathway in colourless strawberry mutant. *Proteomics, 6,* 1574-1587.

Hoehn, E., Baumgartner, D., Gasser, F. and Gabioud, S. (2008). Ripening regulation and consumer expectations. *Acta Horticulturae, 796,* 83-91.

Hoehn, E., Gasser, F., Guggenbühl, B. and Künsch, U. (2003). Efficacy of instrumental measurements or determination of minimum requirements of firmness, soluble solids, and acidity of several apple varieties in comparison to consumer expectations. *Post Harvest Biology and Technology, 27*, 27-37.

Hokanson, S. C., Szewe-McFadden, A. K., Lamboy, W. F. and McFerson, J. R. (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.

Hood, B. L., Veenstra, T. D. and Conrads, T. P. (2004). Mass spectrometry-based proteomics. *International Congress Series, 1266*, 375-380.

Hopkins, M.T., Lampi, Y., Wang, T-W., Liu, Z. and Thompson, J.E. (2008). Eukaryotic translation initiation factor 5A is involved in pathogen-induced cell death and development of disease symptoms in *Arabidopsis*. *Plant physiology, 148*, 479-489.

Hospital, F., Moreau, L., Lacoudre, F., Charcosset, A. and Gallais, A. (1997). More on the efficiency of marker-assisted selection. *Theoretical and Applied Genetics, 95*, 1181-1189.

Hough, L. F., Shay, J. R. and Dayton, D. F. (1953). Apple scab resistance from *Malus* 'Floribunda Sieb'. *Journal of the American Society for Horticultural Science, 62*, 341-347.

Hovmalm, H. A. P., Nybom, H., Barraclough, D., Beuning, L., Bowen, J., Bulley, S. *et al.* (2006). Mal d 1-like allegernic patterns in two apple cultivars. *International Society for Horticultural Science, 123*, 291-294.

Hrazdina, G. (1994). *Genetic engineering of 'McIntosh' apple to prevent softening during storage*. New York Fruit Quarterly, 9-11.

Hrazdina, G., Hemmat, M., Kiss, E., Galli, Z., Aldwinckle, H. S. and Rosenfield, C. L. (2003). Down regulation of ethylene production and softening in McIntosh apples. *Acta Horticulurae, 628,* 239-251.

Hulme, A. C. and Rhodes, M. J. C. (1971). The biochemistry of fruits and their products. Volume 2. Academic Press, London and New York, pp. 620.

Hussey, P. J., Allwood, E. G. and Smertenko, A. P. (2002). Actin-binding proteins in the *Arabidopsis* genome database: properties of functionally distinct plant actin-depolymerizing factors/cofilins. *Philosophical Transactions of The Royal Society of the London Series B-Biological Sciences, 357*, 791-798.

Huxley, A. J., Griffiths, M. Levy, M. (1992). *New Royal Horticultural Society dictionary of gardening.* The MacMillan Press Ltd, London, UK, volume 3, pp. 240.

Igarashi, M., Yoshie Abe, Y., Hatsuyama, Y., Ueda, T., Fukasawa-Akada, T., Kon, T. *et al.* (2008). Linkage maps of the apple (*Malus x domestica* Borkh.) cvs 'Ralls Janet' and 'Delicious' include newly developed EST markers. *Molecular Breeding, 22*, 95-118.

Iglesias, I., Echeverria, G. and Soria, Y. (2008). Differences in fruit colour development, anthocyanin content, fruit quality and consumer acceptability of eight 'Gala' apple strains. *Scientia Horticulturae, 119*, 32-40.

Ito, S., Fishinobu, S., Yoshioka, I., Koga, S., Matsuzawa, H. and Wakagi, T. (2001). Structural basis for the ADP-specificity of a novel glucokinase from a hypertermophilic archaeon. *Structure, 9,* 205-214.

Jaeger, S. R. (2000). Uncovering cultural differences in choice behavior between Samoan and New Zealand consumers: a case study with apples. *Food Quality Preference, 11*, 405-417.

Jaeger, S. R., Andani, Z., Wakeling, I. N. and MacFie, H. J. H. (1998). Consumer preference for fresh and aged apples: a cross-cultural comparison. *Food Quality Preference, 9*, 355-366.

Jahn, R. and Südhof, T.C. (1999). Membrane fusion and exocytosis. *Annual Review in Biochemistry, 68*, 863-911.

Jakubowska, A., Gronwald, J., Menkiszak, J., Górski, B., Huzarski, T., Byrski, T. *et al.* (2007). Integrin ß3 Leu33Pro polymorphism increases BRCA1-associated ovarian cancer risk. *Journal of Medical Genetics, 44*, 408-411.

James, C. M., Clarke, J. B. and Evans, K. M. 2004. Identification of molecular markers linked to the mildew resistance gene *Pl-d* in apple. *Theoretical and Applied Genetics, 110,* 175-181.

Jamet, E., Albenne, C., Boudart, G., Irshad, M., Canut, H. and Pont-Lezica, R. (2008a). Recent advances in plant cell wall proteomics. *Proteomics, 8*, 893-908.

Jamet, E., Boudart, G., Borderies, G., Charmont, S., Lafitte, C., Rossignol, M. *et al.* (2008b). Isolation of plant cell wall proteins. *Methods in Molecular Biology, 425*, 187-201.

Jandus, J., Valentova, O., Kas, J., Daussant, J. and Thévenot, C. (1997). Phospholipase D during tomato fruit ripening. *Plant Physiology and Biochemistry, 35*, 123-128.

Janick, J., Cummings, J. N., Brown, S. K. and Hemmat, M. (1996). Apples. *In:* Janick J. and Moore J. N. (Eds) *Fruit Breeding. Volume I. Tree and Tropical Fruits*. 1st Edition, John Wiley and Sons, New York, USA, 1-79.

Janssen, J. B., Thodey, K., Schaffer, R. J., Alba, R., Balakrishnan, L., Bishop, R. *et al.* (2008). Global gene expression analysis of apple fruit development from the floral bud to ripe fruit. *BMC Plant Biology, 8*, 1-29.

Jansen, J. (2005). Construction of linkage maps in full-sib families of diploid outbreeding species by minimizing the number of recombinations in hidden inheritance vectors. *Genetics, 170*, 2013-2025.

Jansen, J., De Jong, A. G. and Van Ooijen, J. W. (2001). Constructing dense genetic linkage maps. *Theoretical and Applied Genetics, 102*, 1113-1122.

Janzowski, G., Schaefer, S., Fritzen, S., Baum, M., Dietrich, H. and Eisenbrand, G. (2004). Preventive antioxidative effectiveness of polyphenolic apple juice compounds in human colon cells. *Proceedings of the American Association for Cancer Research, 45*, 731.

Jenkins, B, D., Kulhanek, D. J. and Barkan, A. (1997). Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *The Plant Cell, 9*, 283-296.

Jiang, S. Y. and Ramachandran, S. (2004). Identification and molecular characterization of myosin gene Family in *Oryza sativa* Genome. *Journal of Experimental Botany, 45*, 590-599.

Jiang, W., Zhang, M., He, J. and Zhou, L. (2004). Regulation of 1-MCP-Treated banana fruit quality by exogenous ethylene and temperature. *Food Science and Technology International, 10*, 15-20.

Jobling, J. (2002). *Harvest maturity is critical for Pink Lady fruit quality*. http://www.postharvest.com.au/Information_Sheets.htm [May 2009].

Johnston, J. W., Hewett, E. W. and Hertog, M. L. (2002). Postharvest softening of apple (*Malus domestica*) fruit. *Journal of Crop Horticultural Science, 30*, 145-160.

Jones, J. D. G. (2001). Putting knowledge of plant disease resistance genes to work. *Current Opinion Plant Biology, 4,* 281-287.

Jonsson, A. (2003). Consumer evaluation of scab resistant apple varieties. *Eucarpia, 14*, 126-140.

Jorgensen, J. H. (1994). Genetics of powdery mildew resistance in barley. *Critical Review in Plant Science, 13,* 97-119.

Juniper, B. E., Robinson, J., Harris, S. A. and Watkins, R. (2001). Origin of apple (*Malus x domestica* Borkh.). *In*: Reeve, E. C. R. (Ed.), *Encyclopedia of Genetics*. Fitzroy Dearborn, London, UK, 674-677.

Kanamarua, N., Komorib, Y. S., Saitob, M., Katob, H., Takahashib, S., Omurac, M. *et al.* (2004). Transgenic apple transformed by sorbitol-6-phosphate dehydrogenase cDNA: Switch between sorbitol and sucrose supply due to its gene expression. *Plant Science, 167*, 55-61.

Kanayama, Y., Moriguchi, R., Deguchi, M., Kanahama, K. and Yamaki, S. (2005). Effects of environmental stresses and abscisic acid on sorbitol-6-phosphate dehydrogenase expression in Rosaceae fruit trees. *Acta Horticulturae, 738,* 375-381.

Katz, E., Fon, M., Lee, J. L., Phinney, B. A., Sadka, A. and Blumwald, E. (2007). The citrus fruit proteome: insights into citrus fruit metabolism. *Planta, 226*, 989-1005.

Kausch, K. D. and Handa, A. K. (1995). Molecular-cloning and nucleotide-sequence of a lipoxygenase cDNA from ripening tomato fruit. *Plant Physiology, 107*, 669–670.

