Identification of Quantitative Trait Loci Controlling the Requirement for Chilling in Vegetative Budbreak in Apple (*Malus* x *domestica* Borkh.)

by

Maria Magdalena van Dyk

Department of Biotechnology

University of the Western Cape

Supervisors: Prof. D. J. G. Rees

Dr. I.F. Labuschagné

Abstract

The domesticated apple (*Malus* x *domestica* Borkh.) has been distributed into diverse climatic conditions worldwide for commercial production of fruit. Apple trees need exposure to cold temperatures, referred to as chill unit (CU) accumulation during winter, in order for budbreak to occur promptly and uniformly after winter. In warmer production areas the application of dormancy breaking chemicals has enabled successful production of high chilling requiring apple cultivars in suboptimal environmental conditions. In the Western Cape region of South Africa it is common orchard practice to apply dormancy breaking chemicals after winter in order to stimulate vegetative growth. If this is not done prolonged dormancy symptoms (PDS) are experienced which include extended rest, less synchronised breaking of buds and reduced branching. An increasing awareness of both global warming and the negative effects associated with the use of chemical sprays (for both pest and disease resistance and growth regulation) has resulted in the need to breed cultivars better adapted to current and future environmental conditions.

The breeding of new cultivars using conventional breeding methods is a time consuming process, especially in perennial tree species with a long juvenile phase such as apple. The implementation of marker-assisted-breeding (MAB) and selection (MAS) will enable the selection of favourable genotypes at a very early seedling stage. Although markers linked to genes involved in disease resistance for a variety of known apple pathogens have been identified and are already in use in breeding programs, the genetic determinants of dormancy related characteristics residing within the bud itself

(endodormancy) are poorly understood. This hampers the genetic improvement of such characters. Although this study focused on time of initial vegetative budbreak IVB, there are various other characteristics that can be associated with dormancy, such as position and number of budbreak and budbreak duration.

The unravelling of the genetic basis of complex traits such as dormancy, can be done through the construction of a genetic linkage map followed by the identification of genomic regions, known as quantitative trait loci (QTL), that can be ssociated with the trait of interest. This study involved the construction of genetic linkage maps for two mapping pedigrees where the low chilling requiring cultivar 'Anna' was used as common male parent in crosses with the higher chilling requiring 'Golden Delicious' and 'Sharpe's Early'. A third mapping pedigree, with 'Golden Delicious' as female parent and 'Prima' as male parent, was also included. Maps consisted of transferable SSR markers only, facilitating the alignment with the proposed apple reference map (Silfverberg-Dilworth *et al.*, 2006) and adherence to the common LG numbering system now being used for apple genetic linkage maps (Maliepaard *et al.*, 1998). A number of newly developed EST-SSR markers are reported, some of which are candidates for filling large gaps between adjacent SSR markers on the apple reference map. An interactive database was developed to successfully manage the large amount of data generated during this investigation. A selective mapping, or bin mapping strategy (Vision *et al.*, 2000) was developed for two of the three mapping populations in order to facilitate the incorporation and positioning of newly developed markers onto existing genetic linkage maps. This involves the

screening of new markers on a small subset of the population, drastically reducing the cost and time involved.

Genetic linkage maps constructed allowed for the detection of 18 putative QTLs affecting the time of IVB. Four of these QTLs co-localize with previously identified QTLs. A QTL identified on LG 8 confirms a previously identified QTL (Segura *et al.*, 2007), while one of the QTLs identified on LG 9 might coincide with a QTL identified on the corresponding LG 3 of the genetic linkage map constructed by Conner *et al.* (1998). Two QTLs identified on LG 10 might coincide with markers found to co-segregate with time of budbreak in an earlier study conducted by Lawson *et al.* (1995). An additional 14 QTLs involved in time of IVB have been identified. We proposed the testing of four markers in a validation study conducted on a second mapping pedigree derived from a cross between 'Anna' and 'Golden Delicious'. These markers are CH04a12, CH04c06y, CH01h01 and A267. Not only do these markers show significant levels of association with the time of IVB, but segregation of parental alleles from the cultivar 'Anna' for two of these markers, CH04c06y and CH01h01, were found to be associated with the time of IVB in different genetic backgrounds. The identification of markers closely associated with time of IVB will facilitate the implementation of MAS in breeding programs in order to breed cultivars that are better adapted to local climatic conditions.

List of Abbreviations

AFLP	Amplified Fragment Length Polymorphism
APS	Ammonium PeroxodiSulfate
BLAST	Basic Local Alignment Search Tool
bp	Base pair
cDNA	Complementary DeoxyriboNucleic Acid
cM	CentiMorgan
CR	Chilling Requirement
CU	Chilling Units
dATP	Deoxy Adenosine TriPhosphate
dCTP	Deoxy Cytosine TriPhosphate
DFPT	Deciduous Fruit Producers Trust
dGTP	Deoxy Guanosine TriPhosphate
DNA	DeoxyriboNucleic Acid
DNOC	DiNitro-Ortho-Cresol mineral oil
dNTP	Any of dATP, dCTP, dGTP, or dTTP, or a mixture of all four
dTTP	Deoxy Thymidine TriPhosphate
EDTA	Ethylene Diamine Tetra acetic Acid (disodium salt unless specified)
EST	Expressed Sequence Tag
F_1	First filial generation
HTML	HyperText Mark-up Language
LG	Linkage Group

LOD	Logarithm (base 10) of ODds
MAB	Marker Assisted Breeding
MAS	Marker Assisted Selection
Mb	Million base pairs
PAGE	PolyAcrylamide Gel Electrophoresis
PBS	Phosphate-Buffered Saline
PCR	Polymerase Chain Reaction
PDS	Prolonged Dormancy Symptoms
PGR	Plant Growth Regulator
PIC	Polymorphism Information Content
QT	Quantitative Trait
QTL	Quantitative Trait Locus
RAPD	Random Amplified Polymorphic DNA
RFLP	Random Fragment Length Polymorphism
SCAR	Sequence Characterized Amplified Region
SSR	Simple Sequence Repeat
STR	Short Tandem Repeat
STS	Sequence Tagged Site
Taq™	*Thermus aquaticus* DNA polymerase
TEMED	N,N,N',N'-Tetra Methyl-Ethylene Diamine
Tiff	Tagged information file format
T_m	Melting temperature (of oligonucleotide)
Tris (Base)	Tris(hydroxymethyl)aminomethane
WWW	World wide web

Acknowledgements

I would like to especially thank the following people and institutions:

- My heavenly Father for strength, health, patience and perseverance to fulfill my dreams.
- My husband, Riaan, and two lovely daughters, Amanita and Elzerie, for your patience, support and understanding.
- My supervisor, Jasper Rees, for the opportunity to become part of the apple research group at the University of the Western Cape and for all the guidance during my PhD studies. Thank you for all the opportunities to present my work nationally as well as internationally and all the associated opportunities to travel and attend workshops related to my field of interest. Thank you for the financial support that made it possible for me to further my academic career and for the understanding and support when I had to relocate to Pretoria due to personal reasons.
- Prof. Brenda Wingfield at the Department of Genetics at the University of Pretoria for the use of laboratory space as well as other UP and FABI facilities.
- My co-supervisor, Iwan Labuschagnè, for all your patience answering orchard-related questions, guidance and advise.
- All the members of the apple research group at the University of the Western Cape for your friendship, support and helpful discussions. A special thank you to Khashief Soeker for assistance during the initial steps of sequence identification and primer design.

- People in the Department of Genetics at the University of Pretoria and FABI for friendship, useful discussions, help and advice. A special thank you to Vivian Clarens for making sure all my laboratory needs are met.
- Support staff in the Department of Biotechnology at the University of the Western Cape.
- My parents, Lucas and Elsa de Clercq, my parents-in-law, Johnny and Marlene van Dyk, and my sister and brother-in-law, Annake and Danie Botha, for your love and support and all the help provided in caring for my children when I had to put in some extra hours of work or needed to travel internationally.

Declaration

I herewith declare that the work presented in this dissertation is my own and have not previously been submitted to any other institution in order to qualify for a post-graduate degree.

M.M. van Dyk

To Riaan

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1 CHAPTER 1: LITERATURE REVIEW

1.1 GENERAL INTRODUCTION

1.1.1 Apples

Apples belong to the genus *Malus*, which form part of the pome fruits under the subfamily Maloideae in the Rosaceae family. These fruits are characterized by having two to five carpels enclosed in a fleshy covering. The description of the number of species in this genus varies, due to differences between taxonomic viewpoints (Gardiner, *et al.*, 2007), but it has been reported as having 24 primary species (Janick, 2005). The genus also include several species of so-called crabapples, many of which are cultivated as ornamental trees for their profuse blossom and attractive fruits (Janick, *et al.*, 1996).

South Africa has a highly developed and internationally competitive deciduous fruit industry, ranking among the top 20 countries when it comes to production volume. Over the past decade (1996/1997 to 2005/2006 seasons) annual production volumes ranged between 626 000 and 822 000 tons, from a production area estimated at 20 633 hectares in 2006 (source: Deciduous Fruit Producers' Trust (DFPT) http://www.deciduous.co.za). During the 2005/2006 season 42 % of apples produced in South Africa were exported, of which 44 % were exported to the United Kingdom alone (source: Perishable Product Export Control Board (PPECB) http://www.ppecb.com). It is thus clear that the apple industry plays a very important role in South Africa's economy and in the Western Cape agricultural sector.

Apples are popular because of the many ways in which they can be consumed. They may be eaten from the tree or be stored for up to a year. The consumption of apples by the local market was estimated at 58 % of all produce for the 2005/2006 growing season, of which 55 % were dedicated to the fresh produce market, 45 % were processed into sauces, slices, sweets, alcoholic beverages, vinegar or juice and less than 1 % were dried (source: DFPT). Apples became a symbol of wholesomeness with the slogan 'An apple a day keeps the doctor away'.

1.1.1.1 Apple cultivars

A wide variety of apple cultivars are available, each selected and propagated, or cloned, for one or more desirable trait/s, such as disease resistance, fruit quality, colour, flavour, chilling requirement, etc. Apples are grown as composite trees consisting of a clonally replicated fruiting scion budded or grafted onto a rootstock, usually also selected on the basis of a desired trait, such as disease resistance or size control. Occasionally a three-part tree, that includes a genetically distinct trunk or interstem, are used where the scion and rootstock will not form a strong graft union with each other. Apple trees are usually 30 to 40 feet high with short trunks and a fairly round crown of branches, although this varies according to rootstock and training system used. Trees are normally crafted onto dwarfing rootstocks in order to make orchard management and harvesting easier.

Until the latter half of the twentieth century most of the world's apple cultivars were chance seedlings selected by fruit growers, such as the well known cultivar 'Golden Delicious' that was found in West Virginia, America, over 100 years ago. As the fruit growers knowledge and understanding of

genetics, and more specifically the inheritance of traits from parents to offspring, broadened, the breeding of apple cultivars became a highly profitable and competitive industry. Today breeders strive towards breeding apple cultivars with characteristics, such as disease resistance and fruit quality, that are in accordance with consumer demands. Although more than 10 000 cultivars are documented, these consumer demands have resulted in only a few cultivars being grown on a commercial scale worldwide. In South Africa 'Granny Smith' accounts for 25 % of total production followed by 'Golden Delicious' (22 %), 'Royal Gala' (12 %), 'Pink Lady' (7 %), 'Starking' (6 %) and 'Topred' (6 %) (source: DFPT). The appeal for new cultivars is very strong, but the cultivar picture in the industry has changed very little over the past decade and will remain fairly constant according to forecasts made for 2010 and 2015 (source: DFPT). This can mainly be ascribed to the long life of the standard orchard, resulting in a very small fraction of the bearing trees being pulled out and orchards replanted each year. The long juvenile phase during which the farmer has no financial benefit when replanting an orchard with a new cultivar, also adds to the slow release of newer cultivars into the market. In South Africa 31 % of apple orchards are in excess of 25 years and only 8 % of orchards are younger than 3 years (Source: DFPT).

1.1.2 Apple breeding

Breeding of new apple cultivars has proven to be a very long and tedious process requiring more than 20 years, including periods of cross pollination, seedling selection and field trials. Selection processes has always been complicated by the slow growth, the long juvenile phase, the high level of heterozygosity and the strong self-incompatibility present in this species.

These factors has lead to the, fairly recent, release of the first South African bred apple cultivar, 'African Carmine', in 1999, although the breeding program at the Agricultural Research Council (ARC) has been running for almost 30 years. These time constraints often make conventional breeding, or the conventional selection methods, impractical and this has stimulated an interest in the apple genome and molecular marker techniques in order to apply Marker Assisted Breeding (MAB). These techniques have been used successfully in breeding character specific cultivars of annual crop species, such as maize. These new technologies will not replace, but will complement conventional breeding in order to produce cultivars with desired traits after a shorter period of time and with less cost involved in maintaining trees that will only show their 'undesirable' characteristics after years of costly field maintenance and evaluation.

1.1.2.1 Structure of genetic material

Evolutionary studies indicate that changes in the genetic make-up of organisms occur and that these eventually result in speciation. Since hybrids between species are usually sterile due to a lack of chromosome homology and uneven segregation of chromosomes during meiosis, the mere hybridization of two species do not result in the formation of a new species. However, chromosome doubling in the zygote, containing a complete set of chromosomes from each parent, may result in the formation of a tetraploid with two sets of homologous chromosomes in which paring and crossing over occur normally. Such a plant is known as an allotetraploid and is completely fertile. The Maloideae are believed to be allopolyploids, which is not a rare

phenomenon in the plant kingdom and usually results in larger and more vigorous plants.

The Rosaceae family has four sub-families each with a specific basic chromosome number, the Rosoideae has a base chromosome number of x=7, the Prunoideae x=8, the Spiroideae x=9 and the Maloideae (including *Malus* and *Pyrus*) x=17. It is hypothesized that the latter have originated through an ancient hybridization event between the Prunoideae and the Spiroideae. At present the binominal *Malus* x *domestica* has been generally accepted as the appropriate scientific name for the cultivated apple (Gardiner, *et al.*, 2007, Korban and Skirvin, 1994).

The homoeologous regions in the apple genome, identifiable through the use of RFLPs reported to map to two map positions (Maliepaard, *et al.*, 1998), thus originated from the two different ancestors. The diversity of these two genomes originally responsible for the formation of apple, can be illustrated by the fact that the majority of SSR markers used by Liebhard *et al.* (2003b) and Silverberg-Dilworth *et al.* (2006), were found to map to a single position. The fact that some linkage groups are not entirely homoeologous to any one other linkage group, but rather to several segments of different linkage groups (Gardiner, *et al.*, 2007) is a further indication of the diversity of the two genomes.

Most *Malus* species, including *M.* x *domestica*, are diploids (2n=34), as are all of the cultivars used in this study, but higher ploidy levels have been reported for wild apples (e.g. *M. hupehensis* and *M. sargenti*).

1.1.3 Dormancy

Woody perennial plants grown in temperate zones undergo annual periods of dormancy. This enables the plant to survive under sub-optimal winter conditions and can be considered a form of strategy to survive. Dormancy is a complex phenomenon, considered as the regulation imposed on the progressing growth processes, such as vegetative growth and flower formation (Okubo, 1996) at various stages.

The onset of dormancy is marked by the cessation of stem elongation and the formation of terminal buds, also known as autumn bud set. Okubo (1996) defined the induction of dormancy as the change of the primordial that cease growing for a while or that initiate special organs instead of producing shoots. For example, after bud initiation (=induction of dormancy) buds will start swelling (accumulation of reserves) which is the process of the buds deepening in their dormancy (Okubo, 1996). Buds need to set soon enough to prevent freezing and dehydration stress during cold winter months, but Cook and Jacobs (2000) found autumn conditions in warmer areas, such as in the Western Cape region, are not cold enough for normal entry into dormancy. Bud development will only proceed normally during spring if, and when, they have been exposed to cold temperatures during the dormant period, a process also known as vernalization. The amount of cold exposure needed varies between species and between cultivars. Normal bud development also depends on temperatures favourable for growth following the required amount of chilling. Bud flush, or the swelling and emerging of new leaves, should not occur too early in spring as this might cause the growing tissue to be killed by late frost. If bud flush however occurs too late in spring, the growing season

will be shortened, resulting in reduced competitive ability and growth potential (Frewen, *et al.*, 2000). Timing of bud set and bud flush are an indication of whether trees are adapted to their environment.

The genetic mechanisms involved in reproductive bud development, or flowering, and the transition from vegetative to reproductive development have been studied in the model plant *Arabidopsis thaliana*, where the expression of vernalization-specific regulator genes were found to initiate flowering (Iqbal, *et al.*, 2007, Michaels and Amasino, 1999, Reeves and Coupland, 2000, Sheldon, *et al.*, 2000, Sung and Amisino, 2004). The expression of regulator genes is thought to be caused by changes in DNA methylation brought about by prolonged growing under cold conditions (Burn, *et al.*, 1993).

Vegetative bud development is far less understood than reproductive bud development. Studies relating to growth manipulations such as pruning, manipulation of environmental conditions (Cook, *et al.*, 2005, Cook and Bellstedt, 2001, Cook and Jacobs, 2000: Cook, 2001 #196), the application of hormonal growth stimulators (Cook, *et al.*, 2001) and changes in protein levels (Arora, *et al.*, 2003) have aided in our understanding and the ability to manipulate, to a certain extent, budbreak in apple trees. The biochemical pathways underlying dormancy release have been studied by Yakovlev *et al.* (2006) using transcriptomics to show that late-, compared to early-, flushing Norway spruce can be associated with active transcription of genes that actively delay the time of flushing. This suggests that chromatin remodeling through acetylation and deacetylation happens in preparation for bud burst, similar to that of flowering, and this is brought about by artificial (chemical)

and/or natural (chilling) stimuli. The development and implementation of statistical tools for the analysis of complex quantitative traits (Falconer and Mackay, 1996) showed that bud phenology is under strong genetic control (Labuschagné, *et al.*, 2002b, Yakovlev, *et al.*, 2006) (section 1.1.5.1) and this enabled further investigations into the underlying genes through a process known as quantitative trait loci mapping (section 1.2.4).

1.1.3.1 Dormancy and dormancy-related considerations in apple trees

Deciduous fruit trees, such as apples, peaches, nectarines, plums, and apricots drop their leaves during autumn and are bare and dormant over winter until growth resumes in spring. Fruit trees need a minimum amount of cold or chilling hours during the winter in order to break dormancy, grow adequately, and produce flowers and fruits in spring and summer. Apples have been distributed into almost all parts of the world and their genetic variability has allowed adapted types to be selected for different environments. Selection continues for new types to expand apple production in both colder and warmer environments and also to keep production areas viable in the face of global warming, a phenomenon already influencing production practices for a lot of crop species in South Africa.

The Western Cape region of South Africa is very suitable for apple production in terms of the dry summers, which reduce the severity of fungal disease attack, but the moderate winter and low chilling conditions experienced in this region may lead to growth abnormalities referred to as prolonged dormancy symptoms (PDS). Apple trees require between 200 and 1100 hours of chilling (below 7 °C) (Linsley-Noakes, *et al.*, 1994) in order for

budbreak to occur promptly and uniformly during spring. Apple cultivars grown by local breeders and farmers vary in terms of the amount of chilling required for dormancy release after winter. 'Anna' is one of the cultivars that occurs locally that has a low chilling requirement (CR), leading to bud flush in early spring, but it produces poor quality fruit. The well known 'Golden Delicious' on the other hand, has very high fruit quality but has a high CR, causing bud flush to occur late in spring. If the CR is not met during the rest period, dormancy release will not be successful and this can be characterized by absent, reduced or irregular lateral vegetative and reproductive budbreak. This will lead to an abnormal and undesirable tree structure, known as acrotony, referring to a phenomenon where only distal or apical buds burst, as opposed to basitony where more proximal buds burst to form shoots resulting in a better tree structure. Reduced or delayed reproductive budbreak will lead to poor fruit set.

In practice, the CR of locally produced apple cultivars are regularly not met in most orchards in South Africa and this is becoming an increasing phenomenon due to the effects of global warming. Orchards are treated with chemical sprays, not only for disease and pest management, but also in order to 'break' dormancy and stimulate vegetative growth after winter. Due to environmental concerns and health issues, the use of these chemicals are not desirable and an alternative is to select and breed apple cultivars that are better adapted to local conditions, with a lower CR, without compromise in fruit quality, disease resistance and yield. Europe is the main destination of all fruits exported from South Africa and due to European legislation prohibiting the use of certain chemicals and restricting chemical residues of others,

certain chemicals, including the dormancy breaking chemical DiNitro-Ortho-Cresol mineral oil (DNOC), have been phased out. The stimulation of vegetative growth by applying plant growth regulators (PGRs) is still common practice in apple orchards in the Western Cape region of South Africa, by application of other mineral oil compounds of which the chemical residue limits are in accordance with European legislation.

1.1.4 Global warming

Global warming refers to the increase in the average temperature of the earth and the surrounding atmosphere. This phenomenon, caused by the increase in greenhouse gasses through human activity such as the burning of fossil fuels, land clearing and agriculture, has caused the near surface temperature to rise with approximately 0.6 °C during the 20th century (http://en.wikipedia.org/wiki/Image:InstrumentalTemperatureRecord.png). Models based on the regression observed over the past decade predict that global temperatures are likely to increase by between 1.1 to 6.4 °C between 1990 and 2100 (http://en.wikipedia.org/wiki/Globalwarming) and it is predicted that temperatures in the Western Cape region of South Africa will increase by a minimum of 1 °C by the late 2030's (www.csag.uct.ac.za) (Wand, 2007). Further predictions include an increase in unpredictable weather patterns during seasonal transitions, reduced rainfall, increased wind speeds and shorter winters accompanied with a reduction in chilling units (Wand, 2007).

Despite scientific uncertainties regarding impact and the political and public debate regarding causes and solutions, it is clear that this is a phenomenon that needs to be kept in consideration when breeding new cultivars for especially long term crops like apple.

1.1.5 Quantitative traits

Many of the phenotypic variations within populations are quantitative in nature and are known as quantitative traits. These traits are thus not inherited in a simple Mendelian fashion, which means that the variation between individuals, or seedlings within the same population, does not fall into discrete classes, but rather a wide continuous range of expression of the specific character is observed in the seedlings. This continuous distribution is caused by the simultaneous segregation of two or more loci affecting the trait of interest. Fruit size and shape are just two examples of the many characters that, together with dormancy, are under polygenic (multiple gene) control. The range of variation is related to the expression of the character in the parents, with the progeny mean always related to the parental mean. Any deviation from the parental mean will be an indication of a dominant gene, or genes, being present (Falconer and Mackay, 1996). The tendency of the variation found in offspring to resemble that of parental values, for specific traits, is evident of some genetic component to this variation.

1.1.5.1 Classical versus neo-classical quantitative genetics

In the past classical quantitative genetics were used to study these complex (polygenic) characteristics (Falconer and Mackay, 1996). Statistical techniques were used in order to analyze properties like additive and non-additive gene action, gene x gene (epistasis) and genotype x environment interaction, heritability, dominance and heterosis. This is done by fitting

experimental results with predicted genetic models of population means, variances and covariances (Falconer and Mackay, 1996).

Despite considerable progress in advancing our knowledge of genetics, an attempt to move away from such statistical uncertainty towards an actual understanding of the underlying genes, caused quantitative genetics to undergo a transformation, referred to as neo-classical quantitative genetics. This new approach facilitates the dissection of polygenic characters into discrete genetic loci, defining the roles of individual genes (Young, 1996). This is done by the joint analysis of segregation of marker genotypes and phenotypic values of individuals. An effective approach for studying the role of individual genes in complex and polygenic characters is known as 'Quantitative Trait Locus' (QTL) mapping (Young, 1996). This approach is based on the use of pedigrees or mapping populations, that segregate for phenotypic traits of basic and/or commercial interest, and identification of genes by linkage with molecular markers. The process can be divided into four stages: (1) phenotypic trait assessment, (2) molecular marker implementation, (3) linkage map construction and (4) QTL mapping.

Environmental influences are superimposed upon (and may interact with) the genetic contributions to variation, however, in plant populations environmental influences can largely be eliminated by the use of clonal replicates of individuals. This was done by Bradshaw and Stettler (1995) showing that the timing of spring bud flush in *Populus* is among the most highly heritable traits measured in forest trees, with up to 98 % of the total phenotypic variance explained by genetic factors and only 2 % by environmental variables. Bradshaw and Stettler (1995) used the above

mentioned neo-classical quantitative genetic principles and found that as few as five QTLs explain 85 % of the genetic variation in the time of spring bud flush in *Populus*. A similar study conducted by Chen *et al.* (2000) revealed three and six QTLs affecting bud set and bud flush in *Populus* with mean heritabilities estimated at 91 % and 94 % respectively. This confirms the general assumption by neo-classical quantitative genetics that a smaller number of QTLs explain the genetic variation observed, than previously assumed by classical quantitative genetics.

1.1.5.2 Classical quantitative genetic studies related to dormancy in apple trees

During the early 1990's Hauagge and Cummins (1991) made the first attempt to understand the control of endodormancy in apple. They estimated broad sense heritabilities for the duration of bud dormancy in apple, indicating that chilling requirement has a strong genetic component. During a study using characters such as the date/time of initial vegetative- (IVB) and reproductive budbreak (IRB), the total number of vegetative and reproductive budbreak (NB) and flowering duration as criteria quantifying seedling reaction to sub-optimal winter chilling conditions, as experienced in the Western Cape region of South-Africa, Labuschagné *et al.* (2002b) studied the variance structure within and between eight seedling families, with the primary interest of estimating the underlying causal components of variance (environmental and/or genetic). His experimental design allowed for parents and offspring to be grafted onto the same rootstock, giving them a similar physiological status. The total variance of the measurements can then be divided into three components: a between families (genetic), a between seedlings within

families (genetic) and a within clones (environmental) component. Broad sense heritabilities calculated for IRB and IVB (0.75 and 0.69 respectively) far exceeds that calculated for NB (0.30) and suggests that the variation found between seedlings for time of budbreak can be attributed primarily to genetic factors (Labuschagné, *et al.*, 2002a). This makes time of budbreak an ideal criteria to use for genetic improvement through selection. Although Labuschagné *et al.* (2002a, b, 2003) found a significant positive association between time of budbreak and number of budbreak, selection for early budbreak will not necessarily identify seedlings with an increased NB. This might be explained by the fact that selection for very early IVB time will increase the risk for buds being damaged by late frost.

1.2 MOLECULAR MARKER-ASSISTED BREEDING

1.2.1 Molecular Markers

The development of new molecular marker systems has been one of the most dynamic areas in applied molecular genetics. A molecular genetic marker can be considered a fixed point (locus) on a chromosome. The availability of many molecular markers makes it possible to construct genetic linkage maps which can be used for the detection of QTLs for the subsequent application of marker-assisted selection (MAS) for molecular breeding.

Different molecular marker techniques can be used to generate molecular marker data. Each of these techniques has its own advantages and disadvantages when it comes to the practice of genetic linkage analysis performed in order to construct genetic linkage maps. These techniques include Restriction Fragment Length Polymorphisms (RFLPs), Randomly

Amplified Polymorphic DNA (RAPDs), Sequence Characterized Amplified Regions (SCARs), Amplified Fragment Length Polymorphisms (AFLPs), Simple Sequence Repeats (SSRs) and Single Nucleotide Polymorphisms (SNPs). With the exception of RFLP's, all these techniques make use of the Polymerase Chain Reaction (PCR) technique to amplify specific parts of the DNA molecule. SSRs, also known as microsatellites, short tandem repeats (STRs) or simple sequence length polymorphisms (SSLPs), are becoming one of the most important molecular markers in both animals and plants. SSRs provide hyper-variable co-dominant markers, which means that all alleles of a marker can be identified in heterozygotes, making SSRs very favourable for genetic linkage analysis.

1.2.1.1 SSRs or microsatellite markers

SSRs are stretches of 1 to 6 nucleotide units that are repeated in tandem and are randomly distributed trough the genome of eukaryotes (Fujimori, *et al.*, 2003). Differences in the number of repeats are thought to be caused by errors in DNA replication – the DNA polymerase 'slips' when copying the repeat region, changing the number of repeats. The mutation rate of SSRs is estimated at between 5×10^{-4} and 10^{-5} which, according to Hearne *et al.* (1992), are sufficiently stable to use in genetic analysis, rendering them ideal markers for construction of high-resolution genetic maps.

Since microsatellites themselves are usually less than 100 bp long and are embedded in DNA with a unique sequence, they can be amplified in vitro using the polymerase chain reaction (PCR) (Hearne, *et al.*, 1992). Length polymorphisms are easily detected by the use of specially designed primer pairs flanking the repeat. After PCR amplification fragments may be separated

on high resolution polyacrylamide gels, followed by silver staining. Since the resolution of the polyacrylamide gels limits the detection of polymorphisms, especially for larger fragments, and the fact that the use of these high resolution gels are quite labour intensive, primers can be fluorescently labeled in order to make it possible to detect fragment lengths using fluorescence-based DNA detection systems, such as an ABI genetic analyzer (Applied Biosystems). These detection systems use capillary electrophoresis (CE), enabling the separation of amplification products in an automated fashion (Butler, *et al.*, 2001).

Since most SSRs are very polymorphic due to the large variation in the number of repeats, unique genotypic profiles can be generated that permits individual identification (Goldstein, *et al.*, 1999). SSRs have been used extensively in diversity studies and the ability of the method to differentiate individuals when a combination of loci is examined makes the technique very useful for gene-flow experiments and paternity analysis (Coart, *et al.*, 2003, Hokanson, *et al.*, 1998., Oraguzie, *et al.*, 2005). Besides the number of alleles detected, one of the most important characteristics of a locus is its heterozygosity, the probability that an individual has two different alleles at a given locus. From the perspective of genetic linkage map construction, a locus with heterozygosity higher than 70 % is commonly considered a highly informative marker (Liu, *et al.*, 1996).

The only disadvantage, regarding the use of SSRs, that remains is the large initial effort required to identify, clone and sequence microsatellite flanking sequences. This can be reduced by (1) the use of primers designed from sequences derived from closely related species as well as (2) the use of

sequences that are publicly available (e.g. Genbank) for SSR identification and primer design. Cross-species-surveys with microsatellite markers has not only been done for apple and pear (Yamamoto, *et al.*, 2004, Yamamoto, *et al.*, 2001), but has also been done Kuleung *et al.* (2004.) between wheat, rye and triticale. These surveys have shown that it is possible to amplify SSRs from species other than those used during primer design. The extent of the cross-species amplification appears to be correlated with taxonomic distance and the knowledge that some loci amplify across species has stimulated some phylogenetic studies. The fact that *Malus* ssp. (apple) and *Pyrus* spp, (pear) are both studied in our laboratory and that they are taxonomically closely related (Campbell, *et al.*, 1995) may allow the use of SSRs across species, as was done by Pierantoni *et al* (2004), Silverberg-Dilworth *et al.* (2005) and Celton *et al.* (2007).

The number of apple sequences available in Genbank increased dramatically after expressed sequence tags (ESTs) were made publicly available (Korban, *et al.*, 2005, Newcomb, *et al.*, 2006). A survey revealed that these ESTs contain a substantial number of SSRs that can be exploited for the construction of a reference map consisting of transferable, codominant markers. SSRs derived from ESTs (EST-SSRs) were reported to have lower levels of polymorphism than genomic SSRs in domesticated cereal crops (Nicot, *et al.*, 2004), as would be expected due to the conservation of transcribed regions, this is however not the case for tree species such as pines where no significant difference in polymorphic information content were observed (Chagne, *et al.*, 2004). Several reports also indicate that SSRs derived from genes (EST-SSRs or genic-SSRs) have higher quality

amplification and more distinctive peaks than genomic SSRs (Nicot, *et al.*, 2004, Varshney, *et al.*, 2005.). EST-SSRs also has the advantage of tight linkage to known genes (Da Silva, 2001), reflecting the distribution of genes along the genetic map (Varshney, *et al.*, 2005.). Should the repeat motif reside in the gene of interest, it will be more useful in QTL mapping and MAS. EST-SSRs have been shown to be more transferable between different pine species compared to genomic SSRs (Chagne, *et al.*, 2004, Gupta, *et al.*, 2003) and should thus also be more transferable between distantly related species, due to conservation of genes. This phenomenon renders genic-SSRs more useful in the construction of consensus maps and for comparative mapping studies. A study conducted by Yamamoto *et al.* (2004) made use of SSRs, derived from genomic sequences, in order to study the synteny between apple and pear. Although they were able to successfully align the apple and pear maps (only 2 out of 36 SSRs tested were assigned to different linkage groups) the same results may not have been obtained when applying this approach to more distantly related species. Apple SSR markers that have been published up to date are predominantly derived from genomic sequences (genomic-SSRs) (Guilford, *et al.*, 1997, Liebhard, *et al.*, 2002, Liebhard, *et al.*, 2003b, Silfverberg-Dilworth, *et al.*, 2006).

1.2.2 Linkage Analyses and map construction

With the advent of DNA markers, we are in the position of being able to analyze a large number of recognizable loci segregating simultaneously in the same mapping population. When looking at a large number of these loci, the order in which they occur on the genomic material can be predicted on the basis of 'co-segregation'.

In most agricultural crops, inheritance patterns are studied in the F_2 progeny derived from the selfing of F_1 individuals that were obtained from a cross between two homozygous parents. These homozygous parents are either self-fertilizing or inbreeding can be carried out without severe inbreeding depression (Maliepaard, *et al.*, 1997). For any given locus a maximum of two alleles can be present in the F_2 progeny. Both these alleles are inherited from a single, selfed parent. Using molecular markers when doing linkage analysis and map construction is far more complex in a full-sib family of an outbreeding plant species (Maliepaard, *et al.*, 1997). The reason for this being the number of segregating alleles per locus (up to four), segregation of alleles from two parental genotypes and the fact that the linkage phase, or grandparental origin, of the alleles at any given locus is usually unknown. Maliepaard *et al.* (1997) recommended multi-allelic markers, such as SSRs and/or RFLPs, when the intent is to integrate the homologous linkage groups of the respective parents, since recombination can be estimated separately in male and female parents. Another advantage of these markers is the high probability that they can be used over a wide range of crosses (Maliepaard, *et al.*, 1997), and even be transferred to closely related species, enabling comparative mapping and synteny analysis.

A variety of computer software programs have been developed to handle large numbers of segregating/polymorphic loci, in order to establish the best overall genetic map. Some of the most widely used programs include Mapmaker (Lander, *et al.*, 1987), GMendel (Echt, *et al.*, 1992) and JoinMap (Van Ooijen, 2006). These software programs use the genetic principle that the position of, and distance between, loci on the genome is related to the

frequency of recombination between the alleles of different loci. Linkage maps allow studies of the genome structure and the localization of genes of interest, and permit the identification of quantitative trait loci (QTLs) as a first step towards understanding the biological basis of complex traits.

1.2.2.1 Steps involved in genomic map construction

The processes involved in genomic map construction can be divided into five steps. The first of these is to create a segregating mapping population. This is followed by single-locus analysis (determining the genotypes of all loci in all individuals), two-locus analysis (estimating the recombination frequencies between all pairs of loci), the establishment of linkage groups and marker ordering.

1.2.2.1.1 *Creating a segregating population*

Mapping populations commonly used in genomic research are generated by controlled crosses between parents, on the basis of variation at the phenotypic level for a given trait or a trait of interest. When the parents used are heterozygous and outbreeding, as is the case with apple, genetic variation exists between the parents, which is essential in order to trace recombination events.

1.2.2.1.2 *Single-locus analysis*

DNA analysis of individuals in any given mapping population involves the implementation of molecular markers in order to determine the genotypic code for each individual, at each locus. Seedlings derived from a cross between two diploid, heterozygous parents may have up to 4 alleles

segregating at any given locus and these alleles are expected to be passed onto the progeny in equal ratios (Table 1).

Table 1. Allelic frequencies expected in the seedlings derived from a cross between two diploid parents

Cross	Number of alleles	Frequency			
		A1	A2	A3	A4
ab x cd	4	0,25	0,25	0,25	0,25
ab x cc	3	0,25	0,25	0,5	0
ab x ab	2	0,5	0,5	0	0
aa x aa	1	1	0	0	0

Any deviation from the expected segregation pattern might be an indication of low quality data, non-random sampling or insufficient sampling. The possibility of a specific allele, or the trait that it is linked to, rendering the individual lethal or less fit to survive also exists. It is thus clear that segregation distortion does occur in nature and although markers with distorted ratios have been eliminated from linkage analysis by various groups in the past, they are now generally included in analysis protocols, as was done in this study.

1.2.2.1.3 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, phenomena always present in molecular data sets, are essential for genetic linkage map construction. Liebhard *et al.* (2003b) 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 will result in recombination frequencies being estimated from smaller data sets, being less accurate, this can be considered to be more acceptable than genotypic errors. It is thus 'safer' to opt for 'missing data' in cases where there is any uncertainty as to the exact genotype.

1.2.2.1.3.1 *Linkage and Recombination*

Construction of a genetic linkage map requires knowledge of how often loci are inherited together or become separated by genetic recombination. Genetic recombination can be divided into inter-chromosomal recombination, which takes place between non-homologous chromosomes through a process known as independent re-assortment, and intra-chromosomal recombination, which takes place between homologous chromosomes through a process of crossing over.

The recombination frequency between 2 loci depends on the distance between them. The closer loci lie on the same chromosome, the more likely they will inherit together, and the recombination frequency, when studying a number of seedlings, will move closer to zero. A recombination frequency of zero does however not imply that the two loci are geographically next to each other on the actual genome, they might still be some distance apart, it merely is an indication that within the mapping population used, there was no recombination and it can be assumed that they lie close to each other on the

genome. The further apart loci lie on a chromosome the closer the recombination frequency will resemble that of loci lying on different chromosomes, with a recombination frequency (r) of 0.5. A marker pair is considered to be linked when the marker frequencies obtained in the progeny are significantly different from the expected frequencies in the absence of linkage ($r = 0.5$) (Maliepaard, *et al.*, 1997).

There is no known relation between genome size and the frequency of crossing over. Crossing over occurs roughly once or twice per pair of homologous chromosomes commonly giving a linkage group length of approximately 100 cM, due to structural and/or other interference mechanisms. Evidence of genetic control of crossing over, such as genes controlling the frequency of recombination, has been reported, confirming that there is no correlation between physical distance (bp) and map distance (cM) within any given genome and the ratio varies greatly according to genome location.

1.2.2.1.3.2 Linkage phase

Linkage phase is the term used when referring to chromatid associations of alleles of linked loci. When analyzing the F1 progeny of a cross between two heterozygous individuals, the linkage phase, or the grandparental origin, of the alleles are usually unknown. By constructing a genetic linkage maps where the segregation of alleles from both parents are viewed simultaneously or independently, the linkage phase of alleles from the individual parents can be estimated (Maliepaard, *et al.*, 1997) and it is usually done in such a way to reduce possible recombination events.

1.2.2.1.4 Establishing linkage groups

Linkage groups are calculated based on the fact that alleles of markers on different chromosomes segregate independently, while alleles of markers on the same chromosome pair segregate more often in the same combinations. A linkage group is thus a group of loci that have a recombination frequency of less than 50 %, as would be expected for independent, or unlinked loci. If no false linkage is assumed, then the number of genetic linkage groups obtained based on statistics should be the same as the haploid number of chromosomes. The number of linkage groups can however exceed the number of chromosomes if loci on a large segment of a particular chromosome are not observed, causing a 'chromosome break' in the linkage group.

JoinMap®2 and JoinMap®3, that were used for the construction of the most recently published apple genetic linkage maps, uses the 'Log of the odds' (LOD) score test as statistical criterion to test linkage, identified as the base 10 logarithm of the likelihood ratio. Often a LOD value of 3 is used as the significance threshold, meaning linkage is 1000 times more likely than independent segregation. As a chi-square test, this value corresponds to a significance of 0,0002. This high level of stringency is needed because many pairs of markers are usually tested (Maliepaard, *et al.*, 1997). Liebhard *et al.* (2003b) and Silverberg-Dilworth *et al.* (2006) used more stringent conditions, using a LOD score of 4 and 5 respectively, to group markers belonging to the same linkage group, compared to a LOD score of 3 as previously used by Hemmat *et al.* (1994) and Seglias and Gessler (1997). In a data set

containing 44 individuals (as used by Liebhard *et al.* (2003b)) a LOD score of 4.0 represents a recombination frequency of 0.19.

1.2.2.1.5 Gene ordering.

The order of markers on the genetic linkage map can be determined by observing the recombination frequencies between markers. Missing data and errors in the observations will disturb the ordering process, and should be avoided. Double recombination events can also disturb the ordering process, since they often lead to contradictory data.

The number of possible orders for any given linkage group increases as the number of markers increases and the JoinMap algorithm (Stam, 1993) is based on the sequential addition of markers in a systematic way. Jansen *et al.* (2001) described a combination of techniques in order to establish the quality of the data and the maps produced during the construction of dense genetic linkage maps. First the best 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 map position and in order to identify problematic markers. Although Liebhard *et al.* (2003b) used a LOD threshold as high as 4 for the initial grouping of markers into linkage groups, very 'loose' thresholds (LOD = 0.001: REC = 0.499) were used for the ordering of markers within each linkage group. Using such low thresholds implies that all the markers on a given linkage group are influenced by one another, subsequently resulting in larger distances between markers.

1.2.3 Genetic linkage maps for apple

Saturated and high-density genetic linkage maps are very useful in fundamental and applied genetic research. Genetic linkage maps have been made available for *Malus* by Hemmat *et al.* (1994), Conner *et al.* (1997), Maliepaard *et al.* (1998), Liebhard *et al.* (2003b) and Silverberg-Dilworth *et al.* (2006), with the latter having the lowest average marker distance and the best genome coverage. These maps were initially constructed by using mainly RFLPs, isozymes and RAPDs and more recently with the addition of microsatellite markers (SSRs) (Liebhard, *et al.*, 2003b, Silfverberg-Dilworth, *et al.*, 2006). The most complete maps are considered to be the ones organized into 17 linkage groups (LGs) (Liebhard, *et al.*, 2003b, Maliepaard, *et al.*, 1998, Silfverberg-Dilworth, *et al.*, 2006), corresponding to the 17 chromosomes of apple. Liebhard *et al.* (2003b) used a total of 840 molecular markers (AFLP, RAPD, SCAR and SSR), including 115 SSRs, in the construction of genetic linkage maps derived from a controlled cross between 'Fiesta' and 'Discovery'. This linkage map was further saturated by Silverberg-Dilworth *et al.* (2006) through the addition of an additional 148 SSR markers. This brings the total number of polymorphic SSRs to 263, of which 205 and 215 segregated and mapped for 'Fiesta' and 'Discovery' respectively with 145 SSRs present in both parental maps. The final maps covered 1145.3 cM in 'Fiesta' and 1417.1 cM in 'Discovery'. The fact that the map length did not change substantially with the resent addition of new markers, indicate that their present map is very close to full genome coverage. The addition of more markers will thus enable the reduction of average marker distances leading to high-density maps. The use of expressions like 'saturated' and 'high-density',

when referring to genetic linkage maps, are used somewhat inconsistent and should be interpreted as referring to good coverage of the linkage groups rather than gaps. It also needs to be kept in mind that linkage maps tend to be denser closer to the centromere than to the telomeric ends.

The transferability of linkage maps to other crosses relies completely on co-dominant markers such as SSRs (Liebhard, *et al.*, 2003b). Because of the high number of alleles, SSRs can be mapped in more populations, becoming the framework of all linkage maps. Due to the relatively high information content of these markers and the ease with which they can be implemented and analyzed, the time required for the construction of new linkage maps will be considerably reduced. Many groups worldwide are implementing these highly informative transferable markers in order to construct genetic linkage maps to be used in QTL identification studies involving a wide range of phenotypically important characteristics. Silverberg-Dilworth *et al.* (2006) proposed the use of 100 SSR markers, with an average distance of 15 cM between markers (Figure 1), to be used as a framework when new apple mapping populations are used Unfortunately there are regions of the genome for which no SSR markers have been mapped up to date and thus 86 SSR markers were selected for the 15 cM framework map, covering around 85 % of the apple genome (Silfverberg-Dilworth, *et al.*, 2006). All reference marker sequences have been published (Liebhard, *et al.*, 2002, Silfverberg-Dilworth, *et al.*, 2006) with the exception of two, AG11 on LG 1 and HB03AT on LG 6, that have been included in the 15 cM reference map and these could for this reason not be included in this study.

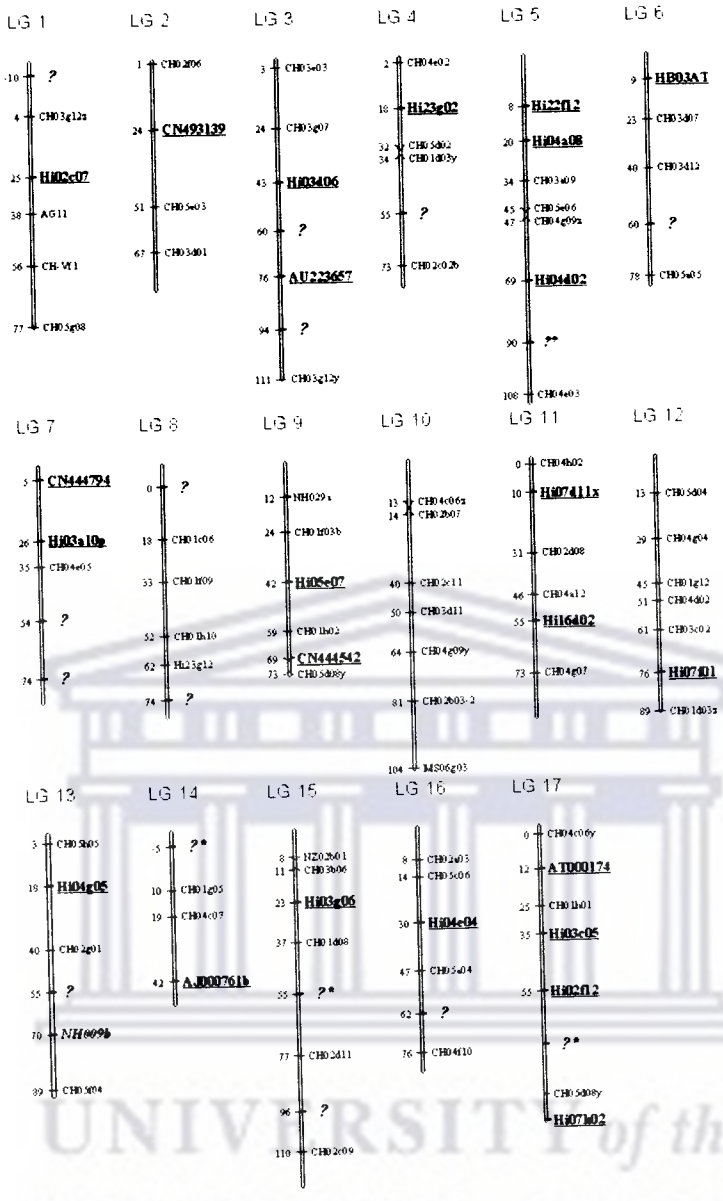


Figure 1. Set of 102 SSR primer pairs for global coverage of the apple genome. Map positions (in cM) are alligned to the 'Discovery' maps. Grey filled segments indicate linkage group segments covered by the 'Fiesta' x 'Discovery' maps. Open bar segments indicate linkage group segments revealed by other, unpublished maps. For 16 loci, indicated with the symbol '?', no primer pairs are publically available yet. The symbol '?*' marks positions of unpublished SSR markers which are expected to become available in the near future.

1.2.4 Quantitative Trait Loci (QTLs) mapping

The theory of using a marker to select for an associated trait was first described by Sax in 1923, when he noted that seed size in bean (a complex trait) could be associated with seed coat colour (a simple, monogenic trait) (Young, 1996). This association was based upon phenotype and the gene affecting seed coat colour was probably closely associated with a major gene affecting seed size. Most complex quantitative traits are controlled by numerous genes (see section 1.1.5) and today modern QTL mapping make use of defined sequences of DNA, each acting as a linked monogenic marker for one gene involved in the determination of the phenotypic trait.

Once a framework map consisting of evenly spaced markers, as proposed by Silfverberg-Dilworth *et al.* (2006), is available, QTL mapping involves the testing of all the DNA markers for the likelihood that they are associated with the quantitative trait distribution. Individuals in a suitable mapping population thus have to be analyzed for the phenotype of interest and in terms of DNA marker genotypes (Young, 1996). Statistical analysis is performed using computer software and if a significant relationship is found between the DNA marker and the phenotype of interest the DNA marker is probably linked to a QTL.

This process has proven to be much more complex than outlined in the paragraph above and although it can be said that successful QTL mapping relies heavily on a good phenotypic scoring method and a genetic linkage map covering the entire genome, there are a lot of other factors that need to be considered. First, and most importantly, it needs to be kept in mind that there are limitations to the amount of recombination that occurs during

meiosis. The recombination frequency per homologous chromosome pair is estimated at 1.5 recombination events per meiosis. This means that the mapping population needs to be sufficiently large in order to study a large number of meiosis, resulting in enough crossovers to map the QTLs with accuracy. A second problem is that it would be difficult to distinguish two or more QTLs that are situated in close proximity on the same chromosome. Two or more such genes that are linked in coupling may be mistaken for one gene, or if linked in repulsion, no QTL would be identified. A third problem is one of statistics and translates into how critical values are estimated for the acceptance or elimination of putative QTLs. During QTL mapping one would strive to limit the occurrence of Type I (false positives) and Type II (missed QTLs) errors (Jansen, 1994).

World wide, different apple research groups, such as groups forming part of the European projects 'Durable Apple Resistance in Europe' (DARE) (Lespinasse and Durel, 1999) and 'High Quality Disease Resistant Apples for a Sustainable Agriculture' (HIDRAS) (Gianfranceschi and Soglio, 2004), HortResearch in New Zealand and various research groups from the United States of America, are using different mapping populations in order to identify specific genes and QTLs responsible for a variety of economically important traits. Summaries regarding genetic linkage maps available for apple (Arus, *et al.*, 2006, Gardiner, *et al.*, 2007), as well as the various genes and QTLs that have been identified by specific research groups (Gardiner, *et al.*, 2007), have been published. Although a number of candidate genes (Table 2) and QTLs (Table 3) have been identified, the larger number of QTLs suggest that, for most of the traits for which candidate genes have been mapped, there are

more genes playing a role in the determination of the expression of the trait in the seedlings. The use of low LOD-thresholds during QTL analyses may however result in the identification of various QTLs, increasing the change of including artifacts. QTLs have also been identified for a number of economically important traits for which no known genes have been mapped yet. The identification QTLs, whether novel or in addition to known and mapped candidate genes, are the first step towards unraveling complex traits into all the contributing genetic factors. The next step will be the identification of markers that can be linked to these QTLs and that can be used in MAS in breeding programs, where the ultimate goal is the pyramiding of favourable genes. The detection of QTLs in the same region in different genetic backgrounds and/or over more than one year of phenotypic trait analysis will increase the possible power of the QTL when it is implemented in MAS.

Table 2. Summary of a variety of phenotypic traits for which candidate genes have been mapped. The linkage groups containing these candidate genes and research articles reporting on their map position.

Trait	Linkage group															References
	1	2	4	5	6	7	8	9	10	11	12	13	15	16	17	
Major Scab resistance	Vf Va	Vr2 Vh2 Vh4 Vh8 Vbj Vt57			Vdr1		Vfh		Vd		Vg Vb				Vm	Manganaris *et al.*, 1994 Vinatzer *et al.*, 2001 Xu and Korban, 2000 Erdin *et al.*, 2006 Hemmat *et al.*, 2003 Patocchi *et al.*, 2004, Patocchi *et al.*, 2005 Bus *et al.*, 2004, 2005a,b Gygax *et al.*, 2004 Tartarini *et al.*, 2004 Calenge *et al.*, 2004, Freslon *et al.*, 2006, Durel *et al.*, 2003, 2004, 2006.
Powdery mildew resistance							Pl-w			Pl-2	Pl-d Pl-1					Maliepaard *et al.*, 1998 James *et al.*, 2004 Lessemann and Dunemann 2006 Liebhard *et al.*, 2002

Woolly apple aphid resistance							Er-1 Er-3										Gardiner *et al.*, 2007
Rosy leaf curly aphid (*Dysaphis divecta*) resistance						Sd1 Sd2											Maliepaard *et al.*, 1998 Cevik and King 2002
Rosy apple aphid (*Dysaphis plantaginea*) resistance							Dp-fl										Durel *et al.*, 2006
Malic acid – fruit acidity														Ma			Maliepaard *et al.*, 1998
Fruit skin colour								Rf									Maliepaard *et al.*, 1998
Self incompa-tibility																SI	Maliepaard *et al.*, 1998
Non-specific lipid transfer protein		Mal d 4	Mal d 3		Mal d 1		Mal d 4	Mal d 2 Mal d 4			Mal d 3	Mal d 1		Mal d 1			Gao *et al.*, 2005
Ethylene production									Md-ACO 1				Md-ACS 1				Costa *et al.*, 2005
Rootsucker formation																Rs	Weeden *et al.*, 1994
Columnar growth									Co								Maliepaard *et al.*, 1998
Dwarfing				Dw1													Celton *et al.*, 2006

Table 3. Summary of a variety of phenotypic traits for which QTLs have been identified. The linkage groups containing such QTLs and the research articles reporting their map position.

Trait	Linkage group																	References
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Scab resistance	X	X			X	X	X			X	X	X	X		X		X	Durel *et al.*, 2003 Liebhard *et al.*, 2003 Calenge *et al.*, 2004
Fire blight resistance			X		X		X					X	X					Calenge *et al.*, 2005 Khan *et al.*, 2006 Peil *et al.*, 2006
Powdery mildew resistance	X	X						X		X		X	X	X	X		X	Stankiewicz-Kosyl *et al.*, 2005 Calenge and Durel, 2006
Fruit harvest date			X															Liebhard *et al.*, 2003
Fruit Flesh Firmness	X					X		X		X	X	X		X				Liebhard *et al.*, 2003 King *et al.*, 2000
Fruit weight	X		X	X		X		X		X		X			X	X		Liebhard *et al.*, 2003
Fruit acidity								X								X		Liebhard *et al.*, 2003
Sugar content			X			X		X	X					X				Liebhard *et al.*, 2003
Number of fruit					X										X	X		Liebhard *et al.*, 2003
Fruit sensory	X	X			X	X				X		X	X			X		King *et al.*, 2000

descriptors																		
Fruit texture descriptors	X		X	X		X	X	X				X	X		X	X		King *et al.*, 2001
Blooming time							X			X							X	Liebhard *et al.*, 2003
Leaf size									X								X	Liebhard *et al.*, 2003
Height increment			X		X			X			X		X				X	Liebhard *et al.*, 2003
Juvenile phase length			X												X			Liebhard *et al.*, 2003
Number of bunches								X							X			Liebhard *et al.*, 2003
Stem diameter	X	X	X					X			X		X	X	X		X	Segura *et al.*, 2007
Tree geometry		X	X				X	X		X						X		Segura *et al.*, 2007
Tree topology	X		X					X	X		X		X			X		Segura *et al.*, 2007
Date of budbreak						X		X										

Several SSRs, especially on LG 2, have been shown to co-localize with mapped genes, enabling MAS. CH05e03 (Liebhard, *et al.*, 2002) has been found to be approximately 0.6 cM from *Vbj* (Gygax, *et al.*, 2004). Ch05e03 and CH02c06 (Liebhard, *et al.*, 2002) were found to be on opposite sides of *Vr* (Hemmat, *et al.*, 2002) and CH02c02a (Liebhard, *et al.*, 2002) were found to be linked to *Vr2* (Patocchi, *et al.*, 2004). On LG 8, CH05a02 (Liebhard, *et al.*, 2002) was mapped 13 cM above the powdery mildew resistance gene *Pl-w* (Evans and James, 2003). On LG 10 CH02b07 (Liebhard, *et al.*, 2002) was mapped 7 cM below the scab resistance gene *Vd* (Tartarini, *et al.*, 2004). CH02d12 (Liebhard, *et al.*, 2002) on LG 11 were found to map in close proximity to a powdery mildew resistance gene, *Pl-2* (Seglias and Gessler, 1997). On LG 12 the apple scab resistance gene *Vb* were mapped (Erdin, *et al.*, 2006) between the SSR markers Hi02d05 and Hi07f01 (Silfverberg-Dilworth, *et al.*, 2006).

Lawson et al. (1995) first reported the detection of two markers, on LG 6, of which the segregation can be associated with differences in timing of vegetative budbreak. It was also reported that the time of vegetative budbreak can be closely associated with the morphological character of terminal bearing. Since the linkage group numbering, more recently defined by Maliepaard *et al.* (1998) and Liebhard *et al.* (2002, 2003b), are used as a reference map, Kenis and Keulemans (2005) found through the use of SSRs, that LG 6 of the genetic linkage map obtained by Lawson *et al.* (1995) corresponds to LG 10 of the reference map. Conner *et al.* (1998) identify 8 putative QTLs influencing 'leaf break', localized on 7 different linkage groups (LG 3, 6, 7, 9, 11, 12 and 15 of un-aligned map) collectively explaining 42 %

of the observed phenotypic variation. Although transferable markers were absent on the genetic linkage map constructed by Conner *et al.* (1998) a study conducted by Maliepaard *et al.* (1998) revealed that LGs 3, 7 and 9 correspond to LGs 9, 8 and 7 respectively, of the now more widely used reference map. Segura *et al,* (2007) identified two QTLs, on LGs 6 and 8 respectively, for the date of budbreak observed in a F1 population, consisting of 121 seedlings, derived from a cross made between two cultivars with the aim of studying tree architecture. The fact that only 2 QTLs were detected compared to the 8 reported by Conner *et al.* (1998) might be ascribed to the fact that the two cultivars used as parentals do not differ significantly in terms of the phenotypic trait under investigation, resulting in a 'narrower' distribution of the trait in the resulting progeny. This is however speculative as the exact phenotypic values are not reported.

Kenis and Keulemans (2004) studied different growth characteristics in a mapping population derived from a cross between 'Telamon' and 'Braeburn'. The QTL instability that they observed over the two year period, render the identified QTLs less useful for prediction of future growth characteristics. The instability might be caused by 1) the fact that growth is a very complex characteristic for which the genetic control changes as the tree matures, 2) influences from environmental factors and 3) differences in the roots of individual plants. The solution to this would be the use of clonal replicates and grafting onto the same rootstock, as was done by Segura *et al.* (2007) studying tree architecture in a 1-year old apple progeny derived from a cross made between 'Starkrimson' and 'Granny Smith'. The identified QTLs do however need to be confirmed in subsequent growing years.

1.2.5 Marker assisted selection (MAS)

Linkage maps and the identification of QTLs affecting economically important traits enable marker-assisted breeding (MAB) and selection (MAS) for complex characteristics. In conventional breeding programs the recurrent selection of individuals with desired phenotypes in each generation is a time consuming, long term and expensive process. MAB is especially promising in perennial tree crops, like apple, where many important traits are expressed only after years of costly field maintenance. Using MAS the presence of favorable genes, or alleles, can be determined at an early stage and the initial population size under investigation can be drastically reduced before seedlings are planted in the field. MAS should preferably be done using multi-allelic reproducible markers, which require minimal amounts of DNA and a minimal amount of DNA isolation and purification steps, so that large numbers of plants can be screened in a short period of time. One of the uses of MAB is the marker-steered introgression of valuable single genes from exotic donors to enhance elite breeding material. MAS allows faster recovery of the recipients of favourable genes than the conventional recurrent backcrossing (Yin, *et al.*, 2003) and 'genetic drag' can be reduced through selection of individuals where a recombination event accured between the favourable allele from one locus and an undesirable allele for a second locus.

Although several markers have been identified mapping close to genes and QTLs of interest (see section 1.2.4), the efficiency of applying these markers for selection purposes, will depend on the distance between the marker and the gene of interest. The closer the marker to the gene of interest, the smaller the recombination frequency between the marker and the actual

gene and the more efficient the selection. It is thus clear that, unless the SSR lies in the actual causal gene itself, the use of MAS can not guarantee the elimination of unwanted alleles, but it will most definitely eliminate a vast majority. A further option is to use SSR markers residing on opposite sides of the gene of interest and only selecting for those individuals with no recombination occurring between the selected markers.

1.3 OBJECTIVES OF THIS STUDY

Increasing demands for the growing of organic food crops have stimulated an international tendency towards research concentrating on the identification of genes responsible for a variety of disease resistance traits, while a lot of research also focus on the identification of genes related to production volumes and fruit quality in a variety of crops. The selection for these favourable genes during any breeding program will not only have health and environmental advantages, but will also increase profitability for farmers.

Besides the application of pesticides for various diseases, apple orchards in the Western Cape region of South Africa require the application of dormancy breaking chemicals in order for budbreak to occur after mild winter conditions. The breeding and selection process for developing a new apple variety, that has a low CR and does not require the application of dormancy breaking reagents, for local production is an arduous process that may take many years using conventional plant breeding.

The power to select desirable individuals in a breeding program based on genotypic configuration is an extremely powerful application of DNA markers and QTL mapping. This means breeders can select on the basis of

genotype rather than phenotype. Which is especially helpful if a target trait is time-consuming to score. By simply eliminating the need for field trials early in a breeding program, MAB may revolutionize the process of cultivar development.

The aim in this study is to construct genetic linkage maps, using published and newly developed SSR markers, and to use these in order to identify regions of the genome that contain putative QTLs involved in the regulation of time of IVB. Phenotypic data recorded during a study conducted by Labuschagné *et al.* (2002a, b), identifying time of IVB as a phenotypic character with a high level of heritability (see section 1.1.5.2), will be used during QTL analysis. Once QTLs have been identified, the possible linkage of SSR markers to the trait of interest can be determined as well as the efficiency of these markers when they are to be used in MAS. This would open the way for MAS for low CR-genotypes at early stages of the breeding program before phenotypic evaluation is possible.

2 CHAPTER 2: TIME OF INITIAL VEGETATIVE BUDBREAK (IVB)

2.1 ABSTRACT

Phenotypic data were analyzed in order to confirm the quantitative nature of time of IVB, a character closely related to prolonged dormancy symptoms experienced in apple orchards in the Western Cape region of South Africa, as described in previous studies by Labuschagné *et al.* (2002a, b). The distribution of the trait were graphically represented keeping three variables in mind: the distribution between full-sibs for each year of investigation, the stability of ranking between individuals from a full-sib family and the stability between clonal replicates. All three these considerations confirmed the quantitative nature of the trait as well as the hypothesis that the trait has a high level of genetic influence.

2.2 INTRODUCTION

Time of IVB is a reflection of the accumulated chilling and heat requirement (Hauagge and Cummins, 1991) needed for dormancy release after winter. In studies conducted by Labuschagné *et al.* (2002a, b) it has been found that time of IVB showed significant levels of variation among eight apple populations. During this study the focus were placed on two of the populations used during these studies by Labuschagné *et al.* (2002a, b), a cross where 'Anna' has been used as a male parent with 'Sharpe's Early' ('Anna' x 'Sharpe's Early') and a cross where 'Golden Delicious' has been used as a

female parent with 'Prima' ('Prima' x 'Golden Delicious'). A third mapping pedigree, derived from a cross between 'Anna' and 'Golden Delicious' ('Anna' x 'Golden Delicious'), for which data regarding time of IVB were collected by Dr. I. F. Labuschagné (unpublished data) were added.

During an initial study conducted by Labuschagné *et al.* (2002b), 60 seedlings from the 'Anna' x 'Sharpe's Early' and 60 seedlings form the 'Prima' x 'Golden Delicious' populations were randomly selected from an orchard on Drostersnes (34°4'15" S 19°04'47" E) in the Western Cape region of South Africa, characterized by low winter chilling. These seedlings, initially in the 5th and 7th growing season respectively, were phenotyped for a 3 year period, from 1996-1998, for time of IVB. A wide range of variation between seedlings were observed for the time of IVB over the three year period (Table 4) (minimum and maximum values). This is evidence that the phenotypic trait is controlled by a number of loci, known as quantitative loci. Table 4 also lists the mean time of IVB over the three year period. The tendency of the observed mean to resemble the expected mean when calculated from the minimum and maximum values listed, is an indication that the trait has not been subjected to selection pressures, as would be expected in a mapping population.

Table 4. Variation observed in the time of IVB during an initial study conducted on adult trees. Time of IVB (averaged over a 3 year period ranging from 1996-1998) is indicated as the day of the year, starting 1st of January, of the first sight of new leaves emerging from the vegetative bud.

Family	Time of IVB		
	Min	Max	Mean
Sharpe's Early X Anna	196	288	243
Golden Delicious X Prima	267	304	283

During a further study Labuschagné *et al.* (2002a) used seven clonal replicates of 100 seedlings from the 'Anna' x 'Sharpe's Early' family and 60 seedlings from the 'Prima' x 'Golden Delicious' family, as well as replicates of the parental cultivars. These were all budded onto M793 rootstocks and planted, in 1998 and 1997 respectively, in 7 randomized blocks in another orchard in the Elgin district (34°8'21" S 19°0'44" E). Tree spacing was 1 m within rows and 3 m between rows. Orchard management, as in the case of the initial experimentation done with adult trees, was typical of commercial practice, except that no pruning or tree growth manipulations (including chemical treatments) were applied. The means for time of IVB for the four cultivars used as parents, as mentioned above, as well as for the two F_1 mapping populations were calculated by Labuschagné *et al.* (2002a) (Table 5).

The resemblance between the observed means between parental and offspring values is also indicative of a quantitative trait not subjected to selective pressure.

Table 5. Summary of the mean time of IVB observed in clonally replicated juvenile trees over a 3 year period (1998-2000 for 'Anna' x 'Sharpe's Early' and 1997-1999 for 'Prima' x 'Golden Delicious'). Time of IVB is indicated as the day of the year, starting 1st of January, of the first sight of new leaves emerging from the vegetative bud.

	Mean IVB
Anna	220
Sharpe's Early	276
Prima	295
Golden Delicious	300
Sharpe's Early X Anna	255
Golden Delicious X Prima	296

The third mapping population, where 'Anna' was used as the male or pollen parent and 'Golden Delicious' was used as the female parent, was added to this study in order to link the two crosses mentioned above through their common parentage which will enable the testing of QTL alleles in different genetic backgrounds. This is in accordance with a new approach to QTL identification and allele mining proposed by Van de Weg *et al.* (2003). This mapping population has however not been included in the studies conducted by Labuschagné *et al.* (2004, 2002a, b) and the observed time of

IVB recorded will be reported and analyzed for the first time during this study. Parental cultivars were not included In this population and were thus not phenotyped under the same environmental conditions as the derived mapping population.

2.3 MATERIALS AND METHODS

2.3.1 Mapping populations

Three QTL mapping pedigrees were founded by interspecific hybridization between four apple cultivars. The low chilling 'Anna' was used as a common male parent in crosses with 'Golden Delicious' and 'Sharpe's Early'. Similarly 'Golden Delicious' was used as a female parent in a cross with 'Prima'.

Two mapping populations, 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious', have been planted in an orchard on Drostersnes (34°4'15" S 19°4'47" E) in the Western Cape region of South Africa during 1991 and 1989 respectively. Seven clonal replicates of a selection of these seedlings were planted in an orchard in the Elgin district (34°8'21" S 19°0'44" E) in 1998 and 1997 respectively. Seedlings derived from the cross between 'Anna' and 'Golden Delicious' were not planted in the orchard, but were kept in bags under shade netting on the Bien Donne experimental farm (33°50'36" S 18°58'39" E), with no clonal replicates being available.

2.3.2 Assessment of phenotypic traits

All phenotypic trait assessment was done and raw data supplied by Dr. I. F. Labuschagné and co-workers at the Agricultural Research Council, Infruitec Nietvoorbij, Stellenbosch. Data obtained from two mapping pedigrees, 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious', were used for previous studies involving classical quantitative genetic principles. Adult trees of the 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious' populations (initially in the 5th and 7th growing season respectively) were phenotyped for a 3 year period, from 1996-1998. The seven clonal replicates made from a selection of these seedlings were also phenotyped for a period of 3 years, from 1998-2000 and from 1997-1999 respectively, when the seedling trees were in the second, third and fourth growing season. Seedlings derived from the cross between 'Anna' and 'Golden Delicious' were first phenotyped in their first growing year, 1999, and then again in 2000, 2002 and 2004.

The date of initial vegetative budbreak (IVB) was scored as the day of the year when the first green leaves, emerging from any vegetative bud, were observed. Days of the year started with day one on the 1st of January and ended with day 365 on the 31st of December.

2.3.3 Processing raw phenotypic data

The phenotypic data, obtained from Dr. I. F. Labuschagné, were subjected to the following basic analysis processes using Microsoft Excel.

2.3.3.1 Verification of the quantitative nature of time of IVB

In order to verify that time of IVB is truly a quantitative trait the distribution of the phenotype in the seedlings from each mapping population, for each year, were studied. For this purpose the average time of IVB were calculated for clonal replicates made for each seedling.

2.3.3.2 Year-to-year variation

The data collected from each seedling over the 3 or 4 year period studied, were plotted on a simple line graph in order to illustrate the overall tendency in differences observed in time of IVB between seedlings, on a year-to-year basis. In the case of clonal replicates, the average between the observations made were used.

2.3.3.3 Clonal similarity analysis

The similarities between the phenotypic data obtained from the 7 clonal replicates made from an individual seedling over the 3 years of phenotypic trait assessment, in the case of 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious', were illustrated.

2.4 RESULTS AND DISCUSSION

2.4.1 Verification of the quantitative nature of time of IVB

The distribution of trait values in each of the seedling population studied, showed a wide (especially in the 'Anna' x 'Sharpe's Early' populations) and continuous distribution, evident of a polygenic trait under the

control of several to multiple genes. This was true for adult trees in the orchard (Figure 2) (Figure 3), averages of clonally replicated juvenile trees in the orchard (Figure 4) (Figure 5) and seedlings kept in seedling bags under shade netting (Figure 6).

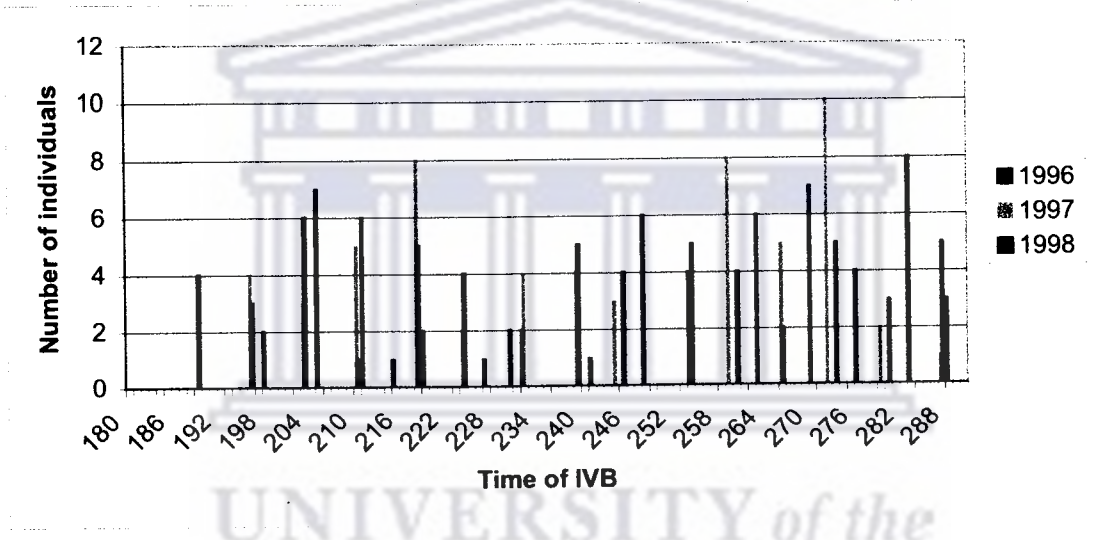


Figure 2. Graphical representation of the distribution of time of IVB observed in 'Anna' x 'Sharpe's Early' adult trees.

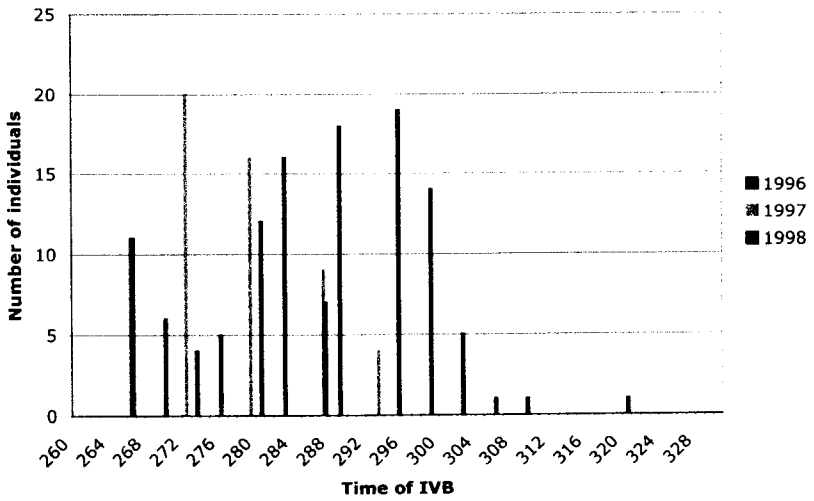


Figure 3. Graphical representation of the distribution of time of IVB observed in 'Prima' x 'Golden Delicious' adult trees

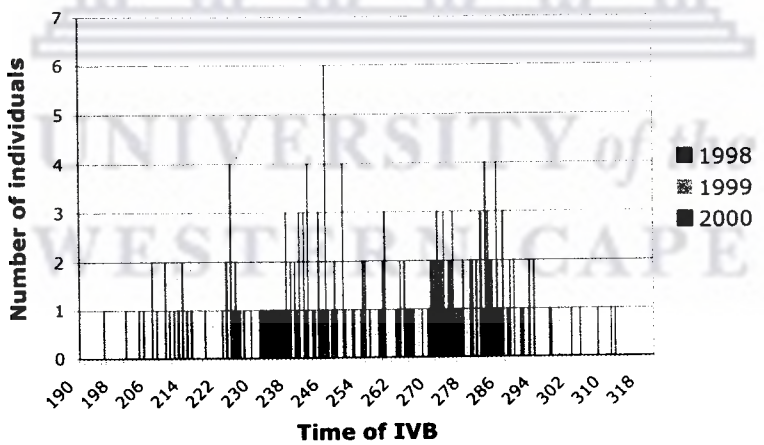


Figure 4. Graphical representation of the distribution of time of IVB observed in clonally replicated juvenile trees from the controlled cross between 'Anna' and 'Sharpe's Early'. The average time of IVB obtained from the 7 clonal replicates of each seedling were used.

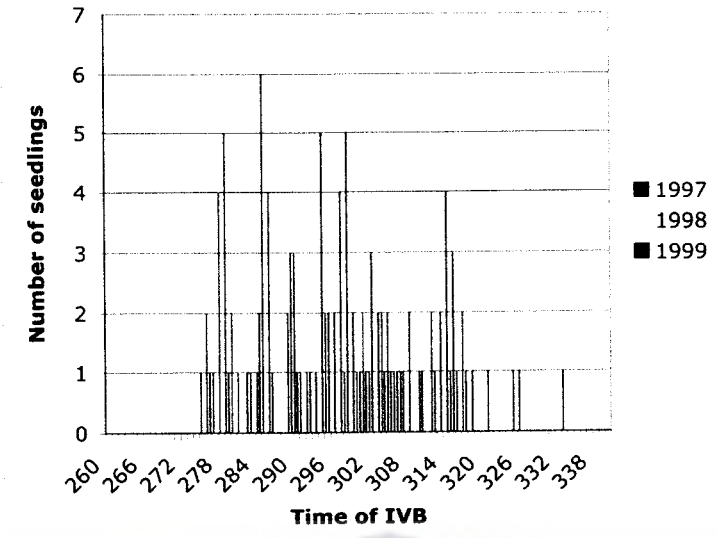


Figure 5. Graphical representation of the distribution of time of IVB observed in clonally replicated juvenile trees from the controlled cross between 'Prima' and 'Golden Delicious'. The average time of IVB obtained from the 7 clonal replicates of each seedling were used.

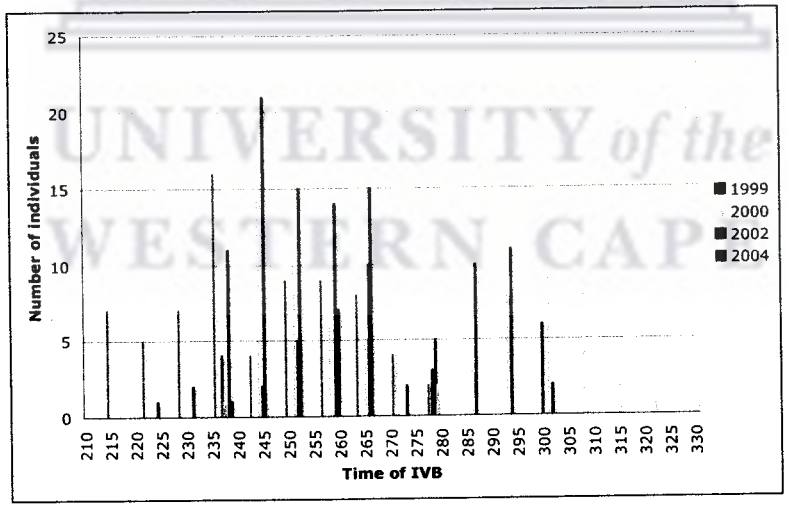


Figure 6. Graphical representation of the distribution of time of IVB observed in seedlings derived from a controlled cross between 'Anna' and 'Golden Delicious'

The high level of variation for time of IVB observed between seedlings, or full-sibs, from the same mapping population during all years of phenotypic trait assessment on all the mapping populations used, confirm that apple cultivars are highly heterozygous and serves as an indication that the trait is amenable to genetic improvement by selection. The high numbers of seedlings in the middle classes suggests additive gene action (Labuschagné, *et al.*, 2002b). Although the range of variation observed for time of IVB stayed relatively constant for both 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious' adult and juvenile populations (Figures 2-5), there has been a tendency towards a more rapid and uniform budbreak in 'Anna' x 'Golden Delicious' seedlings during 2002 and 2004 (Figure 6). This could be explained by data regarding the chilling unit accumulation at the three different locations, Drostersnes (Figure 7), Elgin (Figure 8) and Bien Donne (Figure 9), during the years of phenotypic trait assessment. (Data obtained from Ritha Wentzel at AgroMet-ISCW Institute of the Agricultural Research Council.) The amount of chilling units accumulated on Bien Donne during the period between June and August were higher during 2002 and 2004 than during the previous years of trait assessment and when the CR is met budbreak occurs more promptly and uniformly upon conditions favouring vegetative growth. Although data regarding the accumulation of chilling units in Elgin (Figure 8) were not available for all the months of 1999 and absent for 2000, the overall tendency of budbreak to occur earlier in 'Prima' x 'Golden Delicious' juvenile trees in 1997 (Figure 5) could be ascribed to the fact that chilling units accumulated much earlier during 1997 compared to other years (Figure 8). Bi-modality of

phenotypic distributions, observed in both 'Anna' x 'Sharpe's Early' (Figure 2) (Figure 4) and 'Anna' x 'Golden Delicious' (Figure 6) populations could be explained by the lack of chilling unit accumilation caused by periods of higher tempeature that result in the continuous distribution being split into two prominent peaks.

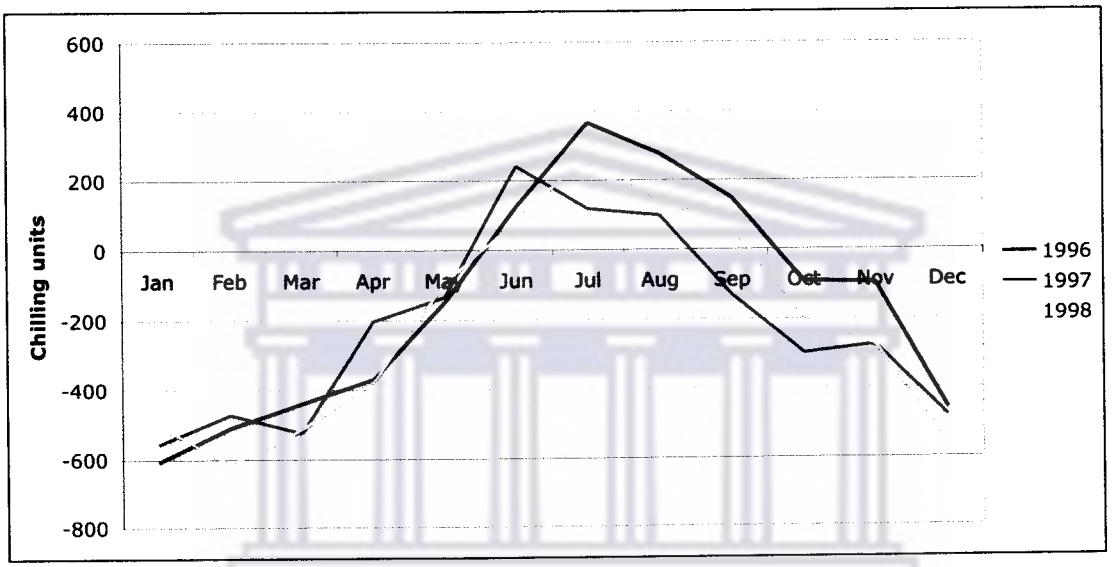


Figure 7. Graphical representation of chilling units accumulated at Drostersnes from 1996 to 1998.

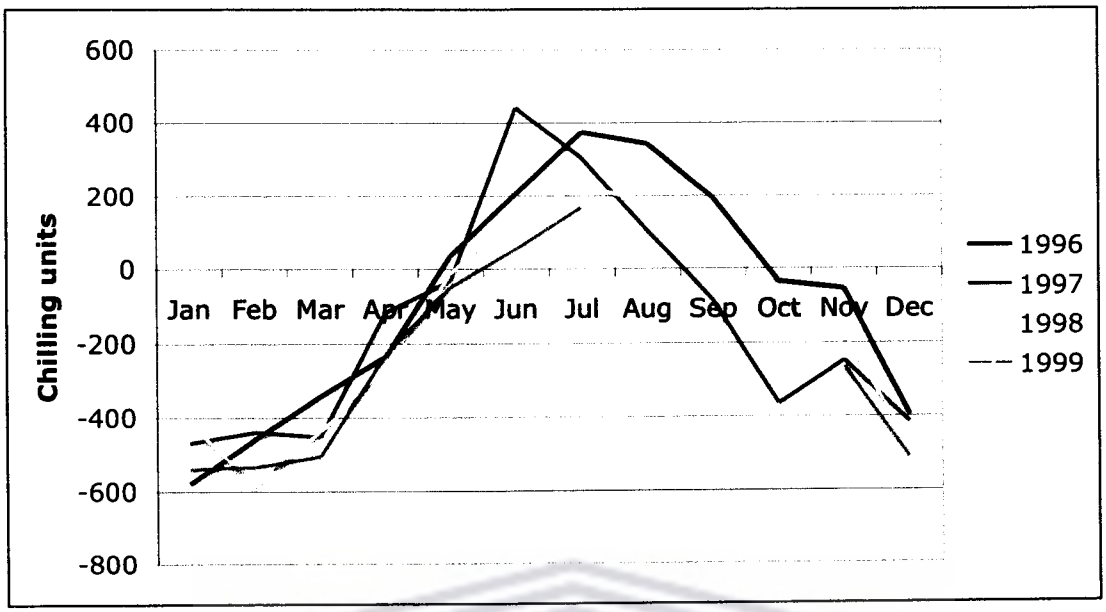


Figure 8. Graphical representation of chilling units accumulated in Elgin from 1997 to 1999.

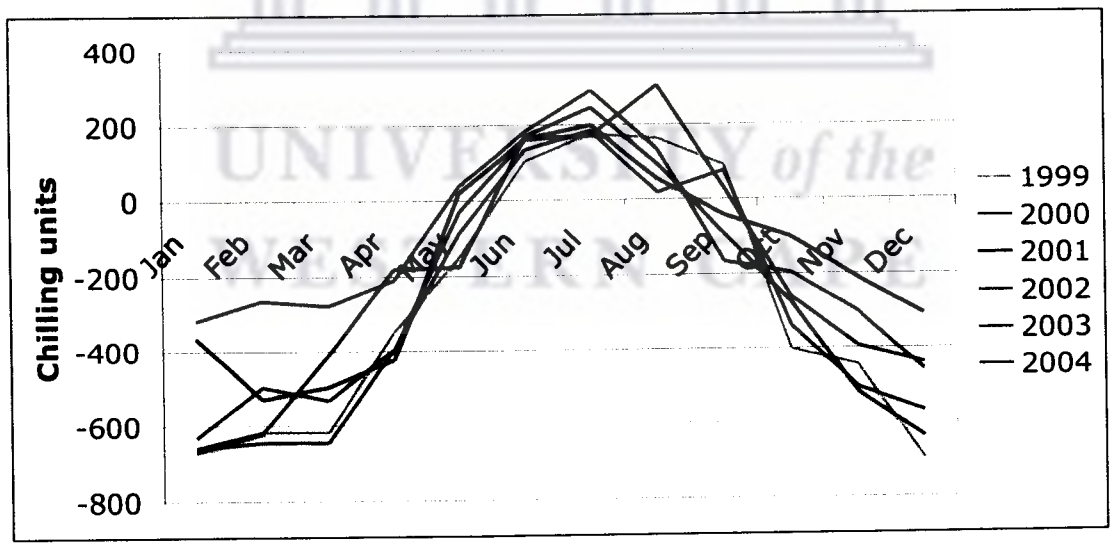


Figure 9. Graphical representation of chilling units accumulated on Bien Donne from 1999 to 2004.

2.4.2 Year-to-year variation

Each individual studied showed a certain amount of consistency when observed over 3 or 4 years, relative to other seedlings derived from the same cross. Seedlings that showed budbreak early in the first year of study, relative to other full-siblings, tend to always break bud earlier than the rest. This was true for adult trees grown in the orchard (Figure 10) (Figure 11), orchard grown juvenile trees (Figure 12) (Figure 13) as well as for seedlings grown in bags under shade netting (Figure 14).

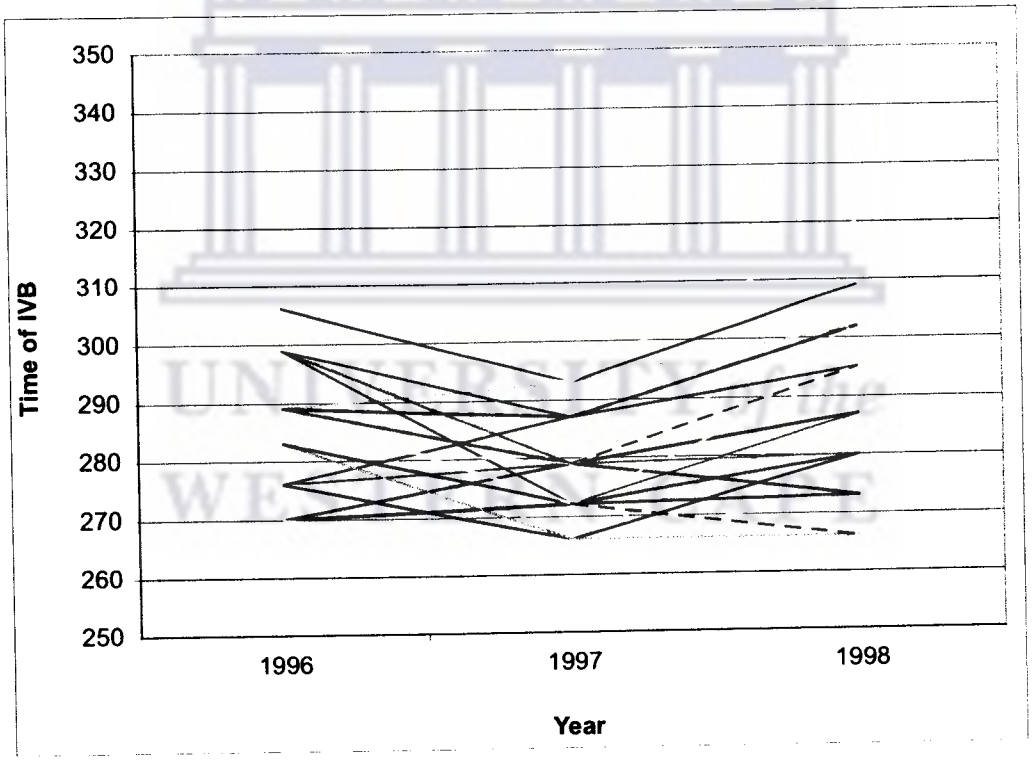


Figure 10. Graphical representation of the year-to-year variation observed in time of IVB in 'Prima' x 'Golden Delicious' adult trees.

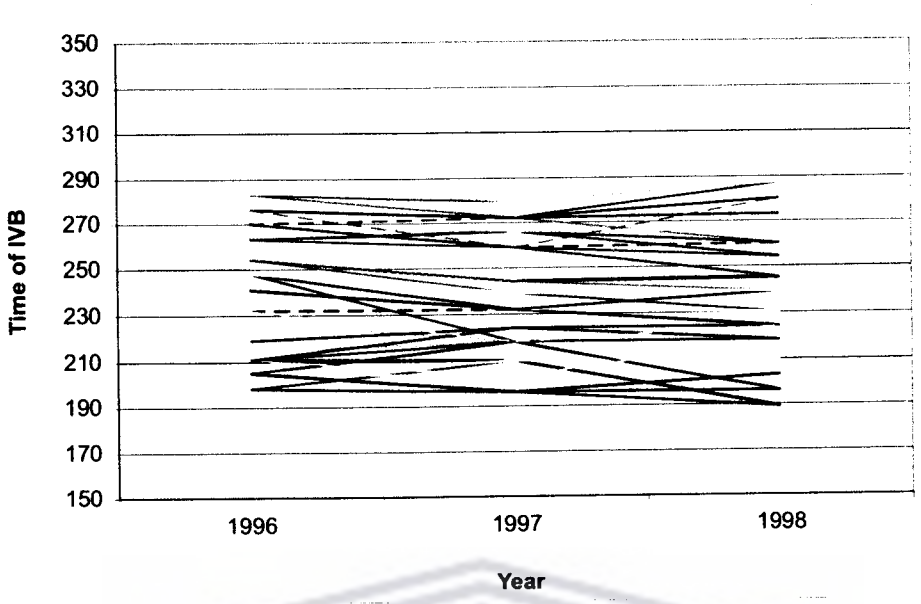


Figure 11. Graphical representation of the year-to-year variation observed in time of IVB in 'Anna' x 'Sharpe's Early' adult trees.

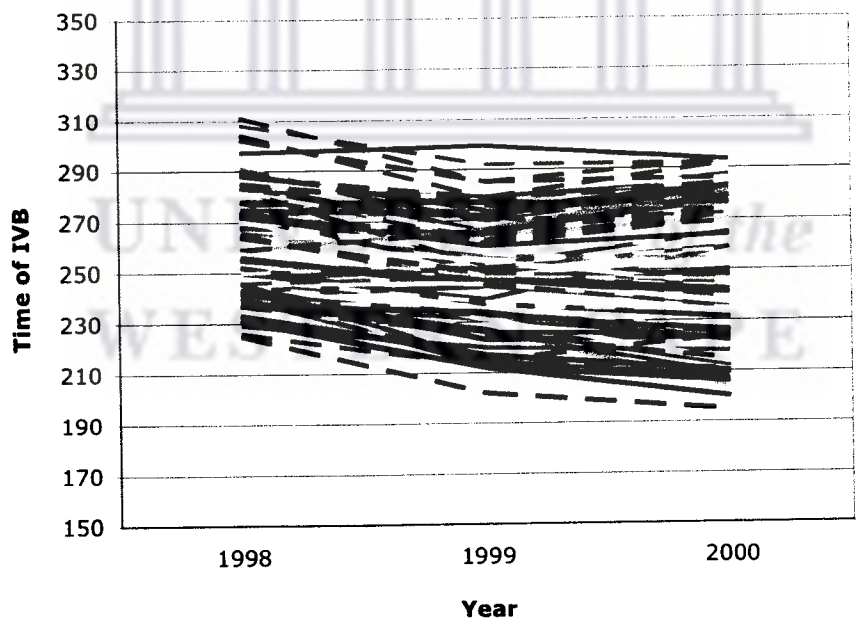


Figure 12. Graphical representation of the year-to-year variation observed in time of IVB in 'Anna' x 'Sharpe's Early' juvenile trees. Averages, obtained from the seven clonal replicates of each tree, were used.

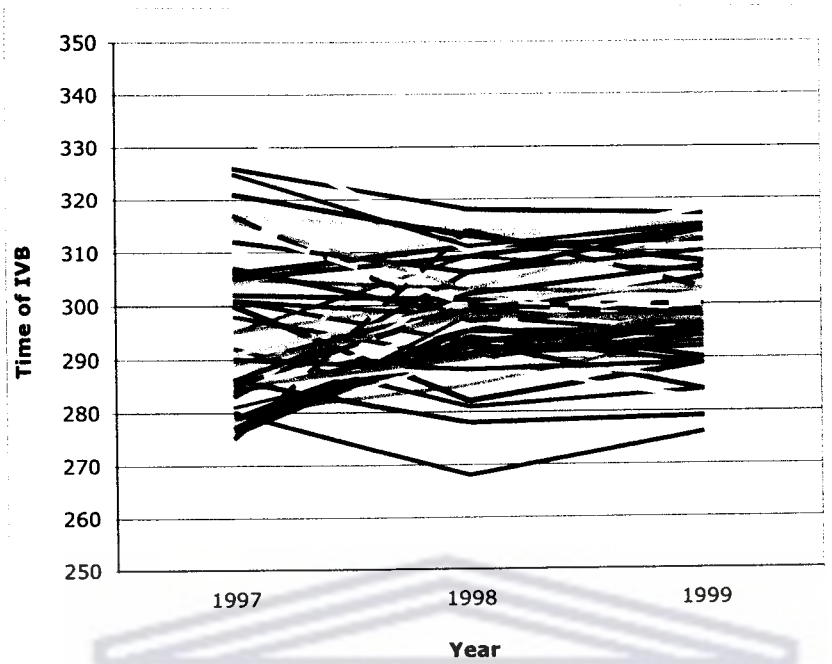


Figure 13. Graphical representation of the year-to-year variation in time of IVB observed in 'Prima' x 'Golden Delicious' juvenile trees, Averages, obtained from the seven clonal replicates of each tree, were used.

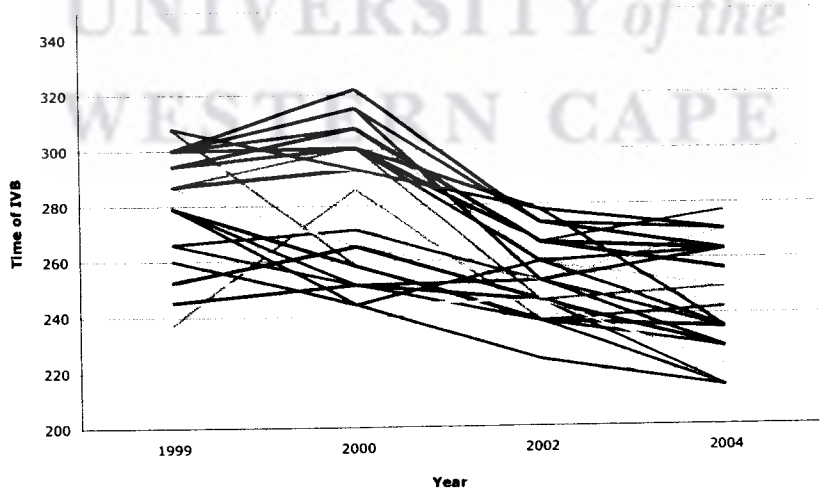


Figure 14. Graphical representation of the year-to-year variation in time of IVB observed in 'Anna' x 'Golden Delicious' seedlings.

The overall tendency for budbreak to occur earlier in 1997 in both 'Anna' x 'Sharpe's Early' and 'Prima x 'Golden Delicious' adult populations (Figure 10)(Figure 11) planted and phenotyped on Drostersnes can be attributed to the earlier accumulation of chilling units during 1997 (Figure 7). This is much more prominent in the 'Prima' x 'Golden Delicious' population (Figure 10) than in the 'Anna' x 'Sharpe's Early' population (Figure 11). The absence of chilling unit data from Elgin (Figure 8) makes it difficult to make sound conclusions regarding the year to year variations observed in juvenile trees (Figure 12) (Figure 13), while the earlier budbreak observed in the 'Anna' x 'Golden Delicious' population during 2002 and 2004 (Figure 14) can be explained by the higher chilling units accumulated early in the growing season (June – July) (Figure 9).

2.4.3 Clonal similarity analysis

Clonal replication is a technique that can be successfully used in plant genomics in order to minimize the influence of environmental factors on any given seedling. These environmental factors may be ascribed to differences in nutrition, differences in the amount of water received, differences in soil composition between the different clonal replicates, although planted in the same orchard, and differences due to human error during phenotypic trait assessment. Since a high level of heritability has been calculated for time of IVB during previous studies (Labuschagné, *et al.*, 2002a), the expectation that clonal replicates will have a similar phenological status were confirmed in both the 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious' population (Figure 15). The wide range of phenotypic trait variation observed in the 'Anna'

x 'Sharpe's Early population (Figure 4) has enabled the selection of 5 seedlings that differed in their phenological status. The phenological status, or ranking, of each of these seedlings remained consistent in terms of clonal replicate used and in terms of different years during which the study were conducted (Figure 15). Although there is no significant correlation between observations made from the same clonal replicate over the 3 year period, the ranking of individuals as being early or late in terms of time of IVB remains consistent. The narrower range of trait distribution observed in the 'Prima' x 'Golden Delicious' population (Figure 5) resulted in the selection of 5 seedlings that do not differ much in their phenological status. The level of variation in time of IVB observed between clonal replicates is similar between seedlings from both populations, with the largest difference between clonal replicates being 30 days.

'Anna' x 'Sharpe's Early' 'Prima' x 'Golden Delicious'

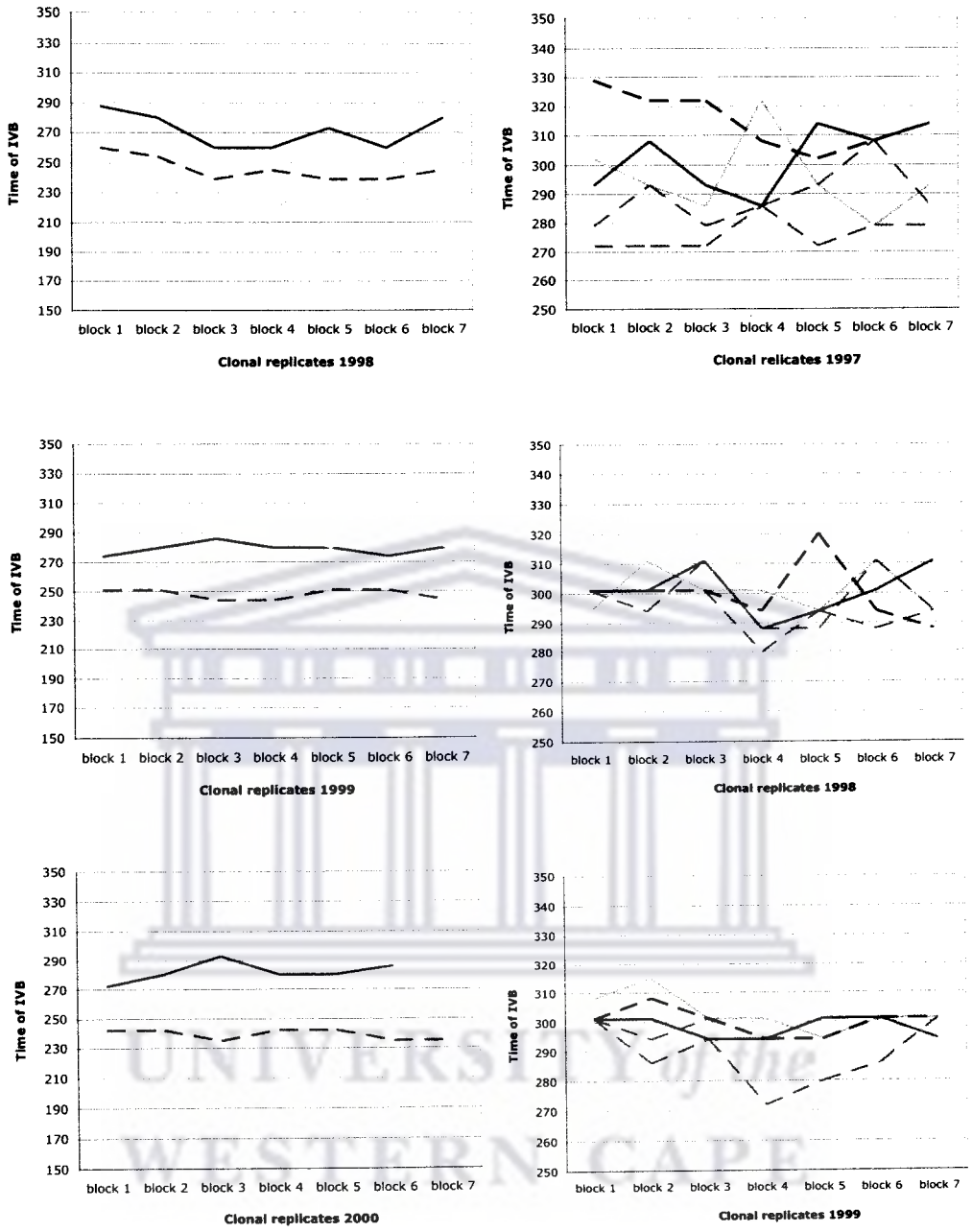


Figure 15. Graphical representation of time of IVB observed over a three year period in the seven clonal replications of five randomly selected juvenile trees from the 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious' mapping populations respectively.

2.5 CONCLUSION

The different ways in which the data have been viewed in this section confirms the fact, described by Labuschagné *et al.* (2002a) who calculated high heritability values for IVB (h^2 = 0.69), that the variation between seedlings can be primarily ascribed to genetic factors. The continuous phenotypic variation observed in time of IVB (Figures 2-6) are typical of a quantitative trait under the control of several to many genes (Labuschagné, *et al.*, 2002a). The stability of the ranking of a specific seedling as either being 'early' or 'late' relative to other seedlings derived from the same cross throughout all the years of study (Figures 10-14) is another indication that the continuous distribution of the trait is not merely coincidental. The phenotypic values obtained from the seven clonal replicates of seedlings (Figure 15) are also indicative of the low level of influence of environmental factors on the ranking order relative to genetic determination.

3 CHAPTER 3: SSR MOLECULAR MARKERS: DEVELOPMENT, INFORMATION CONTENT ANALYSIS AND DATA MANAGEMENT

3.1 ABSTRACT

Although a large number of SSR markers (288 in total), have been developed, primarily from genomic sequence data, mapped and published (Guilford, *et al.*, 1997, Liebhard, *et al.*, 2002, Liebhard, *et al.*, 2003b, Silfverberg-Dilworth, *et al.*, 2006), the recent publication of the most elaborated apple genetic linkage map (Silfverberg-Dilworth, *et al.*, 2006) revealed large segments of the genome that are still poorly represented by SSR markers. SSR markers have become the marker of choice in the construction of genetic linkage maps, especially for the purposes of QTL identification, as they allow the alignment of linkage groups obtained from different mapping populations/cultivars and the identification of the same QTL in different genetic backgrounds due to localization on the same position within a linkage group, as well as newly identified QTLs.

The focus of this study was the development of SSR markers, utilizing the large numbers of Expressed Sequence Tags (ESTs sequences) that have been made publicly available (Korban, *et al.*, 2005, Newcomb, *et al.*, 2006). EST-SSRs have the advantage of representing functional genes and their positioning on the genetic linkage map will not only shed light on the distribution of expressed genes, but the fact that the actual gene might be

involved in the regulation, or direct expression, of a phenotypic trait might render them very useful in the mapping of QTLs.

The design, testing and implementation of markers is an ongoing process, involving several research projects within the apple breeding program. The results obtained after testing a subset (245 out of 322 markers) of the total number of SSR markers designed, have revealed a amplification success of 83 %, with 82 % of these showing some level of polymorphism between cultivars tested. 119 Markers showing some level of heterozygosity between the 4 cultivars used as parentals during this investigation (Chapter 2) have been identified, rendering them ideal for genetic linkage map construction in order to identify QTLs

3.2 INTRODUCTION

Although SSRs spanning the apple genome have been developed, mapped and made publicly available by a number of research groups (Guilford, *et al.*, 1997, Liebhard, *et al.*, 2002, Liebhard, *et al.*, 2003b, Silfverberg-Dilworth, *et al.*, 2006), there are still large gaps between SSRs as described in Chapter 1 and highlighted by Silverberg-Dilworth *et al.* (2006). The characteristics of SSRs, highlighted in Chapter 1, make them the marker of choice in genetic linkage map construction. The focus of this project was the development of SSR markers using EST sequences.

Many projects of cDNA and EST sequencing are underway and a large amount of data have been accumulated (Korban, *et al.*, 2005, Newcomb, *et al.*, 2006) that can be used for the identification of SSRs. Although ESTs are generally shorter (300-1000 bp) than genomic sequences, they have the

advantage of representing functional genes expressed in a given tissue or at a developmental stage. EST-SSRs, or genic SSRs are thus present in expressed regions of the genome and primers are designed from more conserved coding regions of the genome (Varshney, *et al.*, 2005.). EST-SSRs are more frequently found in plants than in mammals (Fujimori, *et al.*, 2003) and have been identified in many plant species, such as *Arabidopsis*, cotton, grapes, soybean, spruce, barley, maize, sorghum and wheat (Varshney, *et al.*, 2005.). Trinucleotide repeats are more abundant in the coding region of genes (Chagne, *et al.*, 2004, Varshney, *et al.*, 2005.), as would be expected due to the 3 base pair reading frame, and dinucleotide repeats and tetranucleotide repeats are more abundant in the 5' and 3' untranslated regions (UTRs). Repeats in the coding region may lead to inactivation or activation of genes or the truncation of proteins, while repeats in 3' and/or 5' UTRs may affect the transcription and/or translation of genes. Fujimori *et al.* (2003) found a gradient of SSRs along the direction of transcription with SSRs being more prominent in the 5' UTR of rice and *Arabidopsis*, suggesting that they may be involved in the regulation of gene expression.