Kawano, S., Abe, H. and Iwamoto, M. (1995). Development of a calibration equation with temperature compensation for determining the brix value in intact peaches. *Journal of Near Infrared Spectroscopy, 3*, 211-218.

Kellogg, E. and Juliano, N. (1997). The structure and function of RuBisCO and their implications for systematic studies. *American Journal of Botany, 84*, 413.

Kendrich, R. E. and Kronenberg, G. H. M. (1994). *Photomorphogenesis in plants,* 2nd Edition. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 828.

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

Kenis, K., Keulemans, J. and Davey, M. W. (2006). Identification and stability of QTLs for frut traits in apple. *Tree Genetics and Genomes, 4*, 647-661.

Kenis, K. and Keulemans, J. (2008). Study of the tree architecture of apple (*Malus x domestica* Borkh.) by QTL analysis of growth traits. *Molecular Breeding, 19*, 193-208.

Ketelaar, T., Allwood, E. G. and Hussey, P. J. (2007). Actin organization and root hair development are disrupted by ethanol-induced overexpression of *Arabidopsis* actin interacting protein 1 (AIP1). *New Phytologist, 174*, 57-62.

Kilcast, D. and Fillion, L. (2001). Understanding consumer requirements for fruit and vegetable texture. *Nutrition and Food Science, 31*, 221-225.

Kim, M. H., Cooper, D. R., Oleksy, A., Devedjiev, Y., Derewenda, U., Reiner, O. *et al.* (2004). The structure of the N-terminal domain of the product of the lissencephaly gene Lis1 and its functional implications. *Structure, 12*, 987-998.

Kim, H., Park, M., Kim, S. J. and Hwang, I. (2005). Actin filaments play a critical role in vacuolar trafficking at the golgi complex in plant cells. *The Plant Cell, 17*, 888-902.

Kim, Y-O., Pan, S. O., Jung, C-H. and Kang, H. (2007). A zinc finger-containing lycine-rich RNA-binding protein, atRZ-1a, has a negative impact on seed germination and seedling growth of *Arabidopsis thaliana* under salt or drought stress conditions. *Plant and Cell Physiology, 48,* 1170-1181.

King, G. J., Lynn, J. R., Dover, C. J., Evans, K. M. and Seymour, G. B. (2001). Resolution of quantitative trait loci for mechanical measures accounting for genetic variation in fruit texture of apple (*Malus pumila* Mill.). *Theoretical and Applied Genetics, 102*, 1227-1235.

King, G. J., Maliepaard, C., Lynn, J. R., Alston, F. H., Durel, C. E., Evans, K. M. *et al.* (2000). Quantitative genetic analysis and comparison of physical and sensory descriptors relating to fruit flesh firmness in apple (*Malus pumila* Mill.). *Theoretical and Applied Genetics, 100*, 1074-1084.

Kjemtrup, S., Boyes, D. S., Christensen, C., McCaskill, A. J., Hylton, M. and Davis, K. (2003). Growth stage-based phenotypic profiling of plants. *In* Grotewold, E. (Ed.), *Plant functional genomics*, Humana Press Inc, Totawa, NJ, USA, pp. 236.

Klein, D., Fink, B., Arold, B., Eisenreich, W. and Schwab, W. (2007). Functional characterization of enone oxidoreductases from strawberry and tomato fruit. *Journal of Agriculture and Food Chemistry, 55*, 6705-6711.

Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis in mouse tissues. A novel approach to testing for induced point mutations in mammals. *Human Genetik, 26*, 231-243.

Knaff, D. B. and Hirasawa, M. (1991). Ferredoxin-dependent chloroplast enzymes. *Biochimica et Biophysica Acta, 1056*, 93-125.

Knight, R. L., Briggs, J. B., Massee, A. M. and Tydeman, H. M. (1962). The inheritance of resistance to woolly aphid, *Eriosoma lanigerium* (Hausmn.) in the apple. *Journal of Horticultural Science of America, 37*, 207-218.

Kokubun, N., Ishida, H., Makino, A. and Mae, T. (2002). The degradation of the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase into the 44-kDa fragment in the lysates of chloroplasts incubated in darkness. *Plant and Cell Physiology, 43*, 1390-1395.

Konishi, M. and Yanagisawa, S. (2005). Signaling crosstalk between ethylene and other molecules. *Plant Biology, 22*, 401-407.

Korban, S. S. and Skirvin, R. M. (1994). Nomenclature of the cultivated apple. *Horticulturae Science, 19*, 177-180.

Korban, S. S., Vodkin, L. O., Liu, L., Aldwinckle, H. S., Ksenija, G., Gonzales, D. O. *et al.* (2005). *Large-scale analysis of EST sequences in the apple genome.* Paper presented at the Plant & Animal Genomes conference abstracts. XIII. Town & Country Convention Center, San Diego, California.

Kosova, K., Vitamvas, P. and Prasil, I. T. (2007). The role of dehydrins in plant response to cold. *Biologia Plantarum, 51*, 601-617.

Kotoda, N., Wada, K., Masuda, T. and Soejima, J. (2003). The break-through in the reduction of juvenile phase in apple using transgenic approaches. *Acta Horticulturae, 622*, 679-687.

Kramer, A. and Twigg, B. A. (1970). The number of assessors required for taste panels basker. *Chemical Senses, 2*, 493-496.

Kruger, W. M., Szabo, L. J. and Zeyen, R. J. (2003). Transcription of the defense response genes chitinase IIb, PAL and peroxidase is induced by the barley powdery mildew fungus and is only indirectly modulated by *R* genes. *Physiological and Molecular Plant Pathology, 63,* 167-178.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, 227,* 680-685.

Lander, E. S. and Botstein, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics, 121*, 185-199.

Lawrence, G. J., Finnegan, E. J., Ayliffe, M. A. and Ellis, J. G. (1995). The L6 gene for flax rust resistance is related to the *Arabidopsis* bacterial resistance gene RPS2 and the tobacco viral resistance gene. *The Plant Cell, 7*, 1195-1206.

Lea, A. G. H. and Drilleau, J.-F. (2003). Cider making. In: Lea, A. G. H. and Piggot, R. R. (Eds). *Fermented beverage production*. Blackie Academic, London, 66-96.

Lecomte, L., Duffe, P., Buret, M., Servin, B., Hospital, F. and Causse, M. (2004). Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. *Theoretical and Applied Genetics, 109*, 658-668.

Lesemann, S. and Dunemann, F. (2006). Recent findings on the biodiversity of the apple powdery mildew pathogen. *Gesunde-Pflanzen, 58,* 117-123.

Li, Z. F., Wan, J. M., Xia, J. F. and Zhai, H. Q. (2003). Mapping quantitative trait loci underlying appearance quality of rice grains (*Oryza sativa* L.). *Yi Chuan Xue Bao, 30*, 251-259.

Liavonchanka, A. and Feussner, I. (2006). Lipoxygenases: occurrence, functions and catalysis. *Journal of Plant Physiology, 163*, 348-357.

Liebhard, R., Gianfranceschi, L., Koller, B., Ryder, C. D., Tarchini, R., Van de Weg, E. *et al.* (2002). Development and characterisation of 140 new microsatellites in apple (*Malus* x *domestica* Borkh.). *Molecular Breeding, 10*, 217-241.

Liebhard, R., Kellerhals, M., Plammatter, W., Jertmni, M. and Gessler, C. (2003a). Mapping quantitative physiological traits in apple (*Malus* x *domestica* Borkh.). *Plant Molecular Biology, 52*, 511-526.

Liebhard, R., Koller, B., Gianfranceschi, L. and Gessler, C. (2003b). Creating a saturated reference map for the apple (*Malus x domestica* Borkh.) genome. *Theoretical and Applied Genetics, 106*, 1497-1508.

Liebhard, R., Koller, B., Patocchi, A., Kellerhals, M., Pfammatter, W., Jermini, M. *et al.* (2003c). Mapping quantitative field resistance against apple scab in a 'Fiesta' x 'Discovery' progeny. *Phytopathology, 93*, 493-501.

Liere, K. and Link, G. (1995). RNA-binding activity of the matK protein encoded by the chloroplast trnk intron from mustard (*Sinapis alba* L.). *Nucleic Acids Research, 23,* 917-921.

Lin, D., Tabb, D. L. and Yates, J. R. I. (2003). Large-scale protein identification using mass spectrometry. *Biochimica et Biophysica Acta, 1646*, 1-10.

Lindhout, P. (2002). The perspectives of polygenic resistance in breeding for durable disease resistance. *Euphytica, 124*, 217–226.

Liu, F. W. and King, M. M. (1978). Consumer evaluation of McIntosh apple firmness. *Horticulturae Science, 13*, 162-163.

Ljanabi, S. M., Forget, L. and Dookun, A. (1999). An improved and rapid protocol for the isolation of polysaccharide- and polyphenol-free sugarcane DNA. *Plant Molecular Biology Reporter, 17*, 1-8.

López-Gómez, R., Cabrera-Ponce, J. L., Saucedo-Arias, L. J., Carreto-Montoy, L., Villanueva-Arce, R., Díaz-Perez, J. C. *et al.* (2008). Ripening in papaya fruit is altered by ACC oxidase cosuppression. *Transgenic Research, 18*, 89-97.