During this study SSRs, containing di-, tri- and tetranucleotide repeats, have been identified from a total of 322 unique EST-unigene sets and primers were designed from the conserved sequences flanking these repeats. Primer modification, by the addition of a 'pig-tail' (Brownstein, *et al.*, 1996), was also explored in order to improve the accuracy of genotyping. Primer pairs were tested for their ability to amplify the targeted fragment, their polymorphic information content and heterozygosity. From the 245 markers that have been tested, 203 markers, or primer pairs, generated amplification products. Each

of these were used to screen a small set of 4 cultivars, used as parentals in the generation of 3 controlled crosses included in this investigation for the purpose of studying dormancy related traits. It must be kept in mind that the small number of parental cultivars tested are not sufficient for accurate statistical calculations of polymorphic information content and heterozygosity, but a good indication of the probable use of these markers in different genetic backgrounds can be obtained. Differences observed in the amplification success of published SSRs are most likely due to changes in the unique sequence from which the primers were designed, which will result in no amplification product being present. Even in the presence of an amplification product, differences were observed in the level of polymorphism and/or heterozygosity. Homozygous markers may be used in cultivar identification studies, as long as they are polymorphic between the cultivars studied, but markers need to be heterozygous in at least one of the parental cultivars used in order to be useful in studies related to genetic linkage map construction and identification of QTLs. Markers that showed some level of heterozygosity were also included in the development of a number of multiplexed reactions, where more than one primer pair were added to a PCR reaction, in order to save time and cost involved in the screening of mapping populations.

The large amounts of data generated during the design step, as well as subsequent steps of testing and implementation of markers to mapping populations, has resulted in the need to develop of a database containing data regarding each primer set.

3.3 MATERIALS AND METHODS

3.3.1 DNA Isolation

Leaf material was collected from apple cultivars frequently used as parentals in local breeding programs, as well as from seedlings obtained from controlled crosses made between 'Anna' and 'Sharpe's Early', 'Prima' and 'Golden Delicious' and 'Anna' and 'Golden Delicious' (section 2.3.1.).

DNA isolation was done using the CTAB method described by Doyle and Doyle (1990) with the addition of PVP (Kim, *et al.*, 1997) in order to bind secondary plant products such as polyphenolics.

3.3.2 In silico SSR detection, primer design and primer synthesis

Unigene sets obtained from public available ESTs (>200 000) (Naik, *et al.*, 2006, Newcomb, *et al.*, 2006) for *Malus*, were searched for simple sequence repeats (SSRs) using the tandem repeats database (TRDB) (http://tandem.bu.edu/cgi-bin/trdb/trdb.exe) which utilizes the tandem repeats finder algorithm (Benson, 1999) to search through sequence data for SSRs, according to specified criteria. These criteria included length of the repeat, which was set at di-, tri- or tetranucleotide repeats and copy number of the repeat unit, set at more than 10 as it was hypothesized that higher numbers of repeats will allow for more variation. The length of the first index was also set at more than 20bp in order to allow for the inclusion of a short sequence before the repeat unit that can be used for primer design. Single base

substitutions and insertions or deletions within the repetitive sequence were minimized by defining a 95 % or more match.

Primer pairs, flanking SSR containing regions, were designed by visual inspection of the conserved sequences flanking repeats. Primers were chosen in such a way that the resulting amplimers vary in size, ranging from 80 bp to 450 bp. Primers had a GC-content of between 40 and 60 % and an ideal melting temperature (T_m) of 60 °C.

Primers were also synthesized for published pear (Yamamoto, *et al.*, 2002a, Yamamoto, *et al.*, 2002b, Yamamoto, *et al.*, 2002c) and apple (Guilford, *et al.*, 1997, Liebhard, *et al.*, 2002, Liebhard, *et al.*, 2003b, Silfverberg-Dilworth, *et al.*, 2006) SSRs. All Primer pairs used during this study were synthesized at Applied Biosystems (Foster City CA, USA) and the primer closest to the repeat were labeled with one of four fluorescent dye colours (viz. 6-FAM™, VIC™, NED™ and PET™).

3.3.3 PCR amplification

An initial single locus (simplex) screen of microsatellite markers were done in order to test the ability of specific primer pairs to amplify target DNA and generate amplification products or fragments. Primer pairs generating such products were then assessed on cultivars used a parentals in order to determine polymorphic information content and heterozygosity. On the basis of above mentioned criteria, loci can then be selected for multiplexing.

3.3.3.1 Simplex

Simplex amplifications were performed in volumes of 20 µl with 1 unit Taq polymerase, 0,2 µM Tris-HCl (pH8.3), 1 µM KCl, 0,07 µM $MgCl_2$, 50 µM

each dNTP's, 0.016 μM each primer and 10 ng DNA template. PCR reactions were optimized, in order to obtain the correct annealing temperature for a specific primer pair, using a 'touch down' approach on an Eppendorf Mastercycler® gradient PCR machine (Eppendorf-Netheler-Hinz GmbH, Hamburg, Germany). The thermal cycling conditions were as follows: (1) 96 °C for 5 min, (2) 10 cycles: 94 °C for 40 sec, (65 °C - 55 °C) touch down to (55 °C - 45 °C) for 40 sec, 72 °C for 2 min, (3) 30 cycles: 94 °C for 40 sec, (55 °C - 45 °C) for 40 sec, 72 °C for 2 min, (4) 72 °C for 45 min and (5) 4 °C hold. Once the optimum annealing temperature was identified, individual primers were implemented on a selection of parental cultivars. Thermal cycling conditions were as mentioned above with the exception that no gradient was used for the annealing temperature, and amplification was done on a 2720-Thermal Cycler (Applied Biosystems, Foster City CA, USA).

3.3.3.2 Multiplex

Primer pairs labeled with the same fluorescent dye, but amplifying differently sized fragments, were selected, pooled and amplified in the same PCR reaction. Multiplex amplifications were performed in volumes of 20 μl with 1 unit Taq polymerase, 0,2 μM Tris-HCl (pH8.3), 1 μM KCl, 0,07 μM $MgCl_2$, 10 μg bovine serum albumin, 50 μM each dNTP's, 0.016 μM each primer and 10 ng DNA template. The thermal cycling conditions were as follows: (1) 96 °C for 5 min, (2) 40 cycles: 94 °C for 40 sec, 55 °C for 1 min, 65 °C for 2 min, (3) 65 °C for 45 min and (4) 4 °C hold, and amplification was done in a 2720-Thermal Cycler (Applied Biosystems, Foster City CA, USA).

3.3.4 Fragment detection

3.3.4.1 Gel Electrophoresis

Gel electrophoresis was performed to verify the amplification success of primers during the initial steps of testing and thereafter to verify amplification before automated fragment analysis. Fragments were resolved on 6 % polyacrylamide (19:1 acrylamide:bis acrylamide) gels (80 mm x 100 mm x 1.5 mm) in 1X TBE. Prior to loading of samples, 2.5x v/v 90 % deionized formamide loading buffer (containing 0.1 % w/v Bromophenol Blue, 0.1 % w/v Xylene Cyanol and TBE), was added to samples after which they were denatured at 95 °C for 5 min. Gels were run in TBE buffer at 15V/ cM for 70 min. Gels were visualized using a shortened silver-staining method: gels were soaked in a 0.1 % w/v $AgNO_3$ solution for 10 min, rinsed with water three times, soaked in 1.5 % w/v NaOH, 0.15 % v/v formaldehyde and 0.01 % w/v $NaBH_4$, until bands appeared and finally rinsed three times with water in order to stop the staining reaction.

3.3.4.2 Automated fragment analysis.

Since actual fragment size determination and differentiation between larger fragments and dinucleotide repeats are difficult to accomplish with the use of gel electrophoresis, the ABI Prism 310 and 3100 (16-capillary array system) Genetic Analyzers (Applied Biosystems, Foster City CA, USA) were used. Size determination of 6-FAM™, VIC™, NED™ and PET™ labeled primers was done using an internal size standard, labeled with either ROX™ (GeneScan™–500 ROX™) or LIZ™ (GeneScan™–500 LIZ™) fluorescent

dyes, POP-4 sieving polymer matrix, 1X Genetic analyzer buffer with EDTA and 35 cM X 50 μm uncoated capillaries.

Samples were prepared by adding 3 μl of a 1:10 diluted PCR product to 10 μl Hi-Di formamide (Applied Biosystems) containing 0.15 μl size standard. In cases where PCR products were pooled to maximize throughput, 1:10 PCR product dilutions were pooled in the ratio 6-FAM™:VIC™:NED™:PET™ = 1:1:3:2 to equalize the fluorescent signal. The samples were heat denatured at 96 °C for 5 min and then snap cooled on ice prior to loading them into the autosampler tray. Samples were injected for 15 s at 15,000 V and separated at 15,000 V for 24 min with a run temperature of 60 °C. The resulting data can be displayed as a electropherogram using GeneScan® software (Applied Biosystems, Foster City, CA)

3.3.5 Primer modification

The addition of a 'pig-tail' to the 5' end of the reverse primer to facilitate genotyping, as described by Brownstein *et al.* (1996) and positively confirmed by Silverberg-Dilworth *et al.* (2006), were tested using three published SSR markers (Table 6). CH05a05 and CH03g12, a single- and multilocus marker respectively, were selected in order to try and improve the amplification product yield. CH03b01 was selected in order to try and eliminate or reduce the high level of stutter.

Table 6. Three primers selected for the testing of amplification results after the addition of a 'pig-tail' to the 5' end of the reverse primer.

Marker	Forward primer	Reverse primer	Pigg-tail sequence
CH05a05	TGT ATC AGT GGT TTG CAT GAA C	GCA ACT CCC AAC TCT TCT TTC T	GTTT
CH03g12	GCG CTG AAA AAG GTC AGT TT	CCA GGA TGC GCA TGT ATT TG	GTTT
CH03b01	ACA AGG TAA CGT ACA ACT CTC TC	GTC ACA AAA CCG CCA GAT G	GTTT

3.3.6 Database management

A database was developed using the FileMaker Pro 8.5v1 software package (FileMaker Inc, Santa Clara, USA) (http://www.filemaker.com). Use was made of different tables containing data regarding primer design, sequence information, observed fragment size distribution and implementation on mapping populations. Defining relationships between each of these data tables enabled the interactive use of the database, in order to get information regarding certain aspects of any selected primer pair, through the click of a button.

3.4 RESULTS AND DISCUSSION

3.4.1 In silico SSR detection, primer design and primer synthesis

SSRs were identified from sequence data using criteria such as the repeat motif, the copy number of the repeat and the length of the first index, specified in TRDB (Figure 16). Di-, tri- and tetranucleotide repeats were identified and for each of these the number of copies of the repeat was set at larger than 10. This value was chosen as it was assumed that the higher the number of copies, the more likely the occurrence of polymorphisms. The length of the first index, or the sequence length available before the start of the repeat, was set at larger than 20 bp, as this will allow the possible positioning of a primer before the start of the repeat. The percentage of matches was set at larger than 95 % as this would eliminate sequences where many insertions, deletions or single base substitutions are present within the repetitive region.

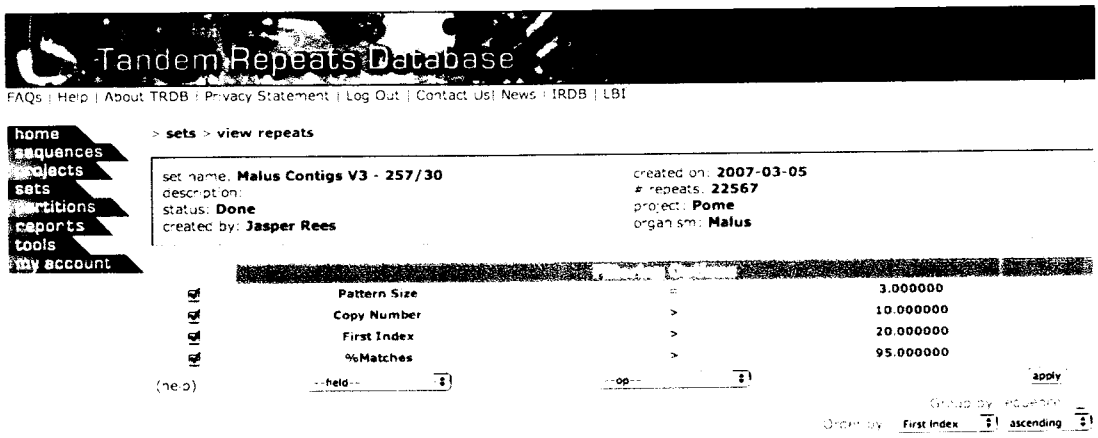


Figure 16. The user interface of the Tandem Repeats Database. In this example a predefined set of Malus contigs with an initial number of 22,567 repeats were filtered for repeats with a pattern size of 3, a copy number greater than 10 and a first index greater than 20. This resulted in the reduction of the number of repeats to 371.

TRDB generates a list of sequences containing repeats (Figure 17), that conform to the set criteria, which can also be viewed graphically (Figure 18) or just in more detail (Figure 19), in order to identify sequences from the more conserved flanking regions that can be used for the design of sequence specific primer pairs. The repetitive sequence selected (Figure 18) has a match percentage of 97 % due to the presence of a single base substitution in the eleventh repeat.

☑	27--63 [browser]	3	12.300000	100	0	0	0	32	0	67	74	37	Contig20593
☑	44--82 [browser]	3	13.000000	97	3	0	35	33	0	30	71	39	Contig8968
☑	49--80 [browser]	3	10.300000	96	1	3	68	0	31	0	55	32	Contig10267
☑	50--84 [browser]	3	12.000000	97	1	2	0	34	0	65	63	35	Contig16808
☑	67--123 [browser]	3	19.000000	98	2	0	64	0	35	0	107	57	Contig21657
☑	74--122 [browser]	3	16.299999	97	3	0	0	65	0	34	91	49	Contig281
☑	77--107 [browser]	3	10.300000	96	4	0	32	64	0	3	55	31	Contig16327
☑	84--114 [browser]	3	10.300000	96	4	0	35	29	32	3	55	31	Contig20367
☑	84--119 [browser]	3	12.000000	100	0	0	33	66	0	0	72	36	Contig12573
☑	86--121 [browser]	3	11.700000	97	1	2	0	36	0	63	63	36	Contig6489
☑	122--152 [browser]	3	10.300000	100	0	0	32	35	32	0	62	31	Contig13042
☑	125--155 [browser]	3	10.300000	100	0	0	67	0	32	0	62	31	Contig15999
☑	134--176 [browser]	3	14.300000	100	0	0	67	0	32	0	86	43	Contig13802
☑	147--179 [browser]	3	11.000000	100	0	0	0	66	0	33	66	33	Contig11537
☑	148--179 [browser]	3	10.700000	100	0	0	34	34	0	31	64	32	Contig19664
☑	168--200 [browser]	3	11.000000	100	0	0	0	33	0	66	66	33	Contig333
☑	169--205 [browser]	3	12.300000	97	3	0	0	64	0	35	67	37	Contig1253
☑	185--228 [browser]	3	14.700000	100	0	0	0	31	0	68	88	44	Contig20191

Figure 17. Repeats conforming to the specified filtering options in TRDB are listed in tabular form. Pattern size is restricted to trinucleotide repeats, while the copy number and % matches are above the specified limits of 10 and 95 % respectively. Information regarding repeat motif and Fasta header are also indicated.

44--82	3	13.000000	97	3	0	35	33	0	30	71	39	Contig8968

Figure 18. Graphical representation of a selected sequence showing flanking sequences as well as the repeat sequence. The trinucleotide is repeated 13 times with 97 % of the nucleotides matching the repetitive motif. A single base substitution in the 11th repeat accounts for 3 % mismatches.

left flanking sequence:

GCACCAGCAGCAGCAGCAGCTTCTTCTTCACACCCTTCAATCC

pattern:

CAT

sequence:

CATCATCATCATCATCATCATCATCATCATCAACATCAT

right flanking sequence:

GATGTTGCTGTTGATTTTCTCCCAACAAATCCCCAAAAAAACAAGTTGAATTCCACCAAACAATGTCTTC
GCTCTTCAGCTCTCCCCCCCAACCCCGACGACGACGACCCCGCGTGATGTAGAACAACCCAACCAACCCCAA
AACGACCTGTCGTTTCTCCGCCCAAACCCTCCGCCCGCCAAATCCCCCGCCCTCCCCACCTTC

Figure 19. Detailed report of a selected sequence showing the repeat and flanking sequences.

Primers were designed using 367 EST sequences, containing 198 dinucleotide, 132 trinucleotide and 37 tetranucleotide repeats. After further investigation and the assembling of EST sequences into contigs, it was clear that 12 dinucleotide, 8 trinucleotide and 6 tetranucleotide repeats were designed from sequences representing contigs already used for primer design. A further investigation and the publication of a recent set of SSR markers by Silverberg-Dilworth *et al.* (2006), revealed 14 dinucleotide and 5 trinucleotide repeats that were identified from sequences, or sequences from the same contig assembly, than used for the design of published markers (Liebhard, *et al.*, 2003b, Silfverberg-Dilworth, *et al.*, 2006). This reduces the number of primer pairs, designed from 'unique' contigs, to 322, including 172 dinucleotide repeats, 119 trinucleotide repeats and 31 tetranucleotide repeats (Table 7).

Table 7. Summary of primer pairs developed during this investigation and the total number of primers developed from unique sequences. Primers are grouped according to the type of repeat.

Type of repeat	Primers synthesized	Primer duplication (internal)	Primer duplication (external)		Number of unique primer pairs
			Number	Publication	
Di	198	12	3	Liebhard et al. 2003	172
			11	Silverberg-Dilworth et al. 2006	
Tri	132	8	5	Silverberg-Dilworth et al. 2006	119
Tetra	37	6			31
TOTAL	367	26	19		322

The primers were designed from flanking sequences such that one primer is situated closer to the repeat. This primer was then synthesized with a fluorescent label attached to the 5'end. The second primer was designed some distance away from the repeat and the first primer, in order to allow for a variety of fragment lengths when the targeted sequence is to be amplified to facilitate the development of multiplexes. Should the resulting fragment be of a different size than expected, the unlabelled primer could be re-designed to be closer to, or further away from, the repeat. This facilitates the amplification of a different fragment length without the re-synthesis of the more expensive fluorescently labeled primer.

3.4.2 PCR amplification

3.4.2.1 Simplex

From the 322 unique SSR markers designed, 245 were tested during this project, of which 203 gave amplification products (Table 8). The success

rate for amplification is thus 83 %, corresponding to the 60-90 % success rate reported by Varshney *et al.* (2005.). The failure of PCR amplification of some SSR containing regions might be ascribed to the use poor quality sequence data in the primer designing step, or it may be caused by the presence of introns in genomic DNA. Chee *et al.* (2004) reported up to 4 introns, ranging from 77 to 611 bp in length, working with cotton EST primer pairs. The latter could also result in the amplification of fragments larger than expected, resulting in fragments not being detected due to the separation ability of gels used. Failure to amplify was confirmed by designing a second unlabelled primer closer to the repeat unit. A total of 30 primer pairs that failed to yield amplification products during initial testing were found to yield amplification products after the designing of a second primer closer to the repeat. After identifying 203 primer pairs yielding amplification products, 163 were subjected to polymorphism testing and 133 (82 %) primer pairs (Table 9), including 68 dinucleotide, 50 trinucleotide and 15 tetranucleotide repeats, yielding polymorphic fragments when used on a selection of apple cultivars, were identified.

Table 8. Summary of the results obtained during the testing of a subset of the total primer pairs designed.

Primer pairs designed	Primer pairs used during this study	Testing workability of primers	Testing polymorphism	Type of polymorphic repeat
322	77 Not included			
	245	42 Not working		
		203	40 Polymorphism not determined	
			30 Not polymorphic	
			133 Polymorphic	68 Dinucleotide
				50 Trinucleotide
				15 Tetranucleotide

Table 9. Polymorphic SSRs, accessions used for the identification of repeats and primer sequences designed from conserved flanking sequences.

Primer nr	Accession nr	Forward primer	Reverse primer
A29	AT000141	GAA ATA AAC ACC GAG TAA ACA G	TGC TAT CTG GTT TTC TTT TAG C
A30	AT000400	CGT ATC GAA GTA GAA CGA CG	CAG GGT TGT ACG GAT TCA CG
A180	CN444111	TGA GGC CAC CTA AAT ATC AC	CAG GAT GAG AGT TCT TGA GC
A182	CN445253	TGC AAG AAT CAT CCA CTT CC	TTG GAC CTG TGA GGA CTC C
A183	CN488421	TTG GAC CTG TGA GGA CTC C	GAT GCC GAG TGT CTG TAT GC
A184	CN489175	AGC CCT CTC CAA TAC CAA CC	TTT CCT GGA AGA GAT TGA CG
A186	CN490349	GTA CTA TCA GCA GAA ACT GG	GAT TTG AGC ACA ACA TAC GG
A188	CN490740	AGG ATC CTT CCT CGA TTT GC	GGC ATT GAG GTT CTT GAT CC
A192	CN491993	AAG CAG TCG CAG CAG GTG	AAC AAC CGT TCG GAT TCT CG
A193	CN492206	ACA TAC TGG AGT CTG CGA GC	CAA TAC GCT AGT GAA GAC GC
A195	CN492475	ACT CAC CCC CTT CCT TTC C	GAA GAA AGG TAG GGG TCA GC
A196	CN492626	TGC AGG TTG AGA TGG TTT GG	GAC CCA AGA ACA ACA AAA CC
A197	CN492735	GAC ATG GCT AAC CAG GAT GC	GGC AAG TCT TAG GGT TCG G
A200	CN493925	TCT CCT TCA CTT CCC ATT CC	TGG TGA TGG CAT ACA CAT CC
A202	CN494248	ACC TCT CTT CAT TCT TCT CC	GAA GAG CAT AGA AGA ACA CC

Primer nr	Accession nr	Forward primer	Reverse primer
A204	CN494928	AAT TAT ATC CGT CCG ACT CCA	TTA CTG CTA CCT GAT GAT CC
A207	CN495433	ACA AGA GCA GCA GCA TTT CG	GTA GCG TGT TTC AGG CAG TC
A208	CN495651	CTT CTC CCA GAA CTG ACT GC	TCT ACA ACC GCA AAC ACG AG
A209	CN495857	TCA AAA CCC ACC TCA TAT TGC	TGA GCT GGG AGG AGT AAG C
A212	CN496144	CTC AGA CTC CTG CTG CAC C	TAC TGC CTG GTG TTT CTT CC
A213	CN496756	TCG GTG GAA GAC CAA GCA G	CAT GAT CAT GTG GCG CCG T
A217	CN579502	TCG TGA AGT GCC AAG TAT CG	TGG CGG ACT GCT CAA TTG C
A219	CN580620	TGC GGT CAA CGA TGT CTT CG	AAG GTA CAA GCC CGC AAA GG
A221	CN580954	TCT CTT GTC AAG GAT GGA CC	GAG CAT TTC TGG ACC TCC G
A222	CN581649	AGC CCT GAT CTT CCT CTA GC	GAC AAT CTT CTG AAA GTC TGG
A225	CN581979	CCC ACT TAG TGG GAA AAG GC	GCG GTG AGA GGC AAT GAA C
A227	CN493171	TCT TAC TTC GTC GGT GGA CC	TGT GTG GCT ATT ACC TGA GG
A229	CN496966	GGA GGA GAA TAT GTG ATT TTG AG	ATG ACG GGA GAT CGT CTT CC
A230	CN497136	GGT GAC TGT AAT CGG TTA CG	TTG TAC AAC ACT ACG ATG GAG
A231	CN580271	TCT GGC TCT CAT CGG TTT GC	TCG ATG CCC TTG TAA CGC C
A232	CN492903	TAT GCT CCA ATA GCA GTT ACG	CTC CTT GAA GCG AGC CTG G
A233	CN903950	TTT CCC TTT TGG CCA GTG CA	GTT TGG GCC TCG ATG ATG G
A234	CN938125	GCC TTC ATC CCC CCT TGA	GGT GTA TAG GAA TCT TGG AG
A235	CN881550	ATC CAA ACA ACC CCA TTG CG	AGT CGA TGT TGA ACG CTC CA
A236	CN910036	GAG AAA CCG TTT GAT TAC AGC	CTC CAT CCC CAA TCA CAC C
A238	CN865016	TTC TTC ACA CCC TTC AAT CC	AAA GCG CCT GCG ATT GCG
A242	CN905641	GGA AGG TTT CAA AGC ACT CA	CAC ATA CAG GTG CGC TTC G
A244	CN947446	CCG TTA CAG CTA TCC AAA CC	ATA ATG GCC ATT CTG TTC AGC
A245	CN943613	TAG CAG AAA CCA GCA GAT GG	GAA GGA CCC GAA TTG GAG C
A247	CN879152	CGT TGG AGA TGA TCA GTA CG	ACC TAC AAT AGT AGT GGA GAC
A249	CN891581	CCA AAA CTC CCA CGA CCG C	CCA GAG CTT GTA GGA CTC G
A253	CO540769	TCC TAG GGT CGG AGA GCA G	CTC AAG AAT CAC CAA CAA TGC
A254	CN933736	TGG CAG CTC CAC CAC AAT C	GCC AGA TTC ACA CGA AAG C
A256	CN868958	TGT ACT GCT TCA GCT TAT TGG	ACT GAG GGA AGA AGA TCC AG
A259	CN904905	GGT CAT GGA ACT GTT GGT GG	TTC TGA TGA ATG AAA GCA CCT
A262	CO865955	TAC TCA TGG CGG CAA CTC C	CTG CAG GAG GTT GCT GGC
A266	CN851624	AAC TGT AGA AAA AAC ACT CCC	GGT CCT CCT TTC ACA AAT GC
A267	CN910302	TTT TCA GGC ATC ACT GTC CC	ATC AGG ATT TCC AAC AGC GC
A268	CN916087	CAG TCG CTC TCA CCC TGC	GAA CTC TAG CCA ACC ACC G
A269	CN889098	GAG TTC GAG TGT TGT AAG GC	CAG CCA TCA CAG TCC ATC C
A274	CN925672	ACA CGG TAA ACA CTA CCA CC	GCG AAC TTC ACC TTC GCA AA
A277	CN866018	TTC CTC TCA TCT ATC CTT TCG	GAG GTG ACA GAC AAA TTC GG
A279	CN887525	TAG TAG CTA CAC ACT CTT TCC	GCA TTG CCT TGA GCT CCA G
A281	CN870040	CCT CAG CAT CAT CAA CCC C	GGA AAT GCG ATT TCG AAC CC
A283	CN921216	CGC ACA CCC CCA AAT GCG	AGA GCT TGT CGC CCT CGG
A284	CO752155	TGC CTA AGA ATC CAT CTG GC	TCT CGA ACT TAC TAA CTA GGC
A285	CO753983	CAC CGT ACG AAG GTT ACC C	ACA GCA GCA TAG TGC AGC C
A286	CN917681	CTC CAT TTA CAA CCA AAG GG	TCC GTG GCA GAC TTT TCG G
A290	CN864595	CTC TGC AAA CTA CCA CCG C	TCC TCC TCA ACA GCG GGG
A293	CN944444	TAG TGC AAG TAC TGG GGC C	CAT CGA TAG AAT AGG ACG GC
A294	CN946851	AAT GAC TCA AGC GAT CAG GG	CCG ATC CAA GTA GTT AAC GG
A296	CN880881	ATA GCT CAT ACC GCT TCT CC	GTG ACG AAA ACC AAG AAC CC
A298	CN943252	TCC CAC TGA CAC TAT CAC C	TGC AGG AAA TGA GAA TGC GC
A300	CN939907	ATC CGC AGA ACT GAA GGC G	ACT GGT CGG TTA TCG ACG G
A301	Z71981	ATAATTGGGGTATGGATGAGG	CTTGTTGGGATTAAATCCGGC
A304	AJ291492	GCG AAC TCC AGT GAG TGG	TAA GCA CTA AAC CAC GGT GC
A307	CN445290	TCACTTTCTCAGTTGCTCTGG	ATGGAAGCTTACTCTTTTCCG

Primer nr	Accession nr	Forward primer	Reverse primer
A308	CN444942	GCT CTC AAA GTC TCT CCA GC	TAC GGA CTC TCT TTG GGG C
A310	AU301301	GGC ATA GCA ATG CTT GAA GG	GAA TAG CAC AAA GGA GGT TGC
A311	AU301254	TCC CGG AAA TTT TTC AAC GC	AAC GCT AGG GAT TGG TCG C
A314	CN495924	CTC TCA ATG AGT CCC CTG C	AGG AAC AAG ACT TGC ATG GC
A315	CN496099	ACC TCT ACC CGA ACT CCA CC	GCA CCA CTT TCT TCC AGG C
A318	CN580227	GAC GTA AAA TCC CTA ATT CCC	TCA TCC CAG TCG TCT TCC C
A319	AF527800	TTGGTCAGACATACACTGGG	GGTGGTAAATCTCCACTACC
A320	CN580637	ACA ACA GCT GAC GAA CAA GC	CTA CTC GTC GAA GTA CGC C
A326	U50187	ACCTGAGAGAGCTCCAAACG	GTGCGCCACGTCAAATACG
A329	CN496002	AGCAGCAGCTAGGCTAGAGC	AAATTGCCTTGCCAGATTAGC
A331	AB162040	GGAGTGCTATTAGCTCCTCC	TCCTTGAATCTCAACTCTAGG
A332	CN488733	CAC AAC CAT TCC ACC AAG TC	CAG CCG GAG CAG TCT ACC
A334	CN444542	AAGCCAGGCCACCAAATCC	GAGAGCTGCATTATTTGGTCC
A335	CO052033	TTG CCA ATC CGC ATT CGC C	TGA GGT TCC CGC CCT TGC
A336	CO168310	GTC GAC TTC GCC CGA AGC	ACG ACC AGG TTC ATG AAC TG
A337	CO898678	CCC AAG TGC ACC ACA TAC AG	AGC TTC TGG CAG CAA GTG C
A339	CO066563	ACA AAG GAA CAG TGA AGA CTC	TAC TTG CTC TGC ATA GTT TGG
A340	CO416051	CCT CAC TAA ACG CAT TGC AC	CGG TAC GAT GAG GAT CAT CC
A341	CO723148	CGG TGG TGA CTA GTA TCA GC	TAT GGA GGA AGA AAC TGA GGC
A343	CV084260	CAA AGC AAA ACA GAG GAT TTG	GGA GCG CAT GAA ATT ACT GC
A344	CO905375	AGT CTC TGT TTT TGC TCG TTC	GAA CGC CGG GTC CCT GC
A352	CO866862	CAT ACG CAG CTC CCA CAC G	AGG AAC TTC TCC AGT GAG G
A355	CO903877	AACAGGCGCCATTATTTGCC	GAA ATC AAA GCC GCT TGA G
A363	CO052202	AAG AAG GGA AAG AAA ACC AGC	AGG GTT GAA GAT TGG GGG C
A365	CO903680	CAG CAG TTG CAA CAA GTC C	GTG GAA ATG GCT AAG CAA GC
A367	CO417416	GAC CTC AGT CCA AGT CTC C	CTG AAT CTC TGC CGG AAG G
A368	CO723511	CTG TCG GGA TTC ATT GTT GC	CCG AGT AGA AGG CTG AAG C
A369	CO865608	CAA CAA GTG TGC CTC TGT GG	AGC AAG CAA CAG ATC AAG CC
A372	CO052555	GAA GTT CTC ATC AAG TCT TGC	GCT TCT GCA CAA TGG CTG G
A376	CO867345	TAC ATC CAC CAT GGA AAG ATC	CTG GTC GGA CAG GTT AAC G
A377	CO068842	TGG TTG GAG ATG TTC CAT GG	ACC AGC TAG ATT ATC TTC TGC
A378	CO753033	ACA CAG TCA TTG CTT CCT CC	ACC CAG CAT GTG GTC GAA G
A379	CO865207	TGC ACC AAA TAA GCC GAT CC	CAA GAA GTG CAA CCA GTC GA
A380	CO866737	AGC AGC TTC CGT TTC CCT G	AAA CAA CCC ACG CTC GGA G
A381	CO751676	TGT GGC TCT GGA TGG TTC C	TAC CAG TCC ATC CGT ATA GC
A383	CO903298	TTG AGA AGC AAT GCT GCC TC	TGC CAC AGT TGG AAG GTG G
A385	CO865258	CTC CTG TGA ATC TGC CAC C	AGA AGC AGC TCT GGC AGG
A386	CO901343	CAC CTC TTC CCT CAT CAG TC	CGA CAA AGG AGA CTG AGA GG
A388	CN494395	CAA TTT CTT CTT CGC CTT CC	TCT CAG CCG CGT GTC TCC
A389	CN581002	TGG AGG GAA AGG AGA AGC AG	CTT GGA AGC TTT CTG TCA GC
A390	CN544851	TTG TCG GAT TTG TAA CCC TAG	TTC CAT ATC AGT TTG GAC ACC
A392	CN445562	CAC AAA CCA ACC GTC TAA CC	GCT CTT GAT CAT AGG CGT GG
A393	CN490103	CCT CCG ATA AAC GTC AAT CG	GGC CAC CTT CGC CAC CG
A395	CN495393	TCC CAA GCT CCC AAC AAA CC	CTA TCT GGG TCG GCC AGG
A396	CN496160	CTT TGC TTA AGC CCC GTC G	TGA AGC GCT TCA CAT ACT GC
A397	CN491038	GCT CTG TCT CGT TGA TCG G	AGC TGC TTC ACC CTC TTG C
A398	CN490644	ATC TCA CAC CTC AGC AGT GA	CTT CTG CCC AAT TCA AGA CC
A399	CN494405	TGC TGA GCA TCG GTA TCT CC	GTG CTC ATC GGG GAT GAC G
A408	CN445331	CGA GAG CAC ACT CCC AAC C	GGG AGA GGA ACA CGG AGC
A412	CN492999	ATG AGA GAG AGC TAC CTC AC	GTA CAA GTT CAG CAG TGA CC
A414	CN489062	ACA ACT TGG TTA CGC GAC AC	GAA CAG ATT AGG GTC GCT GG
A416	CO168103	CTCA AAA CAA GAA CAA TGA GCC	CCC AAA AGG TTT TCC ACA CG

Primer nr	Accession nr	Forward primer	Reverse primer
A417	CV128959	AAA TAG TGT GGA AGA CGC GG	CAA TAT ACT AAT GAG TCC TTC G
A418	CV150384	ACA AAC CAC CAC CAA TTC CC	CCT GAG AGA GCC AAT TGA GC
A419	CO755991	AAT CTC TCG TCT GCA AAC CC	GTA TGA GTA TCC AGC ACC CG
A424	CO415353	ATG AAC AGT CAC AGA CTA TGC	AAC GAA GCA AAG GAA GAC GG
A425	CO756781	ATA AGT TTA GGC TCA TCT GCC	AAA CCC ATC CCA CTT AAG GC
A428	CO902639	CTC CTT TAT CTC TTT CCT CCC	TTG TCG TCC CAA ATC AAG CC
A430	CV656755	AAA CAG CAG AGT GTT GCA G	TCC ATC TAC CAG GAG AGG G
A440	CO416477	CCA CAC AAC ACA AAC CAA CC	TGT GGT CAT TTG GTG AGT CC
A452	CO900827	ACC TTG GTG GCC AAG TAG C	CTT GCG TAT CAA AGC TGC CG
A461	DT000945	AGT TGA CTA CCT CCT CCG C	GTA AGC GAT GAA ACT GAT GC
A466	DT040421	GGC AGA GCA GAT GCA GAT AA	TAT AAG ATG GAA GCC AAT GCC
A477	DT000551	CTA ACC CCA ACC ACC AAC C	GTT TGT CGA GCG TCA TTG TCC G
A494	DT001786	TTC TCT GTC TGT GAA ATT GCG	GTT AAC TGA GCT CCT GGT ATT CC
A497	DT041964	AGC TCT CAG ATT GAG GCC C	GTT TGC CGG AGA ATC GAG AAG G
A531	CN943946	GTC TAC TTC CAG AAC TTG CC	GAT CTC ACC ACA AAA TGC ACT

Marker A334 is also included in the list of newly developed polymorphic markers (Table 9), resulting in a total of 134 markers that are polymorphic. A334 is however designed from the same sequence as was used by Silfverberg-Dilworth *et al.* (2006) (CN444542) and although all other markers designed from sequences, or sequences within the same contig assembly, than published markers, were excluded from further analysis this marker was retained due to the fact that the accession forms part of the proposed 15 cM reference map (Silfverberg-Dilworth, *et al.*, 2006). When implementing A334 and the 133 newly developed markers, reported to be polymorphic, on the 4 cultivars used in this study (Table 10), 119 marker pairs yielded heterozygous amplification products in one or more of these cultivars, rendering them ideal to be used for the purposes of genetic linkage map construction. The remaining 15 marker pairs, although not heterozygous on any of the 4

parental cultivars used in this study, are either heterozygous in other cultivars frequently used as parentals in local breeding programs or can be used for cultivar identification studies as the single homozygous allele observed for each cultivar differs from that observed for another. One marker, A208, yielded fragments larger than 500 bp when used on automated systems. This marker was therefore not accurately scoreable on parental cultivars and will also not be accurately scoreable on a mapping population. Redesigning of the unlabelled marker closer to the repeat sequence will result in a shorter, scoreable fragment and will then render this marker useable for implementation on a mapping population for the purpose of genetic linkage map construction.

Table 10. Fragment sizes obtained after the implementation of newly developed polymorphic markers on four cultivars used as parentals during this study.

Primer	Anna	Golden Delicious	Prima	Sharpe's Early
A29	86	94	86	86
A30	175-181	175-179	nd	nd
A180	405	405	353-405	405
A182	478-494	483-494	nd	nd
A183	275	275-293	274	nd
A184	236-239	239	239-242	239
A186	200	200-206	200	200
A188	192	190	193-211	193
A192	282-284	282-284	284	284
A193	398	398-471 or 471-481	471-481	nd
A195	175-181	175	175-184	175-181
A196	306-319	306-319	306-319	306-319
A197	nd	nd	222	139-222
A200	nd	nd	359-405	359-405
A202	314	314	371	nd
A204	226	209-219	209-223	209-215
A207	±564	±564	±564	±564
A208	1 > 500	1 > 500	1 > 500	2 > 500
A209	145-148	145-148	148	145-155
A212	nd	nd	303-334-349	nd

Primer	Anna	Golden Delicious	Prima	Sharpe's Early
A213	468	nd	nd	479
A217	280-282-288	280-282-288	nd	280-282-288
A219	380	377-380	376-382	333-379
A221	106-111-118	106-115-118	106-118	106-115-118
A222	172-181	172-184	173-187	173-184
A225	311-327	396-401	nd	nd
A227	345	345	345	345-385
A229	nd	167	167	167-171
A230	441	441	441	441
A232	169	169	169-176	169
A233	319	297-319	297-319	nd
A234	338-345	345-352	341-352	341-354
A235	346-348	356	340-343	339
A236	235-241	232-241	325-244	220-241
A238	340-345	334-340	340	340-345
A242	nd	nd	178	182
A244	178-181	181-184	185-188	181-187
A245	165-174	174	174	174
A247	256	243-256	256-265	256
A249	143-294-297	143-294	143-294	143-294
A253	240-264-279	240-260-265-279	250-265-279	250-265-279
A254	334	291-334	291-334	nd
A256	186	180	nd	nd
A259	114-120-138	114-120-138	114-120-138	114-120-138
A262	200-214	200-214	200-214	200-214
A266	248-301-307	248	246	248
A267	426	426	426	426-474
A268	194-198	170-194	174-194	nd
A269	284-326	285-326	284-326	nd
A274	303	298-303	303-307	300-303
A277	195-221	195-221-225	195-223	nd
A279	207	207-214	207-214	214-221
A281	300-302	300-302	nd	300-302
A283	366	374	366	366-374
A284	192	192	nd	189-192
A285	200	197	nd	203
A286	401-427	nd	nd	410-427
A290	362	334-362	362	346-362
A293	374-394	372-378-394-429	372-378-394-429	375-391
A294	105-242	105-242	nd	105-242
A296	427-430	406	nd	406-427
A298	172-194	197	196-198	nd
A300	302	302	302	302
A301	340-345	331-337	nd	331-337
A304	344-350-367-418	344-350-367-418	344-350-367-418	344-350-367-418
A307	338	338-349	338-348	338
A308	273	273	nd	273
A310	228-243	223-241	240-250	229
A311	244-249	232-237	232	232-244
A314	148-175	150-173	148-185	148-175
A315	210-222	210	210	210
A318	273	273	269	273
A319	330	nd	nd	330

Primer	Anna	Golden Delicious	Prima	Sharpe's Early
A320	418	406-418	406	416
A326	160	149-162	149-164	149-162
A329	208-213	208	208-213	208
A331	266	272	272	266-272
A332	127-131-171	131-142-171	142-160-171	127-142-171
A334	190-215	190-209-225	190-215	190-205-215
A335	188	188-196	196	188-196
A336	474	nd	474	386-474
A337	239	235-241	239-243	235-239
A339	420-430	425	430-438	420-430
A340	120-134	120-130	130	120-129
A341	142 or 152-157	147-151	147-151	152-157
A343	266-256	262	262	226-264
A344	407	407-427	427-435	407-435
A352	184-186	174-186	174-186	186-198
A355	230	222-226	222-226	232
A363	227	227	225-227	217-227
A365	229-250	242-244	244-248	244-248
A367	214-228	208-232	nd	204
A368	356-434	356-434	356-434	356-434
A369	160-168	160-164	nd	168
A372	232-234	236	232-236	232
A376	365-440	365-440	nd	365-440
A377	455	401-447	435-447	415-455
A378	273	273-296	273-284	273-288
A379	134-138	120	134-138	120-134-138
A380	239	239	239-253	239-251
A381	233 or 233-235	218	214-218	218-228
A383	344	344-350	344-356	342-344
A385	170-174	170-190	nd	174-190
A386	214-227	208-233	nd	227-231
A388	323	332	323-332	323
A389	253-267	241-253	nd	nd
A390	228-250	242	nd	244-248
A392	139-154	150-154	150-170	139-154
A393	135-143-162	135-158-162	135-158-162	135-143-158
A395	200	219	nd	200-210
A396	144	144-151	nd	144-151
A397	498-510	510	502	498-510
A398	262-266	262-266	nd	262
A399	216-230	216-230	nd	216-230
A401	301-303	303-305	301-303	301-303
A408	458-500	500	500	488-500
A412	215	215	nd	215-219
A414	296	282	294	282-296
A416	192	192	192	187-192
A417	240	232-242	nd	nd
A418	235-243	235	nd	235-248
A419	154	150-154	nd	150-154
A424	329-333	329-333	329-333	329-333
A428	343	343	343	343-348
A430	216-218	221	nd	nd
A440	218	218-224	nd	218

Primer	Anna	Golden Delicious	Prima	Sharpe's Early
A452	443	443	443	443
A461	368-433	368	368	368
A466	325-345	325	325-338-345	325-339-347
A477	154-156	154	nd	nd
A494	143-147	141	nd	nd
A497	168-180	168	nd	nd
A531	327-337	327-341	327-341	288-316-337

Primer pairs amplifying more than two fragments per individual were encountered and are probably multilocus markers where the primers anneal to more than one site. The complexity of these multilocus markers depend largely on the number of loci amplified (two or more) as well as the difference in fragment sizes obtained from the two (or more) different loci. The ease with which these markers can be used will only become clear during segregation analysis of alleles in mapping populations in preparation for genetic linkage map construction,

Size variations, usually where the fragments obtained are much larger than expected, were observed in the amplification products of EST derived SSRs. This was probably due to the presence of introns in genomic DNA. In some cases these size variations resulted in fragments much larger than the 500 bp that can be analyzed using automated fragment detection systems, and the primers had to be redesigned to regions closer to the repeat motif. Amplification products smaller than the expected size could be explained by small deletions and non-specific primer annealing.

3.4.2.2 Multiplex

Several multiplexed PCR reactions (Table 11), consisting of newly developed polymorphic SSRs and previously published SSRs, have been

optimized. Obtaining successful co-amplification with well-balanced PCR product yields, minimal primer dimers appearing and the absence of non-specific amplicons, sometimes required extreme optimization experiments adjusting primer concentrations. Multiplexing was done using primers labeled with the same fluorescent dyes and producing non-overlapping PCR product sizes.

Table 11. PCR primer pairs included in 24 multiplexed reactions. Equal volumes of each primer were used, unless otherwise indicated. *One multiplex included a pear primer. **3 primer pairs did not form part of the newly developed set of 322, as they have been developed from sequences also used by Silfverberg-Dilworth et al. (A307 - CN445290; A329 - CN496002; A319 - AF527800).

Multiplex	Primer	Size range	Ratio	Multiplex	Primer	Size range	Ratio
F1	CH02d08	210–254		N6	CH05e05	127-160	
	CH03h03	72-120			CH02c09	229-258	
	CH02g04	132-198			CH05d11	167-211	
F2	CH05g11	201-259			A200	318-405	
	CH02b03b1	74-109		N7	CH01h01$_1$	97-134	1
	CHo1f09$_1$	114-160			CH02e10	128-177	2.7
	CH04g04	168-186			CH01d08	237-290	2.7
F3	NZ05g08	118-125			A234	335-354	2.8
	Cola	218-256			A452	427-443	2.7
	Ch04e02	143-163		N8	NZ28f4	90-110	1
	CH02d12$_1$	177-230			MS02a01	168-194	1.4
	A283	366-374			A279	207-221	2.18
F4	CH05e06	125–222			A293	372-429	1.68
	CH01h02$_1$	236–256			A398	261-371	1.25
	CH03d01	95-115		N9	CH04d02	106-164	1
	A193	348-471			A227	337-385	1.5
N1	CH03d07	163-226			A318	269-273	1.5
	CH01h10$_1$	88-120			A329**	208-229	1
	CH04g10	124-168			A335	118-196	1
N10	A300	296-308		P1	CH01c09	87-108	
	A307**	338-349			CH01e09b	118–140	
	A310	217-248			CH02h07	214-240	
	A416	187-196			CH05a09	150-200	
	NH009b*	134-166		P7	CH01b07	91-121	
N11	A180	353-405	1.05		A235	339-356	
	A186	196-200	1		A352	174-199	
	A266	248-308	1		A369	159-168	
	A340	120-134	1.26	P8	CH01b12$_1$	122-178	
	A372	232-238	1		CH02a04	66-112	
	A424	329-333	1		A344	407-435	
N12	A196	306-335			A417	235-261	
	A213	468-479		P9	CH04g12	141-186	
	A222	173-187			A236	201-245	
	A377	383-455			A341	142-151	
N2	CH01c06	146–188			A425	309-346	
	CH05d02	194-241		V1	CH02c06$_1$	216–254	
	CH04c07	85-135			CH02c02b	78-137	
N3	CH02g09	98-144			CH01f03b	109-190	
	CH05e03	158-190			A319**	330-343	
	CH02d10a	215-242		V2	CH01f02$_1$	174–206	
N4	CH05g03	132-192			CH02f06$_1$	134-164	
	CH02c11	194-239			CH03a04	92–124	
	CH03b10	96-121			A381	214-235	
N5	NZ02b1	216-240	1.33	V3	MS01a03	209-253	
	CH05c07	107-149	1		MS06g03	138-190	
	CH04f06	159-186	1.44		CH02b12$_1$	101-143	
	A253	240-279	1.44		A215	182-207	
	A376	365-440	1.44	V4	CH01g05	134-188	
					A202	278-371	

High quality primers are essential for successful multiplex amplification reactions. 'Dye blobs' can occur when fluorescent dyes are no longer linked to their respective primers. The migration of free dye in the capillary during electrophoresis can interfere with the detection of true alleles present from other loci amplified in a multiplex.

Multiplexing is the only efficient way to reduce the cost for utilizing the markers for the construction of genetic linkage maps. A reduction in the amplification success of individual primer pairs was observed as the number of markers in the multiplex increased, as was the case in a study conducted by Lin *et al.* (1996). The scale of multiplex amplification has been largely limited due to the selection of non-overlapping fragment size distributions, as well as primer-primer interaction when multiple sets of primers are present in the same reaction mixture. The success of multiplexing depends on the principle that primers should have comparable annealing temperatures and that the primer sequences should not contain excessive regions of complementarity (Butler, *et al.*, 2001), which can lead to primers binding to each other rather than to template DNA leading to the formation of primer-dimers (Schoske, *et al.*, 2003).

3.4.3 Fragment detection

3.4.3.1 Gel electrophoresis

Polyacrylamide gel electrophoresis was successfully used for the initial steps of primer testing in order to determine optimal PCR conditions. Gel electrophoresis were also used throughout the study in order to verify the

presence of amplification products, after simplex and multiplex (Figure 20) reactions, before making use of much more expensive automated systems for accurate fragment size determination.

Figure 20. Multiplex consisting of 3 primers (labeled with the same fluorescent dye colour) separated on 6 % polyacrylamide gel and visualized using silver staining.

3.4.3.2 Automated fragment analysis

Traditional methods used for analysis of PCR products, e.g. agarose or polyacrylamide gel electrophoresis, has been replaced with capillary electrophoresis (CE) with advantages including high resolution, high throughput, automatic operation and on-line detection with automated data acquisition. CE permitted an evaluation of the primer to product ratio, which is an indication of primer efficiency. Unspecific peaks at the beginning of an

electropherogram are an indication that a large amount of primer is remaining after the PCR amplification, indicating that the PCR conditions has not been well optimized.

Although CE can be used for the analysis of a single marker (simplex) (Figure 21) at a time, or for the analysis of a multiplexed reaction (Figure 22), the most cost effective approach is the 'pooling' of different multiplexed reactions (Figure 23) before CE. Multiplexes were pooled in the ration 6-FAM™:VIC™:NED™:PET™ = 1:1:3:2 in order to adjust for differences in fluorescent signal strength.

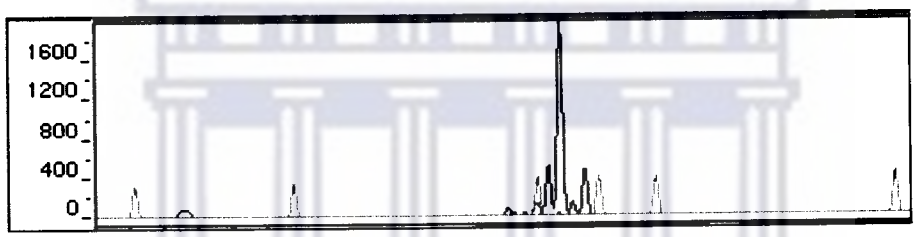


Figure 21. Electropherogram obtained after performing a simplex reaction, using primer pair A494, on the apple cultivar 'Anna'.

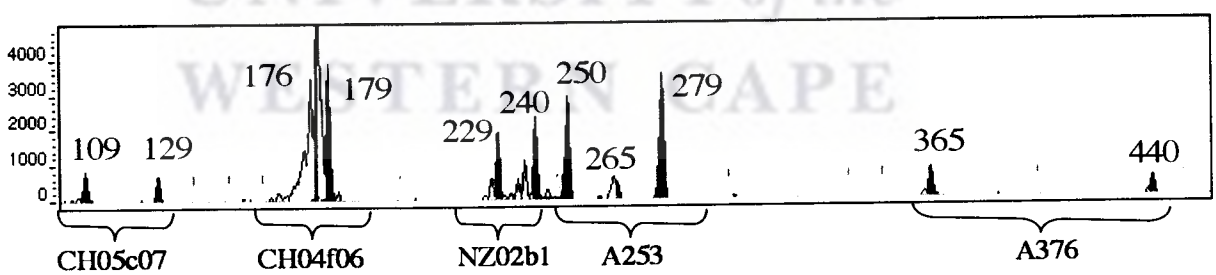


Figure 22. Electropherogram obtained using the multiplex N5 on the cultivar 'Sharpe's Early'

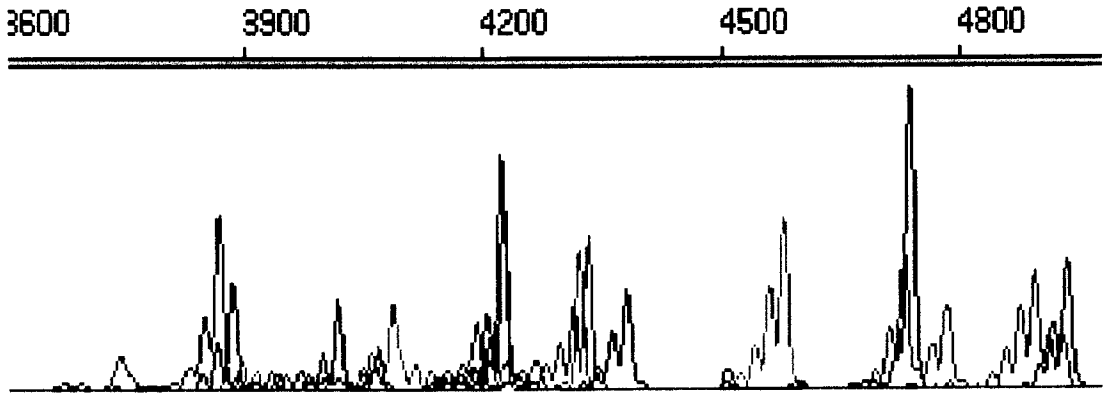


Figure 23. An electropherogram obtained after the pooling of four multiplexes, labeled with four different fluorescent dyes.

3.4.4 Primer modification

Exact fragment size determination, even on an automated DNA detection system such as the ABI genetic analyzers, is sometimes difficult especially when analyzing di-nucleotide repeats. It has been reported that the addition of a 'pig-tail' sequence (GTTT) to the 5' end of the reverse primer facilitates genotyping (Brownstein, *et al.*, 1996). The addition of such a sequence to the reverse primer of three published primer pairs for the amplification of dinucleotide repeats had no positive or significant influence on yield or stutter (Figure 24).

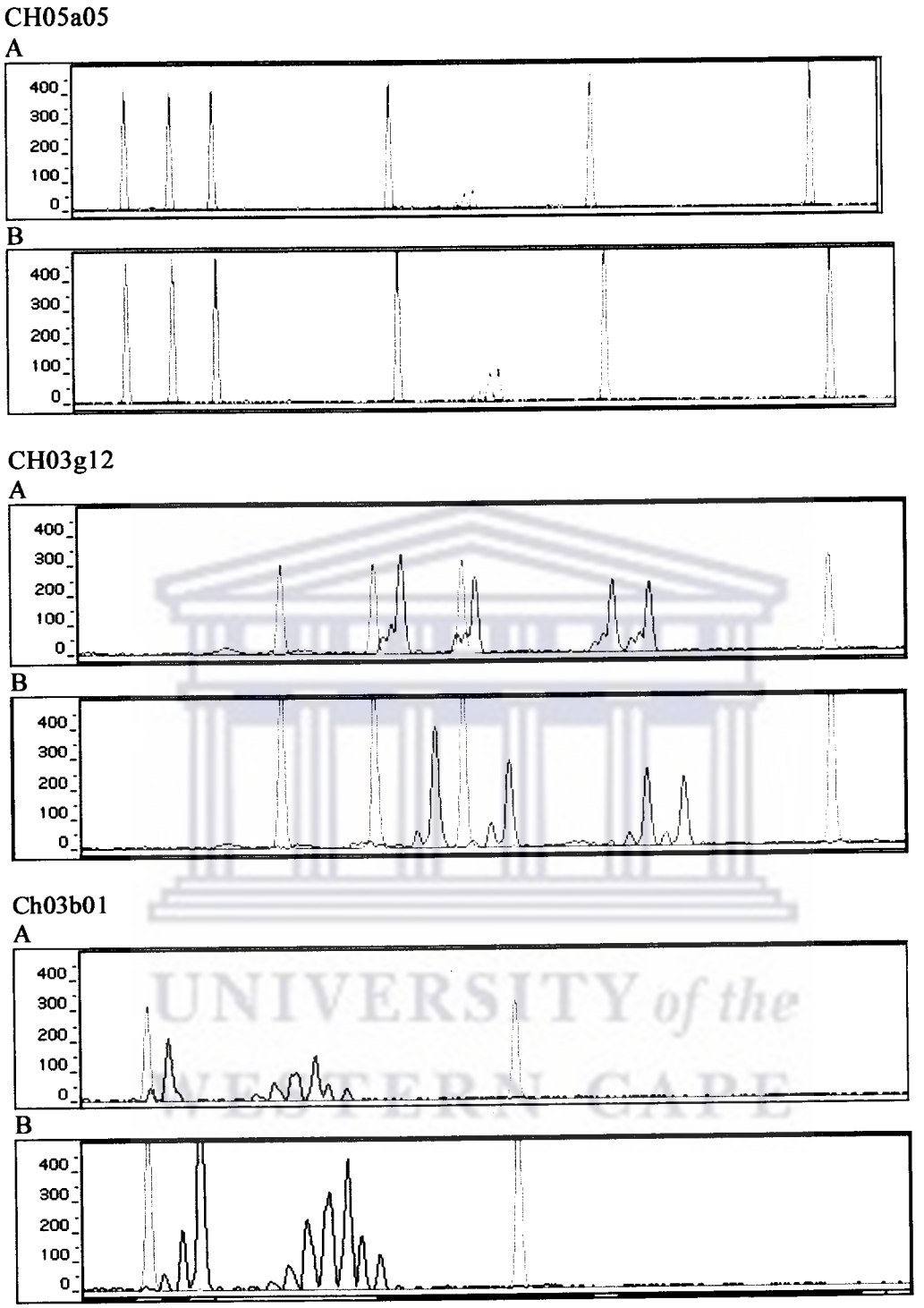


Figure 24. Electropherograms illustrating amplification results obtained using (A) primer pairs without any modification and (B) primer pairs where a 'pig-tail' has been added to the 5' end of the reverse primer. Results are shown for three published SSR primer pairs, CH05a05, CH03g12 and CH03b01 (Liebhard et al., 2002).

The stutter observed in electropherograms (Figures 21-24) is most likely caused by the 'slippage' of the polymerase, leading to products that differ by approximately 1-5 repeat units from the expected product and are usually less intense than the desired product. On an electropherogram obtained during electrokinetic injection the stutter normally appear to the left of the actual peak. If the products of heterozygous individuals overlap it is sometimes difficult to differentiate between 'true' and 'slippage' products, especially when considering heterozygotes for dinucleotide repeats where the two alleles differ with one repeat unit only (e.g. 150 bp and 152 bp). When studying the segregation of the alleles in the progeny the presence of the alleles can easily be confirmed, but the wrong interpretation might be detrimental in cultivar identification studies. The increase in the fragment sizes obtained when using 'pig-tailed' primers is due to the addition of 4 bp to the reverse primer sequence and the subsequent 4bp addition to the amplification product.

3.4.5 Database management

An interactive and highly informative database (Figure 25) (Figure 26) (Figure 27) (Figure 28) (Figure 29) was created, using the 'FileMaker Pro 8.5v1' software package (http://www.filemaker.com), in order to facilitate access to all aspects of primer design, as well as to record and share progress made in the testing and implementation of markers.

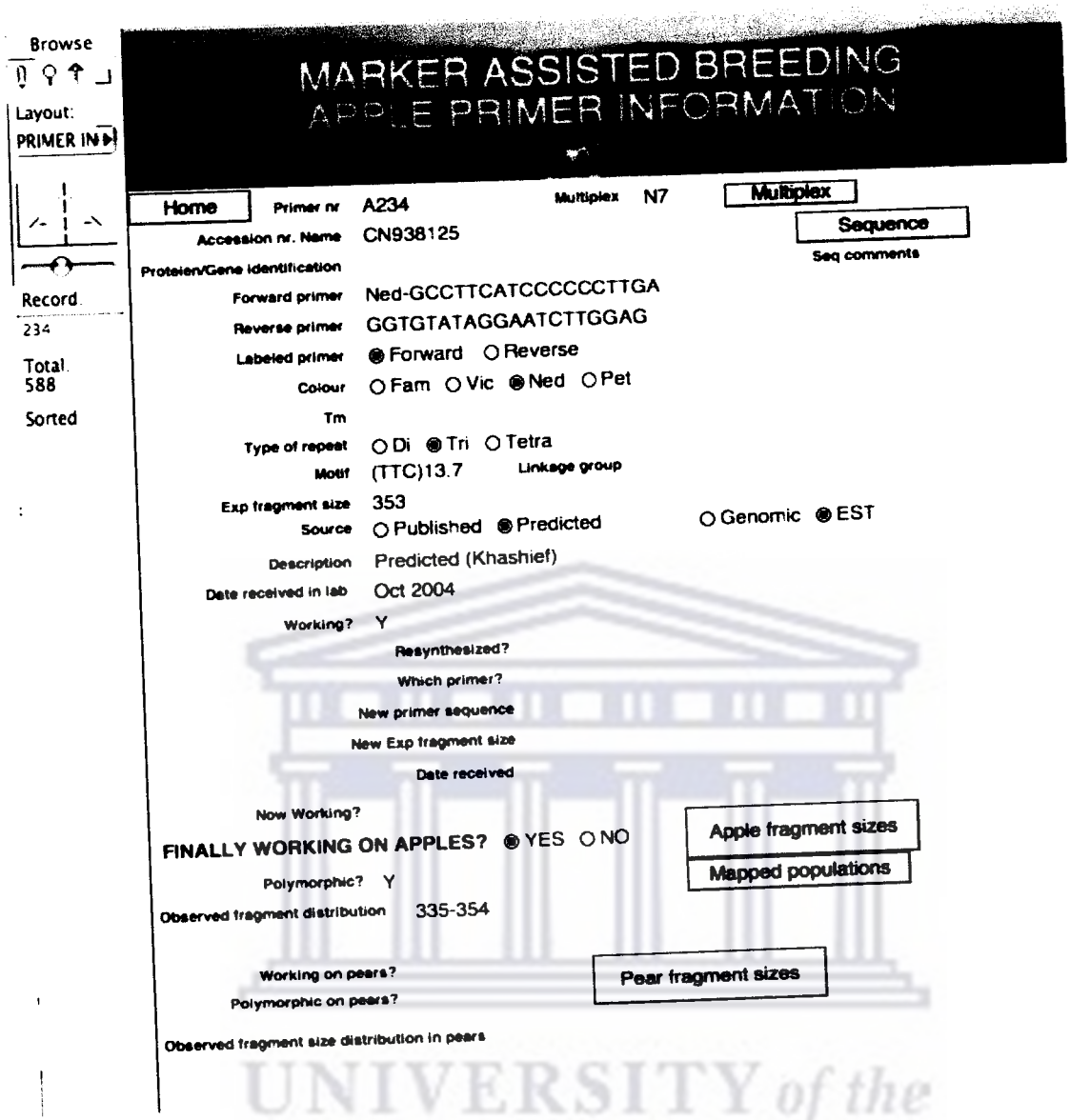


Figure 25. User interface of the database developed using FileMaker Pro 8.5v1. General information regarding the specific primer pair selected is given, with buttons linking to additional information.

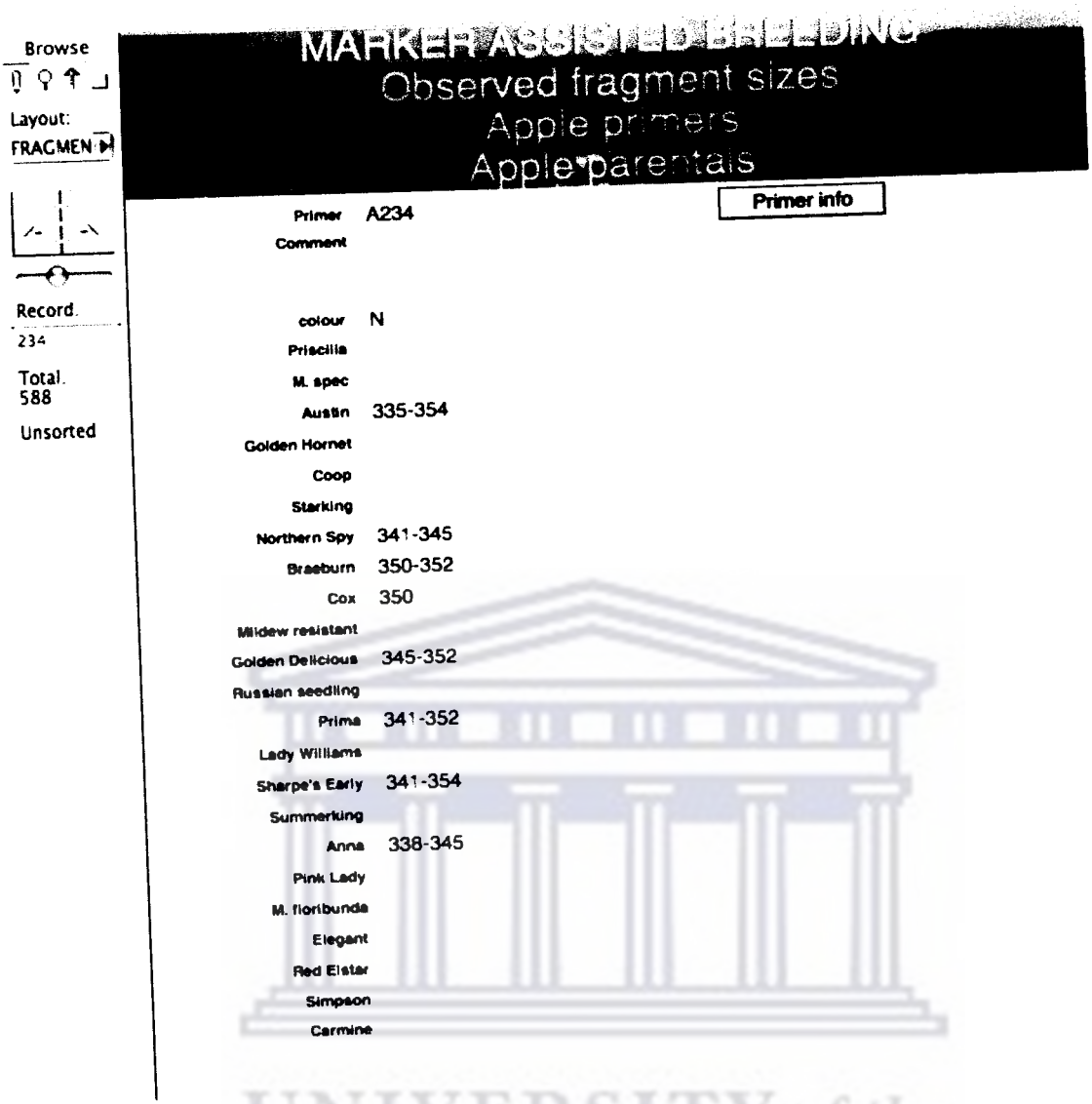


Figure 26. Fragment sizes obtained when a specific primer (A234) is screened across a selection of apple cultivars.

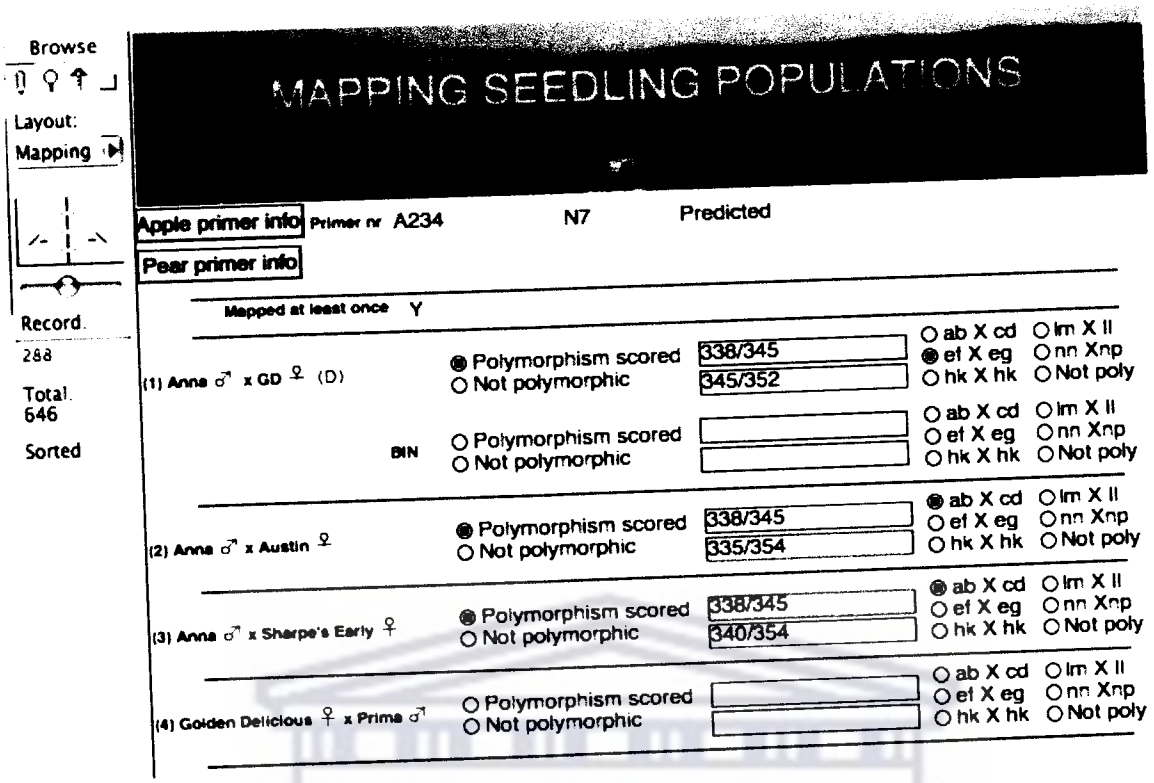


Figure 27. Information regarding the implementation of a specific primer pair (A234) across a selection of different mapping populations used in the apple breeding program.

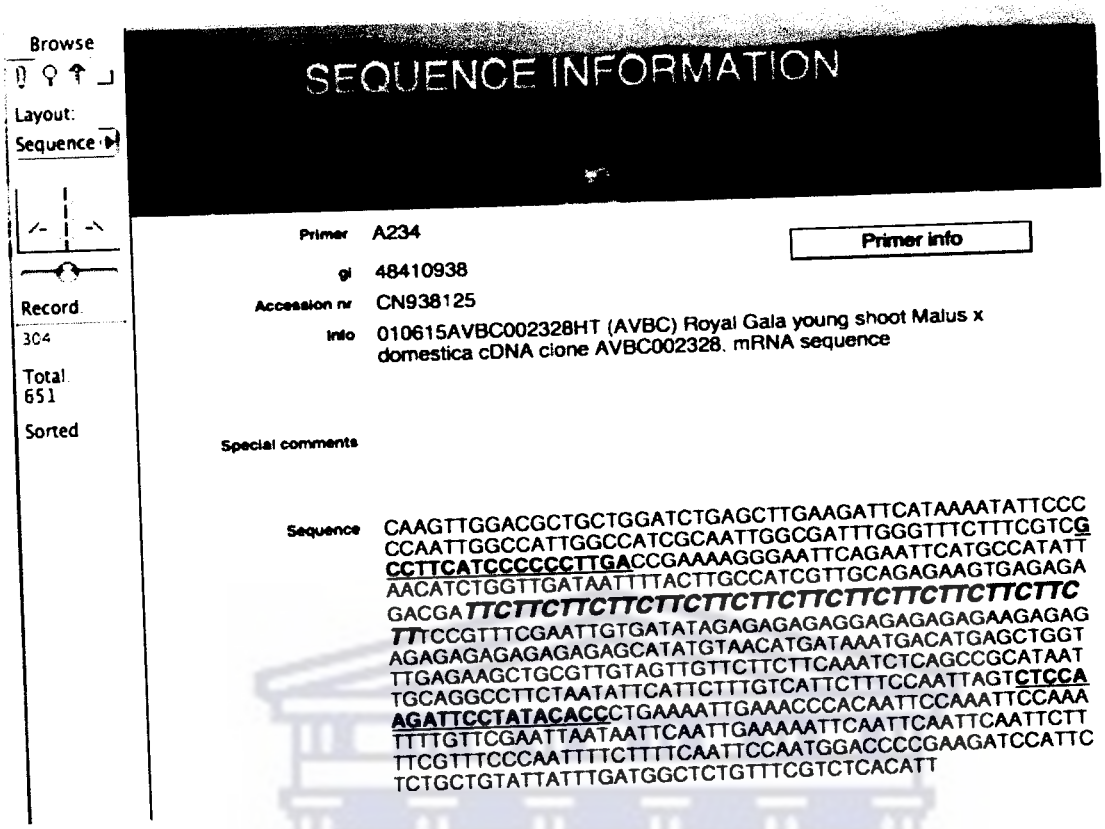


Figure 28. Information regarding the sequence data that were used for the identification of a SSR and the position of the primers relative to the SSR.

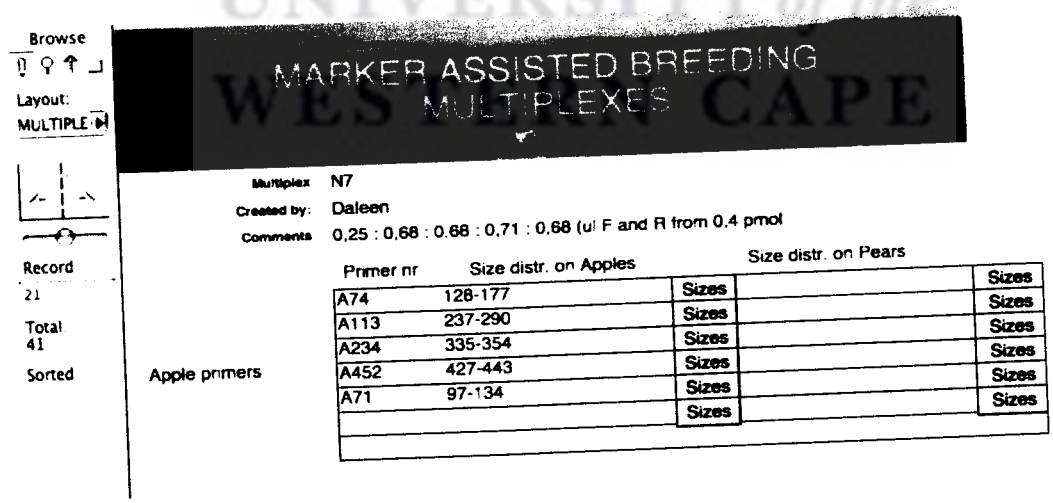


Figure 29. A selection of primer pairs have also been used for the development of multiplexed reaction as indicated by this NED multiplex containing primer pairs A71, A74, A113, A234 and A452.

3.5 CONCLUSION

The identification and testing of new SSR markers has resulted in a high level of amplification success (83 %) as well as a high level of polymorphism detection (82 %). A total of 119 of these polymorphic markers had some level of heterozygosity when tested on the four cultivars used as parentals during this study and can for this reason be used for the purposes of genetic linkage map construction.

Multiplexing of primer pairs labeled with the same fluorescent dye colour, although done successfully for up to 6 markers, mostly involved 3 to 4 markers and was largely limited by the size distribution of individual markers. This, together with the ability to pool different multiplexed reaction before electrokinetic injection, will enable a much more cost effective way of analyzing the segregation of a large number of markers on mapping populations for the purpose of genetic linkage map construction.

Management of large amounts of data regarding all aspects of primer design, amplification results and subsequent steps involving segregation analysis and positioning on genetic linkage maps, have been performed successfully with the database designed utilizing the software 'FileMaker Pro 8.5v1' (Http://www.filemaker.com).

4 CHAPTER 4: MARKER IMPLEMENTATION, SEGREGATION ANALYSIS AND GENETIC LINKAGE MAP CONSTRUCTION

4.1 ABSTRACT

Segregation analysis was performed and genetic linkage maps were constructed using a subset of the available published and newly developed SSR markers implemented on three F1 mapping pedigrees derived from interspecific hybridization of four parental cultivars. 'Anna' was used as a common male parent in crossings with 'Golden Delicious' and 'Sharpe's Early' respectively and in a third population 'Golden Delicious' was used as a female parent in a cross made with 'Prima'. The use of published and previously mapped, highly transferable SSR markers enabled alignment of the obtained LGs with that of the apple reference map (Silfverberg-Dilworth, *et al.*, 2006). All 17 LGs were represented in maps obtained from the mapping pedigree derived from 'Anna' and 'Golden Delicious' while some LGs were absent from maps obtained from the other two mapping pedigrees. The absence of LGs can mostly be ascribed to the low frequency of SSR markers already mapped onto these LGs together with the fact that previously published and mapped SSR markers might be homozygous in the cultivars used as parents. The LGs, or segments thereof, that are available can be successfully used for the

identification of regions containing putative QTLs, the ultimate goal for genetic linkage map construction during this study.

4.2 INTRODUCTION

Seedlings derived from 3 mapping pedigrees (chapter 2) were genotyped using published (Guilford, *et al.*, 1997, Liebhard, *et al.*, 2002, Liebhard, *et al.*, 2003b, Maliepaard, *et al.*, 1998, Silfverberg-Dilworth, *et al.*, 2006, Yamamoto, *et al.*, 2002a, Yamamoto, *et al.*, 2002b) as well as newly developed (chapter 3) SSR markers, that have been shown to be heterozygous in either one, or both, of the parental cultivars (chapter 3). The segregation of the two parental alleles at any given locus, is studied in the progeny, ultimately providing the basis for the construction of a genetic linkage map.

Genetic linkage map construction is far more complex in full-sib families of an outcrossing species than in species derived from pure lines. This can be explained by the following problems stated by Maliepaard *et al.* (1997): 1) markers may differ with regard to the number of segregating alleles (2, 3 or 4), 2) markers may not be heterozygous in both parents, 3) markers that are heterozygous in both parents may be identical, 4) markers may have null alleles and 5) the linkage phase, or grandparental origin, of markers are, in most cases, unknown.

The five segregation types encountered, when working with an outbreeding species, like apple, is best explained by means of genotypic

codes used by the mapping software JoinMap® (Van Ooijen, 2006) (Table 12).

Table 12. Different classes of segregation types encountered when working with a full-sib family, derived from an outbreeding species, as described by JoinMap® 4 codes.

Class	Segregation type	Number of alleles	Segregating alleles		F1	
			Parent 1	Parent 2	genotypic codes	Expected ratio
1	ab x cd	4	Yes	Yes	ac; ad; bc; bd	1:1:1:1
	ef x eg	3	Yes	Yes	ee; ef; eg	2:1:1
2	hk x hk	2	Yes	Yes	hh; hk; kk	1:2:1
3	nn x np	2 or 3	No	Yes	nn; np	1:1
	lm x ll	2 or 3	Yes	No	lm; ll	1:1

Class 1 segregation involves those loci that are heterozygous in both parents and, from a mapping perspective, are fully informative for both parental meiosis. Class 2 segregation includes those loci that, although heterozygous in both parents, are heterozygous for exactly the same two alleles, making the determination of the parental origin of alleles present in a heterozygous seedling impossible when viewing segregation from a single parent. These loci can however be statistically positioned on a genetic linkage map derived from the F1 population where both parental meiosis are viewed simultaneously. Class 3 segregation refers to those loci that are heterozygous in one parent only. The other parent might be homozygous for a corresponding allele or any other allele, thus resulting in the presence of either 2 or 3 alleles. For mapping purposes the allele from the homozygous parent is

uninformative. Although information from all 3 classes of loci can be used for the construction of a F1 genetic linkage map, only class 1 and 3 can be used for the construction of parental genetic linkage maps. It would be expected that the 2 alleles at any given locus will be equally represented in the progeny with a segregation ratio of 1:1. Any deviation from this expected ratio results in segregation distortion, possibly due to sample error when working with a small population size, but also possibly through the occurrence of natural selection against a given allele or a genotype that it occurs in linkage disequilibrium with.

Since the grandparental origin, also known as the linkage phase, of alleles forms the basis of linkage analysis, the first step in map construction is the determination of inheritance vectors (0 or 1) as described by Jansen (2005). JoinMap® 4 (Van Ooijen, 2006), the software used for the construction of genetic linkage maps during this investigation, is able to handle all the computational difficulties related to outbreeding species and assign inheritance vectors to markers in such a way that the number of recombinations between adjacent markers are minimized.

4.3 MATERIALS AND METHODS

4.3.1 Mapping populations

Three mapping pedigrees were founded through the interspecific hybridization of four cultivars, viz. 'Anna' x 'Golden Delicious', 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious'. 'Anna' was used as a

common male parent in the first two crosses and 'Golden Delicious' as a common female parent in crosses 1 and 3.

Leaf material was collected from seedlings obtained from each of these controlled crosses (Table 13) and DNA isolation was done using the CTAB method described by Doyle and Doyle (1990) with the addition of PVP (Kim, *et al.*, 1997) in order to bind secondary plant products such as polyphenolics.

Table 13. Numbers of seedlings from each mapping pedigree used for the construction of genetic linkage maps

Mapping pedigree	Number of seedlings
'Anna' x 'Golden Delicious'	87
'Anna' x 'Sharpe's Early'	94
'Golden Delicious' x 'Prima'	92

4.3.2 Markers used for segregation analysis

Markers forming part of the proposed 15 cM reference map (Silfverberg-Dilworth, *et al.*, 2006), as well as other published (Guilford, *et al.*, 1997, Liebhard, *et al.*, 2002, Silfverberg-Dilworth, *et al.*, 2006, Yamamoto, *et al.*, 2002a, Yamamoto, *et al.*, 2002b) and newly developed SSR markers (Chapter 3) were used to screen the seedlings derived from the 3 controlled crosses (Table 13). Although not all makers were used on all 3 crosses due to a lack of polymorphism detection and/or a lack in time and cost involved, a number of these markers have been genotyped on all three, or in some cases two out of three, mapping populations in order to facilitate the alignment of LGs with each other and with that of the 15 cM reference map.

4.3.3 Fragment analysis

Automated DNA detection was used for segregation analysis of all markers on mapping populations. GeneScan® software (Applied Biosystems, Foster City CA, USA) was used for the collection of data after CE. Genescan® output files were analyzed using Genotyper® 2.5.2 software (Applied Biosystems, Foster City CA, USA). Seedlings were genotyped for each specific locus, using the JoinMap® 4 (Van Ooijen, 2006) coding system (Table 12), according to preset criteria regarding fragment size and intensity.

4.3.4 Genetic Linkage Map Construction

Tables containing seedling genotypes, as inferred by the fragment detection process described above, were exported to Excel (Microsoft Office) and with minor modification imported directly into JoinMap® 4 (Van Ooijen, 2006) for the construction of genetic linkage maps.

In the case of all three mapping populations, 'Anna' x 'Golden Delicious', 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious', seedlings that had missing data points at 25 % of loci tested were removed from any further analysis. Markers that, after the removal of 'weak' seedlings, had missing data observations at 40 % of the remaining seedlings were excluded from the initial LG determination process. LGs were defined using recombination frequencies observed between marker pairs. A threshold value of 0.2 was generally used in order to assign markers to the same LG. This value was in some cases increased to 0.25 in order to enable the grouping of reference markers, known to reside on the same LG, to the same group. After

LG determination, markers that were excluded, as well as markers that did not show sufficient linkage with any other marker/s, were assigned to LGs based on Strongest Cross Linked information (SCL values), a new feature offered in the JoinMap® 4 (Van Ooijen, 2006) software package.

Calculation of marker order and distances between markers were done separately for each group, a point during which some of the markers may have been excluded as their presence lead to 'insufficient linkage' between the markers within a certain group, or their incorporation into the LG resulted in a high 'mean chi-square value' indicating that a lot of double crossover events occur that's highly unlikely, as crossovers are generally limited to one or two per chromosome during meiosis. Regression mapping and Kosambi's mapping algorithm (Kosambi, 1994) were used for the determination of marker order within each group.

Integrated genetic linkage maps were constructed for the F1 progeny derived from each cross, as well as separate parental maps. Graphical representation of genetic linkage maps as well as the alignment of different maps were achieved with the use of MapChart© (Voorrips, 2002).

4.4 RESULTS AND DISCUSSION

4.4.1 Marker implementation and segregation analysis

Published as well as newly developed SSR markers were implemented on all three mapping populations (Table 14). The proposed 15 cM framework map (Silfverberg-Dilworth, *et al.*, 2006) consists of 85 SSR markers amplifying a total of 89 loci (4 markers amplifying 2 loci each). The first focus was the implementation of these reference markers on the mapping populations in order to get a framework map. Segregation analysis has been successfully done using 68 framework markers on the three mapping populations (between 1 and 3 populations each). Primer sequences for two of these markers, AG11 and HBO3AT, are not publicly available and 15 were either not polymorphic on the mapping pedigrees used (CH03d07, CH04e05), were polymorphic but not implemented on the full mapping population (CH04e03, CH05f04, CH02a03 (heterozygous on 'Sharpe's Early' only), Hi16d02 (heterozygous on 'Sharpe's Early' only), Hi04a08, Hi04e04), were difficult to analyze ($CH02d11_1$, CH05c06), did not form amplification products (CN493139-SSR, NH029a) or have not been subjected to testing yet (AU223657-SSR, U78949-SSR, AT000174-SSR).

Table 14. Primer and accession number (newly developed markers) and primer names (published markers) of markers included in the 15 cM reference map and markers implemented on three mapping populations. Observed fragment sizes and segregation types are given in cases where markers have been used for segregation analysis.

Primer nr	Accession nr. or primer name	15 cM Framework	'Anna' X 'Golden Delicious'	'Anna' X 'Sharpe's Early'	'Prima' x 'Golden Delicious'	'Anna' X 'Golden Delicious'			'Anna' X 'Sharpe's Early'			'Prima' x 'Golden Delicious'		
						'Anna'	'Golden Delicious'	Seg. type	'Anna'	'Sharpe's Early'	Seg type	'Golden Delicious'	'Prima'	Seg type
	AG11	X												
	AT000174-SSR	X												
	AU223657-SSR	X												
	CH01c06	X	X	X		159/161	155/161	ef X eg	158/160	160	lm X ll			
	CH01d03	X	X	X		138/142	159/-	ab X cd	137/142	144	lm X ll			
	CH01d08	X	X	X		253	249/270	nn Xnp	253	238/249	nn Xnp			
	CH01f03b	X	X	X	X	138	138/171	nn Xnp	159	157/171	nn Xnp	138/171	138/160	ef X eg
	CH01f091	X			X							120	136	hk X hk
	CH01g05	X	X		X	153/155	137/143	ab X cd				137/144	137/155	ef X eg
	CH01g12$_1$	X	X	X	X	105/128	101/143	ab X cd	105/128	103/130	ab X cd	101/142	105/147	ab X cd
	CH01h011	X	X	X		102/118	114/116	ab X cd	102/118	116/128	ab X cd			
	CH01h021	X		X					237	237/245	nn Xnp			
	CH01h101	X	X	X		89/99	99/107	ef X eg	88/99	88/99	hk X hk			
	CH02a03	X												
	CH02b07	X	X			104/110	104/112	ef X eg						
	CH02c02b	X	X	X	X	109/121	114/121	ef X eg	110/119	110	lm X ll	114/121	109/114	ef X eg
	CH02c09	X	X	X		240/248	240/255	ef X eg	241/247	229/245	ab X cd			
	CH02c11	X	X			225/233	217/231	ab X cd						

Primer nr	Accession nr. or primer name	15 cM Framework	'Anna' X 'Golden Delicious'	'Anna' X 'Sharpe's Early'	'Prima' x 'Golden Delicious'	'Anna' X 'Golden Delicious'			'Anna' X 'Sharpe's Early'			'Prima' x 'Golden Delicious'		
						'Anna'	'Golden Delicious'	Seg. type	'Anna'	'Sharpe's Early'	Seg type	'Golden Delicious'	'Prima'	Seg type
	CH02d08	X	X	X	X	226	224/226	nn Xnp	228	215/223	nn Xnp	224/226	226	lm X ll
	CH02d11$_1$	X												
	CH02f061	X	X	X	X	136/158	142/158	ef X eg	136/158	136/145	ef X eg	142/157	157	lm X ll
	CH02g01	X	X			199/229	199/219	ef X eg						
	CH03a09	X	X	X	X	122/130	126/130	ef X eg	122/130	126/130	ef X eg	126/130	126/130	hk X hk
	CH03b06	X		X					112	112/118	nn Xnp			
	CH03c02	X		X					122/124	122/124	hk X hk			
	CH03d01	X	X	X		96/109	96/109	hk X hk	96/109	96/109	hk X hk			
	CH03d07	X												
	CH03d11	X	X			125/142	117	lm X ll						
	CH03d12	X	X	X		111/119	119	lm X ll	111/119	117	lm X ll			
	CH03e03	X	X	X	X	216/218	214/216	ef X eg	216/218	211/213	ab X cd	213/216	216	lm X ll
	CH03g07	X	X			123/125	115/125	ef X eg						
	CH03g12y	X	X	X		177/181	183/198	ab X cd	177/181	181/186	ef X eg			
	CH03g12z	X	X	X		153/161	161/172	ef X eg	153/161	153/170	ef X eg			
	CH04a12	X	X			186/-	176/182	ab X cd						
	CH04c06y	X	X	X	X	177/187	175/179	ab X cd	177/187	179/187	ef X eg	175/179	175/179	hk X hk
	CH04c07	X	X	X	X	86/96	86/96	hk X hk	-/96	104/110	ab X cd	95	105/107	nn Xnp
	CH04d02	X	X			118	110/130	nn Xnp						
	CH04e02	X	X	X	X	152/162	156	lm X ll	152/162	162/164	ef X eg	157	157/165	nn Xnp
	CH04e03	X												
	CH04e05	X												
	CH04f10	X	X			185/222	185	lm X ll						
	CH04g04	X		X	X				168/172	172/180	ef X eg	172/180	172/180	hk X hk
	CH04g07	X	X	X		168/178	174/?	ab X cd	168/178	144/147	ab X cd			
	CH04g09z	X	X	X		145/147	147/155	ef X eg	145/147	141/147	ef X eg			
	CH04h02	X	X	X		187/191	185	lm X ll	187/191	170/172	ab X cd			

Primer nr	Accession nr. or primer name	15 cM Framework	'Anna' X 'Golden Delicious'	'Anna' X 'Sharpe's Early'	'Prima' x 'Golden Delicious'	'Anna' X 'Golden Delicious'			'Anna' X 'Sharpe's Early'			'Prima' x 'Golden Delicious'		
						'Anna'	'Golden Delicious'	Seg. type	'Anna'	'Sharpe's Early'	Seg type	'Golden Delicious'	'Prima'	Seg type
	CH05a04	X	X	X		166/172	164/-	ab X cd	167/174	164/188	ab X cd			
	CH05a05	X	X	X		219	215/219	nn Xnp	219	202/215	nn Xnp			
	CH05c06	X												
	CH05d02	X	X			194/216	194/223	ef X eg						
	CH05d04	X	X			220/224	222	lm X ll						
	CH05d08y	X		X					123	123/126	nn Xnp			
	CH05e03	X	X	X	X	168/171	177/183	ab X cd	168/171	162/168	ef X eg	178/184	178/184	hk X hk
	CH05e06	X	X	X		137/152	132/145	ab X cd	137/150	129/150	ef X eg			
	CH05f04	X												
	CH05g08	X	X	X	X	174/176	176	lm X ll	174/176	176	lm X ll	176	162/176	nn Xnp
	CH05h05	X	X	X		180/185	169/185	ef X eg	180/185	180	lm X ll			
	CH-Vf1	X	X	X	X	161	137/161	nn Xnp	161	140/161	nn Xnp	137/157	137/157	hk X hk
	CN444794-SSR	X	X	X		257/-	251/253	ab X cd	257/-	257/267	ef X eg			
	CN493139-SSR	X												
	HBO3AT	X												
	Hi02c07	X	X	X	X	111/-	107/113	ab X cd	111/-	113/115	ab X cd	107/113	107/115	ef X eg
	Hi02f12	X		X					149	136/149	nn Xnp			
	Hi03a10	X	X			290	253/290	nn Xnp						
	Hi03c05	X	X	X		191/205	191/205	hk X hk	191/205	187/217	ab X cd			
	Hi03d06	X	X	X		129/137	113/-	ab X cd	129/137	113/-	ab X cd			
	Hi03g06	X	X			172	172/182	nn Xnp						
	Hi04a08	X												
	Hi04d02	X	X	X		174/204	174/219	ef X eg	174/204	174/219	ef X eg			
	Hi04e04	X												
	Hi04g05	X		X					251/255	249/251	ef X eg			
	Hi05e07	X	X	X	X	194/213	213/228	ef X eg	194/213	199/228	ab X cd	213/228	213/228	hk X hk
	Hi07d11	X	X	X		204/214	216	lm X ll	206/214	210	lm X ll			

Primer nr	Accession nr. or primer name	15 cM Framework	'Anna' X 'Golden Delicious'	'Anna' X 'Sharpe's Early'	'Prima' x 'Golden Delicious'	'Anna' X 'Golden Delicious'			'Anna' X 'Sharpe's Early'			'Prima' x 'Golden Delicious'		
						'Anna'	'Golden Delicious'	Seg. type	'Anna'	'Sharpe's Early'	Seg type	'Golden Delicious'	'Prima'	Seg type
	Hi07f01	X	X	X	X	209/213	207/209	ef X eg	209/213	207/209	ef X eg	207/209	207/219	ef X eg
	Hi07h02	X	X	X	X	252/263	244/252	ef X eg	252/263	260/271	ab X cd	244/252	244/252	hk X hk
	Hi16d02	X												
	Hi22f12	X	X	X	X	207/213	200/207	ef X eg	207/213	200/-	ab X cd	200/207	200/-	ef X eg
	Hi23g02	X	X	X		235/250	244/-	ab X cd	235/250	235	lm X ll			
	Hi23g12	X	X	X		231	231/236	nn Xnp	231	236/242	nn Xnp			
	MS02a01	X	X	X	X	168/190	168/190	hk X hk	168/190	168/180	ef X eg	168/189	176/197	ab X cd
	MS06g03	X	X	X	X	156/183	138/156	ef X eg	156/182	156/182	hk X hk	156/158	165/188	ab X cd
	NH009b	X	X	X	X	145/156	140/158	ab X cd	145/156	145/156	hk X hk	140/158	140/158	hk X hk
	NH029a	X												
	NZ02b01	X	X	X		215/235	215/225	ef X eg	214/235	228/235	ef X eg			
	U78949-SSR	X												
	AJ251116-SSR		X	X		163/165	165	lm X ll	163/165	163/165	hk X hk			
	CH01b09b		X	X		177/185	173/183	ab X cd	177/185	172	lm X ll			
	CH01b12$_1$x			X	X				122/125	122/125	hk X hk	122/125	125	lm X ll
	CH01b12$_1$z			X					151/163	151/163	hk X hk			
	CH01e09b		X	X		126/136	136	lm X ll	125/135	121/125	ab X cd			
	CH01e121		X	X	X	248/250	244/250	ef X eg	249/251	251	lm X ll	244/250	244/248	ef X eg
	CH01f021		X	X	X	170/184	170/180	ef X eg	170/184	158/184	ef X eg	170/180	180	lm X ll
	CH01f03a			X					221/229	221/229	hk X hk			
	CH01f12		X	X		146/148	143/-	ab X cd	146/148	146/148	hk X hk			
	CH02a04y		X	X	X	105/-	97/101	ab X cd	105/107	107	lm X ll	89/101	89/101	hk X hk
	CH02a04z		X	X		67/90	67/90	hk X hk	67/90	67/90	hk X hk			
	CH02a08		X	X		138/152	152	lm X ll	138/152	145/152	ef X eg			
	CH02a10		X	X		146/152	146	lm X ll	146/152	146/148	ef X eg			
	CH02b03b1			X	X				92	86/92	nn Xnp	74/86	86/92	ef X eg
	CH02b121		X	X		129/135	139	lm X ll	129/135	127/135	ef X eg			

Primer nr	Accession nr. or primer name	15 cM Framework	'Anna' X 'Golden Delicious'	'Anna' X 'Sharpe's Early'	'Prima' x 'Golden Delicious'	'Anna' X 'Golden Delicious'			'Anna' X 'Sharpe's Early'			'Prima' x 'Golden Delicious'		
						'Anna'	'Golden Delicious'	Seg. type	'Anna'	'Sharpe's Early'	Seg type	'Golden Delicious'	'Prima'	Seg type
	CH02c02a			X					140/165	140/155	ef X eg			
	CH02c061		X		X	232/253	237/241	ab X cd				237/241	237/241	hk X hk
	CH02d10a		X	X		214/243	218	lm X ll	214/243	218	lm X ll			
	CH02d121		X	X	X	178/197	178/194	ef X eg	178/197	178/197	hk X hk	194/197	190/197	ef X eg
	CH02g04		X		X	194	179/194	nn Xnp				179/194	194	lm X ll
	CH02h11a		X	X	X	120/130	126/128	ab X cd	120/130	126	lm X ll	126/128	122/128	ef X eg
	CH02h11b		X	X	X	218/220	214/220	ef X eg	218/220	214/218	ef X eg	214/220	220	lm X ll
	CH03a04		X	X		97/113	113	lm X ll	97/113	113	lm X ll			
	CH03b10		X	X	X	96/118	102/118	ef X eg	96/118	100/118	ef X eg	101/118	101/106	ef X eg
	CH03c01		X	X	X	171/173	171	lm X ll	171/173	171/184	ef X eg	171	171/173	nn Xnp
	CH03d08		X	X		140	124/130	nn Xnp	126/140	122/140	ef X eg			
	CH03d10		X			169/-	163/165	ab X cd						
	CH03g04		X	X		126/134	140	lm X ll	126/134	126/140	ef X eg			
	CH03h03		X		X	81/110	88/117	ab X cd				88/117	88/100	ef X eg
	CH04d07		X			128/134	115/128	ef X eg						
	CH04f04		X			145/151	145/151	hk X hk						
	CH04f06		X	X		176/180	159/180	ef X eg	176/180	176/180	hk X hk			
	CH04g10				X							132	120/132	nn X np
	CH05a02y		X			132/135	110/135	ef X eg						
	CH05a02z		X	X		115/125	115/125	hk X hk	115	110/115	nn Xnp			
	CH05a03		X			189/191	189/191	hk X hk						
	CH05b06x		X			185/191	171	lm X ll						
	CH05b06y		X			211/215	211/215	hk X hk						
	CH05c02		X		X	173	171/177	nn Xnp				171/177	171/179	ef X eg
	CH05c07		X	X	X	137	135/137	nn Xnp	137	109/135	nn Xnp	135/137	137/144	ef X eg
	CH05d11		X			171/181	167/171	ef X eg						
	CH05e04		X		X	159/165	157/165	ef X eg				157/165	157/165	hk X hk

Primer nr	Accession nr. or primer name	15 cM Framework	'Anna' X 'Golden Delicious'	'Anna' X 'Sharpe's Early'	'Prima' x 'Golden Delicious'	'Anna' X 'Golden Delicious'			'Anna' X 'Sharpe's Early'			'Prima' x 'Golden Delicious'		
						'Anna'	'Golden Delicious'	Seg. type	'Anna'	'Sharpe's Early'	Seg type	'Golden Delicious'	'Prima'	Seg type
	CH05f06		X			168/183	174/183	ef X eg						
	CH05g03		X	X	X	132/160	132/164	ef X eg	132/160	132/160	hk X hk	132/164	160/164	ef X eg
	CH05g07y		X	X		147/156	156	lm X ll	147/156	149/?	ab X cd			
	CH05g07z			X					167	167/173	nn Xnp			
	CH05g11		X	X	X	212/248	234/244	ab X cd	214/252	238/252	ef X eg	238/249	203/249	ef X eg
	COLa		X	X		231	219/231	nn Xnp	231	219/231	nn X np			
	Hi01d01x		X	X		191/197	191/194	ef X eg	191/197	191	lm X ll			
	Hi01d01y			X					186/188	186/188	hk X hk			
	Hi01e10		X	X		210/-	220	lm X ll	210/-	212/-	ab X cd			
	Hi02b07			X					204/206	216	lm X ll			
	Hi02d04		X	X	X	217/233	217	lm X ll	217/233	223/-	ab X cd	217/243	239/243	ef X eg
	Hi03a03x		X	X		156/178	168	lm X ll	156/178	156/183	ef X eg			
	Hi03a03y		X	X	X	211/219	223/229	ab X cd	211/219	207/223	ab X cd	223/229	221/223	ef X eg
	Hi04b12		X	X		138/145	145/147	ef X eg	138/145	138/145	hk X hk			
	Hi04c10x		X			174/-	178/187	ab X cd						
	Hi04c10y		X	X		204/212	204	lm X ll	204/212	202/204	ef X eg			
	Hi04f09		X			241/252	252	lm X ll						
	Hi05b09		X	X	X	135/137	135	lm X ll	135/137	135/140	ef X eg	135	135/140	nn Xnp
	Hi05d10		X	X		323/345	321	lm X ll	323/345	321/-	ab X cd			
	Hi06b06		X	X	X	257/259	257/259	hk X hk	257/259	241/259	ef X eg	257/259	257/259	hk X hk
	Hi06f09		X	X		271/277	271/277	hk X hk	271/277	271/277	hk X hk			
	Hi07b02		X	X	X	209/213	209/213	hk X hk	209/213	209/213	hk X hk	209/213	200/213	ef X eg
	Hi07b06		X	X	X	220/222	218	lm X ll	220/222	216/218	ab X cd	218	218/222	nn Xnp
	Hi07d08		X		X	213/228	228	lm X ll				228	218/228	nn Xnp
	Hi08h12		X	X		150	150/168	nn Xnp	150	156/171	nn Xnp			
	Hi12c02		X	X	X	166	175/-	nn Xnp	166	166/187	nn Xnp	175/-	166/-	ab X cd
	Hi12f04			X					180/183	180/183	hk X hk			

Primer nr	Accession nr. or primer name	15 cM Framework	'Anna' X 'Golden Delicious'	'Anna' X 'Sharpe's Early'	'Prima' x 'Golden Delicious'	'Anna' X 'Golden Delicious'			'Anna' X 'Sharpe's Early'			'Prima' x 'Golden Delicious'		
						'Anna'	'Golden Delicious'	Seg. type	'Anna'	'Sharpe's Early'	Seg type	'Golden Delicious'	'Prima'	Seg type
	Hi15h12			X	X				222	222/225	nn Xnp	222	222/225	nn Xnp
	Hi20b03		X	X	X	215/224	215/238	ef X eg	215/224	215/238	ef X eg	215/238	224/-	ab X cd
	Hi21g05		X	X	X	155/158	158	lm X ll	155/158	155/158	hk X hk	158	155/158	nn Xnp
	Hi24f04		X	X	X	147	144/147	nn Xnp	147	144/147	nn Xnp	144/147	144/147	hk X hk
	KA4b		X	X		135/137	135/137	hk X hk	135/137	133/135	ef X eg			
	MS01a03			X	X				207	207/240	nn X np	209/240	209	lm X ll
	MS01a05			X					170/174	156/174	ef X eg			
	NZ05g08		X		X	118/125	118	lm X ll				119	113/125	nn X np
	NZ23g04		X	X		101/107	80/84	ab X cd	108/110	106/108	ef X eg			
	NZ28f04		X	X	X	99/107	92/107	ef X eg	99/107	99/107	hk X hk	92/107	92/107	hk X hk
A334	CN444542	X		X	X				215	205/215	nn Xnp	209/225	216	lm X ll
A30	AT000400.1		X			175/181	175/179	ef X eg						
A186	CN490349		X		X	200	200/206	nn Xnp				200/206	200	lm X ll
A188	CN490740				X							195	193/211	nn Xnp
A193	CN492206		X		X	398	398/471	nn Xnp				471/481	471/481	hk X hk
A209	CN495857			X	X				145/148	145/155	ef X eg	145/148	148	lm X ll
A219	CN580620		X			380	377/380	nn Xnp						
A234	CN938125		X	X		338/345	345/352	ef X eg	338/345	340/354	ab X cd			
A236	CN910036			X					235/241	220/241	ef X eg			
A238	CN865016		X	X	X	340/345	334/340	ef X eg	340/345	340/345	hk X hk	334/340	340	lm X ll
A244	CN947446		X	X		178/181	181/184	ef X eg	178/181	181/187	ef X eg			
A253	CO540769		X			265	260/265	nn Xnp						
A279	CN887525		X			207	207/214	nn Xnp						
A307	CN445290		X		X	338	338/349	nn Xnp				338/349	338/349	hk X hk
A310	AU301301		X	X	X	228/243	223/241	ab X cd	218/229	229	lm X ll	223/240	240/250	ef X eg
A319	AF527800		X			330/-	-/-	lm X ll						
A320	CN580637		X			418	406/418	nn Xnp						

Primer nr	Accession nr. or primer name	15 cM Framework	'Anna' X 'Golden Delicious'	'Anna' X 'Sharpe's Early'	'Prima' x 'Golden Delicious'	'Anna' X 'Golden Delicious'			'Anna' X 'Sharpe's Early'			'Prima' x 'Golden Delicious'		
						'Anna'	'Golden Delicious'	Seg. type	'Anna'	'Sharpe's Early'	Seg type	'Golden Delicious'	'Prima'	Seg type
A326	U50187		X	X		160/-	149/162	ab X cd	160/?	149/162	ab X cd			
A331	AB162040			X					266	266/272	nn Xnp			
A335	CO052033		X			188	188/196	nn Xnp						
A340	CO416051		X	X		120/133	120/130	ef X eg	120/134	120/129	ef X eg			
A341	CO723148			X					152/157	152/157	hk X hk			
A343	CV084260		X	X		226/256	262	lm X ll	226/256	226/264	ef X eg			
A372	CO052555		X	X	X	232/234	236	lm X ll	232/234	232	lm X ll	236	232/236	nn Xnp
A380	CO866737			X					239	239/251	nn Xnp			
A381	CO751676			X	X				233/235	218/228	ab X cd	218	214/218	nn Xnp
A383	CO903298		X	X		344	344/350	nn Xnp	344	342/344	nn Xnp			
A398	CN490644		X			262/266	262/266	hk X hk						
A401	CN544835		X	X	X	301/303	303/305	ef X eg	301/303	301/303	hk X hk	303/305	301/303	ef X eg
A417	CV128959		X			240	232/242	nn Xnp						
A424	CO415353		X	X		329/333	329/333	hk X hk	329/333	329/333	hk X hk			

Newly developed marker A334 was designed from the same accession used by Silfverberg-Dilworth *et al.* (2006) for the design and mapping of the reference marker CN444542-SSR, and is therefore indicated as part of the 15 cM framework markers (Table 14). In addition to the reference markers, segregation of another 79 published SSR markers (Guilford, *et al.*, 1997, Liebhard, *et al.*, 2002, Silfverberg-Dilworth, *et al.*, 2006) and 29 newly developed SSR markers (Table 14) have been studied in the three mapping populations used, or a subset thereof.

It is interesting to note the high level of 'allele sharing' between 'Golden Delicious' and 'Prima', caused by the fact that 'Golden Delicious' is a grandparent of 'Prima' (http://www.hort.purdue.edu/newcrop/pri/coop02.html).

Segregation of alleles from all three classes of loci (Table 12) were easily studied through the interpretation of electropherograms obtained from automated genetic analyzers (Figure 30).