Lortze, E. and Bergh, O. (2004). Early prediction of harvest fruit size distribution of an apple and pear cultivar. *Scientia Horticulturae, 101*, 281-290.

Lu, R. (2004). Hyperspectral Imaging for Nondestructive Assessment of Fruit Quality. *Michigan State University Department of Agricultural Engineering Newsletter*, pp. 24.

Lu, R., Guyer, D. E. and Beaudry, R. M. (2000). Determination of firmness and sugar content of apples using near-infrared diffuse reflectance. *Journal of Texture Studies, 31*, 615-630.

Lukowitz, W., Mayer, U. and Jürgens, G. (1996). Cytokinesis in the *Arabidopsis* embryo involves the syntaxin-related *Knolle* gene product. *Cell, 84*, 61-71.

Luo, Z. W., Tao, S. H. and Zeng, Z. B. (2000). Inferring linkage disequilibrium between a polymorphic marker locus and a trait locus in natural populations. *Genetics, 156*, 457-467.

Mace, E. M., Xia, L., Jordan, D. R., Halloran, K., Parh, D. K., Huttner, E. *et al.* (2008). DArT markers: Diversity analyses and mapping in *Sorghum bicolour*. *Genomics, 9*, 26.

Maharaj, R. (2007). Genetic analysis of resistance to apple scab (*Venturia inaequalis*) in apple (*Malus x domestica* Borkh.). MSc Thesis, University of the Western Cape, South Africa, pp. 127.

Mahmoud, A. A., Sukumar, S. and Krishnan, H. B. (2008). Interspecific rice hybrid of *Oryza sativa x Oryza nivara* reveals a significant increase in seed protein content. *Journal of Agriculture and Food Chemistry, 56*, 476-482.

Mahner, M. and Kary, M. (1997). What exactly are genomes, genotypes and phenotypes? And what about phenomes? *Journal of Theoretical Biology, 186*, 55-63.

Malanczuk, M. L. (2005). Training system and fruit quality in the apple cultivar 'Jonagold'. *Postharvest Biology Technology, 14*, 214-217.

Maliepaard, C., Sillanpaa, M. J., van Ooijen, J. W., Jansen, R. C., Arjas, E. (2001). Bayesian versus frequentist analysis of multiple quantitative trait loci with an application to an outbred apple cross. *Theoretical and Applied Genetics, 103,* 1243-1253.

Maliepaard, C., Alston, F. H., Van Arkel, G., Brown, L. M., Chevreau, E., Dunemann, F. *et al.* (1998). Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *Theoretical and Applied Genetics, 97*, 60-73.

Maliepaard, C., Jansen, J. and Van Ooijen, J. W. (1997). Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences of applications. *Genetic Research, 70*, 237-250.

Malladi, A. (2005). *Genetic control of fruit size in apple*. PhD thesis. Purdue University, pp. 244.

Mallik, R. and Gross, S. P. (2004). Molecular motors: strategies to get along. *Current Biology, 14*, 971-982.

Manganaris, A. G. and Alston, F. H. (1989). Glutamate oxaloacetate transaminase isoenzymes in apple cultivars and rootstocks. *Journal of Horticultural Science, 64*, 9-15.

Manganaris, A. S. G., Alston, F. H., Aldwinckle, H. S., Gustafson, H. L., Brown, S. K. and Weeden, N. F. (1994). Isozyme locus Pgm-1 is tightly linked to a gene (Vf) for scab resistance in apple. *Journal of the American Society for Horticultural Science, 119*, 1286-1288.

Manly, K. F. and Olson, J. M. (1999). Overview of QTL mapping software and introduction to Map Manager QT. *Mammalian Genome, 10*, 327-334.

Manuell, A. L., Yamaguchi, K., Haynes, P. A., Milligan, R. A. and Mayfield, S. P. (2005). Composition and structure of the 80S ribosome from the green alga *Chlamydomonas reinhardtii*: 80S ribosomes are conserved in plants and animals. *Journal of Molecular Biology, 351*, 266-279.

Maple, J. and Moller, S. G. (2005). An emerging picture of plastid division in higher plants. *Planta, 223*, 1-4.

Markoulatos, P., Siafakas, N. and Moncany, M. (2002). Multiplex polymerase chain reaction: a practical approach. *Journal of Clinical Laboratory Analysis, 16*, 47-51.

Markussen, T., Krüger, J., Schmidt, H. and Dunemann, F. (1995). Identification of PCR-based markers linked to the powdery-mildew-resistance gene Pl1 from *Malus robusta* in cultivated apple. *Plant Breeding, 114*, 530-534.

Martínez, G. A., Chaves, A. R. and Añónb, M. C. (2001). Characterization of peroxidase-mediated chlorophyll bleaching in strawberry fruit. *Phytochemistry, 58*, 379-387.

Maruyama, A., Saito, K. and Ishizawa, K. (2001). β-Cyanoalanine synthase and Cys synthase from potato: molecular cloning, biochemical characterization, and spatial and hormonal regulation. *Plant Molecular Biology, 46*, 749-760.

Matic, S., Geisler, D. A., Møller, I. M., Widell, S. and Rasmusson, A. G. (2005). Alamethicin permeabilizes the plasma membrane and mitochondria but not the tonoplast in tobacco (*Nicotiana tabacum* L. cv Bright Yellow) suspension cells. *Biochemistry Journal, 389*, 695-704.

Matsubara, H. and Hase, T. (1983). *Phylogenetic consideration of ferredoxin sequences in plants, particularly algae*. In U. Jensen, DE Fairbrothers, eds, Proteins and nucleic acids in plant systematics. Springer-Verlag, Berlin, 168-181.

Matteson-Heidenreich, M. C., Corral-Garcia, E. A., Momol, E. A. and Burr, T. J. (1997). Russet of apple fruit caused by *Aureobasidium pullulans* and *Rhodotorula glutinis*. *Plant Disease, 81*, 337-342.

Mattevi, A., Fraaije, M. W., Mozzarelli, A., Olivi, L., Coda, A. and van Berkel, W. J. H. (1997). Crystal structures and inhibitor binding in the octameric flavoenzyme vanillyl-alcohol oxidase: the shape of the active-site cavity controls substrate specificity. *Structure, 5*, 907-920.

McCormack, A. L., Schieltz, D. M., Goode, B. and Yang, S. (1997). Direct analysis and identification of proteins in mixtures by LC-MS/MS and database searching at the low-femtomole level. *Analytical Chemistry, 69*, 767-776.

McCouch, S. R., Chen, X., Panaud, O., Temnykh, S., Xu, Y., Cho, Y. G. *et al.* (1997). Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Molecular Biology, 35*, 89-99.

McDonald, W. H., Ohi, R., Miyamoto, D. T., Mitchison, T. J. and Yates III, J. R. (2002). Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: single dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *International Journal of Mass Spectrometry, 219*, 245-251.

McHughen, A. (2006). Plant genetic engineering and regulation in the United States. *Agricultural Biotechnology in California Series, 8179,* 1-5.

Milanesi, L. and Boland, R. (2004). Presence of estrogen receptor (ER)-like proteins and endogenous ligands for ER in *Solanaceae*. *Plant Science, 166*, 397-404.

Missiaggia, A. A., Piacezzi, A. L. and Grattapaglia, D. (2005). Genetic mapping of *Eef1*, a major effect QTL for early flowering in *Eucalyptus grandis*. *Tree Genetics and Genomes, 1*, 79-84.

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

Mohr, G., Perlman, P. S. and Lambowitz, A. M. (1993). Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function. *Nucleic Acids Research, 21*, 4991-4997.

Monteoliva, L. and Albar, J. P. (2004). Differential proteomics: an overview of gel and non-gel based approaches. *Brief Functional Genomics Proteomics, 3*, 220-239.

Mooney, B. P. and Thelen, J. J. (2004). High throughput peptide mass-fingerprinting of soybean seed proteins: automated workflow and utility of UniGene espressed sequence tags database for protein identification. *Phytochemistry, 65*, 1733-1744.

Motulsky, H. (1995). Intuitive biostatistics. Oxford University Press, USA, pp. 386.

Mowry, J. B. and Dayton, D. F. (1964). Inheritance of susceptibility to apple blotch. *Journal of Heredity, 55*, 129-132.

Mullis, K. B. (1990). The unusual origin of the polymerase chain reaction. *Science, 262*, 36-43.

Muratani, M., and Tansey, W.P. (2003). How the ubiquitin-proteasome system controls transcription. *National Review in Molecular Cell Biology, 4,* 192–201.

Murooka, Y., Yukiko, M. and Hayashi, M. (2002). Variation of the amino acid content of *Arabidopsis* seeds by expressing soybean aspartate aminotransferase gene. *Journal of Bioscience and Bioengineering, 94*, 225-230.

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

Nagaraju, J., Kathirvel, M., Subbiah, E. V., Muthulakshmi, M. and Kumar, L. D. (2002). FISSR-PCR: a simple and sensitive assay for high throughput genotyping and genetic mapping. *Molecular Cell Probes, 16*, 67-72.

Naik, S., Hampson, C., Gasic, K., Bakkeren, G. and Korban, S. S. (2006). Development and linkage mapping of E-STS and RGA markers for functional gene homologues in apple. *Genome, 49*, 955-968.