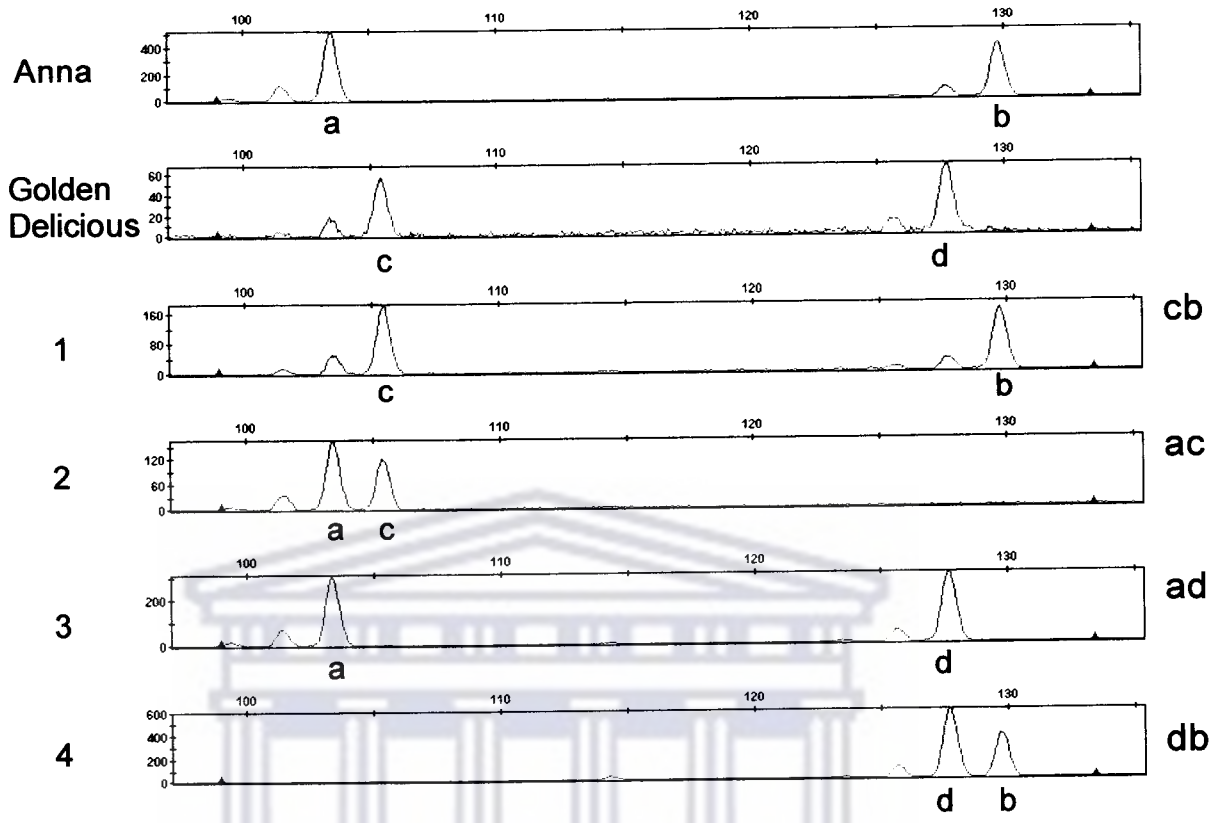


Figure 30. Electropherograms obtained after implementation of the marker CH01g12 on the cultivars 'Anna' and 'Golden Delicious' and the four different classes observed in the F1 population (1-4) derived from a cross between these two cultivars.

In the case of a few markers, one of the parental alleles failed to amplify and was not detected during PCR assays done in order to determine the heterozygosity of the specific marker. Such a marker can easily be misinterpreted as being homozygous and it is only during the study of the segregation of alleles in mapping populations that the presence of a null allele is confirmed.

Possible explanations for the presence of null alleles includes the failure of primer annealing due to nucleotide sequence divergence caused by point mutations, or the preferential amplification of smaller alleles compared to larger alleles. The latter could be overcome by the use of 'touch down' PCR methods (explained in section 3.3.3.1), but this technique was only used to determine optimal PCR conditions and not for the determination of actual fragment sizes. Null alleles may also appear to more frequently encountered when using multiplexed reactions, where there is a lot of competition for available enzyme and nucleotides, resulting in larger fragments being 'absent'. The pooling of various PCR reactions before CE may also lead to the appearance of 'partial nulls' (Dakin and Avise, 2004), due to signal reduction and this might be overcome by loading more sample. False impressions regarding the presence of SSR null alleles are maybe more frequent than would be expected. The presence of null alleles, true or through false interpretation, could be detrimental in studies involving parentage testing or cultivar identification, but when studying a segregating mapping population the presence of a null allele can still be analyzed and scored, resulting in a marker that can be successfully used in the construction of a genetic linkage map. Examples of such markers, assumed to only be useable for segregation analysis in one parent, can be illustrated by the markers Hi02c07, A326 and CN4444794-SSR (Table 15). After implementation of these markers on mapping pedigrees the presence of a null allele was confirmed by the genotypic classes obtained. Markers Hi02c07 (Silfverberg-Dilworth, *et al.*, 2006) (Figure 31) and A326 (Figure 32), yielded 4 seedling genotypic classes each when implemented in a mapping populations where 'Anna', believed to

be homozygous at both loci, was crossed with 'Golden Delicious'. A 105 bp fragment was observed in all seedlings, confirming the possible multi locus nature of the marker Hi02c07 (Silfverberg-Dilworth, *et al.*, 2006) (Figure 31). The same results (data not shown) were obtained using 'Anna' and 'Sharpe's Early' as parental cultivars, confirming the presence of null alleles at these loci in 'Anna'.

Table 15. Examples of markers, believed to be homozygous in 'Anna', that proved to be heterozygous through the existence of null alleles identified through the implementation and screening of markers on segregating mapping populations..

	Hi02c07	A326	CN444794-SSR
'Anna'	111 bp	160 bp	257 bp
'Golden Delicious'	107 & 113 bp	149 & 162 bp	251 & 253 bp
'Sharpe's Early'	113 & 115 bp	149 &162bp	257 & 267 bp

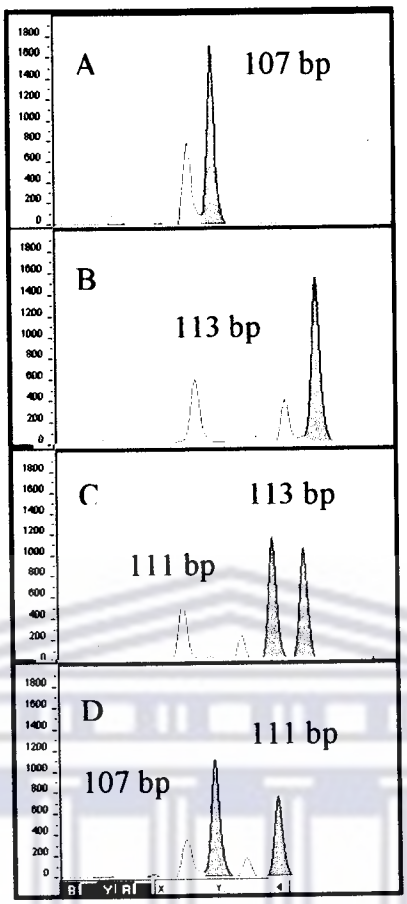


Figure 31. Electropherograms of the four different genotypic classes (A-D) obtained after implementation of the marker Hi02c07 on a F1 mapping population derived from a cross between 'Anna' and 'Golden Delicious'.

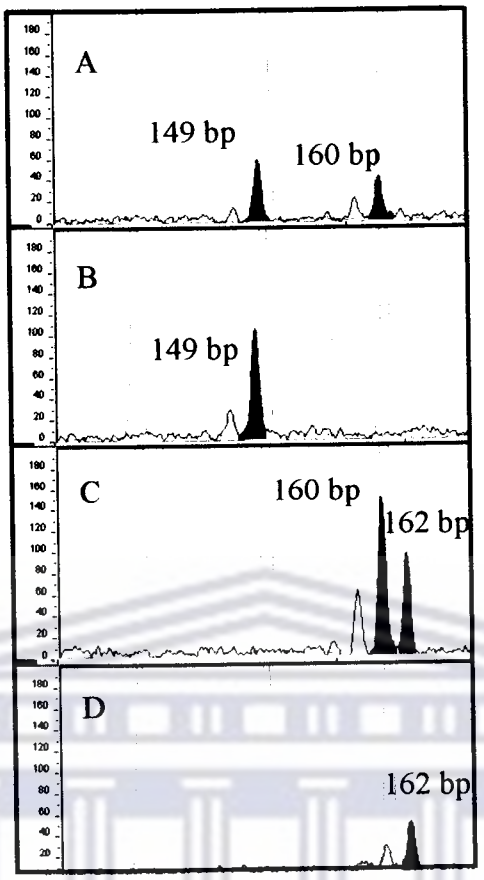


Figure 32. Electropherograms of the four different genotypic classes (A-D) obtained after implementation of the marker A326 on a F1 mapping population derived from a cross between 'Anna' and 'Golden Delicious'.

Marker CN444794-SSR yielded three seedling genotypic classes, instead of the expected two, when implemented on the mapping population derived from 'Anna' x 'Sharpe's Early'. The allele sizes observed in the parental cultivars lead to the expectation that the segregation of this marker would be of the type 'nn x np', since the single fragment observed for 'Anna' leads one to the assumption that it is homozygous. It is only through the study of segregation in the progeny (Figure 33) that the conclusion can be made that there must be a null allele present in the parental cultivar previously assumed to be homozygous and that this cultivar is in fact heterozygous. The segregation type is therefore in fact of the type <efxeg>, but in this case no clear distinction can be made between seedlings containing 2x 257 bp fragments and those containing a single 257 bp fragment and a null allele. These two genotypic classes, expected to be about 50 % of the progeny, thus need to be treated as missing data points, resulting in this marker being highly uninformative on this mapping population and not useable for genetic linkage map construction. The presence of the null allele, when the marker CN444794-SSR is implemented on 'Anna', was also confirmed in a mapping population where 'Anna' was used as a parent together with 'Golden Delicious' (results not shown), but in this case the fragment sizes observed in the second parent, 'Golden Delicious', differed from the fragment observed in the 'Anna' parent, rendering the marker fully informative even though a null allele is present.

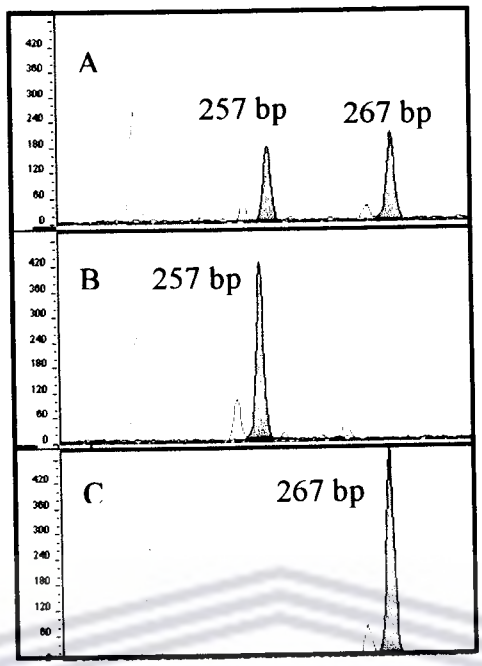


Figure 33. Electropherograms of the three different genotypic classes (A-C) obtained after implementation of the marker CN444974-SSR on a F1 mapping population derived from a cross between 'Anna' and 'Sharpe's Early'.

The presence of a null allele can also lead to the assumption that a marker is problematic. This can be illustrated by A319 that was assumed problematic as it failed to generate amplification products when used on 'Golden Delicious' but generated a single fragment on 'Anna'. When studying the mapping pedigree derived through the crossing of these cultivars it became clear that half the population had no amplification products and the other half had a single fragment. The only possible explanation for this would be the presence of a null allele in 'Anna' and the assumption that 'Golden Delicious' is homozygous for this null allele. This marker did however not show any significant linkage with other markers to enable assignment to a LG

during both F1 and parental genetic linkage map construction (see following section).

Ensuring that observed segregation ratios resemble expected ratios, as outlined in table 12, is thus indicative of whether the fragments observed in parental cultivars are correct. It should be kept in mind that segregation distortion does occur in nature and even more so in mapping populations of limited size, due to sampling error. When observed ratios clearly indicate the presence or absence of a segregation type in the seedlings, caution should be taken and the possibility of the presence of a null allele should be explored.

During the initial steps of genetic linkage map construction, the JoinMap® 4 software (Van Ooijen, 2006) offers the ability to view the segregation ratios observed for each marker analyzed on a specific mapping population, as well as the statistical significance of deviations from these expected ratios (Appendix A).

4.4.2 Genetic linkage map construction

Genetic linkage maps were constructed for the F1 populations derived from each of the three different mapping populations used during this study, 'Anna' x 'Golden Delicious' (Figure 34), 'Anna' x 'Sharpe's Early' (Figure 35) and 'Prima' x 'Golden Delicious' (Figure 36), using JoinMap® 4 (Van Ooijen, 2006). The numbering of LGs is in accordance with Maliepaard *et al.* (1998) and different segments belonging to the same LG were identified through the alignment with reference markers proposed by Silfverberg-Dilworth *et al.* (2006) (Figure 37). Parental genetic linkage maps were constructed for the

two parental cultivars used in each of the three mapping pedigrees (Figure 38) (Figure 39) (Figure 40).

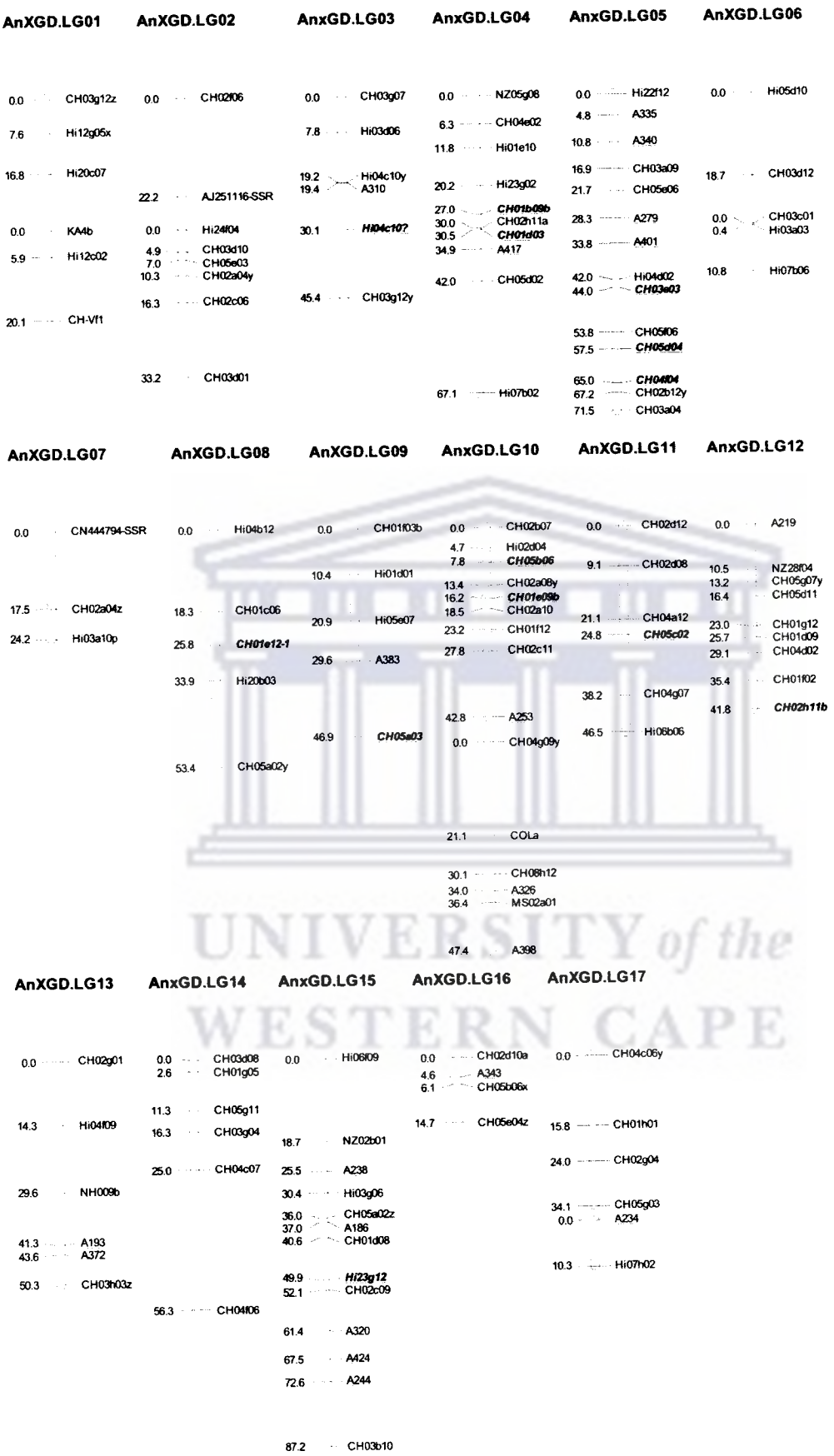


Figure 34. Genetic linkage map constructed using 80 individuals of a F1 population derived from a cross between 'Anna' (male parent) and 'Golden Delicious' (female parent). The 22 linkage groups obtained, are numbered according to the 17 linkage groups obtained by Maliepaard et al. (1998). Newly developed and mapped SSR markers are underlined. Markers that have been published but for which the map positions were unknown, as well as published markers mapping to different positions, are indicated in italics and underlined.

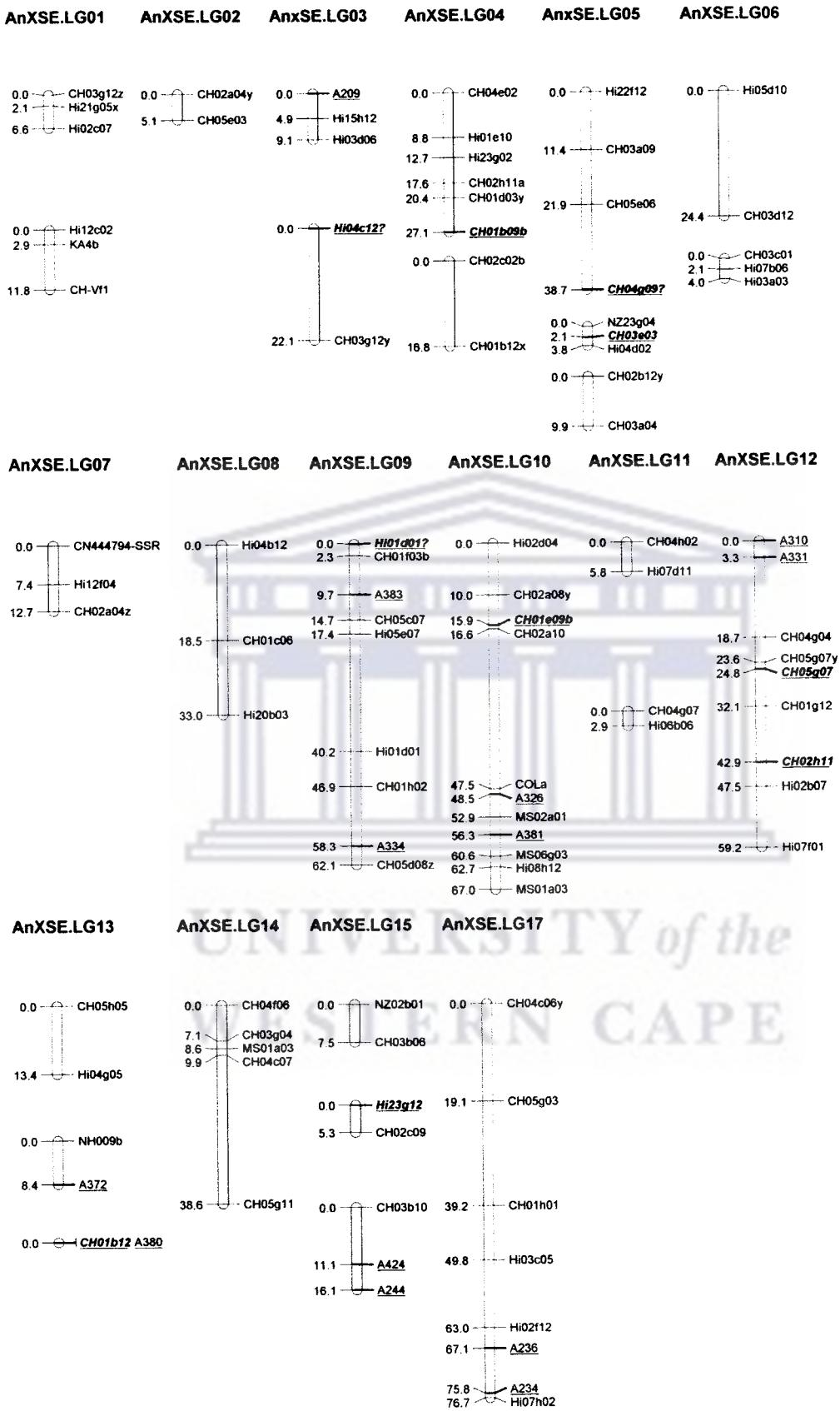


Figure 35. Genetic linkage map constructed using 83 individuals of a F1 population derived from a cross between 'Anna' (male parent) and 'Sharpe's Early' (female parent). The 27 linkage groups obtained, are numbered according to the 17 linkage groups obtained by Maliepaard et al. (1998). Newly developed and mapped SSR markers are underlined. Markers that have been published but for which the map positions were unknown, as well as published markers mapping to different positions, are indicated in italics and underlined.

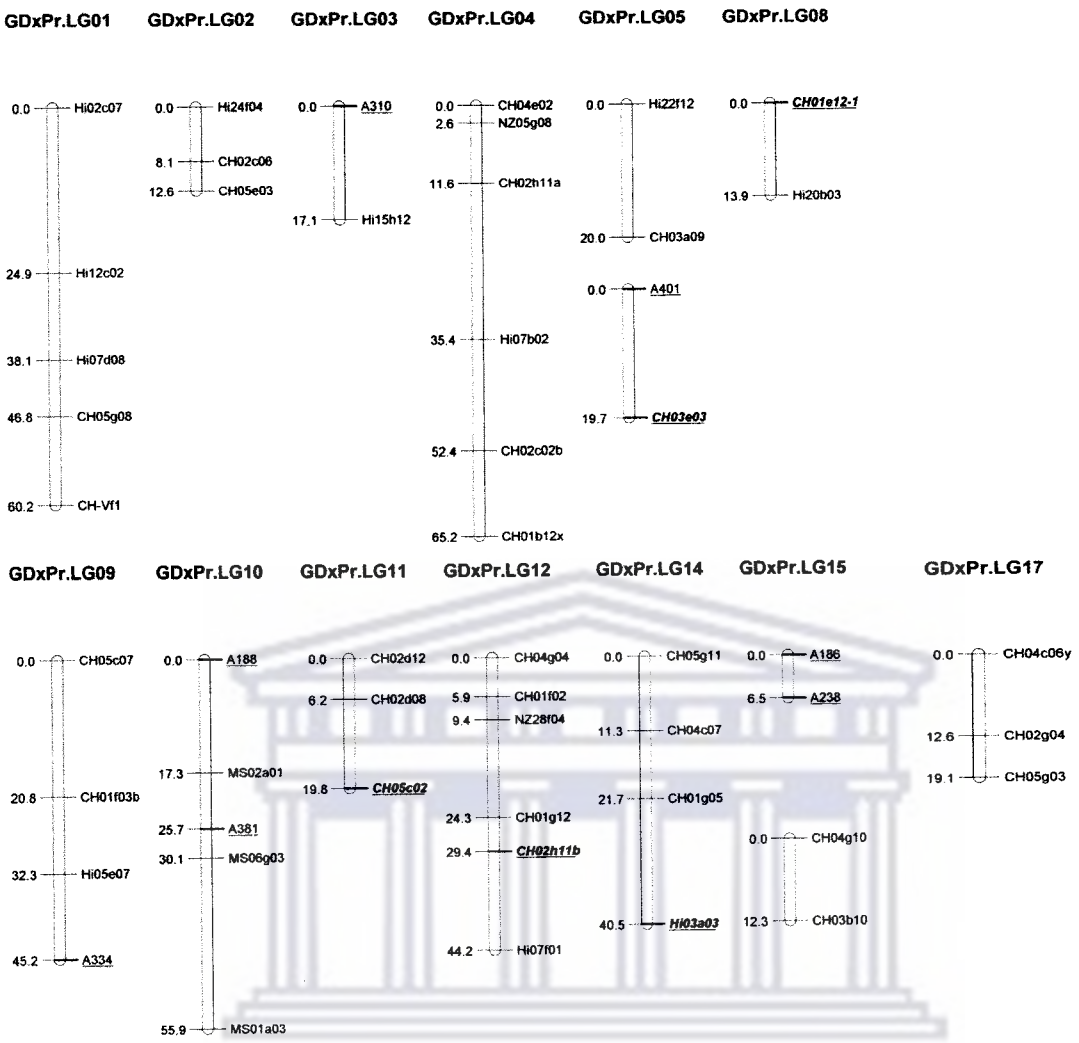
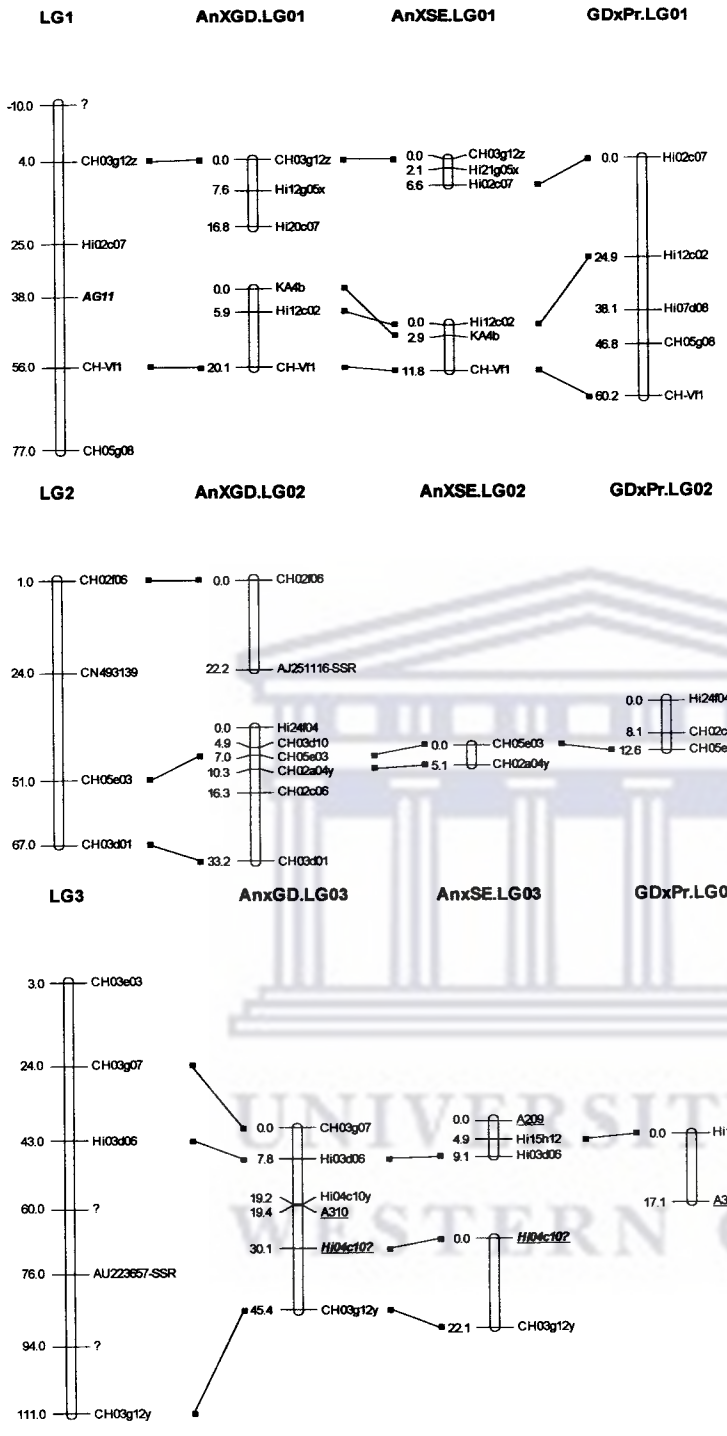
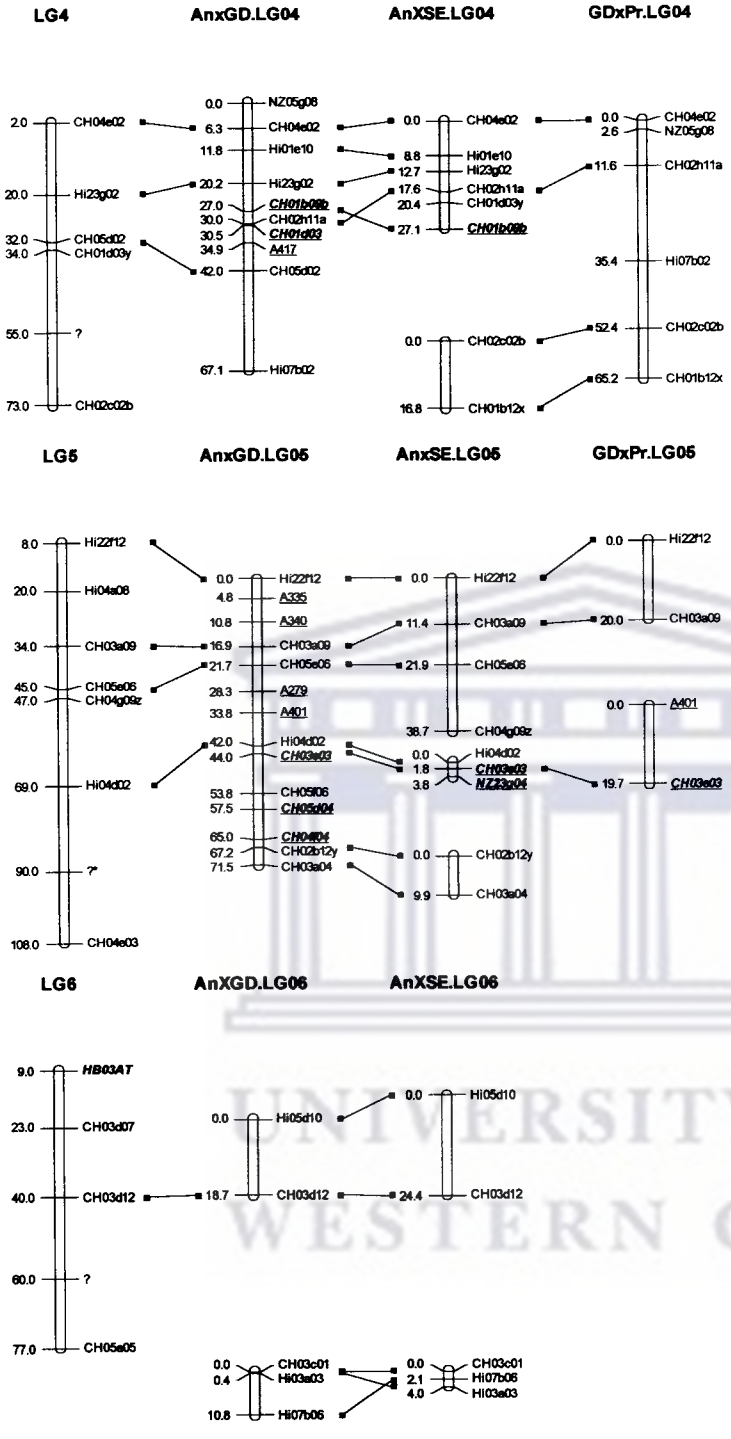
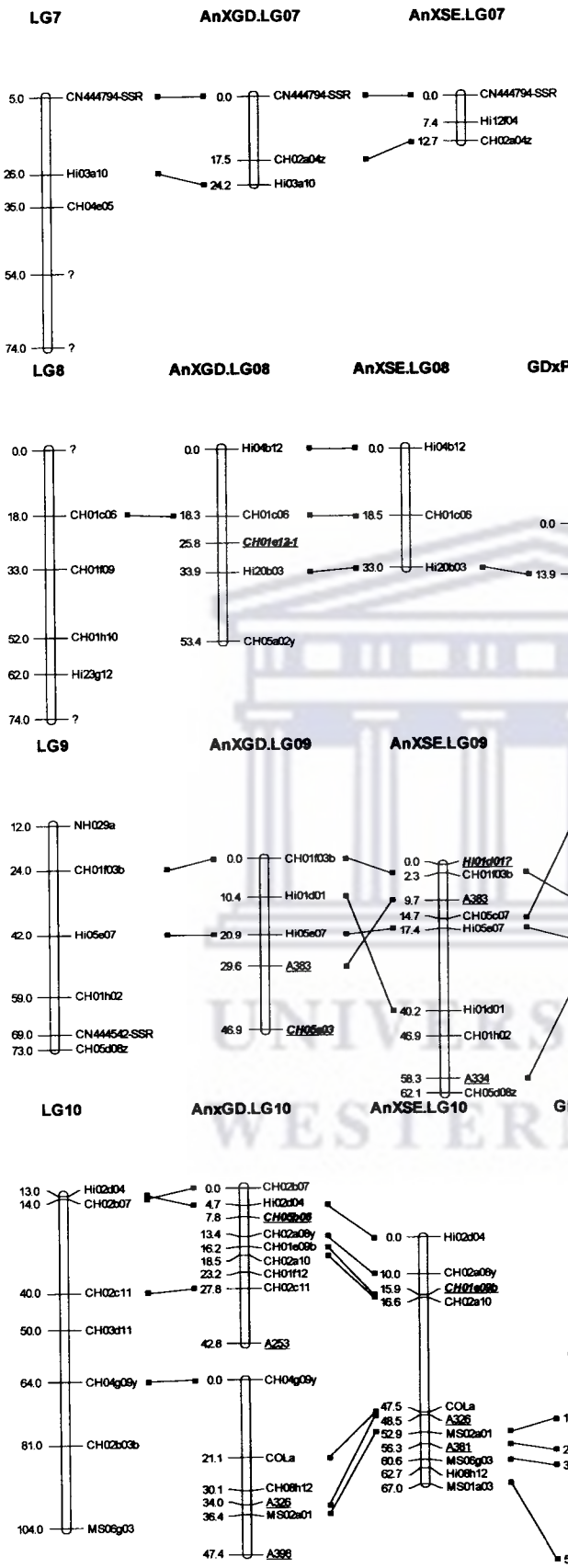
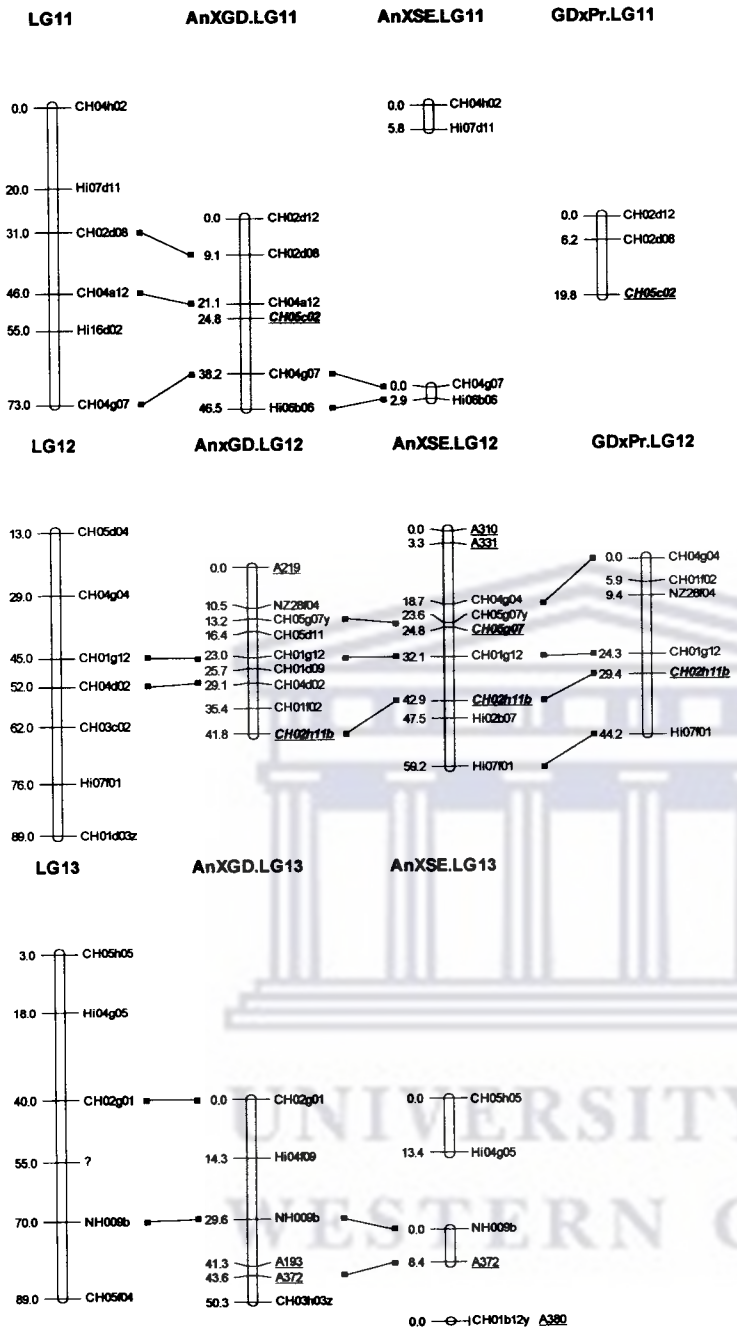


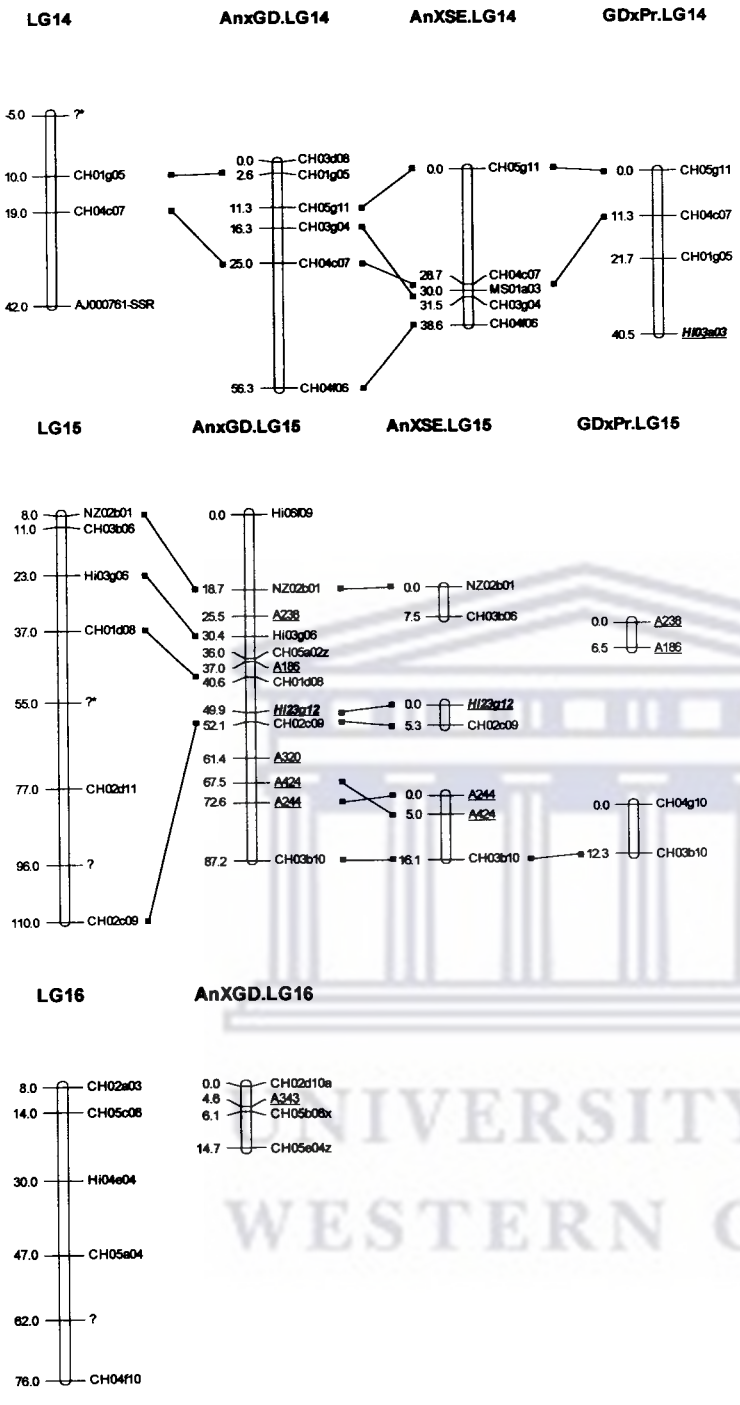
Figure 36. Genetic linkage map constructed using 85 individuals of a F1 population derived from a cross between 'Prima' (male parent) and 'Golden Delicious' (female parent). The 15 linkage groups obtained, are numbered according to the 17 linkage groups obtained by Maliepaard et al. (1998). Newly developed and mapped SSR markers are underlined. Markers that have been published but for which the map positions were unknown, as well as published markers mapping to different positions, are indicated in italics and underlined.











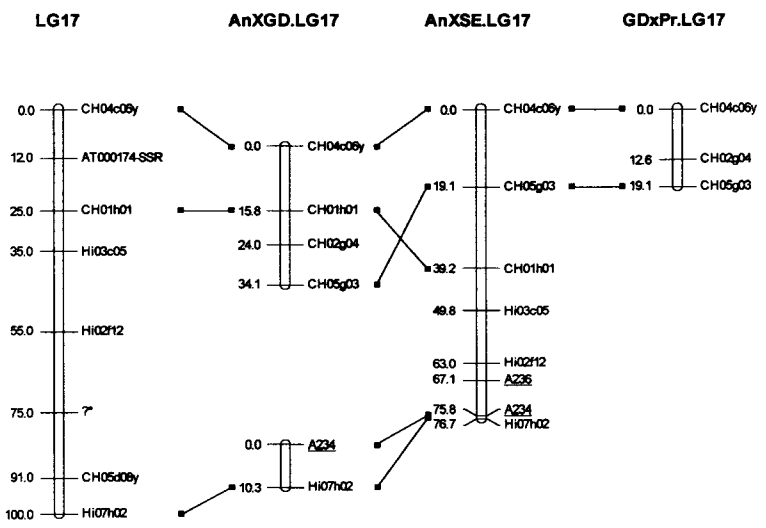
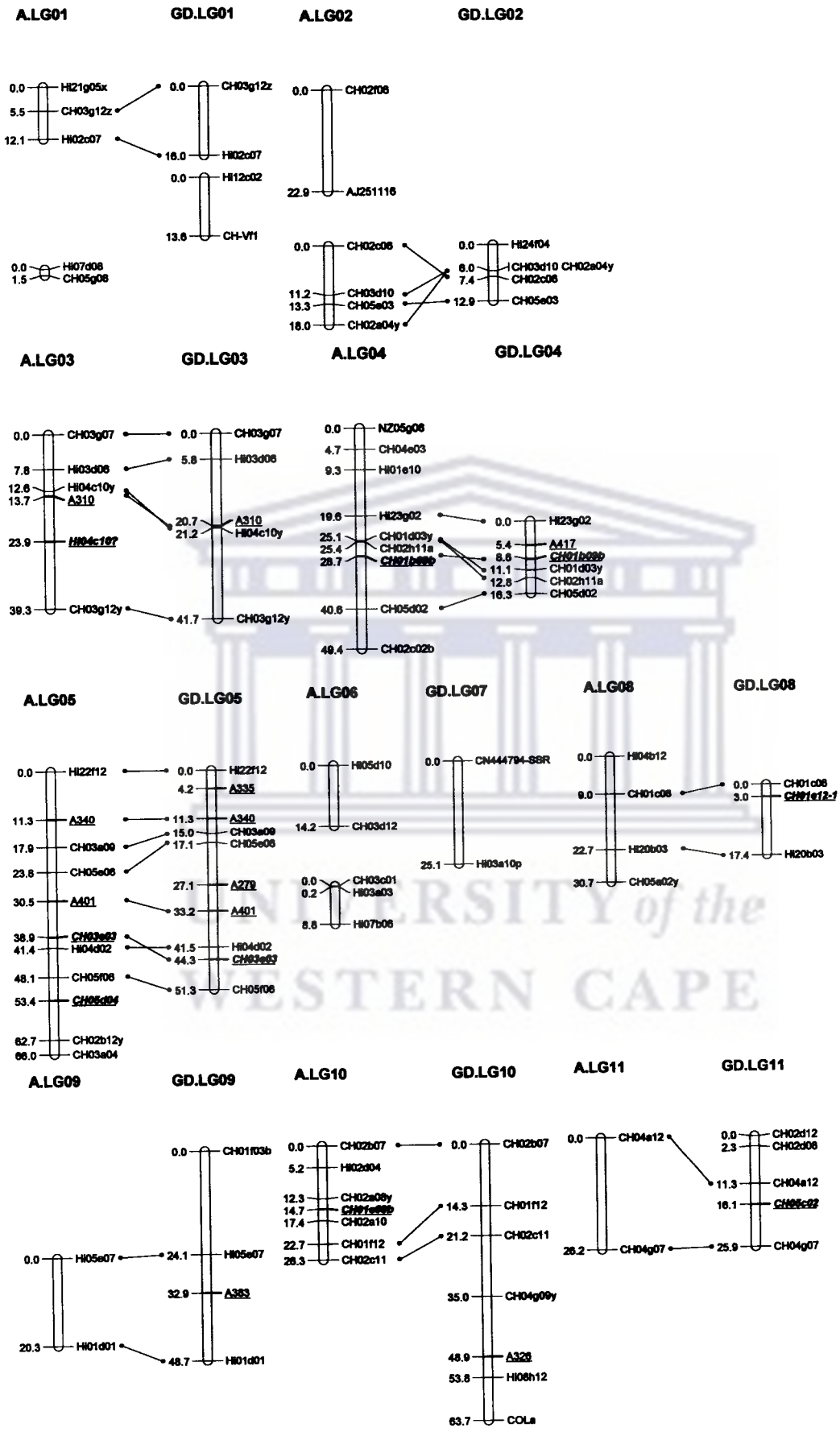


Figure 37. Genetic linkage maps obtained using F1 populations from 3 different mapping pedigrees, aligned to each other as well as to the reference map proposed by Silfverberg-Dilworth et al. (2006). For each linkage group the proposed reference markers are indicated on the left, followed by the alignment of linkage groups, or segments thereof, obtained from each of the three different mapping pedigrees, 'Anna' x 'Golden Delicious' (AnxGD), 'Anna' x 'Sharpe's Early' (AnxSE) and 'Prima' x 'Golden Delicious' (PrxGD). In the case of linkage groups 6, 7, 13 and 16 groupings were not obtained from all three the mapping pedigrees and only those pedigrees with sufficient linkage between markers to form these groups are indicated.



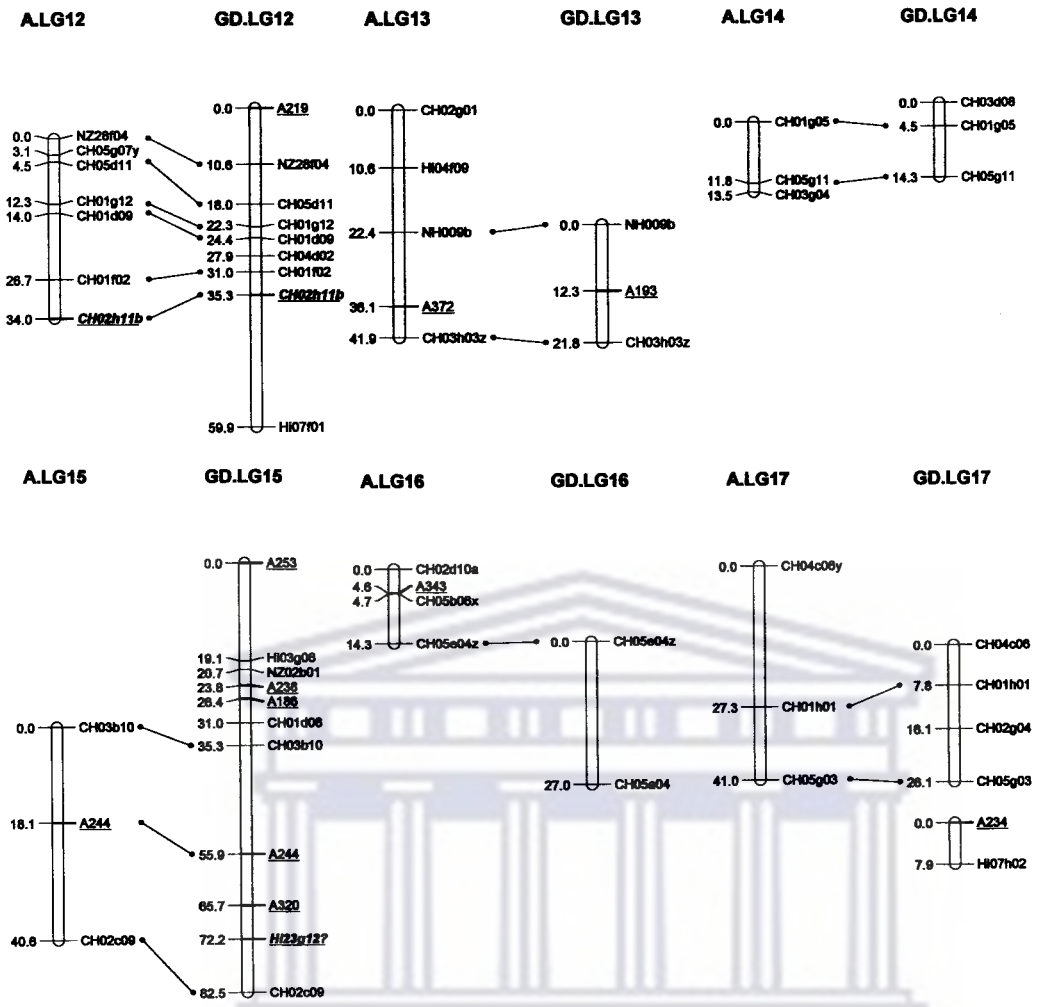
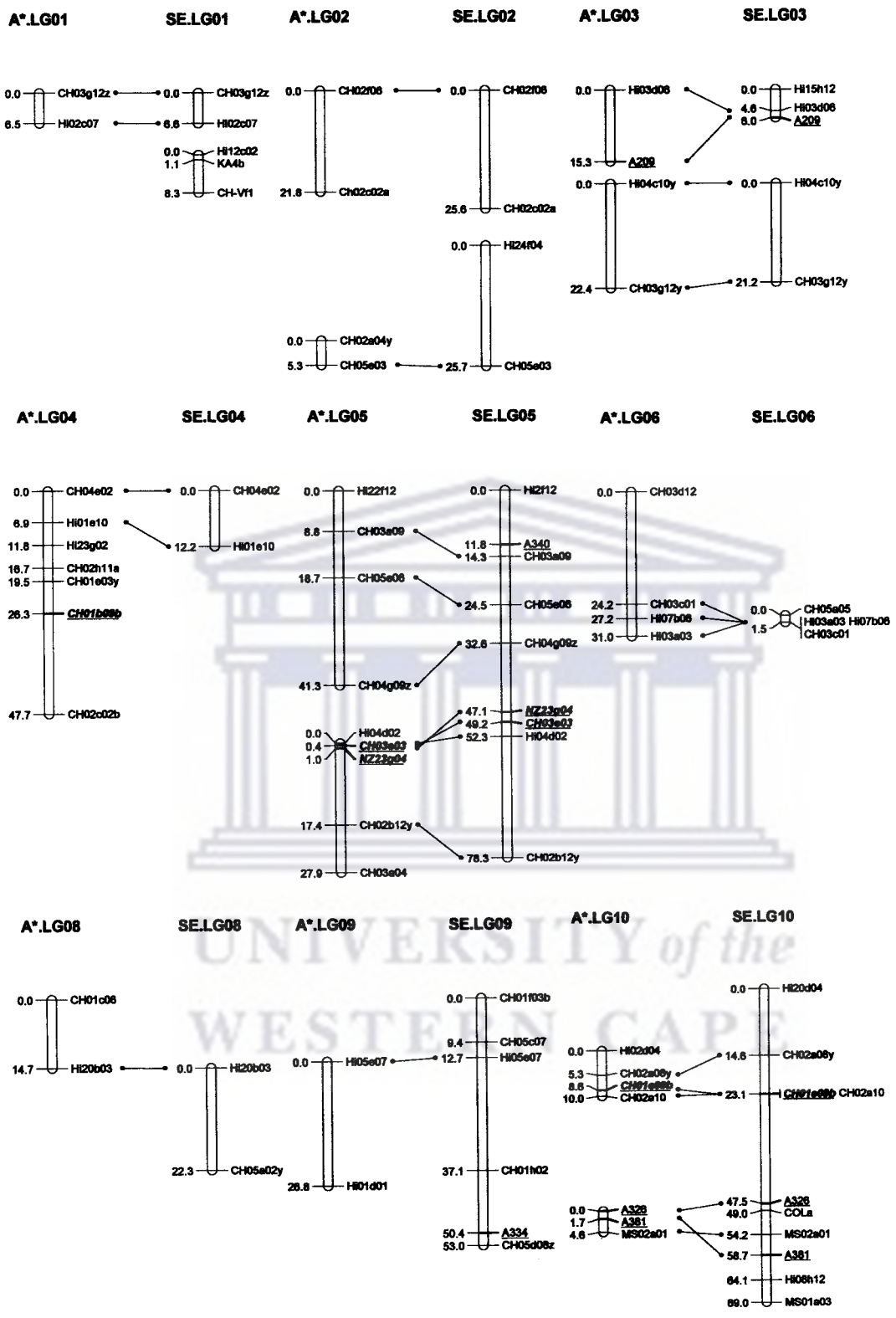


Figure 38. Genetic linkage maps of 'Anna' and 'Golden Delicious'. Numbering of linkage groups are according to Maliepaard et al. (1998). Newly developed and mapped SSR markers are underlined. Markers that have been published but for which the map positions were unknown, as well as published markers mapping to different positions, are indicated in italics and underlined.