Naqvi, S.M., Park, K-S., Yi, S-Y., Lee, H-Y., Bok, S.H. and Choi, D. (1998). A glycine-rich RNA binding protein gene is differentially expressed during acute hypersensitive response following tobacco mosaic virus infection in tobacco. *Plant Molecular Biology, 37*, 571-576.

Ndimba, B. K., Chivasa, S., Simon, W. J. and Slabas, A. R. (2003). Proteomic analysis of changes in the extra-cellular matrix of cellsuspension cultures induced by fungal elicitors. *Proteomics, 3*, 1047-1059.

Ndimba, B. K., Chivasa, S., Simon, W. J. and Slabas, A. R. (2005). Identification of *Arabidopsis* salt and osmotic stress responsive proteins using two- dimensional difference gel electrophoresis and mass spectrometry. *Proteomics, 5*, 4185-4196.

Ndimba, B. K. and Thomas, L. A. (2008). Proteomics in South Africa: Current status, challenges and prospects. *Biotechnology Journal, 3*, 1368-1374.

Neta-Sharira, I., Isaacsonb, T., Luriec, S. and Weiss, D. (2005). Dual role for tomato heat shock protein 21: protecting photosystem II from oxidative stress and promoting colour changes during fruit maturation. *The Plant Cell, 17*, 1829-1838.

Neuhaus, H. and Link, G. (1987). The chloroplast tRNALys (UUU) gene from mustard (*Sinapis alba*) contain a class II intron potentially coding for a maturase-related polypeptide. *Current Genetics, 11,* 251-257.

Newcomb, R. D., Crowhurst, R. N., Gleave, A. P., Rikkerink, E. H. A., Allan, A. C., Beuning, L. L. *et al.* (2006). Analysis of expressed sequence tags from apple (*Malus* x *domestica*). *Plant Physiology Review, 141*, 147-166.

Ngara, R., Rees, J. and Ndimba, B. K. (2008). Establishment of sorghum cell suspension culture system for proteomics studies. *African Journal of Biotechnology, 7*, 744-749.

Nixon, J. (2006). Testing for segregation distortion in genetic scoring data from backcross or doubled haploid populations. *Heredity, 96*, 290-297.

Norelli, J. L., Holleran, H. T., Johnson, W. C., Robinson, T. L. and Aldwinckle, H. S. (2003). Resistance of Geneva and other rootstocks to *Erwinia amylovora*. *Plant disease, 87*, 26-32.

Norton, R. A. (1997). Apples for the 21st century. *Good Fruit Grower, 48*, 39-43.

Nosarzewski, M. and Douglas, D. A. (2007). Tissue-specific expression of sorbitol dehydrogenase in apple fruit during early development. *Journal of Experimental Botany, 58*, 1863-1872.

Ochs, D. (1983). Protein contaminants of sodium dodecyl sulphate-polyacrylamide gels. *Analytical Biochemistry, 135*, 470-474.

Oey, M. L., Vanstreels, E., De Baerdemaeker, J., Tijskens, E., Ramon, H., Hertog, M. L. A. T. M. and Nicolaï, B. (2007). Effect of turgor on micromechanical and structural properties of apple tissue: A quantitative analysis. *Postharvest Biology and Technology, 44,* 240-247.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry, 250*, 4007-4021.

Oilseed proteomics: http://oilseedproteomics.missouri.edu/

Ohkawa, W., Moriya, S., Kanahama, K. and Kanayama, Y. (2008). Re-evaluation sorbitol metabolism in fruit from Rosaceae trees. *Acta Horticulturae, 772*, 159-166.

Okamoto, G., Kuwamura, T. and Hirano, K. (2004). Effects of water deficit stress on leaf and berry ABA and berry ripening in Chardonnay grapevines (*Vitis vinifera*). *Vitis, 43,* 15-17.

Oraguzie, N. C., Iwanami, H., Soejima, J., Harada, T. and Hall, A. (2004). Inheritance of the *Md-ACS1* gene and its relationship to fruit softening in apple (*Malus domstica* Borkh.). *Theoretical and Applied Genetics, 108*, 1526-1533.

Oraguzie, N. C., Yamamoto, T., Soejima, J., Suzuki, T. and De Silva, H. N. (2005). DNA fingerprinting of apple (*Malus* spp.) rootstocks using simple sequence repeats. *Plant Breeding, 124*, 197-202.

Orsi, C. H. and Tanksley, S. D. (2009). Natural variation in an ABC Transporter gene associated with seed size evolution in tomato species. *Public Library of Science Genetics, 5*.

Osborn, T. C., Alexander, D. C. and Fobes, J. F. (1995). Identification of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit. *Theoretical and Applied Genetics, 73*, 350-356.

Pan, H. H. and Shu, Z. H. (2007). Temperature affects colour and quality characteristics of 'Pink' wax apple fruit discs. *Scientia Horticulturae, 112*, 290-296.

Pancholi, V. (2001). Multifunctional α-enolase: its role in diseases. *Cell and Molecular Life Science, 58,* 902-920.

Pantke, U., Volk, T., Smutzler, M., Kox, W. J., Sitte, N., Grune, T. (1999). Oxidised proteins as a marker of oxidative stress during coronary heart surgery. *Free Radical Biology and Medicine, 27,* 1080-1086.

Park, M. H. (2006). The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A. *Journal of Biochemistry, 139*, 161-169.

Paterson, A. H., Damon, S., Hewitt, J. D., Zamir, D., Rabinowitch, H. D., Lincoln, S. E. *et al.* (1991a). Mendelian factors underlying quantitative traits in tomato: Comparison across species, generations and environments. *Genetics, 127*, 181-197.

Paterson, A. H., Tanksley, S. D. and Sorrells, M. E. (1991b). DNA markers in plant improvement. *Advanced Agronomy, 46*, 39-90.

Patocchi, A., Fernández-Fernández, F., Evans, K., Gobbin, D., Rezzonico, F., Boudichevskaia, A. *et al.* (2009). Development and test of 21 multiplex PCRs composed of SSRs spanning most of the apple genome. *Tree Genetics and Genomics, 5*, 211-223.

Patocchi, A., Walser, M., Tartarini, S., Broggini, G. A. L., Gennari, F., Sansavini, S. *et al.* (2005). Identification by genome scanning approach (GSA) of a microsatellite tightly associated with the apple scab resistance gene *Vm. Genome, 48,* 630-636.

Patocchi, A., Bigler, B., Koller, B., Kellerhals, M. and Gessler, C. (2004). *Vr(2)*: a new apple scab resistance gene. *Theoretical and Applied Genetics, 109*, 1087-1092.

Peil, A., Dunemann, F., Garcia, T., Richter, K., Trognitz, B., Hanke, V. *et al.* (2006, March 19-22). *Efforts to elucidate mechanisms and genetics of fire blight resistance in apple.* Paper presented at the 3rd International Rosaceae Genomics Conference.

Peleman, J., Saito, K., Cottyn, B., Engler, G., Seurinck, L., Van Montagu, M. *et al.* (1989). Structure and expression analyses of the S-adenosylmethionine synthetase gene family in *Arabidopsis thaliana*. *Genetics, 84*, 359–369.

Penefsky, H. S. and Cross, R. L. (1991). Structure and mechanism of F1, F0-type ATP synthases and ATPases. *Advanced Enzymology, 64*, 173-214.

Peng, Y. and Lu, R. (2007). Prediction of apple fruit firmness and soluble solids content using characteristics of multispectral scattering images. *Journal of Food Engineering, 82*, 142-152.

Perkins, D. N., Pappin, D. J., Creasy, D. M. and Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis, 20*, 3551-3567.

Peumans, W. J., Barre, A., Derycke ,V., Rouge´, P., Zhang, W., May, G. D. *et al.* (2000). Purification, characterization and structural analysis of an abundant β-1, 3-glucanase from banana fruit. *European Journal of Biochemistry, 267,* 1188-1195.

Pilet-Nayel, M. L., Muehlbauer, F. J., McGee, R. J., Kraft, J. M., Baranger, A. and Coyne, C. J. (2002). Quantitative trait loci for partial resistance to Aphanomyces root rot in pea. *Theoretical and Applied Genetics, 106*, 28–39.

Pinheroa, R. G., Almquista, K. C., Novotnac, Z. and Paliyat, G. (2003). Developmental regulation of phospholipase D in tomato fruits. *Plant Physiology and Biochemistry, 44*, 223-240.

Pirondini, A., Visioli, G., Malcevschi, A. and Marmiroli, N. (2006). A 2-D liquid-phase chromatography for proteomic analysis in plants tissues. *Journal of Chromatography B, 833*, 91-100.

Plotto, A. A., Azarenko, N., Mattheis, J. P. and McDaniel, M. R. (1995). 'Gala', 'Braeburn' and 'Fuji' apples: maturity indices and quality after storage. *Fruit Variety Journal, 49*, 133-142.

Predreschi, R., Vanstreels, E., Carpentier, S., Hertog, M., Lammertyn, J., Robben, J. *et al.* (2007). Proteomic analysis of core breakdown disorder in 'Conference' pears (*Pyrus communis* L.). *Proteomics, 7,* 2083-2099.

Predreschi, R., Hertog, M., Robben, J., Noben, J. P., Nicolaï, B. (2008). Physiological implications of controlled atmosphere storage of 'Conference' (*Pyrus communis* L.): A proteomic approach. *Postharvest Biology and Technology, 50,* 110-116.

Proust, J., Houlné, G., Schantz, M. L. and Schantz, R. (1996). Characterization and gene expression of an annexin during fruit development in *Capsicum annuum*. *Federation of European Biochemical Societies Letters, 383,* 208-212.

Pua, E., Lim, S. S., Pei, L. and Liu, J. (2000). Expression of a UDP-glucose pyrophosphorylase cDNA during fruit ripening of banana (*Musa acuminata*). *Australian Journal of Plant Physiology, 27*, 1151-1159.

Puehringer, H. M., Zinoecker, I., Marzban, G., Katinger, H., Laimer, M. (2003). mdAP a novel protein in apple, is associated with the major allergen Mal d 1. *Gene, 321,* 369-388.

Quail, P. H., Boylan, M. T., Parks, B. M., Short, T. W., Xu, Y and Wagner, D (1995). Phytochrome: photosensory perception and signal transduction. *Science, 268,* 675-680.

Raab, T., López-Ráez, J. A., Klein, D., Caballero, J. L., Moyano, E., Schwab, W. *et al.* (2006). FaQR, required for the biosynthesis of the strawberry flavor compound 4-hydroxy-2,5-dimethyl-3(2H)-furanone, encodes an enone oxidoreductase. *The Plant Cell, 18*, 1023-1037.

Rahimpour, M., Soheili, M., Eslami, N., Fayaz, A., Sardari, S. and Vaziri, B. (2007). Carrier ampholyte isoelectric focusing based two-dimensional electrophoresis in

preliminary screening of differential proteomics analysis. *Chromatographia, 66,* 133-136.

Rakwal, R. and Agrawal, G. K. (2003). Rice proteomics: Current status and future perspectives. *Electrophoresis, 24,* 3378-3389.

Ramakrishna, W., Deng, Z., Ding, C., Handa, A. K. and Ozminkowski, J. R. H. (2003). A novel small heat shock protein gene, vis1, contributes to pectin depolymerization and juice viscosity in tomato fruit. *Plant Physiology, 131*, 725-735.

Reddy, A. S. N. and Day, I. S. (2001). Analysis of the myosins encoded in the recently completed *Arabidopsis thaliana* genome sequence. *Genome Biology, 2*, 2-7.

Reddy, S. A. and Thomas, T. L. (1996). Unexpected appearance of a toxic substance, octadecatetraeic acid, at genetic engineering of a tobacco plant for production of gamma-linolenic acid. *Nature Biotechnology, 14*, 639-642.

Riccardi, F., Gazeau, P., de Vienne, D. and Zivy, M. (1998). Protein changes in response to progressive water deficit in maize. Quantitative variation and polypeptide identification. *Plant Physiology, 117*, 1253-1263.

Robertson, D., Mitchell, G. P., Gilroy, J. S., Gerrish, C., Bolwell, G. P. and Slabas, A. R. (1997). Differential extraction and protein sequencing reveals major differences in patterns of primary cell wall proteins from plants. *Journal of Biological Chemistry, 272*, 15841-15848.

Robinson, J. P., Harris, S. A. and Juniper, B. E. (2001). Taxonomy of the genus *Malus Mill*. (Rosaceae) with emphasis on the cultivated apple, *Malus domestica* Borkh. *Plant Systemantic and Evolution, 226*, 35-58.

Roche, P., Alston, F. H., Maliepaard, C., Evans, K. M., Vrielink, R., Dunemann, F. *et al.* (1997). RFLP and RAPD markers linked to the rosy leaf curling aphid resistance gene (Sd1) in apple. *Theoretical and Applied Genetics, 94*, 528-533.

Rocco, M., D'Ambrosio, C., Arena, S., Faurobert, M., Scaloni, A., Marra, M. (2006). Proteomic analysis of tomato fruits from two ecotypes during ripening. *Proteomics, 6,* 3781-3791.

Ronimus, R. S., de Heus, E. and Morgan, H. W. (2001). Sequencing, expression, characterisation and phylogeny of the ADP-dependent phosphofructokinase from the hyperthermophilic euryarchaeal *Thermococcus zilligii*. *Biochimica et Biophysica Acta, 1517*, 384-391.

Rose, O. and Falush, D. (2005). A threshold size for microsatellite expansin. *Molecular Biology and Evolution, 15*, 613-615.

Royo, J., Leon, J., Vancanneyt, G., Albar, J. P., Rosahl, S., Ortego, F. *et al.* (1999). Antisense-mediated depletion of a potato lipoxygenase reduces wound induction of proteinase inhibitors and increases weight gain of insect pests. *Proceedings of the National Academy of Science of USA, 96*, 1146-1151.

Ruiz, C. and Asins, M. J. (2003). Comparison between poncirus and citrus genetic linkage maps. *Theoretical and Applied Genetics, 106,* 826-836.

Rupasinghe, H. P. V., Murr, D. P., Paliyath, G. and Skog, L. (2000). Inhibitory effect of 1-MCP on ripening and superficial scald development in 'McIntosh' and 'Delicious' apples. *The Journal of Horticultural Science and Biotechnology, 75*, 271-276.

Rusholme-Pilcher, R. L., Celton, J.-M., Gardiner, S. E., and Tustin, D. S. (2008). Genetic markers linked to the dwarfing trait of apple rootstock 'Malling 9'. *Journal of American Society for Horticultural Science, 133,* 100-106.

Sadka, A., Dahan, E., Or, E. and Cohen, L. (2000). NADP(+)-isocitrate dehydrogenase gene expression and isozyme activity during citrus fruit development. *Plant Science, 158*, 173-181.

Salentijn, E. M. J., Aharoni, A., Schaart, J. G., Boone, M. J. and Krens, F. A. (2003). Differential gene expression analysis of strwberry cultivars that differ in fruit-firmness. *Physiologia Plantarum, 118*, 571-578.

Salgado, K. C., Von Pinho, E. V., Guimaraes, C. T., Von Pinho, R. G., Ferreira, C. A. and Andrade, V. (2008). Mapping of quantitative trait locus associated with maize tolerance to high seed drying temperature. *Genetics and Molecular Research, 7*, 1319-1326.

Samimy, C. and Cummins, J. N. (1992). Distinguishing apple rootstocks by isozyme patterns. *Horticulturae Science, 27*, 829-831.

Sancho, A. I., Foxall, R., Rigby, N. M., Browne, T., Zuidmeer, L., van Ree, R. *et al.* (2006). Maturity and storage influence on the apple (*Malus domestica*) allergen Mal d 3, a nonspecific lipid transfer protein. *Journal of Agricultural Food Chemistry, 54*, 5098-5104.

Saravanan, R. S. and Rose, J. K. C. (2004). A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues. *Proteomics, 4*, 2522–2532.

Sargent, D. J., Fernandez-Fernandez, F., Rys, A., Knight, V. H., Simpson, D. W. and Tobutt, K. R. (2007). Mapping of A1 conferring resistance to the aphid *Amphorophora idaei* and dw (dwarfing habit) in red raspberry (*Rubus idaeus* L.) using AFLP and microsatellite markers. *Plant Biology, 7*, 15.

Sarry, J. E., Sommerer, N., Sauvage, F. X., Bergoin, A., Albagnac, G. and Romieu, C. (2004). Grape berry biochemistry revisited upon proteomic analysis of the mesocarp. *Proteomics, 4,* 210-215.

Sauer, J. D. (1993). *Historical geography of crop plants: a select roster*. CRC press. Florida, USA, pp. 309.

Saviani, E. E., da Silva, A. J. and Martins, J. S. (1998). Role of F1F0-ATP synthase and uncoupling protein in mitochondrial respiratory uncoupling of ripened avocado fruits. *Plant Physiology and Biochemistry, 36*, 575-582.

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

Seffens, W. S., Almoguera, C., Wilde, H. D., Von der Haar, R. A. and Thomas, T. L. (1990). Molecular analysis of a phylogenetically conserved carrot gene: developmental and environmental regulation. *Developmental Genetics, 11*, 65-76.

Seglias, N. P. and Gessler, C. (1997). Genetics of apple powdery mildew resistance from *Malus zumi* (Pl2). *International Organisation for Biological and Integrated Control of Noxious Animals and Plants - West Palaearctic Regional Section Bull, 20*, 195-208.

Seo, S., Okamoto, N., Seto, H., Ishizuka, K., Sano, H. and Ohashi, Y. (1995). Tobacco MAP kinase-a possible mediator in wound signal transduction pathways. *Science, 270*, 1988-1992.

Servakaranpalayam, S. S. (2006). *Potential application of hyperspectral imaging for the determination of total soluble solids, water content and firmness in mango.* MSc thesis, McGill University, Canada, pp. 80.

Severson, D. W. and Kassner, V. A. (1995). Analysis of mosquito genome structure using graphical genotyping. *Insect Molecular Biology, 4,* 279-286.

Seymour, G. B., Manning, K., Eriksson, E. M., Popovich, A. H. and King, G. J. (2002). Genetic identification and genomic organization of factors affecting fruit texture. *Journal of Experimental Botany, 53*, 2065–2071.