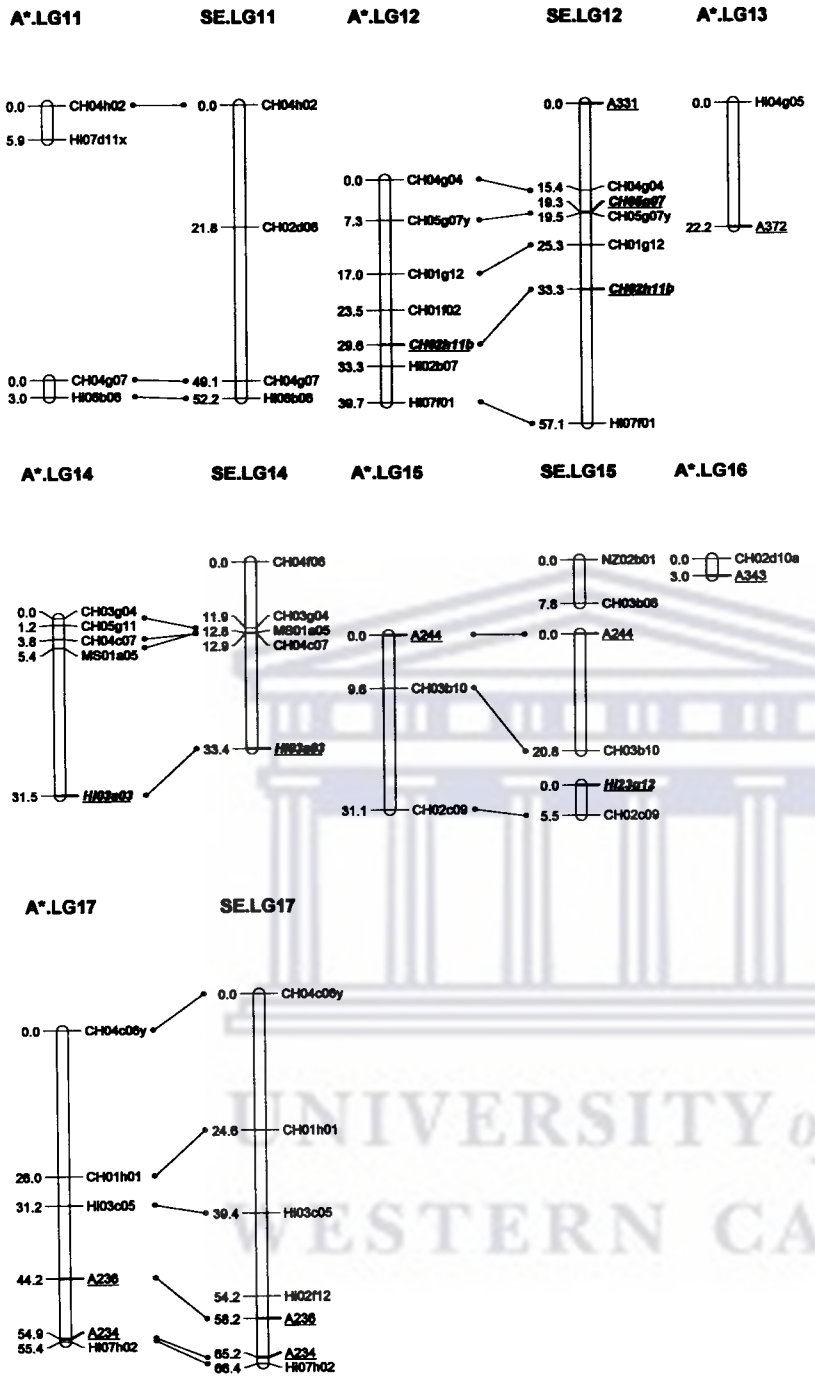


Figure 39. Genetic linkage maps of 'Anna' and 'Sharpe's Early'. Numbering of linkage groups are according to Maliepaard et al. (1998). Newly developed and mapped SSR markers are underlined. Markers that have been published but for which the map positions were unknown, as well as published markers mapping to different positions, are indicated in italics and underlined.

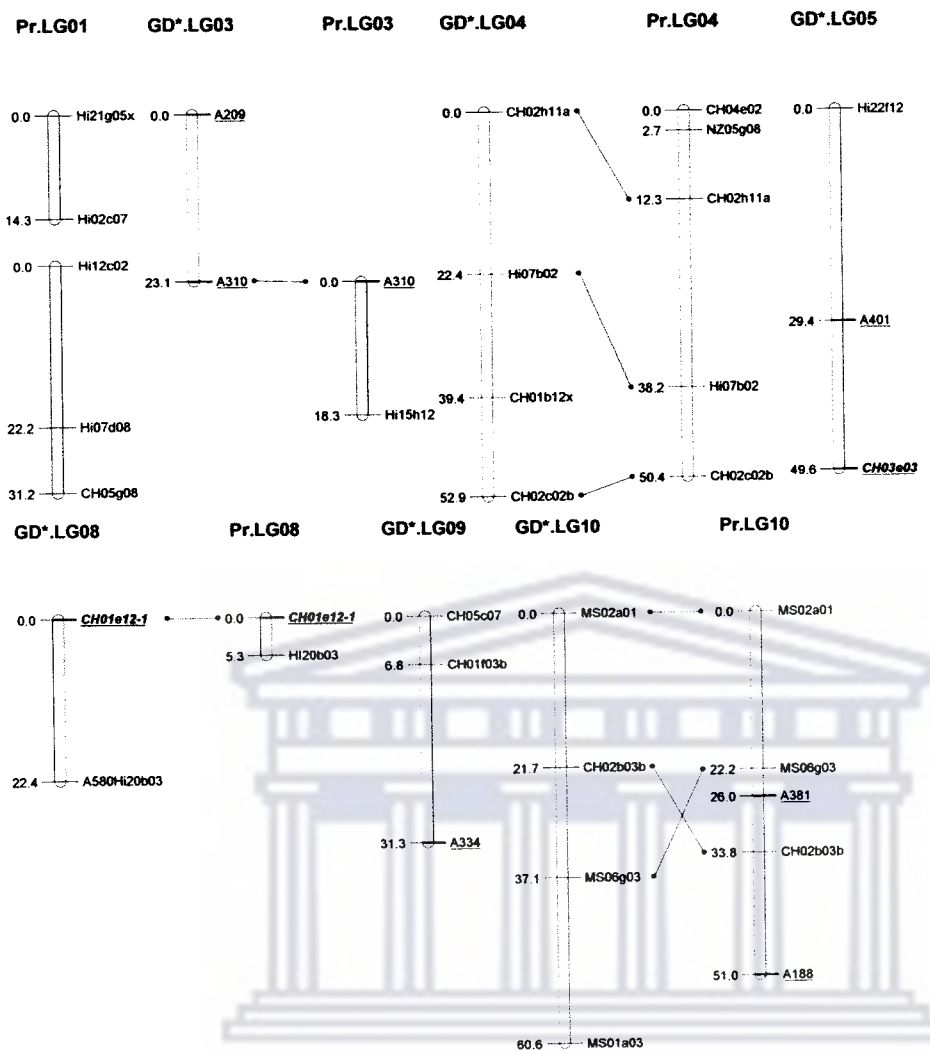


Figure 40. Genetic linkage maps of 'Golden Delicious' and 'Prima'. Numbering of linkage groups are according to Maliepaard et al. (1998). Newly developed and mapped SSR markers are underlined. Markers that have been published but for which the map positions were unknown, as well as published markers mapping to different positions, are indicated in italics and underlined.

The number of seedlings with more than 25 % missing data points, that were removed during the calculation of genetic linkage maps, varied between the different mapping populations studied, as well as between the different parental maps (Table 16). Although markers with more than 40 % missing data points, calculated after the removal of seedlings, were excluded during the determination of genetic LGs, a number of these markers were successfully assigned to groups, based on SCL values (Van Ooijen, 2006). A number of markers that did not show sufficient linkage with the LGs obtained using selected grouping criteria, were also successfully added to LGs using SCL values (Table 16). Markers that remained excluded, after assignment of markers to SCL groups, in F1 populations (Table 17) as well as during the study of separate parental meiosis (Table 18), were not mapped due to large amounts of missing data making the placement of these markers on the map/s very difficult. Markers having enough data to be included in initial LG determination steps but that did not show recombination frequencies with any other markers to enable their assignment to LGs, might be situated so far apart from any other marker/s that the recombination frequency observed between them resemble that of markers residing on different LGs (possible explanation 1). The occurrence of 'recombination hotspots' may also result in the 'breaking up' of LGs into two or more segments. This, together with the observation that markers expected to have sufficient linkage with groups obtained based upon published data, are difficult to explain (possible explanation 3) and might be overcome by the implementation of more markers in order to generate genetic linkage maps that are more saturated. This is an expensive and time consuming exercise and although genetic

linkage maps constructed for each of the three mapping pedigrees used, do not all have the expected number of 17 LGs, the LGs, or segments thereof, that are available (Table 19) can be successfully used for the identification of putative QTLs, the ultimate goal for the construction of linkage maps during this investigation. The absence of certain LGs from the genetic linkage maps obtained can also be explained by the fact that, for some LGs, only one published marker was found to be segregating in the mapping pedigree under investigation (possible explanation 2). Newly developed markers, as well as published markers for which the map positions are unknown, were also found among the set of unlinked markers, probably due to the fact that they are positioned on the ends of LGs (possible explanation 1) or due to the occurrence of 'recombination hotspots' (possible explanation 2), but these assumptions need to be proved and these markers are therefore grouped in a class of their own (possible explanation 4).

Table 16. Summary of the number of individuals and number of segregating markers used during the construction of genetic linkage maps.

		Number of individuals	Total number of markers analysed	Number of individuals excluded	Number of loci excluded	Number of ungrouped loci	Number of excluded loci assigned to SCL groups	Number of ungrouped loci assigned to SCL groups	Number of markers removed due to high Chi-square values etc.	Total number of markers included in genetic linkage map
F1 GENETIC LINKAGE MAPS										
'Anna' x 'Golden Delicious'		87	149	7	10	15	7	3	5	129
'Anna' x 'Sharpe's Early'		94	132	11	7	24	5	4	10	100
'Golden Delicious' x 'Prima'		92	69	7	5	15	2	7	7	51
PARENTAL GENETIC LINKAGE MAPS										
'Anna' x 'Golden Delicious'	**'Anna'**	87	107	9	6	20	2	7	8	76
	'Golden Delicious'	87	101	13	8	12	6	2	4	77
'Anna' x 'Sharpe's Early'	**'Anna'**	94	83	13	3	9	1	2	9	74
	'Sharpe's Early'	94	89	14	5	10	4	1	3	76
'Golden Delicious' x 'Prima'	**'Golden Delicious'**	92	39	9	2	8	2	1	1	31
	'Prima'	92	39	12	4	9	2	1	0	29

Table 17. Summary of markers excluded from genetic linkage map construction and markers found not to have sufficient linkage with obtained linkage groups, during analysis of F1 mapping populations. Possible explanations for markers not showing linkage with any other marker or group of markers are: (1) the marker has been published and are 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 (2) 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 (3) as well as newly developed and published but unmapped markers (4).

	Excluded markers			Unlinked markers		
	Marker	LG	% Missing data	Marker	LG	Possible explanation
'Anna' x 'Golden Delicious'	CH04f10	16	63 %	CH02c02b	4	1
	CH02c09	15	56 %	CH05a05	6	3
	A307		47 %	Hi05b09	7	1
				CH01h101	8	1
				CH03d11	10	3
				MS06g03	10	1
				CH04h02	11	1
				CH05a04	16	1
				Hi03c05	17	3
				A30		4
				A319		4
				Hi03a03	14?	3
'Anna' x 'Sharpe's Early'	CH02d10a	16	45 %	CH05g08	1	1
	A340		56 %	CH02c02a	2	3
				CH02f061	2	
				CH03d01	2	
				Hi24f04	2	
				AJ251116-SSR	2	
				Hi07b02	4	1
				Hi05b09	7	1
				CH05a02y	8	1
				CH02b03b1	10	3
				CH02d08	11	3
				CH02d121	11	
				NZ28f04	12	3
				CH03d08	14	1
				Hi06f09	15	1
				CH05a04	16	2
				Hi03a03	14?	1 or 3
				CH01f03a		4
				A341		4
				A343		4
'Golden Delicious' x 'Prima'	Hi21g05	1	49 %	CH02f061	2	1
	Hi07b06	6	48 %	Hi05b09	7	2
	A307		43 %	CH01f091	8	3
				Hi02d04	10	1
				NH009b	13	3
				CH03h03	13	
				CH05e04	16	2
				Hi07h02	17	1

Table 18. Summary of markers excluded from genetic linkage map construction and markers found not to have sufficient linkage with obtained linkage groups, during analysis of parental meiosis. Possible explanations for markers not showing linkage with any other marker or group of markers are: (1) the marker has been published and are 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 (2) 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 (3) as well as newly developed and published but unmapped markers (4).

Excluded markers			Unlinked markers			Excluded markers			Unlinked markers		
Marker	LG	% Missing data	Marker	LG	Possible explanation	Marker	LG	% Missing data	Marker	LG	Possible explanation
'Anna' x 'Golden Delicious'											
'Anna'						'Golden Delicious'					
CH04d07	11	53 %	CN444794-SSR	7	1	CH05c07	9	54 %	CH02f061	2	1
Hi07f01	12	61 %	Hi05b09	7		CH04d07	11	53 %	CH02c02b	4	1
CH04f10	16	63 %	CH01h101	8	1				CH05a05	6	2
A238		61 %	CH03d11	10	1?				CH01h101	8	1

			MS06g03	10						MS06g03	10	1
			A326	10						CH02g01	13	1
			CH02d121	11	1					CH04f06	14	
			CH04h02	11						A30		4
			CH04f06	14	1					CH05h05		4
			NZ02b01	15	1					Hi03a03	14?	1
			CH05a04	16	1							
			A319		4							
			Hi03a03	14?	1							

'Anna' x 'Sharpe's Early'

'Anna'						'Sharpe's Early'					
CN444794 -SSR	7	56 %	CH05g08	1	1	CN444794-SSR	7	56 %	Hi05d10	6	1
A340		56 %	KA4b	1					Hi05b09	7	2
			Hi05b09	7	2				CH02b03b1	10	3
			CH01e121	8?	3				CH01f021	12	3
			CH03d08	14	1				Hi04g05	13	1
			NZ02b01	15	1				A380	13?	
			CH05a04	16	1				CH05g11	14	3
									CH05a04	16	2
									A343		4

'Prima' x 'Golden Delicious'

'Golden Delicious'						'Prima'					
			Hi02c07	1	1	CH05c07	9	38 %	Hi22f12	5	1
			Hi12c02	1		Hi07b06	6	40 %	A401	5?	
			CH02f061	2	2				CH03c01	6	2
			Hi02d04	10	1				Hi05b09	7	2
			CH03h03	13	1?				CH01f03b	9	2

			A372	13						Hi02d04	10	1
			CH03b10	15	1					CH03h03	13	2
										CH05g03	17	2

Table 19. Summary of linkage groups represented in each of the F1 and parental genetic linkage maps constructed.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
F1 MAPS																		
'Anna' x 'Golden Delicious'		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
'Anna' x 'Sharpe's Early'		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
'Golden Delicious' x 'Prima'		X	X	X	X	X			X	X	X	X	X		X	X		X
PARENTAL MAPS																		
'Anna' x 'Golden Delicious'	**'Anna'**	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X
	'Golden Delicious'	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X
'Anna' x 'Sharpe's Early'	**'Anna'**	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X
	'Sharpe's Early'	X	X	X	X	X	X		X	X	X	X	X		X	X		X
'Golden Delicious' x 'Prima'	**'Golden Delicious'**			X	X	X			X	X	X	X	X		X	X		X
	'Prima'	X		X	X				X		X	X	X		X	X		

After the grouping of markers into LGs the marker order within each group was determined. The mean chi-square contribution of each marker within a LG was studied in order to determine whether the marker fits well between neighbouring loci. A number of markers were removed (Table 20) as their presence lead to unexpected events of double crossovers, or they caused insufficient linkage between markers due to the fact that they are situated far away from neighbouring markers.

Table 20. Markers excluded from genetic linkage groups during the determination of marker order, due to resulting high chi-square values and insufficient linkage determination.

Mapping population		Marker eliminated
F1 genetic linkage maps		
'Anna' x 'Golden Delicious'		NZ23g04 CH05c07 CH05h05 Hi07d11 Hi07f01
'Anna' x 'Sharpe's Early'		CH01f02 CH05a05 CH01f12 CH01h10 CH03c02 CH01d08 CH01e12-1 A238 A401
'Golden Delcious' x 'Prima'		CH02b03b CH02a04 CH03c01 A193 A209 A372 Hi06b06
Separate parental genetic linkage maps		
'Anna' x 'Golden Delicious'	**'Anna'**	NZ23g04 A30 CH01e12-1 CH04g09 CH05h05 A234 Hi07d11 Hi07h02
	'Golden Delicious'	NZ23g04 CH05a02y A307 Hi04b12
'Anna' x 'Sharpe's Early'	**'Anna'**	CH05h05 A310 Hi05d10
	'Sharpe's Early'	CH01d08 A383
'Golden Delicious' x 'Prima'	**'Golden Delicious'**	Hi03a03
	'Prima'	-

LG 1: Reasons for markers expected to form part of LG 1 not showing sufficient linkage (Table 17 & 18) can mostly be ascribed to markers being situated on the end of the LG, or markers situated far apart on the same LG, as illustrated during the construction of the parental map 'Anna', when studying the mapping population derived from a cross between this cultivar and 'Sharpe's Early' (Figure 39). The markers CH05g08 (Liebhard, *et al.*, 2002) and KA4b (Yamamoto, *et al.*, 2002b) are both situated some distance away from the other two markers, CH03g12 (Liebhard, *et al.*, 2002) and Hi02c07 (Silfverberg-Dilworth, *et al.*, 2006), for which segregation ratios were studied in this mapping pedigree. The tendency of this LG to 'break' into two groups due to the genetic distance between markers Hi02c07 and Hi12c02 (Silfverberg-Dilworth, *et al.*, 2006) might be overcome with the implementation of the marker AG11, for which the sequence data has not yet been published.

LG 2: The failure of LG 2 markers to show significant levels of recombination to enable their grouping to the same LG, when studying the F1 mapping population derived from a cross between 'Anna' and 'Sharpe's Early' (Figure 35 & Table 17), is difficult to explain. It is only after construction of the separate parental genetic linkage maps (Figure 36) that the conclusion can be made that these markers are all situated some distance apart and that this, together with the fact that the two parental meiosis are studied simultaneously, caused the recombination frequency observed between them to resemble that of unlinked markers.

LG 3: The newly developed marker, A310, was shown to form part of this LG in two F1 mapping pedigrees (Figure 37), as well as in the genetic

linkage maps constructed studying their separate parental meiosis (Figure 38 & 40). The same marker was mapped onto LG 12 in a genetic linkage map constructed using the F1 population derived from the cross made between 'Anna' and 'Sharpe's Early' (Figure 35), but eliminated from this LG during genetic linkage map construction in the parent 'Anna', due to resulting high chi-square values. During the implementation of the multilocus marker Hi04c10 (Silfverberg-Dilworth, *et al.*, 2006), two loci were also detected and scored when implementing this marker on the 'Anna' x 'Golden Delicious' population (Figure 34), but both these mapped to LG 3, and not to LG 3 and LG 4 as reported. Since it was difficult to determine which of these correspond to the locus Hi04c10y mapped onto LG 3 (Silfverberg-Dilworth, *et al.*, 2006), the locus producing the smaller fragments were labeled Hi04c10y and the other, of which the segregation were also studied and scored in the population derived from a cross between 'Anna' and 'Sharpe's Early' (Figure 35), Hi04c10?. The position of both these were confirmed during the construction of separate parental genetic linkage maps for these populations (Figure 38 & 39).

LG 4: The fact that the marker Hi07b02 do not show sufficient linkage (Table 17) to either of the two segments comprising LG 4 when studying the 'Anna' x 'Sharpe's Early' mapping pedigree (Figure 35), might be ascribed to the genetic distance between this marker and the two segregating markers studied on either side of this marker. A previously published but unmapped marker, CH01b09b (Liebhard, *et al.*, 2002), was assigned to this LG based on the F1 genetic linkage maps (Figure 37) and separate parental genetic linkage maps (Figure 38 & 39) of two mapping pedigrees.

LG 5: Although the marker CH03e03 were reported to map onto LG 3 (Liebhard, *et al.*, 2002), this marker was mapped onto LG 5 in all three F1 mapping populations studied (Figure 37) as well as in their separate parental genetic linkage maps (Figure 38-40). Since the fragments observed are similar in size (211-218 bp) to those reported for 'Fiesta' (198 & 206 bp) and 'Discovery' (198 & 216 bp), it can be concluded that the same locus has been studied and in this case results regarding the position of this marker obtained during this investigation contradicts those reported by Liebhard *et al.* (2002) and Silfverberg-Dilworth *et al.* (2006). The marker NZ23g04 (Guilford, *et al.*, 1997), reported to map to LG 6 (Liebhard, *et al.*, 2002), were also mapped to LG 5 in the 'Anna' x 'Sharpe's Early' mapping population (Figure 35) as well as in the genetic linage maps derived from studying the separate parental meiosis' (Figure 39). Although this marker also mapped to this LG in the F1 and parental maps of the 'Anna' x 'Golden Delicious' population, the resulting chi-square value lead to uncertainty regarding the correct placement of this marker and the marker was therefore eliminated (Table 20). The resulting chi-square values when studying this marker in the genetic linkage maps obtained when studying the 'Anna' x 'Sharpe's Early' mapping population are in accordance with a marker showing a good fit within the LG it has been assigned to. The position of this marker is again in contradiction to the position published. The marker CH05d04 (Liebhard, *et al.*, 2002), although reported to map to LG 12, yielded fragments slightly bigger than expected and the placement of this marker onto LG 5 in both the F1 genetic linage map (Figure 34) and the parental map 'Anna' (Figure 38) when studying the mapping population derived from a cross between 'Anna' and 'Golden

Delicious' could be due to the amplification of a different locus. The previously published but unmapped marker, CH04f04 (Liebhard, *et al.*, 2002), were mapped to LG 5 during genetic linkage map construction of the F1 'Anna' x 'Golden Delicious' mapping population (Figure 34). This marker is heterozygous for the same two alleles in the two parental cultivars involved and can thus not be mapped in the separate parental genetic linkage maps.

LG 6: The genetic linkage maps constructed for the F1 mapping population derived from a cross between 'Golden Delicious' and 'Prima' and their separate parental meiosis did not yield a LG 6 (Figure 36 & 40). This is due to the fact that only one published marker forming part of this LG, CH03c01 (Liebhard, *et al.*, 2002), has been studied in this population. The marker CH05a05 (Liebhard, *et al.*, 2002) did not show sufficient linkage with other markers from this group when implemented on the 'Anna' x 'Golden Delicious' population, resulting in this LG appearing in two segments (Figure 34). From the 5 other published markers on LG 6 that were implemented on this population, only one was found to be heterozygous in 'Golden Delicious', CH03c01 (Liebhard, *et al.*, 2002), but again insufficient linkage was found to enable the grouping of these two markers when studying the parental meiosis of the 'Golden Delicious' cultivar, resulting in LG 6 only being represented by the parental cultivar 'Anna' (Figure 38).

LG 7: This LG is again not represented in the genetic linkage maps constructed for the F1 as well as the separate parental maps resulting from a cross between 'Golden Delicious' and 'Prima', since the marker Hi05b09 (Silfverberg-Dilworth, *et al.*, 2006) is the only segregating marker studied in this population. Although this marker was studied in both the other mapping

pedigrees, it has not been found to have sufficient linkage with other markers from this LG, whether F1 genetic linkage maps (Table 17) or genetic linkage maps studying separate parental meiosis (Table 18) were constructed. This could be due to the genetic distance between this marker and other markers on this LG. The upper part of this LG were successfully constructed for the F1 mapping populations derived from crosses made between 'Anna' and 'Golden Delicious' (Figure 34) and ' Anna' and 'Sharpe's Early' (Figure 35), but the only parental genetic linkage map containing a segment of this LG is the 'Golden Delicious' parental genetic linkage map constructed from the 'Anna' x 'Golden Delicious' mapping population (Figure 38).

LG 8: Previously published but unmapped marker, CH01e12 (Liebhard, *et al.*, 2002), was successfully mapped to LG 8 in the 'Anna' x 'Golden Delicious' and 'Prima' x 'Golden Delicious' F1 genetic linkage maps (Figure 37) as well as in the parental genetic linkage maps (Figure 38 & 40), although it was eliminated during the construction of the parental map 'Anna' in the 'Anna' x 'Golden Delicious' mapping population (Figure 38) due to an increase in the chi-square value.

LG 9: The previously published but unmapped marker, CH05a03 (Liebhard, *et al.*, 2002), was successfully mapped to LG 9 in the genetic linkage map constructed using the F1 mapping population resulting from a cross between 'Anna' and 'Golden Delicious'. The position of this marker could not however be confirmed through the construction of separate parental genetic linkage maps as the parents share the same two parental alleles. The marker Hi01d01 (Silfverberg-Dilworth, *et al.*, 2006), indicated as a possible multilocus marker by Silfverberg-Dilworth *et al.*, amplified two loci, one

mapping to the top of the LG when studying the F1 mapping population derived from a cross between 'Anna' and 'Sharpe's Early' and the other lower down on the same genetic linkage map as well as the F1 genetic linkage map derived from a cross between 'Anna' and 'Golden Delicious'. The marker mapping to the central region of the LG was confirmed by the different parental genetic linkage maps (Figure 38 & 39) but the position of the locus mapping to the top of the LG could not be confirmed as it was heterozygous for the same two alleles in the parentals of the mapping population in which it has been studied.

LG 10: CH05b06 (Liebhard, *et al.*, 2002) is a multilocus marker reported to amplify 3 loci that have been mapped to LG 5 and LG 16. During this study this marker amplified two loci, but the fragment sizes obtained for one of these loci and for which segregation analysis were performed on the 'Anna' x 'Golden Delicious' mapping population, are slightly larger than the fragments reported. It is thus clear that this marker also amplifies a locus mapping to the top of LG 10 (Figure 34), but since the two parentals involved in the formation of this mapping population share the same two alleles, accurate positioning could not be confirmed by studying the separate parental meiosis. The previously published but unmapped marker, CH01e09b (Liebhard, *et al.*, 2002), was successfully assigned to this LG in genetic linkage maps constructed for the F1 population as well as the separate parental meiosis in the 'Anna' x Golden Delicious' (Figure 34 & 38) and 'Anna' x 'Sharpe's Early' (Figure 35 & 39) mapping pedigrees.

LG 11: The previously published but unmapped marker, CH05c02 (Liebhard, *et al.*, 2002), was successfully assigned to this LG in genetic

linkage maps constructed for the F1 population as well as the separate parental meiosis in the 'Anna' x Golden Delicious' (Figure 34 & 38) and 'Prima' x 'Golden Delicious' (Figure 38 & 40) mapping pedigrees.

LG 12: The previously published but unmapped marker, CH02h11b (Liebhard, *et al.*, 2002), was successfully assigned to this LG in genetic linkage maps constructed for the F1 population in all three mapping pedigrees (Figure 37) as well as in the separate parental meiosis (Figure 38-40).

LG 13: This LG is not represented in genetic linkage maps constructed during the study conducted on the F1 mapping population derived from a cross between 'Golden Delicious' and 'Prima' since the only two segregating markers analyzed, NH009b (Yamamoto, *et al.*, 2002b) and CH03h03 (Liebhard, *et al.*, 2002) are situated far apart. The newly developed marker A372 that mapped to this LG when implemented on other mapping pedigrees also did not show sufficient linkage with either of these markers to enable the formation of a genetic LG. This is a little unexpected as A372 has been shown to be located between these two markers in genetic linkage maps derived from both other mapping populations used (Figure 37).

LG 14: The marker Hi03a03 (Silfverberg-Dilworth, *et al.*, 2006) was mapped by Silfverberg-Dilworth *et al.* (2006) to LG 6, but they also indicated this marker as a possible multilocus marker, amplifying more than one locus. When implemented on the three mapping pedigrees under investigation it was clear that this marker amplified 2 loci, one mapping to LG 6, and although the second locus was found to map to LG 14 in the 'Golden Delicious' x 'Prima, mapping population (Figure 36), it remained unlinked in the other two mapping populations. The fact that it did not show linkage in the 'Anna' x

'Golden Delicious' (Figure 34) and 'Anna' x 'Sharpe's Early' (Figure 32) populations, could be explained by the fact that it lies on the end of the LG. During the construction of genetic linkage maps for the separate contributing meiosis in the case of each mapping population, the second locus amplified by the marker Hi03a03 (Silfverberg-Dilworth, *et al.*, 2006) still remained unlinked in the parental maps constructed for the 'Anna' x 'Golden Delicious' population (Figure 38), due to the large distance between this marker and the rest of the LG 14 markers studied on this population, but it showed significant linkage with other markers of LG 14 when constructing parental genetic linkage maps for the population derived from a cross between 'Anna' and 'Sharpe's Early' (Figure 39). The reason for the inclusion of this specific marker in separate parental maps, while excluded from the genetic linkage map constructed when the two meiosis are studied simultaneously, can be ascribed to the computational difficulties when studying a large number of recombinations.

LG 15: Hi23g12 (Silfverberg-Dilworth, *et al.*, 2006) has been indicated as a marker with the possibility of amplifying more than one locus. Although it has been published as mapping onto LG 4, it has been mapped onto LG 15 in both the 'Anna' x 'Golden Delicious' (Figure 34) and 'Anna' x 'Sharpe's Early' (Figure 35) F1 mapping pedigrees and their corresponding parental genetic linkage maps (Figure 38 & 39), probably due to the amplification of another locus. In the genetic linkage map obtained for the F1 mapping pedigree 'Anna' x 'Sharpe's Early', the orientation of the bottom section of LG 15 could not be accurately determined as it contained only one previously published and mapped marker (Figure 35). After comparison with the marker order obtained

for the F1 mapping pedigree 'Anna' x 'Golden Delicious' (Figure 34), the orientation of the bottom section of LG15 obtained for 'Anna' x 'Sharpe's Early' was inverted (Figure 37). During the construction of integrated genetic linkage maps, segregation ratios from both parental genotypes are viewed simultaneously, making it difficult to accurately determine marker order. During the construction of parental genetic linkage maps the segregation from individual parental cultivars are viewed separately, leading to a more accurate determination of marker order. This would explain the differences between marker order determination for the integrated 'Anna' x 'Golden Delicious' map (Figure 34 & 37) and that obtained for the individual parental genetic linkage maps 'Anna' and 'Golden Delicious' (Figure 38), with the latter corresponding to the marker order on the reference map (Silfverberg-Dilworth, *et al.*, 2006)

LG 16: Very few of the genetic linkage maps constructed contained LG 16. The only genetic linkage maps that include this LG, or segments thereof, are those constructed for the F1 mapping pedigree 'Anna' x 'Golden Delicious' (Figure 34) and their resulting parental maps (Figure 38) as well as the parental map for 'Anna' when studying the separate parental meiosis of the population 'Anna' x 'Sharpe's Early' (Figure 39). Reasons for the lack of representation of this LG in genetic linkage maps are due to a combination of weak markers yielding low levels of amplification products making scoring of results difficult (CH04f10 and CH02d10a (Liebhard, *et al.*, 2002)) and low levels of heterozygosity in parentals used (CH05c06 and CH02a03 (Liebhard, *et al.*, 2002)). CH05a04 (Liebhard, *et al.*, 2002) did not show significant linkage to the rest of the markers from this LG in the F1 mapping population of 'Anna' x 'Golden Delicious' and the parental genetic linkage map constructed

for the 'Anna' parent in both this mapping population and the mapping population 'Anna' x 'Sharpe's Early'. The reason for the total lack of representation of this LG in the genetic linkage maps constructed for the F1 population and contributing parentals in the 'Prima' x 'Golden Delicious' mapping pedigree is that only one segregating marker was successfully studied in this population and this marker has exactly the same two alleles in the two parental cultivars making its use in the construction of parental maps impossible, even in the presence of another heterozygous marker.

LG 17: This LG is well represented in all genetic linkage maps constructed (Figures 34-40), except for the 'Prima' parental map constructed from the 'Prima' x 'Golden Delicious' mapping pedigree, due to the fact that the segregation of only one out of the 4 markers residing on this LG can be studied in the progeny. The other three markers are either not heterozygous in this parental cultivar (CH04c06y and CH02g04 (Liebhard, *et al.*, 2002)) or share the same two alleles with the second parental cultivar used (Hi07h02 (Silfverberg-Dilworth, *et al.*, 2006)).

A total of 27 newly developed SSR markers were successfully mapped onto one or more of the genetic linkage maps obtained (Table 21). Marker A310 mapped to LG 3 in 2 of the F1 mapping pedigrees used (Figure 37) but onto LG 12 in the 'Anna' x 'Sharpe's Early' mapping population (Figure 35 & 37). When studying the separate parental meiosis of the parents involved in each of the mapping pedigrees, the position of this marker is confirmed to LG 3 in all 4 parental maps derived from the mapping populations 'Anna' x 'Golden Delicious' (Figure 38) and 'Prima' x 'Golden Delicious' (Figure 40). The marker A310 is homozygous in the cultivar

'Sharpe's Early' and although heterozygous in 'Anna', is was eliminated during the determination of marker order (Table 20) when studying segregation in the parent 'Anna' from the 'Anna' x 'Sharpe's Early' mapping pedigree as it resulted in a high chi-square value.

Table 21. Summary of newly developed markers that have been mapped onto one or more genetic linkage map. * Different map positions were obtained for these two markers using different mapping pedigrees.

Linkage group	Marker	Genetic linkage maps								
		F1 genetic linkage maps			Parental genetic linkage maps					
		'Anna' x Golden Delicious	'Anna' x 'Sharpe's Early'	'Golden Delicious' x 'Prima'	'Anna' x Golden Delicious		'Anna' x 'Sharpe's Early'		'Golden Delicious' x 'Prima'	
					'Anna'	'Golden Delicious'	'Anna'	'Sharpe's Early'	'Golden Delicious'	'Prima'
LG3	A310*	X		X	X	X			X	X
	A209		X				X	X	X	
LG4	A417	X				X				
LG5	A335	X				X				
	A340	X			X	X		X		
	A279	X				X				
	A401	X		X	X	X			X	
LG9	A383	X	X			X				
	A334		X	X				X	X	
LG10	A253*	X								
	A188			X						X
	A326	X	X			X	X	X		
	A381		X	X			X	X		X
	A398	X								
LG12	A310*		X							
	A219	X				X				
	A331		X					X		
LG13	A193	X				X				
	A372	X	X		X		X			
	A380		X							
LG15	A253*					X				
	A238	X		X		X			X	
	A186	X		X		X			X	
	A320	X				X				
	A424	X	X							
	A244	X	X		X	X	X	X		
LG16	A343	X			X		X			
LG17	A234	X	X			X	X	X		
	A236		X				X	X		

Marker A253 mapped to LG 10 when implemented on the F1 mapping population 'Anna' x 'Golden Delicious' (Figure 34 & 37), but when the parental meioses were studied separately it mapped onto LG 15 of the 'Golden Delicious' parent (Figure 38). Although the placement of a marker is theoretically more accurate when studying segregation from one parent only, the position of this marker can not be statistically determined as was the position of marker A310. The position of this marker thus remain unknown and must be confirmed through further investigation.

4.5 CONCLUSION

Genetic linkage maps were successfully constructed for all three the F1 mapping pedigrees under investigation as well as for the separate parental cultivars used in each. The steps followed during the construction of genetic linkage maps led to the successful positioning of between 71 % and 89 % of the segregating markers implemented on the different mapping populations (Table 16) and the assignment of 27 newly developed SSR markers (Table 21) to 10 different LGs.

Despite the fact that not all the LGs are represented in all genetic linkage maps obtained (Table 19) and that some markers needed to be eliminated (Tables 17, 18 & 20) during the construction of these maps, the LGs that are available can be used successfully for an initial attempt to understand the underlying genetic component/s involved in the regulation of time of initial vegetative budbreak, the goal for genetic linkage map construction during this investigation.

5 CHAPTER 5: BIN MAPPING

5.1 ABSTRACT

The availability of large volumes of sequence data in the public domain has enabled the identification of a large number of SSR-containing sequences that can, through the use of primers designed from unique flanking sequences, be used for genetic linkage map construction (Chapters 3 & 4). These genetic linkage maps can be used for localization of specific genes on the genome or can be used in QTL analysis, a first step towards the identification of genes involved in various economically important traits. Not only is the identification and design of sequence specific primers for each of these identified SSRs a time consuming and expensive process, but the implementation of each of these SSR-markers on a full mapping population is also expensive and time consuming and may, especially for those SSRs residing in a genomic region already saturated with a large number of SSRs, have a low information content. Since the identification of QTLs, using genetic linkage maps, requires an average marker distance of 15 cM and markers spanning the whole genome, a technique needed to be implemented to predict the possible positioning of markers on the genome. This will facilitate the selection of only those markers that reside on regions that are poorly represented by SSR markers and those markers that might lead to the extension of the telomeric ends of LGs, to be implemented on the full mapping population for a more accurate determination of marker position. Selective mapping or bin mapping is a technique that has been used successfully for

this purpose in other crop species and it was tested here for the first time on a F1 mapping pedigree derived from a cross between two highly heterozygous individuals.

It has been shown that as few as six seedlings can be used to determine the possible position of a marker, but the accuracy of positioning increased with an increase in seedling numbers. Not only does a balance need to be found between the cost and time involved in the accuracy of position determination, but also between the ease of PCR setup and data management and position determination. A total of 10 seedlings was proposed for a more accurate determination of possible position and these, together with 2 parental cultivars can easily be managed due to the 8x12 nature of PCR-plates used, enabling the amplification of 8 markers in one PCR experiment. Another option is the initial use of 6 seedlings together with 2 parentals to enable the amplification of 12 markers in one PCR experiment, followed by the amplification using an additional 4 seedlings only for those markers for which position determination could not be accurately determined using 6 seedlings only.

5.2 INTRODUCTION

Only a selection of the newly developed markers that are available (Chapter 3) have been used during the construction of genetic linkage maps (Chapter 4) and the positions of these were determined on one or more of the three different mapping pedigrees used. Large segments of the apple reference map (Silfverberg-Dilworth, *et al.*, 2006) and of the genetic linkage maps constructed during this investigation (Chapter 4) do not contain SSRs.

The QTL detection software used (Chapter 6) does not require genetic linkage maps that are very well saturated but rather maps have markers that are evenly spaced throughout the genome. For this reason it was decided to use selective mapping, also known as bin mapping, to determine the position of newly developed and unmapped SSRs. Marker positions are not accurately determined through the use of this technique, but it facilitates the placement of markers on the map within a certain segment. This reduces the time and cost involved in the preliminary screening of markers in order to find markers residing in those segment that are poorly populated with SSR markers. Bin mapping has been used successfully in species such as wheat (Conley, *et al.*, 2004, Qi, *et al.*, 2004), potato (Isidore, *et al.*, 2003) and prunus (Howad, *et al.*, 2005).

Vision *et al.* (2000) first proposed this two-phase mapping approach for whole genome mapping of molecular markers in 2000. It involves the generation of a high confidence framework map, followed by the addition of new markers to this map based on the genotypic results obtained from a selection of individuals. *Bins*, which are intervals along a linkage group within which no breakpoints occur when viewing the subset of individuals selected, are defined and used as criterion for the possible positioning of new markers. The choice of which individuals to use when applying selective mapping can be arbitrary (random), or software packages (Vision, *et al.*, 2000) can be used to select a subset of individuals that will facilitate the minimization of bin length and the maximization of bin number (Vision, *et al.*, 2000). Larger sample sizes will result in higher breakpoint density, but a balance is necessary between cost and mapping effort involved in genotyping the

selected individuals, and the precision with which these markers can be placed on the reference map.

Bin mapping, or selective mapping, were used to roughly determine the position of newly developed markers on the genetic linkage map of apple by making use of the parental genetic linkage maps derived from a cross made between the cultivars 'Anna' and 'Golden Delicious' (Chapter 4). Those markers that localize to linkage groups, or segments thereof, that are not well characterized by SSR markers, or markers that can not be placed on the genetic linkage map due to undefined bin codes, may then be implemented across the whole mapping population in order to accurately map new markers to these regions. Markers that localize to regions where a putative QTL has been identified can also be implemented on the whole mapping population resulting in a denser genetic linkage map in regions of interest that might enable a more accurate estimation of the position of genes of interest.

5.3 MATERIALS AND METHODS

5.3.1 Defining 'BINS'

The parental genetic linkage maps obtained for 'Anna' and 'Golden Delicious', when studying the F1 mapping population derived from a cross made between these two cultivars (see section 4.3.2), were used to define bins for three sets of seedlings containing 6 (bin set A), 8 (bin set B) and 10 (bin set C) individuals respectively (Table 22). Bins were defined using the grandparental origin of the alleles at each locus, for each of the selected seedlings, as criteria. The binary coding system (0's and 1's) used to define bins are difficult to search and for this reason binary codes where converted

to decimal numbers using Microsoft Excel, which offers this function for up to 10 characters only. For this reason the number of seedlings contained in each bin set were limited to ten or less.

Table 22. Summary of seedlings included in the three different bin sets used.

Seedling number	Bin set A	Bin set B	Bin set C
11	X		X
12		X	X
14	X		X
16	X		X
19	X		X
51	X	X	X
102	X	X	X
105		X	X
277		X	X
320		X	X
395		X	
426		X	
Total	6	8	10

5.3.2 Bin mapping new markers

A total of 45 markers showing some level of heterozygosity when implemented on the parental cultivars 'Anna' and 'Golden Delicious', have been implemented on bin mapping seedlings (Table 23) in order to determine their possible positioning on the genetic linkage map. A selection of markers found to be homozygous on both 'Anna' and 'Golden Delicious' (A29, A180, A202, A227, A230, A232, A256, A283, A284, A285, A300, A380, A388, A395,

A412, A416 and A428), were implemented on bin set A in order to confirm their homozygosity and the absence of possible null alleles.

Table 23. Summary of markers selected for bin mapping. Observed fragment sizes when implemented on the parental cultivars 'Anna' and 'Golden Delicious', the segregation type and the bin set/s used are indicated.

Marker	Fragment sizes		Segregation type	Bin set used		
	'Anna'	'Golden Delicious'		Bin set A (6seedlings)	Bin set B (8seedlings)	Bin set C (10seedlings)
A184	236-239	239	lmxll	X		
A204	210-227	210-219	efxeg	X		
A315	210-222	210	lmxll	X		
A329	208-213	208	lmxll	X		
A332	127-131	131-142	efxeg	X		
A341	142	147-151	nnxnp	X		
A344	407	407-427	nnxnp	X		
A389	253-267	241-253	efxeg	X		
A397	448-510	510	lmxll	X		
A419	154	150-154	nnxnp	X		
A430	216-218	221	lmxll	X		
A440	218	218-224	nnxnp	X		
A466	325-345	325	lmxll	X		
A477	154-156	154	lmxll	X		
A494	143-147	141	lmxll	X		
A497	168-180	168	lmxll	X		
A195	175-181	175	lmxll	X	X	
A222	172-181	181-184	efxeg	X	X	
A236	235-241	232-235	efxeg	X	X	
A245	165-174	174	lmxll	X	X	
A247	256	243-256	nnxnp	X	X	
A267	466-484	484	lmxll	X	X	
A277	221	221-225	nnxnp	X	X	
A301	340-345	331-337	abxcd	X	X	
A314	150-173	148-175	abxcd	X	X	
A352	186	174-186	nnxnp	X	X	
A355	230	222-226	nnxnp	X	X	
A365	229-250	242-244	abxcd	X	X	
A367	214-228	208-232	abxcd	X	X	

A369	160-168	160-164	efxeg	X	X	
A377	455	401-447	nnxnp	X	X	
A379	134-138	120	lmxll	X	X	
A392	139-154	150-154	efxeg	X	X	
A182	478-494	491-494	efxeg			X
A233	319	297-319	nnxnp			X
A235	346-348	356	lmxll			X
A249	294-297	294	nnxnp			X
A274	303	298-303	nnxnp			X
A290	362	334-362	nnxnp			X
A296	427-429	406-408	abxcd			X
A311	244-249	232-237	abxcd			X
A334	215	209-225	nnxnp			X
A386	208-221	208-233	efxeg			X
A393	143-162	158-162	efxeg			X
A408	458-500	500	lmxll			X

To determine the possible positioning of a new marker on the genetic linkage map, the genotypic codes obtained for each of the bin mapping seedlings need to be converted to indicated grandparental origin, as was done during the definition of bins (section 5.3.1.). Because the grandparental origin, or linkage phase, of unmapped markers are not known, both possible origins were considered.

5.3.3 Confirming the positions of bin mapped markers

A selection of 8 markers, including (a) 6 with bin definitions not corresponding to defined bin codes and (b) 2 with bin definitions enabling their possible position on the genetic linkage map, were implemented on the whole mapping population derived fro a cross made between 'Anna' and 'Golden Delicious'. This was done in order to determine if markers (a) extend the linkage groups, resulting in an increase in the number of bins defined for a

specific linkage group, and if markers (b) map to the possible positions identified.

5.4 RESULTS AND DISCUSSION

5.4.1 Defining 'BINS'

The definition of bins can be best explained by viewing a specific linkage group and for this purpose LG 12, obtained during the construction of genetic linkage maps using a F1 mapping pedigree derived from a controlled cross made between 'Anna' and 'Golden Delicious' (see Chapter 4), were selected. During genetic linkage map construction the segregation of markers from each individual parental cultivar were viewed simultaneously for the construction of an F1 map as well as separately for the construction of cultivar specific genetic linkage maps (Figure 41) using the JoinMap® 4 (Van Ooijen, 2006) software package. It is worth noting that, although the marker Hi07f01 (Silfverberg-Dilworth, *et al.*, 2006) has been excluded during the initial steps followed during the determination of genetic linkage groups, it has been successfully assigned to this linkage group when studying the segregation of markers from the 'Golden Delicious' cultivar, based on SCL values (see Chapter 4). The fact that it was not assigned to this linkage group when studying the segregation of markers from both parental genomes simultaneously or studying the segregation of markers from the parental cultivar 'Anna' only, can be ascribed to the genetic distance between this marker and other markers on this linkage group leading to a higher recombination frequency.

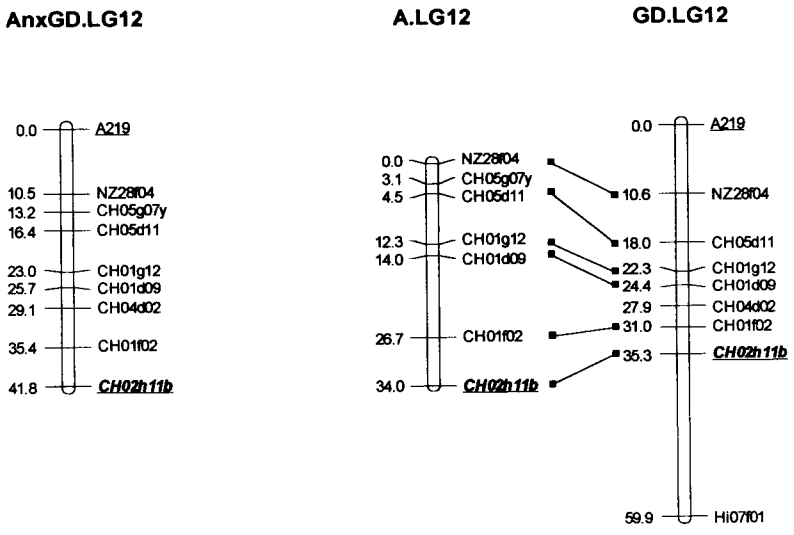


Figure 41. Graphical representation of linkage group 12 obtained during genetic linkage map construction using the F1 mapping pedigree obtained from a controlled cross between 'Anna' and 'Golden Delicious'.

The linkage phase (or grandparental origin) of the markers within each linkage group are, although at first not known, calculated by the JoinMap® 4 (Van Ooijen, 2006) mapping software (Figure 42) (Figure 43) in such a way that the recombinations observed, in the progeny analyzed, are minimal.

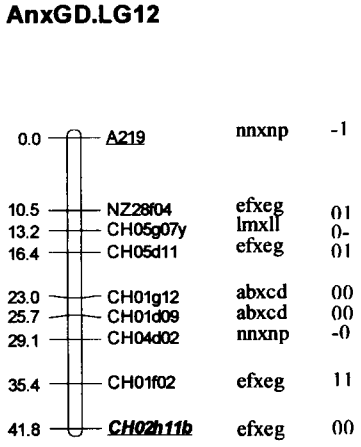


Figure 42. Graphical representation of linkage group 12 obtained from the joint analysis of segregating markers from both parental cultivars. To right of each marker is the segregation type of the specific marker followed by the linkage phase, or grandparental origin, of each marker as calculated by the JoinMap® 4 software package.

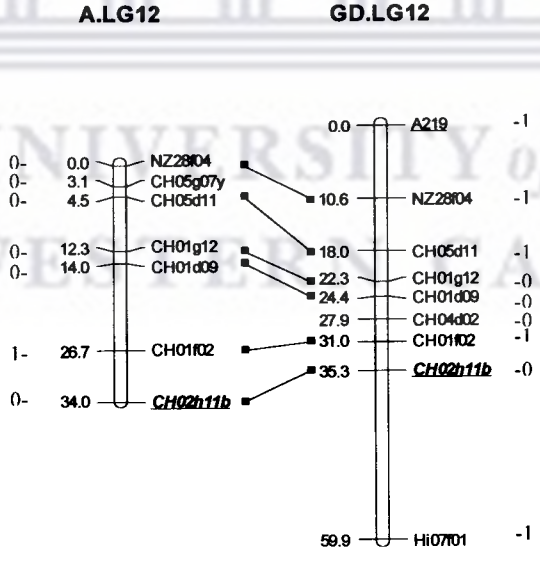


Figure 43. Graphical representation of linkage group 12 obtained during the analysis of markers segregating in each parental meiosis. Next to each marker is the linkage phase, or grandparental origin, of the specific marker as calculated by the JoinMap® 4 software package.

The linkage phase of each marker is indicated as 0 or 1, indicating grandmaternal and grandpaternal origin of the first allele in the combination respectively. A marker with the segregation type abxcd and a linkage phase of [01] thus implies that, when working with the mapping population 'Anna' x 'Golden Delicious', the 'a'-allele is inherited from the maternal parent of 'Anna' and the other allele 'b' is then, by elimination, inherited from the paternal parent of 'Anna'. This is indicated by the '0' in the first position. The '1' in the second position refers to the grandparental origin of alleles in the 'Golden Delicious' parent. The 'c'-allele originates from the paternal parent of 'Golden Delicious' and the 'd'-allele from the maternal parent.

During this investigation seedlings were randomly selected for bin definition purposes, and the resulting 3 bin sets were used to define bins to enable comparison between number of seedlings used and accuracy of bin mapping new markers. After selecting the seedlings to be used for the definition of bins, the genotypic codes of the individual seedlings for each marker on the respective linkage group (Table 24)(Table 25) were changed in order to reflect the grandparental origin of the allele rather than the specific coding system used. This data was then used for the graphical representation of linkage groups using the software package GGT (Graphical Geno Types) (http://www.dpw.wau.nl/pub/ggt/) (Figure 44).

Table 24. Genotypic codes for each of the 12 different seedlings used for bin mapping purposes, when viewing the segregation of alleles on linkage group 12 from the parental cultivar 'Golden Delicious'.

Locus	Segre-gation	Phase	Seedlings											
			11	12	14	16	19	51	102	105	277	320	395	426
A219	<nnxnp>	{-1}	np	np	nn	nn	np	nn	np	np	nn	nn	np	np
NZ28f04	<nnxnp>	{-1}	np	np	nn	np	np	nn	np	np	--	np	np	np
CH05d11	<nnxnp>	{-1}	np	nn	nn	np	np	nn	np	np	nn	np	np	np
CH01g12	<nnxnp>	{-0}	nn	np	np	nn	nn	np	nn	nn	nn	nn	nn	nn
CH01d09	<nnxnp>	{-0}	nn	np	np	nn	nn	np	--	nn	--	--	nn	nn
CH04d02	<nnxnp>	{-0}	nn	np	--	nn	nn	np	nn	nn	--	nn	nn	nn
CH01f02	<nnxnp>	{-1}	np	nn	nn	np	np	nn	nn	np	nn	np	np	np
CH02h11b	<nnxnp>	{-0}	nn	np	--	nn	nn	np	nn	nn	np	nn	nn	nn
Hi07f01	<nnxnp>	{-1}	np	--	--	np	np	nn	np	np	--	--	np	--

Table 25. Genotypic codes for each of the 12 different seedlings used for bin mapping purposes, when viewing the segregation of alleles on linkage group 12 from the parental cultivar 'Anna'.