Shen, Q.-H., Zhou, F., Bieri, S., Haizel, T., Shirasu, K. and Schulze-Lefert, P. (2003). Recognition specificity and *Rar1/Sgt1* dependence in barley *Mla* disease resistance genes to the powdery mildew fungus. *The Plant Cell, 15*, 732-744.

Shibaike, H. (1998). Molecular genetic mapping and plant evolutionary biology. *Journal of Plant Research, 111*, 383-388.

Shinozuka, H., Hisano, H., Yoneyama, S., Shimamoto, Y., Jones, E.S., Forster, J. W. *et al*. (2006). Gene expression genetic mapping analysis of a perenial ryegrass glycine rich RNA-binding protein gene suggest a role in cold adaption. *Plant Molecular Biology, 275*, 399-408.

Silfverberg-Dilworth, E., Matasci, C. L., Van de Weg, W. E., Van Kaauwen, M. P. W., Walser, M., Kodde, L. P. *et al.* (2006). Microsatellite markers spanning the apple (*Malus* x *domestica* Borkh.) genome. *Tree Genetics and Genomes, 2*, 202-204.

Silfverberg-Dilworth, E., Besse, S., Paris, R., Belfanti, E., Tartarini, S., Sansavini, S., Patocchi, A. and Gessler, C. (2005). Identification of functional apple scab resistance gene promoters. *Theoretical and Applied Genetics, 110*, 1119-1126.

Similarity Matrix of Proteins: http://webclu.bio.wzw.tum.de/portal/web/simap/proteinreport [May 2009].

Sineshchekov, V. A. (2005). Polymorphism of phytochrome A and its functional implications. *In* Wada, M., Shimazaki, K. and Iino, M. (Eds.), *Light sensing in plants*. Springer Japan, Japan, 95-102.

Singer, T., Fan, Y., Chang, H.-S., Zhu, T., Hazen, S. P. and Briggs, S. P. (2006). A high-resolution map of *Arabidopsis* recombinant inbred lines by whole-genome exon array hybridization. *Public Library of Science Genetics, 2*, 1-10.

Singh, R. P., Nie, X., Singh, M., Coffin, R. and Duplessis, P. (2002). Sodium sulphite inhibition of potato and cherry phenolics in nucleic acid extraction for virus detection. *Journal of Virology Methods, 99*, 121-131.

Singha, S. T., Baugher, A., Townsend, E. C. and D'Souza, M. C. (1991a). Anthocyanin distribution in 'Delicious' apples and relationship between anthocyanin concentration and chromaticity values. *Fruit Variety Journal, 116*, 497-499.

Singha, S., Townsend, E. C. and Baugher, T. A. (1991b). Relationship between visual rating and chromaticity values in 'Delicious' apple strains. *Fruit Variety Journal, 45*, 33-36.

Slaughter, D. C., Thompson, J. F. and Tan, E. S. (2003). Nondestructive determination of total and soluble solids in fresh prune using near infrared spectroscopy. *Postharvest Biology and Technology, 28*, 437-444.

Smith, E. F. (2002). Regulation of flagellar dynein by calcium and a role for an axonemal calmodulin and calmodulin-dependent kinase. *Molecular Biology of the Cell, 13*, 3303-3313.

Smith, H. (2000). Phytochromes and light signal perception by plants. An emerging synthesis. *Nature, 407*, 585-591.

Stankiewicz-Kosyl, M., Pitera, E. and Gawronski, S. W. (2005). Mapping QTL involved in powdery mildew resistance of the apple clone U211. *Plant Breeding, 124*, 63-66.

Stanley, C. J., Stokes, J. R. and Tustin, D. S. (2001). Early prediction of apple fruit size using environmental indicators. *Acta Horticulturae, 557*, 441-446.

Stebbins, R. L. (1991). Current situation and future trends in apple cultivars in the Pacific Northwest. *Fruit Variety Journal, 45*, 79-83.

Stone, J. M., Trotochaud, A. E., Walker, J. C. and Clark, S. E. (1998). Control of meristem development by *clavata1* receptor kinase and kinase-associated protein phosphatase interactions. *Plant Physiology, 117*, 1217-1225.

Sunako, T. (1999). An allele of the ripening-specific 1-aminocyclopropane- 1-carboxylic acid synthase gene (ACS1) in apple fruit with a long storage life. *Plant Physiology, 119*, 1297-1303.

Suzuki, T., Ueda, A., Kobayashi, N., Yang, J., Tomaru, K., Yamamoto, M. *et al.* (2008). Proteasome-dependent degradation of α-catenin is regulated by interaction with ARMc8α. *Biochemistry Journal, 411*, 581-591.

Swiderski, M. R., Birker, D. and Jones, J. D. G. (2009). The TIR domain of TIR-NB-LRR resistance proteins is a signaling domain involved in cell death induction. *Molecular Plant Microbe Interaction, 22,* 157-165.

Takos, M. A., Jacob, A. S., Bogs, J., Robinson, S. P. and Walker, A. R. (2006). Light-induced expression of a *MYB* gene regulates anthocyanin biosynthesis in red apple. *Plant Physiology, 142*, 1216-1232.

Tang, J., Xia, H., Cao, M., Zhang, X., Zeng, W., Hu, S. *et al.* (2004). A comparison of rice chloroplast genomes. *Plant Physiology, 135*, 412-420.

Tanksley, S. D. (1993). Mapping polygenes. *Annual Review of Genetics, 27*, 205-233.

Tanksley, S. D. and Hewitt, J. (1988). Use of molecular markers in breeding for soluble solids contents in tomato a re-examination. *Theoretical and Applied Genetics, 75*, 811-823.

Tanksley, S. D., Young, N. D., Paterson, A. H. and Bonierbale, M. W. (1989). RFLP mapping in plant breeding: new tools for an old science. *Biotechnology 7*, 257–264

Tartarini, S., Gennari, F., Pratesi, D., Palazzetti, C., Sansavini, S., Parisi, L. *et al.* (2004). *Characterisation and genetic mapping of a major scab resistance gene from the old Italian apple cultivar 'Durello di Forli'*. Paper presented at the Eucarpia symposium on fruit breeding and genetics.

Tartarini, S. and Sansavini, S. (2003). The use of molecular markers in pome fruit breeding. *Acta Horticulturae, 622*, 129-132.

Taylor, D. R. and Knight, J. N. (1986). Russeting and cracking of apple fruit and their control with plant growth regulators. *Acta Horticulturae, 179*, 817-820.

Terryn, N., Van Montagu, M. and Inzé, D. (1993). GTP-binding proteins in plants. *Plant Molecular Biology, 22*, 143-152.

Thomas, L. A. (2008). *Regulation of sulfur assimilation in onion (Allium cepa L.)*. PhD thesis, Massey University, New Zealand, pp. 215.

Thompson, J. M. and Taylor, I. (1971). Genetic susceptibility to *Glomerella* leaf blotch in apple. *Journal of Heredity, 62*, 303-306.

Tijskensa, L. M. M., Rodisb, P. S., Hertoga, M. L. A. T. M., Waldronc, K. W., Inghamc, L., Proxeniab, N. *et al.* (1997). Activity of peroxidase during blanching of peaches, carrots and potatoes. *Journal of Food Engineering, 34*, 355-370.

Tobutt, K. R., Boskovic, R. and Roche, P. (2000). Incompatibility and resistance to woolly apple aphid in apple. *Plant Breeding, 119*, 65-69.

Tong, W. M., Cortes, U. and Wang, Z. Q. (2001). Poly(ADP-ribose) polymerase: a guardian angel protecting the genome and suppressing tumorigenesis. *Biochimica et Biophysica Acta, 1552*, 27-37.

Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J. *et al.* (2001). Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics, 1*, 377-396.

Toraskar, V. M. and Modi, V. V. (1984). Peroxidase and chilling injury in banana fruit. *Journal of Agriculture and Food Chemistry, 32*, 1352-1354.

Tu, K. and de Baerdenaeker, J. (1997). A study of prestorage heat treatment effect on apple texture: Destructive and nondestructive measurements. *Journal of food Processing and Preservation, 21*, 495-506.

TuTiempo: http://www.tutiempo.net/en/Climate/Cape_Town [May 2009].

Unlu, M., Morgan, M. E. and Minden, J. S. (1997). Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. *Electrophoresis, 18,* 2071-2077.

Valderrama, P. and Clemente, E. (2004). Isolation and thermostability of peroxidase isoenzymes from apple cultivars Gala and Fuji. *Food Chemistry, 87*, 601-606.