Locus	Segre-gation	Phase	Seedlings											
			11	12	14	16	19	51	102	105	277	320	395	426
NZ28f04	<lmxll>	{0-}	lm	ll	ll	ll	ll	ll	lm	ll	--	ll	ll	ll
CH05g07y	<lmxll>	{0-}	lm	ll	ll	ll	ll	ll	lm	ll	lm	ll	ll	ll
CH05d11	<lmxll>	{0-}	lm	ll	ll	ll	ll	ll	lm	ll	lm	ll	ll	ll
CH01g12	<lmxll>	{0-}	ll	lm	ll	ll	ll	ll	lm	ll	ll	ll	ll	ll
CH01d09	<lmxll>	{0-}	ll	lm	ll	ll	ll	ll	--	ll	--	--	ll	ll
CH01f02	<lmxll>	{1-}	lm	ll	lm	lm	lm	lm	ll	lm	lm	lm	lm	lm
CH02h11b	<lmxll>	{0-}	ll	lm	--	ll	ll	ll	lm	ll	ll	ll	ll	ll

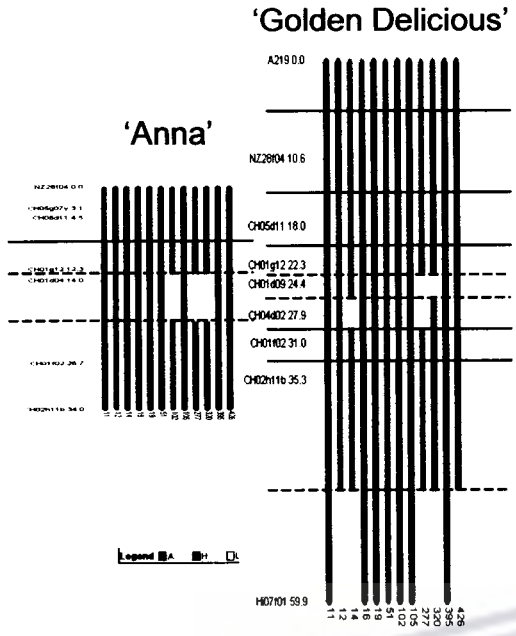


Figure 44. Graphical representation, using the software package GGT, of the grandparental origin of alleles for the 12 seedlings selected to be used during bin definition. Left: Chromosomal segments inherited form the parent 'Anna'. Right: Chromosomal segments inherited from the parent 'Golden Delicious'. In both cases red segments indicate inheritance from the grandmaternal side and green segments indicate inheritance from the grandpaternal side. Missing data point are indicated in grey. Points of recombination during parental meiosis are indicated with a solid horizontal line and, due to missing data, broken horizontal lines are used to indicate possible recombination points.

It is thus clear that, using all 12 seedlings, 2 bins can be defined for the parental meiosis observed In the cultivar 'Anna'. Missing data points at 3 cultivars may result in the possible increase in the number of bins from 2 to 4, but since this will require a double recombination event to occur, which is very unlikely due to chiasma interference, the assumption of 2 bins are most likely correct. For the purposes of bin mapping, all possibilities need to be explored and for this reason the possibility of double recombination will not be excluded. Six bins were defined for the parental meiosis observed in the cultivar 'Golden Delicious'. Missing data points occurring at 3 of the 9 markers mapped to this linkage group could result in an increase of the possible bins from 6 to 9.

Making use of 6 seedlings in bin set A, results in 64 (2^6) possible bin combinations. Parental genetic linkage maps constructed for 'Anna' and 'Golden Delicious' (Chapter 4) consists of 76 and 77 markers respectively and the combination of the segregation patterns observed from both parental meiosis should enable accurate estimation of possible marker positions. When a marker is heterozygous and thus segregating in one parental meiosis only, the chances of more than one possible position increases. It was thus decided to also make use of 8 seedlings (bin set B) and 10 seedlings (bin set C), increasing the possible number of bin combination to 256 (2^8) and 1024 (2^{10}) respectively. Increasing the number of seedlings thus increases the uniqueness of each bin, although not necessarily resulting in an increase in the total number of bins (Table 26).

Table 26. Summary of the number of bins as defined using the three different bin sets on the two parental cultivars 'Anna' and 'Golden Delicious'.

	'Anna'			'Golden Delicious'		
Linkage group	Bin set A (6 seedlings)	Bin set B (8 seedlings)	Bin set C (10 seedlings)	Bin set A (6 seedlings)	Bin set B (8 seedlings)	Bin set C (10 seedlings)
LG1	2	2	2	2	4	4
LG2	3	4	4	3	3	3
LG3	6	6	6	5	5	5
LG4	6	7	7	4	4	4
LG5	10	10	10	9	9	9
LG6	5	5	5	-	-	-
GL7	-	-	-	2	2	2
LG8	4	3	4	3	3	3
LG9	2	2	2	3	4	4
GL10	3	2	3	6	6	7
LG11	2	2	2	5	4	5
LG12	4	4	4	6	9	9
LG13	5	5	5	3	2	3
LG14	3	3	3	1	1	1
LG15	3	2	3	11	11	11
LG16	3	3	4	2	2	2
LG17	3	3	3	4	4	5
TOTAL	64	63	67	68	73	77

Appendix B contains the binary coding used to define the bin sets, as well as values obtained when converting binary codes to decimal numbers.

5.4.2 Bin mapping new markers

5.4.2.1 Determination of possible bin positions

Implementation of newly developed markers on bin mapping seedlings and the subsequent determination of the possible location of these markers are explained by studying the marker A355. When implemented on the parental cultivars 'Anna' and 'Golden Delicious' this marker yielded fragments

of 230 bp and 222-226 bp respectively. During segregation analysis, using both bin sets A and B the presence of a null allele in the parental cultivar 'Anna' became clear. Due to the fact that the grandparental origin (linkage phase) of the alleles from each parental cultivar is unknown, both possible origins need to be considered, meaning that the 230 bp allele inherited from the parental cultivar 'Anna' could have had a grandmaternal (0) or grandpaternal (1) origin. The genotypic coding system used (ab) thus need to be changed in order to resemble possible grandparental origin of allele (Table 27).

Table 27. Summary of the determination of the possible positioning of marker A355 using bin sets A and B.

Cultivar	Fragment sizes (bp)		Coding system	Seedlings								Binary code	Decimal code	Possible position
Bin set A														
				11	14	16	19	51	102					
'Anna'	230-null	Allele	ab	a	a	a	b	a	b					
		Possible grandparental origin	0-	0	0	0	1	0	1			000101	5	LG2,6
			1-	1	1	1	0	1	0			111010	58	X
'Golden Delicious'	222-226	Allele	cd	d	c	d	d	c	d					
		Possible grandparental origin	-0	1	0	1	1	0	1			101101	45	LG2
			-1	0	1	0	0	1	0			010010	18	LG12,15
Bin set B														
				12	51	102	105	277	320	395	426			
'Anna'	230-null	Allele	ab	b	a	b	a	a	a	b	b			
		Possible grandparental origin	0-	1	0	1	0	0	0	1	1	10100011	163	LG2
			1-	0	1	0	1	1	1	0	0	01011100	92	LG5
'Golden Delicious'	222-226	Allele	cd	c	c	d	d	c	c	c	d			
		Possible grandparental origin	-0	0	0	1	1	0	0	0	1	00110001	46	LG2
			-1	1	1	0	0	1	1	1	0	11001110	206	X

The binary code, or the representing decimal number, can then be used to search the bins defined for the parental cultivar involved for possible positioning of the marker on the genetic linkage map. The results obtained when marker A355 (Table 27) is implemented on both bin set A and B, enables the possible positioning of this marker on LG 2 and the determination of linkage phase [00]. The possible positions identified on LG 5, 6, 12 and 15, due to individual bin duplication, are eliminated by bin positions obtained in the other parental cultivar. The possible positioning of this marker could easily have been determined by the use of the smaller number of seedlings contained in bin set A, but the larger number of seedlings in bin set B enabled the more accurate positioning of this marker to the bottom of LG 2, close to CH02a04y. This more accurate positioning are due to the fact that the number of possible bins defined for LG 2 are increased from 3 to 4 when using the larger bin mapping set B.

A total of 16 newly developed markers, proven to be heterozygous in either one or both the parental cultivars used, have been implemented on bin mapping set A (Table 23), enabling the positioning of very few markers to one segment only (Table 28).

Table 28. Summary of markers implemented on bin set A, the fragments observed in the respective parents 'Anna' and 'Golden Delicious' (GD), segregation type, binary codes considering both grandparental origins, the conversion of binary code to decimal as well as possible positions resulting from individual parental meiosis and definite positions being determined through the combination of both parental meiosis or in the case of single parent segregation, unique bins.

Marker	Parent	Fragment sizes	Code	Binary code if linkage phase is 0		Binary code of linkage phase is 1		Decimal		Possible position		Definite position
					Possible interpre-tation of missing data		Possible interpre-tation of missing data					
A184	Anna	236/239	ml	101001		010110		41	22	x	LG 5	LG 5
	GD	239	ll									
A204	Anna	210/227	ef	001101		110010		13	50	LG 13	LG 3,15	LG 15
	GD	210/219	eg	010001		101110		17	46	LG 7,14,15	LG 15	
A315	Anna	210/222	lm	001010		110101		10	53	LG 17	LG 1,16	?
	GD	210	ll									
A329	Anna	208/213	lm	1U0100		0U1011						?
					110100		011011	52	27	LG 11	LG 4	
					100100		001011	36	11	LG 5,8,11	LG 9	
	GD	208	ll									
A332	Anna	127/131	fe	110101		001010		53	10	LG 1,16	LG 17	LG 1

Marker	Parent	Fragment sizes	Code	Binary code if linkage phase is 0		Binary code of linkage phase is 1		Decimal		Possible position		Definite position
					Possible interpre-tation of missing data		Possible interpre-tation of missing data					
	GD	131/142	eg	001000		110111		8	55	LG 1,10,13	LG 5	
A341	Anna	142	nn									
	GD	147/151	np	UU1111		UU0000						
					111111		110000	63	48	LG 9	LG 8,11,15	?
					101111		100000	47	32	LG 9,15	LG 8,11,15	
					011111		010000	31	16	LG 9	LG 11,15	
					001111		000000	15	0	LG 9,15,17	LG 10,11,15	
A344	Anna	407	nn									
	GD	407/427	np	UU1111		UU0000						
					111111		110000	63	48	LG 9	LG 8,11,15	?
					101111		100000	47	32	LG 9,15	LG 8,11,15	
					011111		010000	31	16	LG 9	LG 11,15	
					001111		000000	15	0	LG 9,15,17	LG 10,11,15	
A389	Anna	253/267	ef	U10001		U01110						
					110001		101110	49	46	LG 16	LG 5	
					010001		001110	17	14	LG 14	LG 1,2,8	LG 1 or LG 2
	GD	241/253	eg	U00101		U11010						
					100101		111010	37	58	LG 1,2	x	
					000101		011010	5	26	LG 10	LG 12	
A397	Anna	548/510	lm	000001		111110		1	62	LG 12,14	x	?
	GD	510	ll									

Marker	Parent	Fragment sizes	Code	Binary code if linkage phase is 0	Possible interpre-tation of missing data	Binary code of linkage phase is 1	Possible interpre-tation of missing data	Decimal		Possible position		Definite position
A419	Anna	154	nn									?
	GD	150/154	np	010000		101111		16	47	LG 11,15	LG 9,15	
A430	Anna	216/218	lm	100010		011101		34	29	LG 3,17	x	?
	GD	221	ll									
A440	Anna	218	nn									?
	GD	218/224	np	110111		001000		55	8	LG 5	LG 1,10,13	
A466	Anna	325/345	lm	000000		111111		0	63	LG 12,14	LG 9	?
	GD	325	ll									
A477	Anna	154/156	ef	101110		010001		46	17	LG 5	LG 14	
	GD	154/-	eg	1U110U		0U001U						
					111101		010011	61	19	x	LG 3,5,12,15	LG 5
					111100		010010	60	18	x	LG 12,15	
					101101		000011	45	3	LG 2	LG 3	
					101100		000010	44	2	x	LG 12	
A494	Anna	143/147	lm	100000		011111		32	31	x	LG 2,4,9	?
	GD	141	ll									
A497	Anna	168/180	lm	U00000		U11111						
					000000		111111	0	63	LG 12,14	LG 9	?
					100000		011111	32	31	x	LG 2,4,9	
	GD	168	ll									

The inability to assign markers to a single linkage group using only the six seedlings in bin set A are mostly due to the level of heterozygosity, enabling the study of segregation from one parent only (11 out of 16). Since a lot of bin duplication occurs when using a smaller bin set, the assignment of markers segregating in one parental meiosis only, varied from one possible bin position for A184, to 5 possible bin positions for A329, A341, A344 and A497 respectively. This bin set can however be used to select markers that may have a possible position on a specific linkage group. For example, if a QTL has been identified on LG 17, the densification of this linkage group may enable a more accurate positioning of the QTL and a closer linkage between a marker and the gene of interest, facilitating MAS. When viewing the results obtained in Table 8, four markers (viz. A315, A341, A344 and A430) can be selected for further analysis.

Markers that are heterozygous, and thus segregating, in both parental cultivars can be more accurately assigned to specific chromosomal segments. Three markers were assigned to specific LG s (A204, A332 and A477) while the fourth were assigned to two possible positions (A389) (Table 29). One should however keep in mind that the possibility exists for bin duplication in an unrepresented section of the genome and accurate positioning can only be done using the whole mapping population (see next section).

Marker A477 also represents a case where the presence of a null allele has been confirmed in one of the parental cultivars used. This became clear

when some of the seedlings tested (3 out of 6) had amplification products for the 156 bp allele from the parental cultivar 'Anna' only with no amplified fragment from the parental cultivar 'Golden Delicious'. The segregation type is thus not of the expected type <lmxll>, but rather <efxeg>. The fact that the non-null allele from the 'Golden Delicious' parental cultivar is shared between the two cultivars used as parents, makes the interpretation of results more complicated. Seedlings showing a single fragment of 154bp only can thus be of the genotypic class 'ee' or 'eg'. While this still renders segregation information from the parental cultivar 'Anna' fully informative, these seedlings need to be treated as having missing data when studying segregation from the parental cultivar 'Golden Delicious'. Although the marker has two possible positions when viewing segregation from the parental cultivar 'Anna' only, the positioning on LG 5 can be confirmed by studying all possible interpretations of the missing data from the parental cultivar 'Golden Delicious'.

The level of success for the determination of possible positions of markers on linkage groups was low; using only six bin mapping seedlings (bin set A) only 5 markers were assigned to one or two possible bin positions only (5 out of 16 or 31 %) (Table 29). Therefore another set of 17 markers showing some level of heterozygosity on the two parental cultivars tested, were implemented on both bin sets A and B (Table 23) resulting in more (11 out of 16 or 69 %) of the markers being assigned putative bins (Table 29). Although accurate bin positioning was not always possible using either one of the bin sets, the position of, for example, A195 can be determined by combining the data yielded from both bin sets. Combining the data obtained for marker A222 results in the possible positioning of this markers to LG 14 in the parental

cultivar 'Anna', but the fact that no corresponding bins have been identified through the use of both bin sets studying segregation in the parental cultivar 'Golden Delicious', is an indication that the specific bin in which this marker resides have not been identified for this specific parent. Studying the parental genetic linkage maps depicted in Figure 38 of Chapter 4 together with the bin definitions as listed in Appendix B, leads to the conclusion that this marker must be located close to the marker CH03g04 (see confirmation in next section). Similarly marker A301 could not be positioned at all on the parental map derived for 'Anna', but through the use and comparison of results from 'Golden Delicious' using both bin set A and B, was assigned a possible position to LG 12. Closer investigation into the genetic linkage maps (Figure 38, Chapter 4) and defined bins (Appendix B) lead to the conclusion that this marker must be positioned at the bottom of this linkage group, close to the marker Hi07f01, a region not represented in the map of 'Anna'.

Table 29. Summary of markers implemented on bin sets A and B, the fragments observed in the respective parents 'Anna' and 'Golden Delicious' (GD), segregation type, binary codes considering both grandparental origins, the conversion of binary code to decimal as well as possible positions resulting from individual parental meiosis and definite positions being determined through the combination of both parental meiosis or in the case of single parent segregation, unique bins.

Marker	#	Parent	Fragment sizes	Code	Binary code if linkage phase is 0	Possible interpretation of missing data	Binary code of linkage phase is 1	Possible interpretation of missing data	Decimal		Possible position		Definite position
A195	8	Anna	175/181	lm	01101U10		10010U01						LG 5
						01101110		10010101	110	149	x	LG 5	
						01101010		10010001	106	145	x	LG 10	
		GD	175	ll									
	6	Anna	175/181	lm	010011		101100		19	44	LG 3	LG 5,8,13	
		GD	175	ll									
A222	8	Anna	172/181	fe	00110U00		11001U11						LG 14*
						00110000		11001111	48	207	x	x	
						00110100		11001011	52	203	LG 14	x	
		GD	181/184	eg	00110U10		11001U01						

Marker	#	Parent	Fragment sizes	Code	Binary code if linkage phase is 0		Binary code of linkage phase is 1		Decimal		Possible position		Definite position
						Possible interpre-tation of missing data		Possible interpre-tation of missing data					
						00110110		11001101	54	205	x	LG 12	
						00110010		11001001	50	201	LG 4	LG 12	
	6	Anna	172/181	fe	000001		111110		1	62	LG 12,14	x	
		GD	181/184	eg	010001		101110		17	46	LG 7	LG 15	
	8	Anna	235/241	ef	11111110		00000001		254	1	x	x	
		GD	232/235	ge	00000000		11111111		0	255	x	x	
A236	6	Anna	235/241	ef	111111		000000		63	0	LG 9	LG 12,14	?
		GD	232/235	ge	110100		001011		52	11	LG 8	LG 17	
	8	Anna	165/174	ml	110111UU		001000UU						
						11011111		00100000	223	32	LG 5	x	
						11011101		00100001	221	33	x	x	
						11011100		00100010	220	34	LG 17	x	
A245						11011110		00100011	222	35	LG 5	X	?
		GD	174	ll									
	6	Anna	165/174	ml	111110		000001		62	1	x	LG 12,14	
		GD	174	ll									
A247	8	Anna	256	nn									LG 3 or LG 10
		GD	243/256	pn	1011U011		0100U100						
						10111011		01000100	187	68	x	LG 3,5	
						10110011		01001100	179	76	LG 10	LG 5,12	
	6	Anna	256	nn									

Marker	#	Parent	Fragment sizes	Code	Binary code if linkage phase is 0	Possible interpre-tation of missing data	Binary code of linkage phase is 1	Possible interpre-tation of missing data	Decimal		Possible position		Definite position
		GD	243/256	pn	0U1101		1U0010						
						001101		110010	13	50	LG 13	LG 3,15	
						011101		100010	29	34	LG 10,15	LG 17	
	8	Anna	466/484	ml	01001111		10110000		79	176	x	x	
		GD	484	ll									?
A267	6	Anna	466/484	ml	100010		011101		34	29	LG 3, 17	x	
		GD	484	ll									
	8	Anna	221	nn									
		GD	221/225	np	00100U1U		11011U0U						
						00100010		11011101	34	221	x	x	
						00100110		11011100	38	220	LG 3	x	
A277						00100011		11011001	35	217	x	x	LG 15
						00100111		11011000	39	216	LG 3	LG 15	
	6	Anna	221	nn									
		GD	221/225	np	101101		010010		45	18	LG 2	LG 12,15	
													LG 12*
A301	8	Anna	340/345	ab	U0010U00		U1101U11						
						10010100		11101111	148	239	x	x	
						10010000		11101011	144	235	x	x	
						00010100		01101111	20	111	x	x	
						00010000		01101011	16	107	x	x	
		GD	331/337	cd	U0110U11		U1001U00						

Marker	#	Parent	Fragment sizes	Code	Binary code if linkage phase is 0	Possible interpretation of missing data	Binary code of linkage phase is 1	Possible interpretation of missing data	Decimal		Possible position		Definite position
						00110011		11001100	51	204	x	LG 12	
						00110111		11001000	55	200	x	LG 12,17	
						10110011		01001100	179	76	LG 10	LG 5,12	
						10110111		01001000	183	72	LG 10	LG 12	
	6	Anna	340/345	ab	111000		000111		56	7	x	x	
		GD	331/337	cd	101101		010010		45	18	LG 2,	LG 12,15	
	8	Anna	150/173	ab	11110U1U		00001U0U						
						11110111		00001101	247	13	x	X	
						11110110		00001100	246	12	x	X	
						11110011		00001001	243	9	LG 4	X	
						11110010		00001000	242	8	x	X	
A314		GD	148/175	cd	10110U0U		01001U1U						LG 4* or LG 15*
						10110000		01001111	176	79	LG 15		
						10110001		01001110	177	78	x		
						10110100		01001011	180	75	x	LG 15	
						10110101		01001010	181	74	x	LG 15	
	6	Anna	150/173	ab	011011		100100		27	36	LG 4	LG 5,8,11	
		GD	148/175	cd	001101		110010		13	50	LG 9,10	LG 5,15	
A352	8	Anna	186	nn									LG 15
		GD	174/186	pn	0010101U		1101010U						
						00101011		11010101	43	213	LG 15	x	

Marker	#	Parent	Fragment sizes	Code	Binary code if linkage phase is 0	Possible interpretation of missing data	Binary code of linkage phase is 1	Possible interpretation of missing data	Decimal		Possible position		Definite position
						00101010		11010100	42	212	x	x	
	6	Anna	186	nn									
		GD	174/186	pn	110001		001110		49	14	LG 3,7,15	LG 5	
	8	Anna	229/250	ab	10011001		01100110		153	102	LG 11	x	
		GD	242/244	cd	01100110		10011001		102	153	LG 3	LG 11	
	6	Anna	229/250	ab	1U0100		0U1011						
						110100		001011	52	11	LG 11	LG 9	LG 11
A365						100100		011011	36	27	LG 5,8,11	LG 4	
		GD	242/244	cd	0U1111		1U0000						
						011111		100000	31	32	LG 8	LG 8,11,15	
						001111		110000	15	48	LG 9,15,17	LG 8,11,15	
	8	Anna	214/228	ab	0101111		1010000		47	80	x	LG 3,5	
A367		GD	208/232	cd	0011011		1100100		27	100	x	LG 5	?
	6	Anna	214/228	ab	111110		000001		62	1	x	LG 12,14	
		GD	208/232	cd	101101		010010		45	18	LG 2	LG 12,15	
A369	8	Anna	160/168	ef	00111111		11000000		63	192	LG 1	x	LG 1
		GD	160/164	eg	10001100		01110011		140	115	LG 1	x	
	6	Anna	160/168	ef	1U0U01		0U1U10						
						110101		011110	53	30	LG 1, 16	x	
						110001		011010	49	26	LG 16	LG 4	
						100101		001110	37	14	x	LG 1,2,8	

Marker	#	Parent	Fragment sizes	Code	Binary code if linkage phase is 0	Possible interpretation of missing data	Binary code of linkage phase is 1	Possible interpretation of missing data	Decimal		Possible position		Definite position
						100001		001010	33	10	LG 12	LG 17	
		GD	160/164	eg	0U1U00		1U0U11						
						011100		110111	28	55	LG 5,10,11	LG 5	
						011000		110011	24	51	LG 11	LG 3,5,15	
						001000		100111	8	39	LG 1,10,13	LG 1,10,13	
						001100		100011	12	35	LG 5,10	LG 5,10	
A377	8	Anna	455/-	ab	11010011		00101100		211	44	LG 4	LG 13	LG 13
		GD	401/447	cd	01111001		10000110		121	134	x	LG 13	
	6	Anna	455/-	ab	0U0010		1U1101						
						010010		111101	18	61	LG 3,5	LG 13,16	
						000010		101101	2	45	LG 3,5,17	LG 5,13	
		GD	401/447	cd	1U0111		0U1000						
						100111		011000	39	24	LG 5	LG 11	
						110111		001000	55	8	LG 5	LG 1,10,13	
A379	8	Anna	134/138	lm	1111111U		0000000U						?
						11111111		00000001	255	1	LG 4	x	
						11111110		00000000	254	0	x	x	
		GD	120	ll									
	6	Anna	134/138	lm	001011		110100		11	52	LG 9	LG 11	
		GD	120	ll									
A392	8	Anna	139/154	ef	0000001U		1111110U						?

Marker	#	Parent	Fragment sizes	Code	Binary code if linkage phase is 0		Binary code of linkage phase is 1		Decimal		Possible position		Definite position
						Possible interpre-tation of missing data		Possible interpre-tation of missing data					
						00000010		11111101	2	253	x	x	
						00000011		11111100	3	252	x	x	
		GD	150/154	eg	0101010U		1010101U						
						01010101		10101010	85	170		LG 3,15	
						01010100		10101011	84	171			
	6	Anna	139/154	ef	010000		101111		16	47	LG 14	LG 9	
		GD	150/154	eg	011110		100001		30	33	LG 5	LG 3,7,15	

Marker A314 was assigned to LG 4 when studying the parental cultivar 'Anna', but to LG 15 when studying the parental cultivar 'Golden Delicious'. Studying the genetic linkage maps (Figure 38, Chapter 4) and bin definitions (Appendix B) confirmed both these positions as possible, since segments of LG 4 and LG 15 are not represented in the genetic linkage maps obtained for the parental cultivars 'Golden Delicious' and 'Anna' respectively.

A third set of 12 markers were implemented on bin set C, consisting of 10 individuals (Table 23). Segregation results obtained indicated the presence of a possible null allele in 'Anna' for the marker A233 (Table 30). Since the non-null allele present in this cultivar is shared with 'Golden Delicious' and 8 of the 10 seedlings show a single fragment at 319 bp that could represent either 'ee' or 'ef' genotypic codes, this marker is not suitable for bin mapping in the parental cultivar 'Anna'. Similarly the possible presence of a null allele in the cultivar 'Golden Delicious' and the fact that it shares an allele with 'Anna', has rendered the marker A249 only partially informative when studying the segregation of alleles from the parental cultivar 'Golden Delicious'. The fact that only one seedling had a single fragment of 294 bp enabled the consideration of both possible genotypic codes in order to try and determine a possible position for this marker. Neither of these two markers could however be assigned a possible position on the genetic linkage map and therefore are assumed to localize to regions of the genome for which no bin position has been identified. Segregation results for the remaining ten markers on bin set C (Table 31) enabled the identification of one or two possible positions for 6 markers, as well as the total absence of defined bins for another 3 markers

(11 out of 12 or 92 %). The 6 markers for which no possible bin definitions were available, may thus reside on the edges of linkage groups, facilitating the expansion of the genetic linkage map to cover telomeric regions, or it may reside within previously constructed linkage groups. These markers should be implemented on the full mapping population, as they may ultimately lead to better SSR-coverage of the full genome. The remaining marker was assigned to 3 possible linkage groups due to missing data observations at 3 of the 10 genotypes. A more accurate determination of its position will be possible by the addition of marker information for genotype 9 that will facilitate the elimination of at least one of these possibilities.

Table 30. Summary of markers implemented on bin set C, the fragments observed in the respective parents 'Anna' and 'Golden Delicious' (GD), segregation type, binary codes considering both grandparental origins, the conversion of binary code to decimal as well as possible positions resulting from individual parental meiosis and definite positions being determined through the combination of both parental meiosis or in the case of single parent segregation, unique bins.

Marker	Parent	Fragment sizes	Code	Binary code if linkage phase is 0		Binary code of linkage phase is 1		Decimal		Possible position		Definite position
					Possible interpretation of missing data		Possible interpretation of missing data					
A182	Anna	478/494	fe	0010101011		1101010100		171	-172	X	X	?
	GD	491/494	ge	0000101001		1111010110		41	-42	X	X	
A233	Anna	-/319	fe	U0UUUU1UUU						X	X	?
	GD	297/319	ge	1011110111		0100001000		-265	264	X	X	
A235	Anna	346/348	ml	01U1101100		10U0010011						?
					0111101100		1000010011	492	-493	X	X	
					0101101100		1010010011	364	-365	X	X	
	GD	356	ll									
A249	Anna	294/297	ef	0111011101		1000100010		477	-478	X	X	?
	GD	294/-	eg	U011011101								
					1011011101		0100100010	-291	290	X	X	
					0011011101		1100100010	221	-222	X	X	
A274	Anna	303	nn									LG 1, 10 or 13
	GD	298/303	pn	10U01111UU		01U10000UU						

					1010111111		0101000000	-321	320	X	X	
					1010111110		0101000001	-322	321	X	LG 1,13	
					1010111101		0101000010	-323	322	X	X	
					1010111100		0101000011	-324	323	X	LG 1,10	
					1000111111		0111000000	-449	448	X	X	
					1000111110		0111000001	-450	449	X	X	
					1000111101		0111000010	-451	450	X	X	
					1000111100		0111000011	-452	451	X	X	
	Anna	362	nn									
	GD	334/362	pn	11111U111U		00000U000U						
A290					1111111111		0000000000	-1	0	X	X	LG 12
					1111111110		0000000001	-2	1	X	X	
					1111101110		0000010001	-18	17	X	LG 12	
					1111101111		0000010000	-17	16	X	LG 12	
	Anna	427/429	ab	0U0001111U		1U1110000U						
					0100011111		1011100000	287	-288	X	X	
					0100011110		1011100001	286	-287	X	X	
					0000011111		1111100000	31	-32	LG 14	X	
					0000011110		1111100001	30	-31	X	X	LG 14
A296	GD	406/408	cd	1U0111001U		0U1000110U						
					1101110011		0010001100	-141	140	X	LG 14	
					1101110010		0010001101	-142	141	X	X	
					1001110011		0110001100	-397	396	X	LG 15	
					1001110010		0110001101	-398	397	LG 15	X	
A311	Anna	244/249	ab	U000000000		U111111111						LG 4*
					1000000000		0111111111	-512	511	X	LG 4,9	
					0000000000		1111111111	0	-1	X	LG 9	
	GD	232/237	cd	U000000000		U111111111						
					1000000000		0111111111	-512	511	LG 8	X	

					0000000000		1111111111	0	-1	X	X	
A334	Anna	215	nn									?
	GD	209/225	np	0100111100		1011000011		316	-317	X	X	
A386	Anna	208/221	ef	00U1010100		11U0101011						LG 9 or 16
					0011010100		1100101011	212	-213	X	X	
					0001010100		1110101011	84	-85	X	LG 16	
	GD	208/233	eg	00U0000000		11U1111111						
					0010000000		1101111111	128	-129	X	X	
					0000000000		1111111111	0	-1	X	LG 9	
A393	Anna	143/162	fe	1111101101		0000010010		-19	18	X	LG 3	LG 3
	GD	158/162	ge	0111110100		1000001011		500	-501	X	LG 3	
A408	Anna	458/500	ml	0101000000		1010111111		320	-321	X	X	?
	GD	500	ll									

The implementation of 17 markers, believed to be homozygous in both 'Anna' and 'Golden Delicious', on bin set A revealed two markers, A256 and A395, that yielded null alleles in one of the parental cultivars. These markers were then also implemented on bin set B (Table 31). The presence of the null alleles was concluded as seedlings contained in the bin sets used, did not all show amplification of the expected fragment. A256, expected to yield fragments of 180 and 186 bp in all seedlings, yielded only a 186 bp fragment in some seedlings. The only possible explanation is the presence of a null allele in the cultivar 'Golden Delicious', rendering it heterozygous. Similarly A395 yielded one fragment of 219 bp in some seedlings and not the expected fragment of 200 bp, resulting in the conclusion that 'Anna' is the carrier of a null allele. Comparing possible map positions between the two bin sets used, identifies the position of these two markers on LG 1 and LG 10 respectively.

Table 31. Summary of markers, believed to be homozygous, implemented on bin sets A and B, fragments observed in the respective parents 'Anna' and 'Golden Delicious' (GD), segregation type, binary codes considering both grandparental origins, conversion of binary code to decimal as well as possible positions resulting from individual parental meiosis and definite positions being determined through the combination of both parental meiosis or in the case of single parent segregation, unique bins.

Marker	#	Parent	Fragment sizes	Code	Binary code if linkage phase is 0	Possible interpre-tation of missing data	Binary code of linkage phase is 1	Possible interpre-tation of missing data	Decimal		Possible position		Definite positio n
A256	8	Anna	186	nn									LG1
		GD	180/-	np	01110011		10001100		115	140	x	LG1	
	6	Anna	186	nn									
		GD	180/-	np	1U0111		0U1000						
						110111		011000	55	24	LG5	LG11	
						100111		001000	39	8	LG5	LG1,10,13	
A395	8	Anna	200/-	lm	1U010U01		0U101U10						LG10
						11010101		00101010	213	42	x	x	
						11010001		00101110	209	46	x	x	
						10010101		01101110	149	110	LG5	x	
						10010001		01101010	145	106	LG10	x	
		GD	219	ll									
	6	Anna	200/-	lm	U010U0		U101U1						
						101010		110111	42	55	LG15	LG16	
						101000		110101	40	53	x	LG1,16	
						001010		010111	10	23	LG17	x	
						001000		010101	8	21	LG10	LG6	

5.4.2.2 Confirmation of marker positions

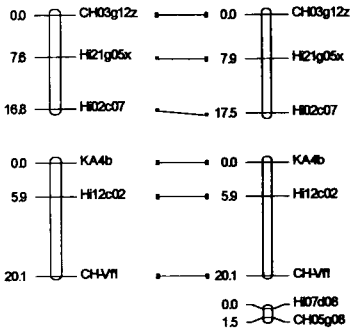
A Selection of 8 markers, implemented on the bin mapping sets mentioned above, were implemented on the full mapping population derived from a cross between the cultivars 'Anna' and 'Golden Delicious' in order to determine their position on the genetic linkage map (Table 32). Map construction for this population was done using criteria as explained in chapter 4, with the exception that the addition of new markers (increasing the total number of markers implemented from 149 to 157) resulted in the exclusion of a different set of individuals due to missing data observations at more than 25 % of the markers used. This accounts for slight differences in the distances observed between adjacent markers on the different linkage groups obtained when studying the segregation from both parental meiosis simultaneously, compared to that from the previous mapping experiment (Figure 45). The ordering of markers on LG 2 and LG 14 are different when comparing the results obtained during the two mapping experiments (Figure 45). The ordering of markers on LG 2 during the first mapping experiment, resulted in higher chisquare values, rendering the order obtained during the second experiment more acceptable due to lower mean chisquare values. None of the markers were excluded during the first mapping experiment due to the fact that they have all been published on this LG during previous studies (Liebhard, *et al.*, 2002, Silfverberg-Dilworth, *et al.*, 2006) and slight differences in marker order also occurred between the published maps of 'Fiesta' and 'Discovery'. The marker order remained exactly the same when studying the separate parental meioses during both mapping experiments (Figure 35 & 46). The ordering of

markers on a large segment from LG 14 were inverted during the second mapping experiment (Figure 45), a phenomenon that could be ascribed to genotyping errors, although it 'reverted' to the order obtained during the first mapping experiment when studying the separate parental meioses (Figure 46). The change in marker order observed between the two integrated maps (Figure 45) may be the result of genotyping errors, but may also be caused by the computational difficulties arising when studying the segregation of alleles from both parents simultaneously. Genetic linkage maps constructed for the separate parental meiosis (Figure 46) have very low chisquare values for all markers involved (between 0.02 and 0.135 for 'Anna' and 0.07 and 0.267 for 'Golden Delicious'). Seven of the 8 newly implemented markers were mapped onto various linkage groups (Table 33) and are indicated in red on the genetic linkage maps derived from the F1 mapping population as well as the separate parental meiosis (Figures 45 & 46). The incorporation of segregation results obtained from the 8 newly implemented markers also made it possible for other markers, that did not previously show significant linkage with any other markers, to be assigned to linkage groups and these markers are indicated in green on the genetic linkage maps (Figures 45 & 46).

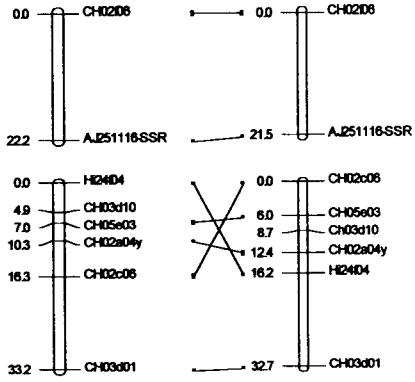
Table 32. Summary of the possible positioning of markers on the genetic linkage map of 'Anna' and 'Golden Delicious', as predicted using bin mapping, and actual positioning as inferred after implementation of markers on the full mapping population.

Marker		Bin mapping results		Mapping results		
	Segregation type	Bin set used	Possible positioning (LG)	F1	'Anna'	'GD'
A222	efxeg	A&B	14	14	14	14
A236	efxeg	A&B	?	17	Not linked	17
A245	lmxll	A&B	?	12	12	
A267	lmxll	A&B	?	16	16	
A314	abxcd	A&B	4/15	Not linked		
A379	lmxll	A&B	?	9	9	-
A392	efxeg	A&B	?	13	13	13
A332	efxeg	A	1	7	7	Not linked

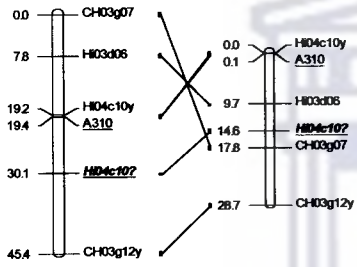
AnXGD.LG01 AnXGD(2).LG01



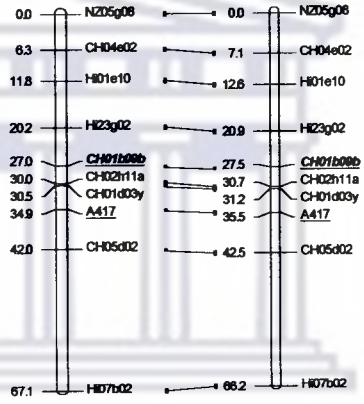
AnXGD.LG02 AnXGD(2).LG02



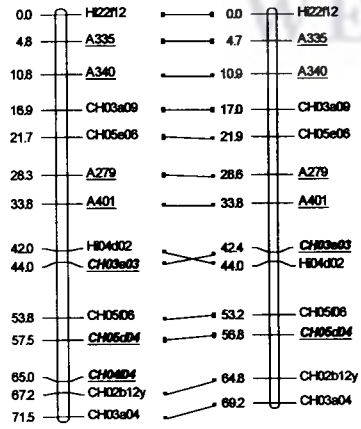
AnxGD.LG03 AnXGD(2).LG03



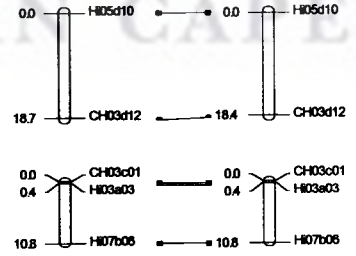
AnxGD.LG04 AnXGD(2).LG04



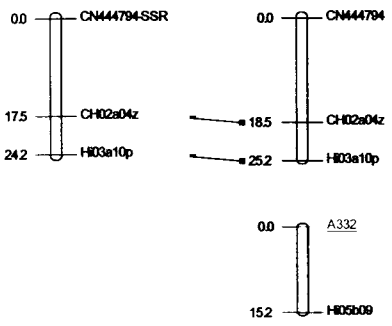
AnxGD.LG05 AnXGD(2).LG05



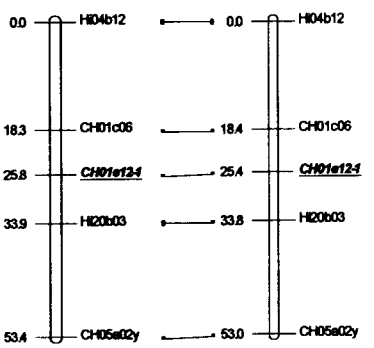
AnXGD.LG06 AnXGD(2).LG06



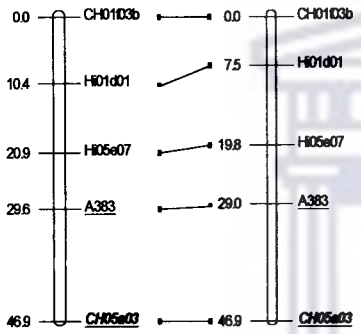
AnXGD.LG07 **AnXGD(2).LG07**



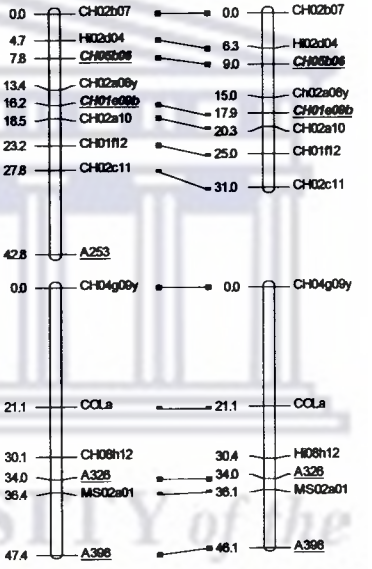
AnXGD.LG08 **AnxGD(2).LG08**



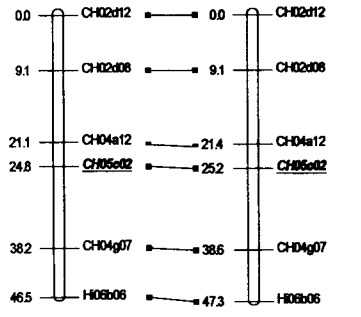
AnXGD.LG09 **AnXGD(2).LG09**



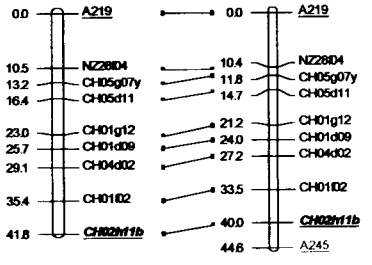
AnxGD.LG10 **AnXGD(2).LG10**



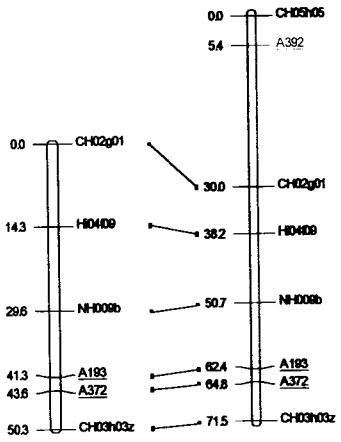
AnXGD.LG11 **AnXGD(2).LG11**



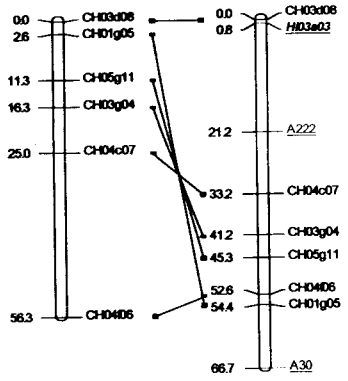
AnxGD.LG12 **AnXGD.LG12**



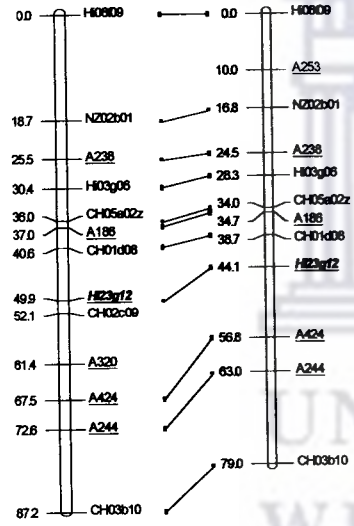
AnXGD.LG13 AnXGD(2).LG13



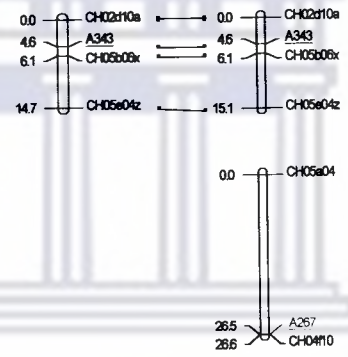
AnxGD.LG14 AnXGD(2).LG14



AnxGD.LG15 AnXGD(2).LG15



AnXGD.LG16 AnXGD(2).LG16



AnXGD.LG17 AnXGD(2).LG17

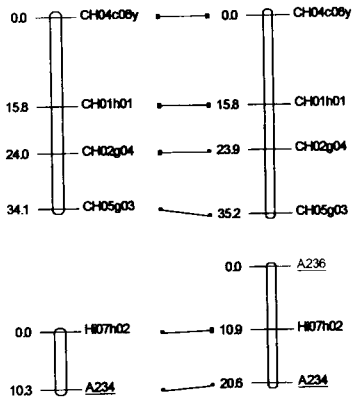
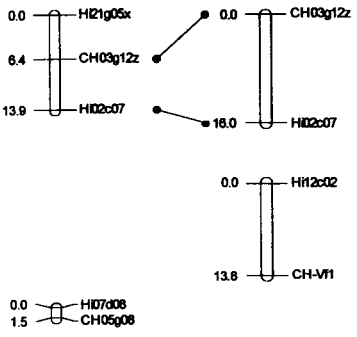
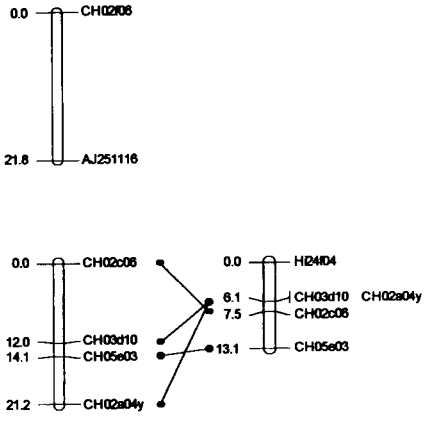


Figure 45. Graphical representation of the alignment of the genetic linkage maps obtained using an F1 populations derived from a cross between 'Anna' and 'Golden Delicious' before (AnxGD) and after (AnxGD(2)) the implementation of an additional 8 markers. Markers indicated in red are the additional markers for which possible bin positions on the genetic linkage map were determined. Markers in green were found not to show significant linkage with other markers to be incorporated into the initial genetic linkage map, but after the addition of segregation results obtained from the additional 8 markers they have been incorporated into the map.

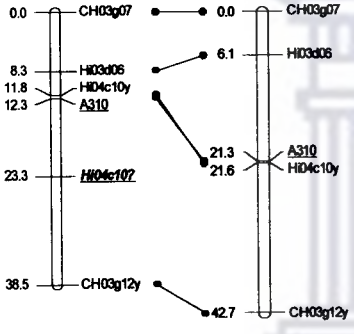
A(2).LG01 GD(2).LG01



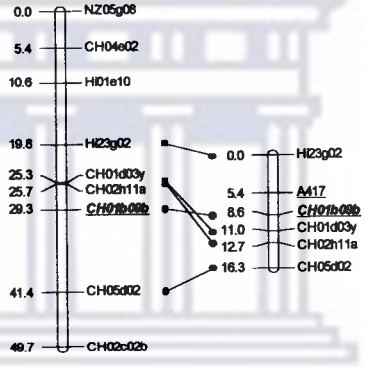
A(2).LG02 GD(2).LG02



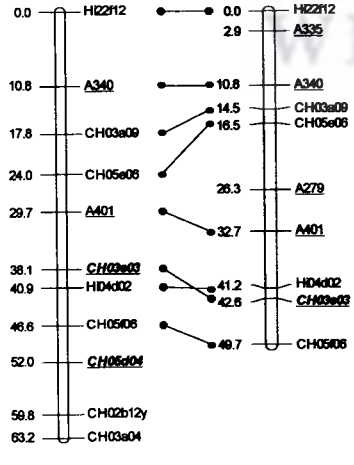
A(2).LG03 GD(2).LG03



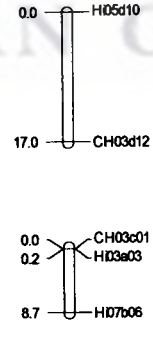
A(2).LG04 GD(2).LG04



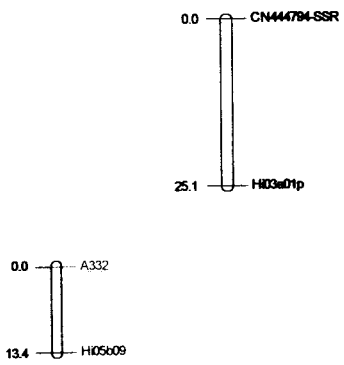
A(2).LG05 GD(2).LG05



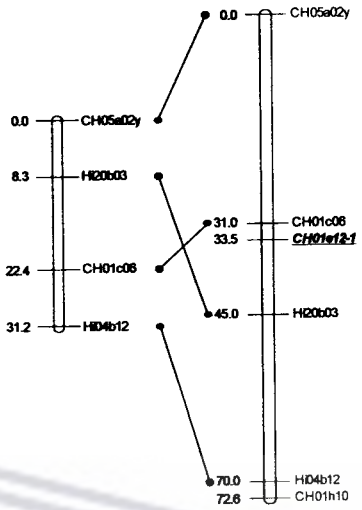
A(2).LG06



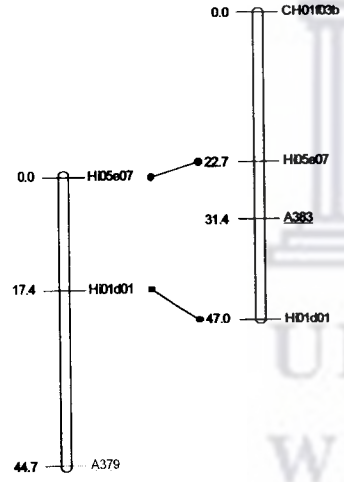
A(2).LG07 GD(2).LG07



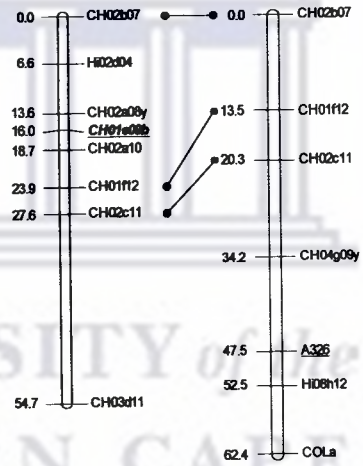
A(2).LG08 GD(2).LG08



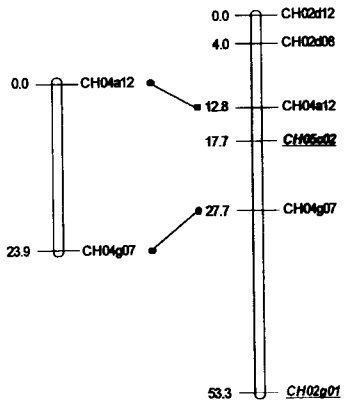
A(2).LG09 GD(2).LG09



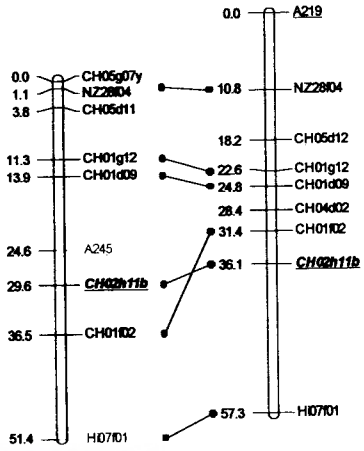
A(2).LG10 GD(2).LG10



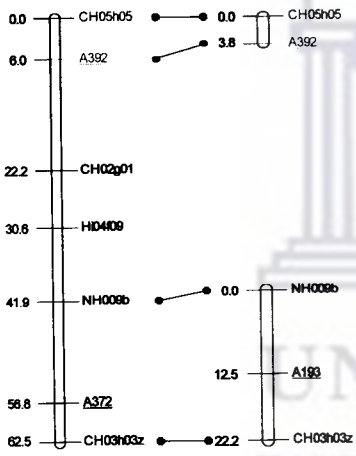
A(2).LG11 GD(2).LG11



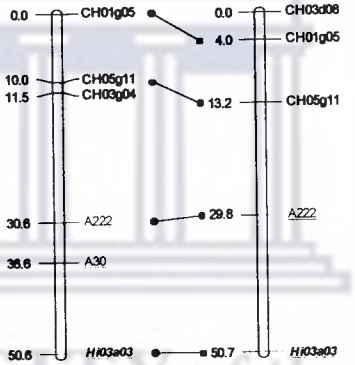
A(2).LG12 GD(2).LG12



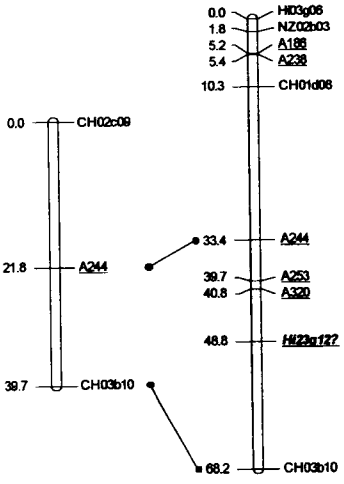
A(2).LG13 GD(2).LG13



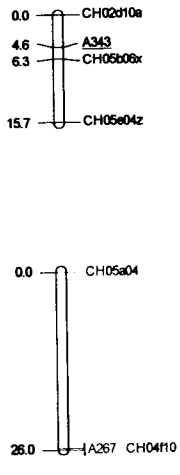
A(2).LG14 GD(2).LG14



A(2).LG15 GD(2).LG15



A(2).LG16



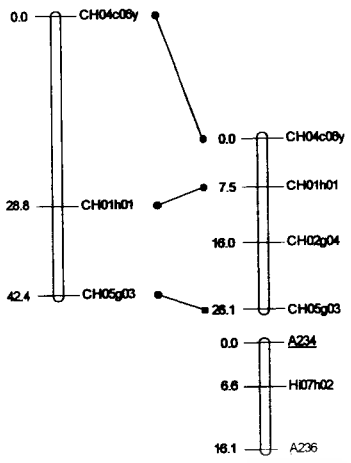


Figure 46. Genetic linkage maps of 'Anna' and 'Golden Delicious', as obtained from the study of segregation results obtained from 157 marker loci, including 8 markers (indicated in red) for which possible positions on the genetic linkage map were determined using selective mapping or bin mapping. Newly developed and mapped SSR markers are underlined. Markers that have been published but for which the map positions were unknown, as well as published markers mapping to different positions, are indicated in italics and underlined. Markers, not previously included in the genetic linkage maps, that were mapped only after the expansion of the genetic linkage maps due to the addition of 8 new markers are indicated in green.