Van Berloo, R. (1999). GGT: software for the display of graphical genotypes. *Journal of Heredity, 90*, 328-329.

van Berloo, R., Zhu, A., Ursem, R., Verbakel, H., Gort, G. and van Eeuwijk, F. A. (2008). Diversity and linkage disequilibrium analysis within a selected set of cultivated tomatoes. *Theoretical and Applied Genetics, 117*, 89-101.

van Der Biezen, E. A. and Jones, J. D. G. (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends in Biochemical Sciences, 23*, 454-456.

van Der Straeten, D., Rodrigues-Pousada, R. A., Goodman, H. M. and Van Montagu, M. (1991). Plant enolase: gene structure, expression, and evolution. *The Plant Cell, 3*, 719-735.

van der Vossen, E. A., van der Voort, J. N., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D. C. *et al.* (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *Plant Journal, 23*, 567-576.

van Dyk, M. M., Labuschagne, I. F. and Rees, D. J. G. (2009). Genetic linkage map construction and identification of QTLs affecting time of initial vegetative budbreak in apple (*Malus x domestica* Borkh.). *Acta Horticulturae, 814*, 585-590.

van Dyk, M. M. (2008). Identification of quantitative trait loci controlling the requirement for chilling in vegetative budbreak in apple (*Malus x domestica* Borkh.). PhD Thesis. University of the Western Cape, South Africa.

van Eenennaam, A. (2004). *Marker-assisted selection backgrounder*. Available: http://animalscience.ucdavis.edu/animalbiotech/Outreach/Marker_Assisted_Selection_in_Beef_Cattle.pdf, [April 2008]

van Ooijen, J. W. (1999). LOD significance thresholds for QTL analysis in experimental populations of diploid species. *Heredity, 83*, 613-624.

van Ooijen, J. W. (2004). MapQTL® 5, Software for the mapping of quantitative triat loci in experimental populations. Wageningen, Netherlands: Kyazma B. V.

van Ooijen, J. W. and Voorrips, R. E. (2001). JoinMap® 3.0, Software for the calculation of genetic linkage maps. Wageningen, The Netherlands: *Plant Research International*.

van Ooijen, J. W. and Voorrips, R. E. (2006). JoinMap® 4.0, Software for the calculation of genetic linkage maps. Wageningen, The Netherlands: *Plant Research International*.

Ventura, M., de Jager, A., de Putter, H. and Roelofs, F. P. M. M. (1998). Non-destructive determination of soluble solids in apple fruit by near infrared spectroscopy (NIRS). *Postharvest Biology Technology, 14*, 21-27.

Verma, D. P. S. (2001). Cytokinesis and building of the cell plate in plants. *Annual Review in Plant Physiology, 52*, 751-784.

Verma, D. P. S. and Hong, Z. (2008). Cell division control in plants, *Plant Science*, *9*, 292-295.

Vernoud, V., Horton, A. C., Yang, Z. and Nielsen, E. (2003). Analysis of the small GTPase gene superfamily of *Arabidopsis*. *Plant Physiology, 131*, 1191-1208.

Veronico, P., Giannino, D., Melillo, M. T., Leone, A., Reyes, A., Kennedy, M., W *et al.* (2006). A novel lipoxygenase in pea roots. Its function in wounding and biotic stress. *Plant Physiology, 141*, 1045-1055.

Vieths, S., Schoning, B. and Jankiewicz, A. (1993). Occurrence of IgE binding allergens during ripening of apple fruits. *Food, Agriculture and Immunology, 5,* 93-105.

Vinatzer, B. A., Patocchi, A., Tartarini, S., Gianfranceschi, L., Sansavini, S. and Gessler, C. (2004). Isolation of two microsatellite markers from BAC clones of the *Vf* scab resistance region and molecular characterization of scab-resistant accessions in *Malus* germplasm. *Plant Breeding, 123*, 321-326.

Vinterhalter, D. V. and James, D. T. (1986). The use of peroxidase polymorphism in the identification of Malling and Malling Merton apple rootstocks. *Journal of Horticultural Science, 110*, 509-515.

Vision, T. J., Brown, D. G., Shmoys, D. B., Durrett, R. T. and Tanksley, S. D. (2000). Selective mapping: a strategy for optimizing the construction of high density linkage maps. *Genetics, 155*, 407-420.

Visser, T., Schaap, A. A. and De Vries, D. P. (1968). Acidity and sweetness in apple and pear. *Euphytica, 17*, 153-167.

Vogel, J., Börner, T. and Hess, W. (1999). Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plants mutants. *Nucleic Acids Research, 27,* 3866-3874.

Vogl, C. and Xu, S. Z. (2000). Multipoint mapping of viability and segregation distorting loci using molecular markers. *Genetics, 155*, 1439-1447.

Volz, R. K., Harker, F. R. and Lang, S. (2003). Firmness decline in `Gala' apple during fruit development. *Journal of American Society for Horticultural Science, 128*, 797-802.

Volz, R. K., Harker, F. R., Hallet, I. C. and Lang, S. (2004). Development of texture in apple fruit-a biophysical perspective. *Acta Horticulturae, 636,* 473-479.

Voorrips, R. E. (2002). MapChart: Software for the graphical resentation of lonkage maps and QTL's. *The Journal of Heredity, 93*, 77-78.

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

Wadsworth, G. J. (2006). The plant aspartate aminotransferase gene family. *Physiology of Plant, 100*, 998-1006.

Wadsworth, G. J., Gebhardt, J. S. and Matthews, B. F. (1995). Characterization of a soybean cDNA clone encoding the mitochondrial isozyme of aspartate aminotransferase, AAT4. *Plant Molecular Biology, 27*, 1085-1095.

Wang, G., Ellendorff, U., Kemp, B., Mansfield, J. W., Forsyth, A., Mitchell, K. *et al.* (2008a). A genome-wide functional investigation into the roles of receptor-like proteins in *Arabidopsis*. *Plant Physiology, 147*, 503-517.

Wang, W., Scali, M., Vignani, R., Spadafora, A., Sensi, E., Mazzuca, S. *et al.* (2003). Protein extraction for two-dimensional electrophoresis from olive leaf, a plant tissue containing high levels of interfering compounds *Electrophoresis, 24*, 2369-2375.

Wang, W., Tai, F. and Chen, S. (2008b). Optimizing protein extraction from plant tissues for enhanced proteomics analysis. *Journal of Separation Science, 31*, 2032-2039.

Wang, W., Vignani, R., Scali, M. and Cresti, M. (2006). A universal and rapid protocol for protein extraction from recalcitrant plant tissues for proteomic analysis. *Electrophoresis, 27*, 2782-2786.

Wang, X. (1999). The role of phospholipase D in signaling cascades. *Plant Physiology, 120*, 645-651.

Wang, X. (2005). Regulatory functions of phospholipase D and phosphatidic acid in plant growth, development and stress responses. *Plant Physiology, 139*, 566-573.

Wang, X., Yang, P., Gao, Q., Liu, X., Kuang, T., Shen, S. *et al.* (2008). Proteomic analysis of the response to high-salinity stress in *Physcomitrella patens*. *Planta, 228*, 167-177.

Wang, X., Xu, Y., Peng, C., Fan, R. and Gao, X. (2009). Ubiquitous distribution and different subcellular localization of sorbitol dehydrogenase in fruit and leaf of apple. *Journal of Experimental Botany, 60,* 1025-1034.

Watanabe, N., Akond, A. S. and Nachit, M. M. (2006). Genetic mapping of the gene affecting polyphenol oxidase activity in tetraploid durum wheat. *Journal of Applied Genetics, 47*, 201-205.

Watkins, G. B., Nock, J. E. and Whitaker, B. D. (2000). Responses of early, mid and late season apple cultivars to postharvest application of 1-methylcyclopropene (1-MCP) under air and controlled atmosphere storage conditions. *Postharvest Biology Technology, 19*, 17-32.

Weber, D. and Helentjaris, T. (1989). Mapping RFLP loci in maize using B-A translocations. *Genetics, 121*, 583-590.

Weebadde, C. K., Wang, D., Finn, C. E., Lewers, K. S., Luby, J. J., Bushakra, J. *et al.* (2008). Using a linkage mapping approach to identify QTL for day-neutrality in the octoploid strawberry. *Plant Breeding, 127*, 94-101.

Weeden, N. F., Hemmat, M., Lawson, D. M., Lodhi, M., Bell, R. L., Manganaris, A. G. *et al.* (1994). Development and application of molecular marker linkage maps in woody fruimt crops. *Euphytica, 77,* 71-75.

Weeden, N. F. (1989). Applications of isozymes: In plant breeding. *Plant Breeding, 6*, 1154-1160.

Weeden, N. F. and Lamb, R. C. (1987). Genetics and linkage analysis of 19 isozyme loci in apple. *Journal of American Society for Horticultural Science, 112*, 865-872.

Wei, F., Wing, R. A, Wise, R. P. (2002). Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley. *The Plant Cell, 14*, 1903-1917.

Wei, F., Gobelman-Werner, K., Morroll, S. M., Kurth, J., Mao, L., Wing, *et al.* (1999). The *Mla* (pwdery mildew) resistance cluster is associated with three NBS-LRR gene families and supresed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics, 153,* 1929-1948.

Weil, C. F. and Wessler, S. R. (1990). The effects of plant transposable element insertion on transcription initiation and RNA processing. *Annual Review of Plant Physiology and Plant Molecular Biology, 41*, 527-552.

Weiss, K. C. and Bisson, L. F. (2001). Optimisation of the Amido Black assay for determination of the protein content of grape juices and wines. *Journal of the Science of Food and Agriculture, 81*, 583-589.

Welsh, J. and McClelland, M. (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research, 8*, 7213-7218.

Wen, C., Wu, M., Goh, C. and Pua, E. (1995). Cloning and nucleotide sequence of a cDNA encoding S-adenosyl-L-methionine synthetase from mustard (*Brassica juncea* L.). *Plant Physiology, 107*, 1021-1022.

Wen, Y. and Wu, W. (2008). Experimental designs and statistical methods for mapping quantitative trait loci underlying triploid endosperm traits without maternal genetic variation. *Journal of Heredity, 99*, 546-551.

Wenzl, P., Caig, V., Carling, J., Cayla, C., Evers, M., Jaccoud, D. *et al.* (2004). *Diversity Arrays Technology, a novel tool for harnessing crop genetic diversity.* Paper presented at the 4th International crop science congress.

Wheeler, N., Payne, P., Hipkins, V., Saich, R., Kenny, S. and Tuskan, G. (2006). Polymix breeding with paternity analysis in Populus: A test for differential reproductive success (DRS) among pollen donors. *Tree Genetics and Genomics, 2*, 53-60.

Wheelwright, N. T. and Logan, B. A. (2004). Previous-year reproduction reduces photosynthetic capacity and slows lifetime growth in females of a neotropical tree. *Proceedings of the National Academy of Sciences of USA, 101*, 8051-8055.

Whitaker, B. D., Smith, D. L. and Green, K. C. (2001). Cloning, characterization and functional expression of a phospholipase D alpha from tomato fruit. *Physiologia Plantarum, 112*, 87-94.

Whitkus, R., Doebley, J. and Lee, M. (1992). Comparative genome mapping of sorghum and maize. *Genetics, 132*, 1119-1130.

Whittaker, D. J., Smith, C. S. and Cardner, R. C. (1995). Three cDNAs encoding S-adenosyl-1-methionine synthetase from *Actinidia chinensis*. *Plant Physiology, 108*, 1307-1308.

Wilkinson, J. Q., Lanahan, M. B., Conner, T. W. and Klee, H. J. (1995). Identification of mRNAs with enhanced expression in ripening strawberry fruit using polymerase chain reaction differential display. *Plant Molecular Biology, 27,* 1097-1108.

Williams, J. L. and Conner, J. K. (2001). Sources of phenotypic variation in floral traits in wild radish, *Raphanus raphanistrum* (Brassicaceae). *American Journal of Botany, 88*, 1577-1581.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. and Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research, 18*, 6531-6535.

Williams, M. N. V., Pande, N., Nair, S., Mohan, M. and Bennett, J. (1991). Restriction fragment length polymorphism analysis of polymerase chain reaction products amplified from mapped loci of rice (*Oryza sativa* L.) genomic DNA. *Theoretical and Applied Genetics, 82*, 489-498.

Wise, R. P. and Ellingboe, A. H. (1983). Infection kinetics of *Erysiphe graminis* f. sp. *hordei* on barley with different alleles at the Ml-a locus. *Phytopathology, 73*, 1220-1222.

Wisniewski, M., Close, T. J., Artlip, T. and Arora, R. (1999). Seasonal patterns of dehydrins and 70-kDa heat-shock proteins in bark tissues of eight species of woody plants. *Physiology of Plants, 96*, 496-505.

Woodhead, M., McCallum, S., Smith, K., Cardle, L., Mazzitelli, L. and Graham, J. (2008). Identification, characterisation and mapping of simple sequence repeat (SSR) markers from raspberry root and bud ESTs. *Molecular Breeding, 22,* 555-

563.

Wu, R. L., Ma, C.-X. and Casella, G. (2007). *Statistical genetics of quantitative traits: Linkage, maps, and QTL*. Springer-Verlag, New York.

Wu, F. S. and Wang, M. Y. (1984). Extraction of proteins for SDS-PAGE. *Analytical Biochemistry, 139,* 100-103.

Wunsch, A. and Hormaza, J. I. (2007). Characterization of variability and genetic similarity of European pear using microsatellite loci development in apple. *Scientia Horticulturae, 113*, 37-43.

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

Xu, C., Garrett, W. M., Sullivan, J., Caperna, T. J. and Natarajan, S. (2006). Separation and identification of soybean leaf proteins by two-dimensional gel electrophoresis and mass spectrometry. *Phytochemistry, 67*, 2431-2440.

Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J. *et al.* (2003). Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science, 302*, 842-846.

Yamaguchi, Y., Nakamura, T., Kusano, T. and Sano, H. (2000). Three *Arabidopsis* genes encoding proteins with differential activities for Cys synthase and β-cyanoalanine synthase. *Plant and Cell Physiology, 41*, 465-476.

Yamaki, S. (1987). ATP-promoted sorbitol transport into vacuoles isolated from apple fruit. *Plant and Cell Physiology, 28*, 557-564.

Yamaki, S. (1980). A sorbitol oxidase that coverts sorbitol to glucose in apple leaf. *Plant and Cell Physiology, 21*, 591-599.

Yamamoto, T., Kimura, T., Sawamura, Y., Kotobuki, K., Ban, Y., Hayashi, T. *et al.* (2001). SSR's 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. *et al.* (2002a). Simple sequence repeats for genetic analysis in pear. *Euphytica, 124*, 129-137.

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

Yang, H. Y., Korban, S. S., Kruger, J. and Schmidt, H. (1997a). A randomly amplified polymorphic DNA (RAPD) marker tightly linked t the scab-resistance gene Vf in apple. *Journal of American Society of Horticulture Science, 122*, 47-52.

Yang, S. F. and Hoffman, N. E. (1984). Ethylene biosysnthesis and its regulation in higher plants. *Annual Review of Plant Physiology, 35*, 155-189.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G. *et al.* (2006). The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology, 60*, 107-124.

Yin, X., Stam, P., Kropff, M. J. and Schapendonk, A. H. C. M. (2005). Crop modeling, QTL mapping and their complementary role in plant breeding. *Agronomy Journal, 95*, 90-98.

Yin, X.-R., Chen, K.-S., Allan, A. C., Wu, R.-M., Zhang, B., Lallu, N. *et al.* (2008). Ethylene-induced modulation of genes associated with the ethylene signalling pathway in ripening kiwifruit. *Journal of Experimental Botany, 59*, 2097-2108.

Yin, Z. J. (2006). Methodology of QTL's linkage analysis for discrete traits controlled by polygenes. *Yi Chuan, 28*, 578-582.

Yoder, J. I. (2001). Host-plant recognition by parasitic Scrophulariaceae. *Plant Biology, 4*, 359-365.

Yokoi, S., Quintero, F. J., Cubero, B., Ruiz, M. T., Bressan, R. A., Hasegawa, P. M. *et al.* (2002). Differential expression and function of Arabidopsis thaliana NHX Na^+/H^+ antiporters in the salt stress response. *Plant Journal, 30*, 529-539.

Young, N. D. (1992). Applications of DNA genetic markers to the study of plant growth and development. *Plant Growth Regulation, 12*, 229-236.

Young, N. D. (1996). QTL mapping and quantitative disease resistance in plants. *Annual Review of Phytopathology, 34*, 479-501.

Zaragoza, S., Trenor, I., Alonso, E., Primo-Millo, E. and Agusti, M. (1992). Treatments to increase the final fruit size on *Satsuma clausellina*. *Proceedings for the International Society of Citriculture, 2,* 725-728.

Zeleny, R., Kolarich, D., Strasser, R. and Altmann, F. (2006). Sialic acid concentrations in plants are in the range of inadvertent contamination. *Planta, 224*, 222-227.

Zhang, C., Tanabe, K., Wang, S., Tamura, F., Yoshida, A. and Matsumoto, K. (2006). The impact of cell division and cell enlargement on the evolution of fruit size in *Pyrus pyrifolia*. *Annals of Botany, 98*, 537-543.

Zhang, D., Chen, S., Peng, Y. and Shen, Y. (2001). Abscisic acid-specific binding sites in the flesh of developing apple fruit. *Journal of Experimental Botany, 52*, 2097-2103.

Zhang, G., Angeles, E. R., Abenes, M. L. P., Khush, G. S. and Huang, N. (1994). Molecular mapping of a bacterial blight resistance gene on chromosome 8 in rice. *Rice Genetics, 11*, 142-144.

Zhang, H. and Cramer, W. A. (2004). Purification and crystallization of the cytochrome b6f complex in oxygenic photosynthesis. *Methods in Molecular Biology, 274*, 67-78.

Zhang, H., Whitelegge, J. P. and Cramer, W. A. (2001). Ferredoxin:$NADP^+$ oxidoreductase is a subunit of the chloroplast cytochrome b6 f complex. *Journal of Biological Chemistry, 276*, 38159-38165.

Zhao, Z., Zhang, W., Stanley, B. A. and Assmann, S. M. (2008). Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. *The Plant Cell, 20*, 3210-3226.

Zhou, F., Kurth, J., Wei, F., Elliot, C., Vale, G., Yahiaoui, N. *et al.* (2001). Cell-autonomous expression of barley *Mla1* confers race-specific resistance to the powdery mildew fungus via a *Rar1*-independent signalling pathway. *The Plant Cell, 13*, 337-350.

Zhu, J. (2003). Regulation of ion homeostasis under salt stress. *Plant Biology, 6*, 441-445.

Zou, W., Aylor, D. L. and Zeng, Z. B. (2007). eQTL Viewer: visualizing how sequence variation affects genome-wide transcription. *BMC Bioinformatics, 8*, 7-11.

Zuo, X. and Speicher, D. W. (2000). A method for global analysis of complex proteomes using sample prefractionation by solution isoelectrofocusing prior to two-dimensional elctrophoresis. *Analytical Biochemistry, 284*, 266-270.