Marker A222 mapped to LG 14, as predicted (Table 30). Although marker A332 did not map to LG 1 as expected (Table 29), it mapped to a position or chromosomal segment not represented in the original map and for which no bin definition was thus available. The identification of LG 1 as a possible position for this marker was based upon 6 seedlings only (bin set A) and it needs to be kept in mind that the possibility for duplicated bins remains. Although segregating in both parental meiosis, A322 is not present in the parental genetic linkage map derived for 'Golden Delicious' due to the homozygotic nature of Hi05b09 in this cultivar. Five of the 6 markers for which no possible position were identified, due to an absence of the specific bin code or due to contradicting results from the two different parental meiosis, were mapped and served to extend the telomeric ends of linkage groups 9, 12, 13, 16 and 17 respectively (Figures 45 & 46). Not only did the addition of new markers aid in the incorporation of markers into these linkage groups, they also facilitated the linkage of previously unlinked markers to various other linkage groups (e.g. LG 8 and LG 10) (Figures 45 & 46), probably as a result of reduced computational difficulty, expanding the total number of markers incorporated in the F1 genetic linkage map from 129 to 140 markers and those in the parental genetic linkage maps from 79 to 96 markers for 'Anna' and 77 to 91 markers for 'Golden Delicious'.

5.5 CONCLUSION

The use of selective mapping, or bin mapping, proved a powerful tool in the identification of the possible position/s of markers on linkage groups. Although precision of marker placement increases with an increase in the

number of seedlings included in the bin set selected, a balance needs to be achieved between the precision needed and the cost and effort involved.

Identification of markers residing on a specific linkage group for the purpose of densifying certain genomic regions, might justify the use of a smaller bin set in order to eliminate all those markers with no possible position identified within the region of interest. Due to the 8 x 12 layout of PCR plates used, the use of 6 seedlings only enabled the incorporation of the 2 parental cultivars in order to amplify 12 markers during one PCR reaction. As the purpose of this investigation relates to cost and time efficiency, the use of 6 individuals is recommended as starting point, as it will enable an initial screening for workability and problem identification (presence of possible null alleles). For more precision in localization the addition of another 4 seedlings/genotypes will increase the bin set to 10 resulting in a 10 digit binary code. The PCR plate can then easily be divided into 2X (4x12) blocks enabling the amplification of 24 loci in one PCR reaction. The use of 8 seedlings, as was done in bin set B, made the plate management slightly more difficult. The power of reduction in the total number of possible positions by increasing the seedling number to 10 (comparing Table 29 and Table 31) justifies the cost involved.

For the purpose of obtaining a genetic linkage map with a better genome coverage and with regularly spaced SSR markers, it is advised that those markers that can not be assigned possible positions on the genetic linkage map due to the absence of specific bin codes, be implemented on the whole mapping population. This might result either in the expansion of genetic linkage groups to also include telomeric ends or the definition of new bins

residing between existing bins. New bin codes can be assigned to expanded maps on a regular basis in order to facilitate the more accurate assignment of possible map positions for subsequent markers. Implementation of the bin mapping technique should thus be seen as a continuous process involving several rounds of genetic linkage map construction and subsequent bin definition. The expectation would be a reduction in the number of markers for which no bin code exists on the genetic linkage map during each round of analysis.

6 CHAPTER 6: QUANTITATIVE TRAIT LOCI (QTL) IDENTIFICATION

6.1 ABSTRACT

Conventional breeding methods have been used for extended periods in many cultivated crop species and although selection of cultivars for breeding are based on phenotypic characteristics, it involves the selection of favourable genes, or more specifically alleles. The level of success when implementing conventional breeding has been much higher for annual plants, or animals with a short generation cycle, than for perennial plants, or animal with a longer generation cycle. The identification of the underlying gene/s affecting a trait of interest, and possible linkage with a molecular marker, will facilitate the selection of phenotypically 'superior' individuals at a very early stage. This is especially promising in perennial tree crops like apple where the phenotypic trait can only be assessed after years of costly field maintenance. The first step towards the identification of such genes is the identification of genomic regions showing some level of association with the trait of interest. This can be done utilizing genetic linkage maps and phenotypic trait assessment data and is known as QTL mapping.

Through the implementation of the maximum likelihood interval mapping approach a total of 18 putative QTLs, associated with the time of initial vegetative budbreak (IVB), were identified in three mapping populations. Fifteen of these QTLs were identified over at least 3 years in any given population, and 4 were identified in different genetic backgrounds.

6.2 INTRODUCTION

Selection of 'naturally' occurring genes have been done for centuries by farmers and breeders of both plants and animals. Only the best individuals, evaluated on the basis of phenotypic traits, are chosen and used as parents for subsequent generations. Through the identification of genes that result in the expression of these desired phenotypic traits, this selection can now be taken one step further to selection of the genotype. Selection is thus done at the DNA level.

Cloning of favourable genes into popular varieties of fruit may lead to minor changes satisfying the ever changing demands of consumers, but the societal issues surrounding the creation of transgenics makes this a much more controversial strategy. Transgenics have been used by various groups in order to confirm the putative function of genes as was done by Kotoda *et al.* (2002) cloning the apple gene *MdMADS5*, a putative homolog of *Arabidopsis* (*AP1*), into *Arabidopsis*, leading to the conclusion that *MdMADS5* is an endogenous apple gene that may be involved in early flowering.

In the past decade studies related to phenotypic traits of economic importance in a variety of apple cultivars, have led to the identification of genes, or regions of the genome that contain genes, contributing to trait variation. Since genetic variation of most phenotypic traits is quantitative in nature and the result of numerous interacting loci, knowledge about the genetic makeup of favourable traits, and being able to determine their status through linkage with molecular markers, enables us to utilize these genes in breeding programs. The genomic regions that contain genes of interest are known as Quantitative Trait Loci (QTL). Any QTL identified is more 'valuable'

from a breeding perspective if it has been identified over several years of study and if it has been studied in different genetic backgrounds. The use of co-dominant microsatellite markers enables the alignment of genetic linkage maps obtained from different cultivars, as well as from closely related species, resulting in more accurate chromosomal position comparison between QTLs identified in different mapping populations (Liebhard, *et al.*, 2003a).

Factors affecting the power of QTL detection include the marker density of the genetic linkage maps, as well as the number of individuals included in the assessment of the phenotypic trait/s under investigation. QTLs can be successfully identified using genetic linkage maps with an average marker distance of 15 cM, as proposed for the apple reference map by Silfverberg-Dilworth *et al.* (2006). Although a higher map resolution will not contribute significantly to the number of QTLs identified, it will facilitate a more accurate, or precise, identification of the genomic regions that contain such a QTL and may even lead to specific gene identification (Seymour, *et al.*, 2002). QTLs can also be identified using partial genetic linkage maps as was done by Stankiewicz-Kosyl *et al.* (2005), using AFLP and SSR markers to screen the F1 progeny derived from a cross between 'Idared' and the clone U 211 in order to study the underlying causal components for resistance to powdery mildew. They were able to align linkage groups with those published by Liebhard *et al.* (2003b) and they identified 10 putative QTLs for powdery mildew resistance. Most of these QTLs could only be associated with resistance in 1 of the 5 years during which phenotypic assessment was performed. This might be explained by different physiological races of powdery mildew being predominant in field isolates in different years or

differential expression during different developmental stages of the plant. One of these QTLs showed acceptable levels of association with phenotypic values over several years and occurs on the same linkage group as previously identified resistance genes. It can thus be a potential locus for MAS when breeding for mildew resistance (Stankiewicz-Kosyl, *et al.*, 2005).

Another factor determining the power of QTL analysis is the number of individuals subjected to phenotypic assessment and it has been reported that the probability of QTL detection increases with an increase in the number of seedlings phenotyped (Van Ooijen, 1992). Using smaller populations will thus only facilitate the detection of QTLs with a very large effect as it increases the difficulty of proving the existence of a segregating QTL.

In this study we make a first attempt towards identifying genomic regions containing QTLs involved in the determination of time of IVB, a character closely associated with dormancy release. This will facilitate the identification of SSR markers linked to favourable alleles in order to be used in MAS to breed apple cultivars that are better adapted to mild winter conditions as experienced in the Western Cape region of South Africa.

6.3 MATERIALS AND METHODS

6.3.1 Genetic linkage maps

Parental genetic linkage maps as well as integrated genetic linkage maps, combining segregating loci from both parentals, were constructed and described in Chapter 4. The linkage maps obtained for the mapping pedigrees 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious', as described in Chapter 4, were used for the identification of QTLs. The linkage maps

obtained for the mapping pedigree 'Anna' x 'Golden Delicious', after the addition of bin mapped markers, as described and constructed in Chapter 5, were also used for the identification of QTLs.

6.3.2 Phenotypic trait data

Phenotypic data, described in Chapter 2, was used for the identification of putative QTLs for time of IVB. In the case of two mapping pedigrees, 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious', data was obtained from both adult trees and the subsequent clonal replicates of these trees (see Chapter 2). These datasets were treated separately in order to facilitate the identification of possible differences between adult and juvenile trees. Data obtained during different years of study were also treated separately throughout this investigation in order to identify QTLs that remain consistent over the period of investigation.

6.3.3 Mapping practice

The software package MapQTL® 5 (Van Ooijen, 2004) was used for the detection of QTLs for each of the 3 different mapping populations used, 'Anna' x 'Golden Delicious', 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious'. QTL analysis was first performed separately for each of the parental maps after which the integrated map was used.

Interval mapping was performed for each year a specific population was subjected to phenotypic assessment. Although LOD significant thresholds of 4.4 and 2.9 are proposed by Van Ooijen (1999) for declaring QTLs identified on integrated (4 QTL alleles) and parental (2 QTL alleles) maps significant, a putative QTL was considered potentially significant during

this study if it had a LOD value above 2 during one or more year/s of investigation.

QTLs were graphically depicted on integrated maps as a bar, using the software MapChart© (Voorrips, 2002). Confidence intervals (CIs) corresponding to a LOD score drop of 0.5 and 1 on either side of the likelihood peak were indicated. QTLs identified, whether from the same population studied over several years, or from different genetic backgrounds, were considered to be the same QTL when they have overlapping CIs.

6.3.4 Nonparametric mapping using Kruskal-Wallis

The nonparametric mapping function used by MapQTL® 5 (Van Ooijen, 2004) is the Kruskal-Wallis test and this was used in order to identify markers in which the different genotypes can be associated with differences in the mean time of IVB. This will enable the identification of SSR markers, and more specifically alleles at these loci, that are good candidates for marker-assisted-selection (MAS), in order to breed cultivars with a lower chilling requirement that are more adapted to local climatic conditions.

6.4 RESULTS AND DISCUSSION

6.4.1 QTL mapping

Due to the fact that a single approach to QTL analysis of a quantitative trait is not enough to understand the genetic control of any specific trait (Asins, 2002), the precision of QTL detection was enhanced by trait assessment over several years and by using different mapping populations

with shared ancestry, resulting in the trait being studied in different developmental stages and in different genetic backgrounds. From a breeding perspective, the use of different populations provides information regarding more alleles from any given locus shown to be associated with the same QTL, as well as the identification of more QTLs affecting the trait of interest.

A total of 7 QTLs were detected in the 'Anna' x 'Golden Delicious' population (Table 33), although significance thresholds were not met in all four years of investigation for any of these QTLs. In cases where significant LOD thresholds were met in at least one of the four years of investigation, the highest LOD value obtained in additional years are listed for comparative purposes. Two QTLs were detected using the parental genetic linkage map of 'Anna' and 1 QTL were detected using the parental genetic linkage map of 'Golden Delicious'. These 3 QTLs and an additional 4 QTLs were detected using the integrated genetic linkage map (Table 33). QTL analysis was done separately for adult trees and clonally propagated juvenile trees for both the 'Anna' x 'Sharpe's Early' (Table 34)(Table 35) and the 'Prima' x 'Golden Delicious' (Table 36)(Table 37) populations. Although LOD significance thresholds were met in all three years of investigation for most of the identified QTLs in both these populations, some exceptions do occur and for these the highest LOD value obtained in additional years are listed for comparative purposes. During QTL analysis conducted on adult trees derived from the cross between 'Anna' and 'Sharpe's Early', 4 QTLs were detected using the parental genetic linage map of 'Anna', 3 using the parental genetic linkage map of 'Sharpe's Early' and an additional 2 using the integrated genetic linkage map. One QTL were identified in both parental genetic linkage maps

resulting in a total of 8 QTLs. QTL analysis conducted on the juvenile trees of the same mapping pedigree resulted in the identification of 3 QTLs using the parental genetic linkage map of 'Anna', 1 QTL using the parental genetic linkage map of 'Sharpe's Early' and an additional 3 QTLs using the integrated genetic linkage map. A total of 7 QTLs were thus identified in the study conducted on juvenile trees and 6 of these QTLs are shared between adult and juvenile 'Anna' x 'Sharpe's Early' trees.

Table 33. Putative QTLs for time of IVB, detected by interval mapping, identified on the integrated genetic linkage map of 'Anna' x 'Golden Delicious'. The linkage group (LG) containing the QTL, the map on which the QTL has been identified, either integrated only (AnxGD) or on a parental as well (An or GD), the maximum LOD score, position on the LG, name of the marker residing at this position or the markers flanking the QTL peak on both sides (positions also indicated), the proportion of phenotypic variation explained by the QTL, the estimated mean of the distribution of the quantitative trait associated with the four different genotypes and the most favourable allele, or allele combination in cases where the QTL was only detected on the integrated map, associated with early IVB.

LG	Map	year	LOD	Position (cM)	Marker/s	% Expl.	Average time of IVB observed in the four different genotypes possible for the QTL				Allele associated with early IVB
							ac	ad	bc	bd	
2	AnxGD	1999	0.82	32.7	CH03d01	6.4	280	277	268	279	b+c
		2000	0.97	32.7	CH03d01	6.1	285	275	270	287	
		2002	2.27	32.7	CH03d01	13	259	253	248	259	
		2004	1.69	32.7	CH03d01	11.7	253	244	237	246	
5	GD	1999	0.67	27.9	CH05e06 (21.9 cM) A279 (28.6 cM)	6	283	270	274	272	d
		2000	1.2	24.9	CH05e06 (21.9 cM) A279 (28.6 cM)	9.2	289	274	275	267	
		2002	3.04	20	CH03a09 (17 cM) CH05e06 (21.9 cM)	17.9	255	248	260	249	

		2004	0.98	21.9	CH05e06	7	249	239	246	238	
9	AnxGD	1999	2.57	45	A383 (29 cM) CH05a03 (46.9 cM)	34.3	289	285	270	255	b+d
		2000	1.55	42	A383 (29 cM) CH05a03 (46.9 cM)	23	294	291	268	263	
		2002	1.44	47	CH05a03	11.9	256	259	252	246	
		2004	3.72	46.9	CH05a03	45.2	264	248	240	224	
11	AnxGD	1999	2.31	7	CH02d12 (0 cM) CH02d08 (9.1 cM)	20.9	263	279	287	272	a+c
		2000	0.96	9	CH02d12 (0 cM) CH02d08 (9.1 cM)	9.7	272	265	287	280	
		2002	2.08	12.1	CH02d08 (9.1 cM) CH04a12 (21.4 cM)	19.7	244	250	260	255	
		2004	1.75	10.1	CH02d08 (9.1 cM) CH04a12 (21.4 cM)	25	233	234	249	251	
16	AnxGD	1999	2.46	26.5	A267	51.7	277	291	285	254	b+d
		2000	7.45	26.5	A267	85.3	304	257	297	251	
		2002	1.84	26.5	A267	35.8	264	244	251	251	
		2004	2.15	26.5	A267	61.3	260	232	252	228	
17	An	1999	1.95	0	CH04c06	15.9	288	282	269	267	b
		2000	1.99	6	Ch04c06 (0 cM) CH01h01 (15.8 cM)	15.5	276	292	266	270	
		2002	1.88	9	Ch04c06 (0 cM) CH01h01 (15.8 cM)	14.5	253	260	251	249	
		2004	2.2	4	Ch04c06 (0 cM) CH01h01 (15.8 cM)	19.2	242	255	235	240	
17	An	1999	3.05	34.9	CH02g04 (23.9 cM) CH05g03 35.2 cM)	24.6	278	286	274	261	b

	2000	1.74	21.8	CH01h01 (15.8 cM) CH02g04 (23.9 cM)	12.4	279	286	260	271
	2002	1.71	29.9	CH02g04 (23.9 cM) CH05g03 35.2 cM)	11.5	255	258	250	248
	2004	2	32.9	CH02g04 (23.9 cM) CH05g03 35.2 cM)	15.5	243	252	240	234

Table 34. Putative QTLs for time of IVB in adult trees, detected by interval mapping, identified on the integrated genetic linkage map of 'Anna' x 'Sharpe's Early'. The linkage group (LG) containing the QTL, the map on which the QTL has been identified, either integrated only (AnxSE) or on a parental as well (An or SE), the maximum LOD score, position on the LG , name of the marker residing at this position or markers flanking the QTL peak on both sides (positions also indicated), the proportion of phenotypic variation explained by the QTL, the estimated mean of the distribution of the quantitative trait associated with the four different genotypes and the most favourable allele, or allele combination in cases where the QTL was only detected on the integrated map, associated with early IVB.

LG	Map	Year	LOD	Position (cM)	Marker/s	% Expl.	Average time of IVB observed in the four different genotypes possible for the QTL	Allele

							ac	ad	bc	bd	associated with early IVB
1	AnxSE	1996	5.76	0	Hi12c02	59.3	225	272	271	230	a+c
		1997	5.54	0	Hi12c02	75.1	219	266	265	223	
		1998	5.09	0	Hi12c02	53.9	216	266	259	223	
8	AnxSE	1996	5.17	21.5	CH01c06 (18.5 cM) Hi20b03 (33 cM)	67.4	228	271	262	215	b+d
		1997	3.89	21.5	CH01c06 (18.5 cM) Hi20b03 (33 cM)	70.2	219	264	259	217	
		1998	4.23	20.5	CH01c06 (18.5 cM) Hi20b03 (33 cM)	73.6	222	267	257	207	
9	An	1996	2.21	17.4	Hi05e07	20.4	229	236	249	261	a
		1997	1.97	17.4	Hi05e07	18.8	229	229	244	254	
		1998	2.94	17.4	Hi05e07	27.3	222	222	241	257	
10	SE	1996	2.2	0	Hi02d04	21.7	261	228	256	238	d
		1997	2.64	0	Hi02d04	24.2	254	226	252	229	
		1998	1.85	0	Hi02d04	18.6	253	220	246	230	
10	SE	1996	1.97	56.3	A381	16.7	259	238	254	227	d
		1997	2.28	56.3	A381	18.7	254	233	247	223	
		1998	2.36	56.3	A381	19.6	251	228	248	217	
14	An	1996	3.86	22.9	CH05g11 (0 cM) CH04c06 (28.7 cM)	82.4	235	265	269	210	a
		1997	1.93	24.9	CH05g11 (0 cM) CH04c06 (28.7 cM)	33.7	260	247	258	221	
		1998	1.74	23.9	CH05g11 (0 cM) CH04c06 (28.7 cM)	34	264	242	259	216	
17	An	1996	1.74	0	CH04c06	17.9	241	226	257	259	a

		1997	1.97	5	CH04c06 (0 cM) CH05g03 (19.1 cM)	12.9	238	213	252	255	
		1998	2.7	0	CH04c06	8.3	228	218	257	252	
17	An & SE	1996	2.92	67.1	A236	23.6	220	252	247	260	a+c
		1997	2.39	64	Hi02f12 (63 cM) A236 (67.1 cM)	20.8	220	244	240	254	
		1998	3.05	64	Hi02f12 (63 cM) A236 (67.1 cM)	27.1	211	238	240	256	

Table 35. Putative QTLs for time of IVB in juvenile trees, detected by interval mapping, identified on the integrated genetic linkage map of 'Anna' x 'Sharpe's Early'. The linkage group (LG) containing the QTL, the map on which the QTL has been identified, either integrated only (AnxSE) or on a parental as well (An or SE), the maximum LOD score, position on the LG , name of the marker residing at this position or markers flanking the QTL peak on both sides (positions also indicated), the proportion of phenotypic variation explained by the QTL, the estimated mean of the distribution of the quantitative trait associated with the four different genotypes and the most favourable allele, or allele combination in cases where the QTL was only detected on the integrated map, associated with early IVB.

LG	Map	year	LOD	Position (cM)	Marker/s	% Expl.	Average time of IVB observed in the four different genotypes possible for the QTL				Allele associated with early IVB
							ac	ad	bc	bd	
1	AnxSE	1998	4.45	0	Hi12c02	30.8	252	276	270	247	b+d
		1999	5.53	0	Hi12c02	38.8	240	268	263	235	
		2000	4.9	0	Hi12c02	34	238	272	264	229	
3	AnxSE	1998	2.92	16	Hi04c10? (0 cM) CH03g12y (22.1 cM)	17.8	263	271	249	254	b
		1999	2.44	10	Hi04c10? (0 cM) CH03g12y (22.1 cM)	16.8	256	258	241	236	
		2000	2.07	0	Hi04c10?	11.7	255	257	248	230	
9	An	1998	2.91	17.4	Hi05e07	18.5	255	250	274	264	a
		1999	3.35	17.4	Hi05e07	20.2	247	236	262	257	
		2000	2.56	17.4	Hi05e07	15.8	245	235	264	258	
10	SE	1998	4.25	2	Hi02 d04 (0 cM) CH02a08y (10 cM)	24.2	273	248	265	253	d
		1999	5.29	0	Hi02d04	28.4	264	236	258	242	
		2000	5.1	0	Hi02d04	27.5	267	232	261	239	
17	An	1998	1.62	7	CH04c06 (0 cM) CH05g03 (19.1 cM)	15.2	252	258	265	272	a
		1999	1.65	0	CH04c06	10.7	243	246	259	258	
		2000	1.85	0	CH04c06	11.9	240	246	263	259	
17	AnxSE	1998	3.33	26.1	CH05g03 (19.1 cM) CH01h01 (39.2 cM)	35.1	250	283	273	257	a+c
		1999	2.1	39.2	CH01h01	11.8	241	256	258	250	
		2000	2.28	39.2	CH01h01	12.8	238	258	261	249	
17	An	1998	1.72	67.1	A236	9.1	253	254	264	268	a

1999	1.79	67.1	A236	9.5	241	245	256	256
2000	1.94	67.1	A236	10.2	238	243	257	259

Table 36. Putative QTLs for time of IVB in adult trees, detected by interval mapping, identified on the integrated genetic linkage map of 'Prima' x 'Golden Delicious'. The linkage group (LG) containing the QTL, the map on which the QTL has been identified, either integrated only (GDxPr) or on a parental as well (GD or Pr), the maximum LOD score, position on the LG , name of the marker residing at this position or markers flanking the QTL peak on both sides (positions also indicated), the proportion of phenotypic variation explained by the QTL, the estimated mean of the distribution of the quantitative trait associated with the four different genotypes and the most favourable allele, or allele combination in cases where the QTL was only detected on the integrated map, associated with early IVB.

LG	Map	year	LOD	Position (cM)	Marker/s	% Expl.	Average time of IVB observed in the four different genotypes possible for the QTL				Allele associated with early IVB
							ac	ad	bc	bd	
3	GD	1996	2.16	17.1	Hi15h12	39.3	279	284	294	290	a
		1997	2.56	17.1	Hi15h12	46.2	270	272	283	280	
		1998	2.75	17.1	Hi15h12	50.1	273	279	296	291	

4	Pr	1996	2.52	35.4	Hi07b02	20.2	281	291	285	291	c
		1997	2.57	35.4	Hi07b02	20.4	2701	279	275	280	
		1998	2.09	26.6	CH02h11a (11.6 cM) Hi07b02 (35.4 cM)	23	275	287	285	293	
9	GD	1996	2.15	20.8	CH01f03b	18.3	288	280	290	290	a
		1997	1.62	20.8	CH01f03b	13.4	276	272	280	277	
		1998	2.06	18	CH05c07 (0 cM) CH01f03b (20.8 cM)	22	288	275	291	288	
9	GDxPr	1996	2.05	45.2	A334	21.6	287	280	292	289.	a+d
		1997	1.58	41.3	Hi05e07 (32.3 cM) A334 (45.2 cM)	19.4	276	272	282	276	
		1998	1.6	45.2	A334	19.6	287	277	294	286	
11	GDxPr	1996	1.91	5	CH02d12 (0 cM) CH02d08 (6.2 cM)	21.9	294	282	285	289	a+d
		1997	2.97	6.2	CH02d08	33	284	272	274	278	
		1998	2.29	6.2	CH02d08	25.8	296	280	281	291	

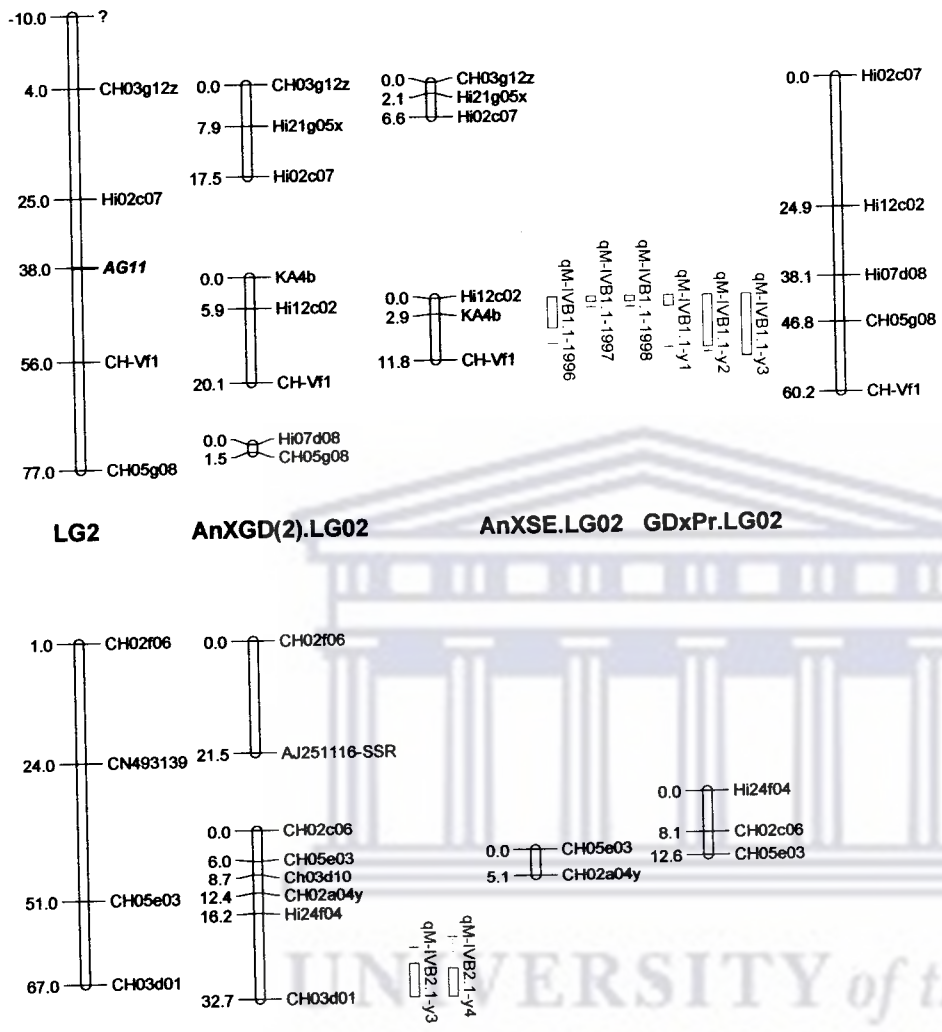
Table 37. Putative QTLs for time of IVB in juvenile trees, detected by interval mapping, identified on the integrated genetic linkage map of 'Prima' x 'Golden Delicious'. The linkage group (LG) containing the QTL, the map on which the QTL has been identified, either integrated only (GDxPr) or on a parental as well (GD or Pr), the maximum LOD score, position on the LG , name of the marker residing at this position or markers flanking the QTL peak on both sides (positions also

indicated), the proportion of phenotypic variation explained by the QTL, the estimated mean of the distribution of the quantitative trait associated with the four different genotypes and the most favourable allele, or allele combination in cases where the QTL was only detected on the integrated map, associated with early IVB.

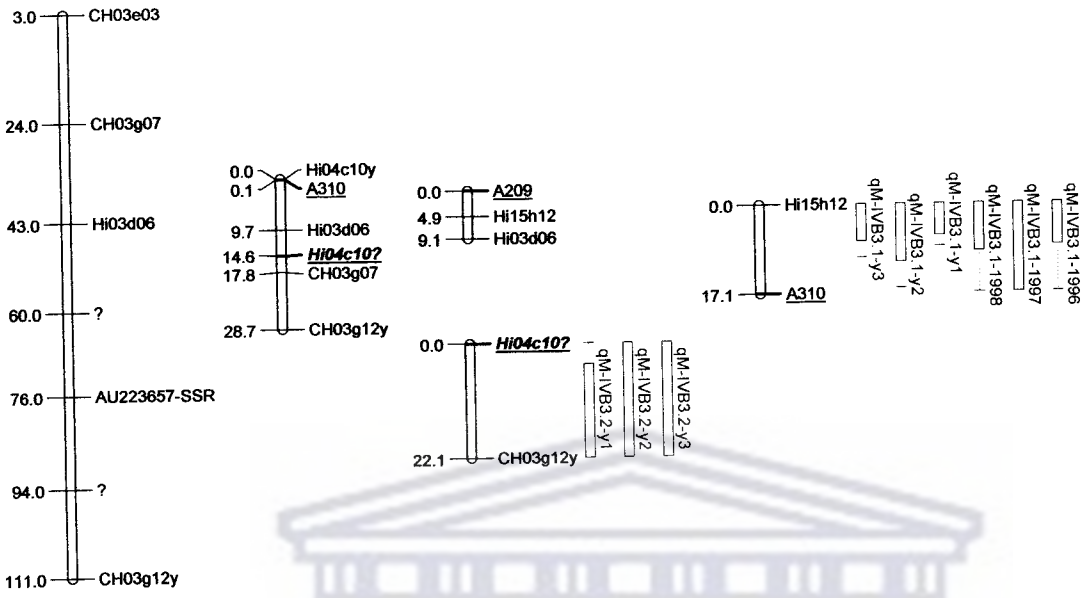
LG	Map	year	LOD	Position (cM)	Marker/s	% Expl.	Average time of IVB observed in the four different genotypes possible for the QTL				Alleles associated with early IVB
							ac	ad	bc	bd	
3	GDxPr	1997	2.94	17.1	Hi15h12	70	284	312	302	284	b+d
		1998	3.15	17.1	Hi15h12	32.1	299	306	302	290	
		1999	3.49	17.1	Hi15h12	47.5	299	307	304	290	
4	GDxPr	1997	2.01	48.4	Hi07b02 (35.4 cM) CH02c02b (52.4 cM)	24	306	290	292	287	a+d
		1998	1.12	48.4	Hi07b02 (35.4 cM) CH02c02b (52.4 cM)	12.2	303	294	300	301	
		1999	0.7	48.4	Hi07b02 (35.4 cM) CH02c02b (52.4 cM)	7.3	303	296	301	302	

A total of 5 QTLs were identified in adult trees derived from a cross between 'Golden Delicious' and 'Prima', 2 of which were also identified in the parental genetic linkage map of 'Golden Delicious' and 1 in the parental genetic linkage map of 'Prima'. QTL analysis conducted on the juvenile trees from the same mapping pedigree revealed only 2 QTLs that were both detected in the integrated genetic linkage map only, with only 1 of these QTLs being shared between adult and juvenile 'Prima' x 'Golden Delicious' trees.

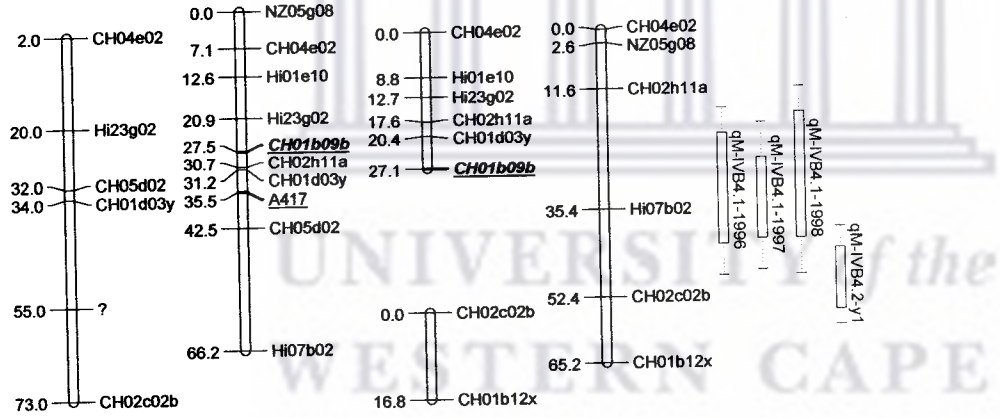
Detected QTLs were graphically represented as bars next to the relevant linkage groups, or segments thereof, indicating CI's corresponding to a LOD score drop of 0.5 and 1 (Figure 47). QTLs where the LOD significance threshold of 2 were not met in all the years during which the investigations were conducted, were also included in the graphical representation if the map position was the same as that for other QTL detected on the specific linkage group and if the LOD threshold obtained was above 1.5.

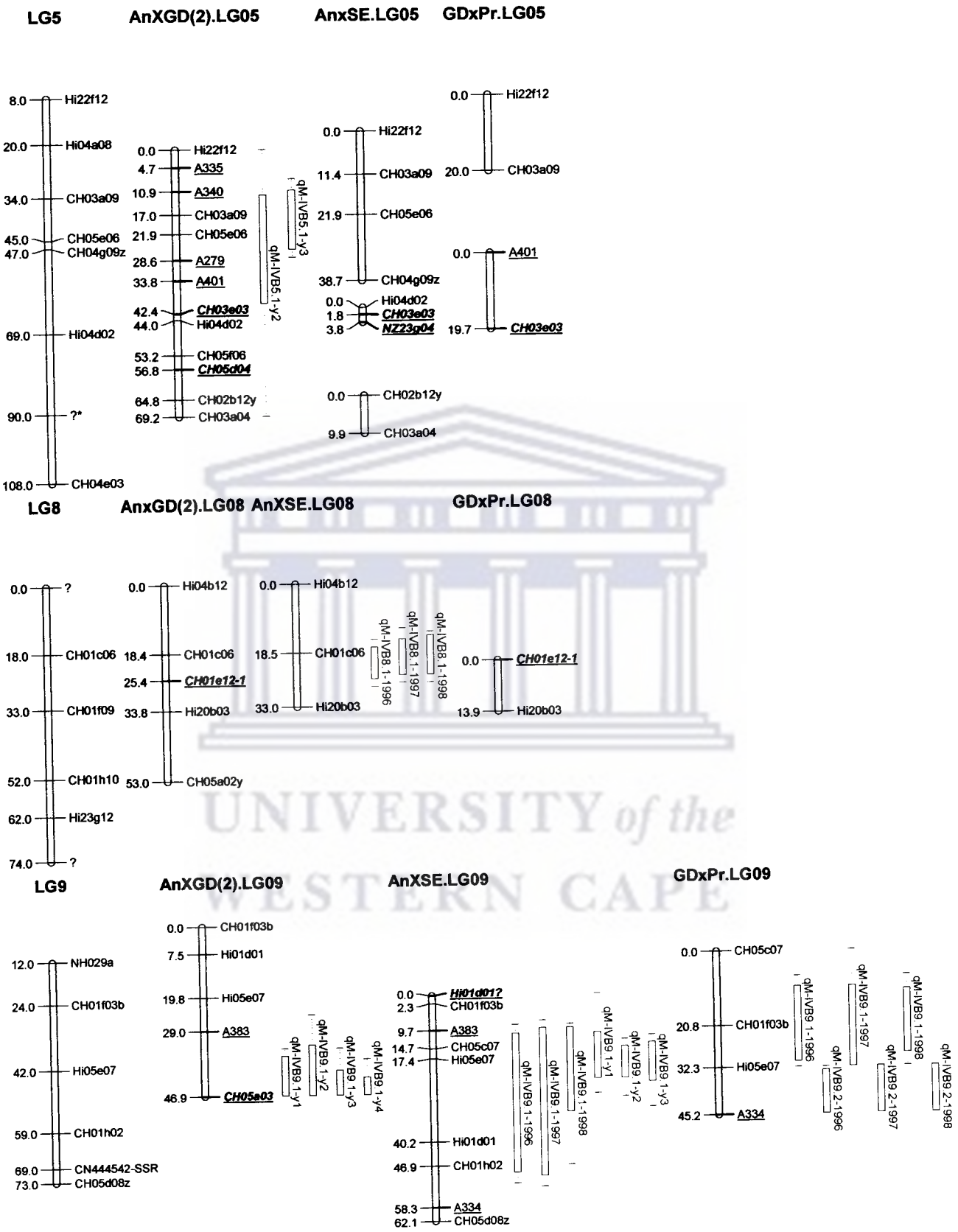


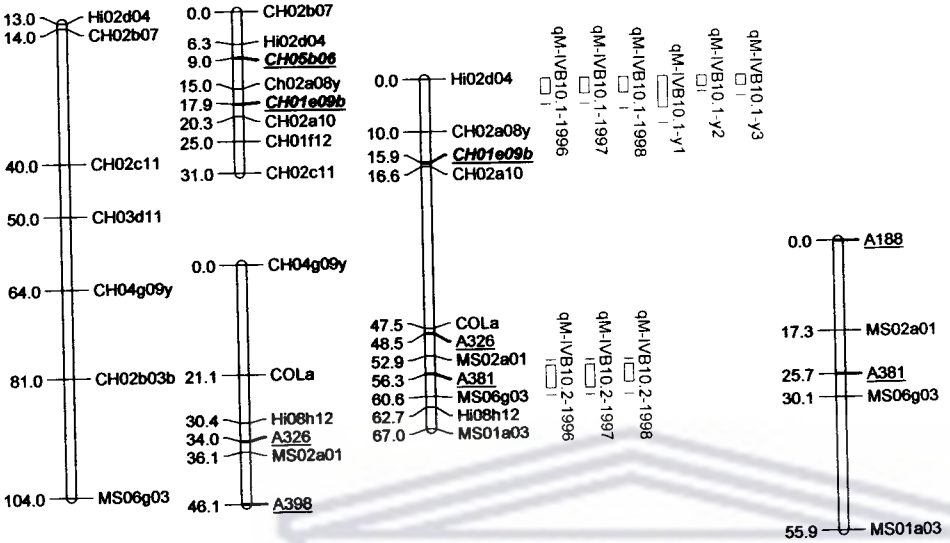
LG3 **AnXGD(2).LG03 AnxSE.LG03** **GDxPr.LG03**

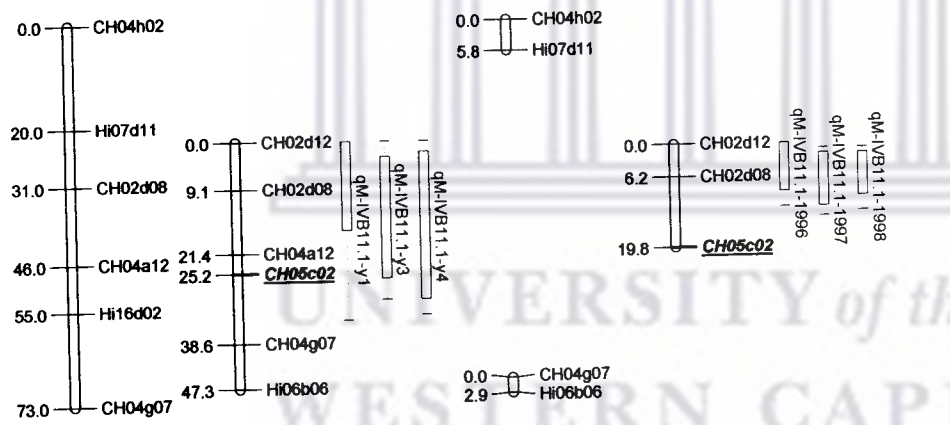


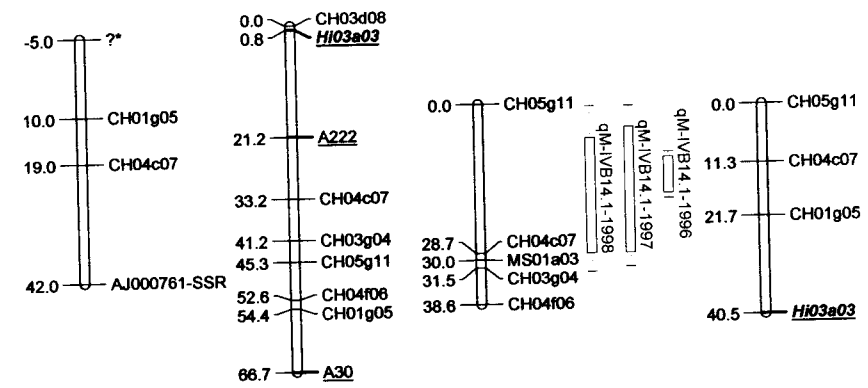
LG4 **AnXGD(2).LG04 AnXSE.LG04** **GDxPr.LG04**



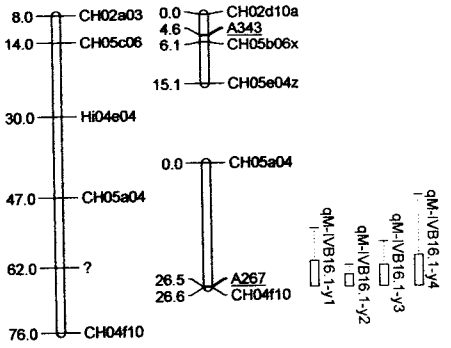








LG16 **AnXGD(2).LG16**



LG17 **AnXGD(2).LG17** **AnXSE.LG17** **GDxPr.LG17**

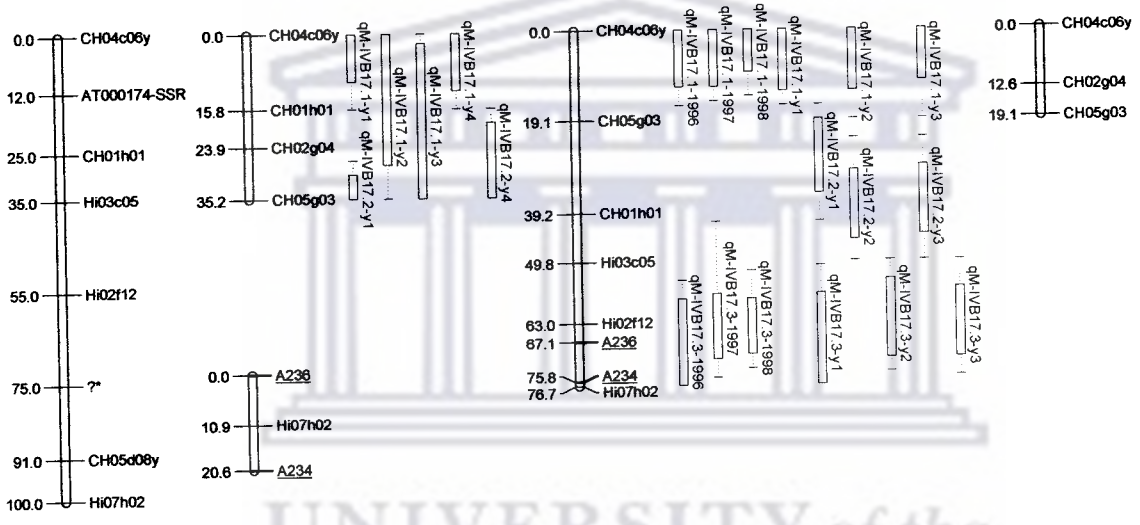


Figure 47. Graphical representation of the genetic positions (in cM) of the time of IVB QTLs identified in three mapping pedigrees using interval mapping. AnxGD(2) represents linkage groups constructed during a second round of genetic linkage map construction for the mapping pedigree 'Anna' x 'Golden Delicious'. AnxSE and GDxPr represents linkage groups constructed for mapping pedigrees 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious' respectively. Only the linkage groups (LGs) carrying QTLs are represented. Notation of LGs are in accordance with Maliepaard et al. (1998) and markers from the proposed 15 cM framework map (Silfverberg-Dilworth et al. (2006)) are included as reference. QTLs are represented by bars indicating 5 % confidence intervals and broken lines indicating 10 % confidence intervals. QTL nomenclature are in accordance with standardized criteria for the Rosaceae (www.rosaceae.org).

On LG 1 a QTL was detected in the 'Anna' x 'Sharpe's Early' mapping pedigree, during all 3 years of study conducted on both adult and juvenile trees. The QTL was in both cases detected on the integrated genetic linkage map only, with LOD significance thresholds ranging from 4.45 to 5.76 (Tables 35 & 36). The phenotypic variation explained by the QTL ranged from 30.8 % to 75.1 %.

On LG 2 a QTL was identified on the integrated genetic linkage map 'Anna' x 'Golden Delicious'. The QTL exceed the LOD threshold of 2 during the 3rd year of study only (Table 34), although both the 3rd and 4th year were graphically depicted (Figure 47). Due to the low LOD scores obtained and the

fact that the potential QTL was not detected when studying two other mapping populations involving either one of these parentals, the QTL is most probably not a true QTL, but a false positive.

On LG 3 two QTLs were identified. The first was identified in the 'Prima' x 'Golden Delicious' mapping population and remained consistent during the 3 years of study conducted on adult trees as well as during the 3 years of study conducted on juvenile trees (Figure 47). Although the QTL was also identified on the parental genetic linkage map of 'Golden Delicious' during analysis performed on adult trees (Table 37), it was only identified on the integrated genetic linkage map during analysis performed on juvenile trees (Table 38). The LOD significance threshold ranged from 2.16 to 3.49 and the phenotypic variance explained between 32.1 and 70 %. The second QTL was identified on the integrated genetic linkage map of the 'Anna' x 'Sharpe's Early' mapping pedigree and though remaining consistent over three years, only reached significant LOD thresholds in the study conducted on juvenile trees (Table 36).

On LG 4 a QTL was identified for adult trees on the parental genetic linkage map 'Prima' in the 'Prima' x 'Golden Delicious' mapping pedigree (Table 37). The identification of a second QTL on this linkage group on the integrated genetic map in juvenile trees, needs to be confirmed as it reached the significant LOD threshold during only one year of investigation (Table 38). The possibility of this being the same QTL needs to be kept in mind.

On LG 5 a QTL was identified on the 'Golden Delicious' parental genetic linkage map in the 'Anna' x 'Golden Delicious' mapping pedigree. The QTL exceed the LOD significance threshold during the 3rd year of study only

(Table 34), although an attempt was also made to plot the results obtained in the 2nd year of investigation.

On LG 8 a QTL was identified on the integrated genetic linkage map of 'Anna' x 'Sharpe's Early' during all 3 years the study was conducted on adult trees (Table 35). The LOD significance values ranged from 3.89 to 5.17 and the phenotypic variance explained between 67.4 and 73.6 %.

On LG 9 a QTL was identified in all three mapping populations used, 'Anna' x 'Golden Delicious' (Table 34), 'Anna' x 'Sharpe's Early' adult (Table 35) and juvenile trees (Table 36) as well as 'Prima' x 'Golden Delicious' adult trees (Table 37). In the 'Anna' x 'Golden Delicious' mapping pedigree the QTL was identified on the integrated genetic linkage map, in the 'Anna' x 'Sharpe's Early' mapping pedigree on the 'Anna' parental genetic linkage map and in the 'Prima' x 'Golden Delicious' mapping pedigree on the 'Golden Delicious' parental genetic linkage map. The existence of a second potential QTL on this linkage group when studying the 'Prima' x 'Golden Delicious' mapping pedigree could not be confirmed since the QTL was identified on the integrated genetic linkage map of this population only and since it exceeded the LOD threshold of 2 during the first year of investigation only. Overlapping CI's when graphically depicted indicates that this might be the same QTL.

On LG 10 two QTLs were identified on the parental genetic linkage map of 'Sharpe's Early'. The first remain constant over all 3 years of investigation conducted on adult trees as well as all 3 years of study conducted on juvenile trees (Tables 35 & 36) with LOD scores of between 4.25 and 5.29 on juvenile trees. The second QTL was only identified during the three years of investigation conducted on adult trees (Table 35).

On LG 11 a QTL was identified on the integrated genetic linkage maps of the 'Anna' x 'Golden Delicious' and the 'Prima' x 'Golden Delicious' mapping pedigrees. The QTL reached significant LOD levels of above 2 in 2 of the 4 years during which a study was conducted on the 'Anna' x 'Golden Delicious' mapping pedigree (Table 34), as well as during all 3 years of investigation on 'Prima' x 'Golden Delicious' adult trees (Table 37).

On LG 14 a QTL was identified on the parental genetic linkage map of 'Anna' in the mapping pedigree 'Anna' x 'Sharpe's Early', during a study conducted on adult trees (Table 35). Although the QTL exceeded the LOD significance threshold during the first year of study only, the LOD scores obtained in subsequent years were high enough for graphical presentation. The amount of phenotypic variance explained ranged between 33.7 and 82.4 %.

On LG 16 a QTL was identified on the integrated genetic linkage map of 'Anna' x 'Golden Delicious' (Table 34). The QTL reached a LOD score of 7.45 during the second year of investigation and although LOD scores were much lower during other years the phenotypic variance explained ranged between 35.8 and 85.3 %.

On LG 17, three QTLs were identified in the mapping pedigree 'Anna' x 'Sharpe's Early'. The first was identified on the parental genetic linkage map of 'Anna' and although the significance threshold of 2 were only exceeded in the 3rd year of study conducted on adult trees (Table 35), the LOD scores obtained in other years were just below this significance threshold. The second was identified on the integrated map and proved to be significant only in juvenile trees (Table 36) with an explained phenotypic variance of between

11.8 and 35.1 %. The third was identified on both parental genetic linkage maps in the study conducted on adult trees, with LOD values between 2.39 and 3.05 (Table 35) and on the parental genetic linkage map of 'Anna' in the study conducted on juvenile trees (Tables 36). While the LOD significance threshold was above 2 for all 3 years of study conducted on adult trees it dropped to slightly below the significance level in juvenile trees, but was included in the graphical representation as it explains between 9.1 and 10.2 % of the phenotypic variation which promotes confirmation of the QTL position. The amount of phenotypic variation explained in adult trees ranged from 16.7 to 19.6 %. The positions of the first 2 QTLs on this linkage group were confirmed by the localization of two QTLs to the same positions on the parental genetic linkage map of 'Anna' in the mapping pedigree 'Anna' x 'Golden Delicious' (Table 34). Although the LOD significance threshold of 2 was only exceeded in the 1st and the 1st and 4th year respectively, LOD scores obtained in additional years were high enough to be plotted. The slightly lower LOD significance thresholds reached in the 2nd and 3rd year of investigation resulted in the 2 QTLs being interpreted as 1 since confidence intervals overlapped.

QTL detection was much more effective (larger LOD scores) and much more consistent (detected over all years of investigation) in the 'Anna' x 'Sharpe's Early' population than those detected in any of the other two populations. This might be the result of a combination of small population sizes as well as amount of observed phenotypic variation. When considering the average time of IVB for the four cultivars used as parentals during this investigation (viz. 'Anna' 220 days, 'Golden Delicious' 300 days, 'Sharpe's

Early 276 days and 'Prima' 295 days), the amount of phenotypic variation expected would be far less for the cross derived between 'Golden Delicious' and 'Prima' and this was indeed true as outlined in Chapter 2. The amount of variation observed in the 'Anna' x 'Sharpe's Early' mapping population exceeded that observed for the 'Anna' x 'Golden Delicious' population (Chapter 2).

6.4.2 Nonparametric mapping

Although the identification of different genomic regions involved in the determination of a specific trait broadens our knowledge and understanding regarding the number and effect of genes involved in the determination of a specific phenotypic character, the identification of markers that can be implemented for MAS is of interest to the breeder. For this reason markers were subjected to a Kruskal-Wallis nonparametric ANOVA in order to identify alleles associated with an early IVB. Significant results obtained during Kruskal-Wallis analysis (Appendix C) were summarized (Table 38) and can be used in order to identify potential candidate markers for MAS. Markers showing significant association with the trait of interest, could be associated with likelihood peaks of identified QTLs. Differences observed between the mean time of IVB associated with specific alleles of allele combinations are also indicated.

Table 38. Summary of markers where segregation of alleles can be associated with differences in the mean time of IVB. The linkage group containing the marker and the associated QTL, the population in which it has been identified, the parental genotype/s associated with phenotypic variation, the allele sizes and the specific allele or allele combination, in cases where both parental genotypes are involved, associated with vegetative budbreak occurring earlier or later than the mean. Differences between the mean time of IVB observed in adult and/or juvenile years are indicated.

LG	Marker	Associated QTL	population	Parental cultivar	Allele sizes	Allele/Allele combination associated with early IVB	Allele associated with late IVB	
1	KA4b	IVB1.1	AnxSE	SE	133-135	135	133	Adults (234 vs. 256) Juveniles (248 vs. 262)
2	CH03d01	IVB2.1	AnxGD	An GD	96-109 96-109	96-96 or 109-109	96-109	4Years (263 and 256 vs. 270)
3	Hi04c10?	IVB3.2	AnxSE	An SE	204-212 202-204	202-212	202-204	Juveniles (239 vs. 262)
	CH03g12y		AnxSE	An SE	181-177 181-186	177-181	181-186	Juveniles (243 vs. 263)
	A209	IVB3.1	GDxPr	GD	148-145	148	145	Adults (278 vs. 287)
	A310		GDxPr	GD	240-223	240	223	Adults (278 vs. 287) Juvenile years 2&3 (297 vs. 304)
4	Hi07b02	IVB4.1	GDxPr	Pr	213-200	213	200	Adults (279 vs. 287) Juvenile year 1 (289 vs. 297)
5	CH03a09	IVB5.1	AnxGD	GD	130-126	126	130	Year 2 & 3 (257 vs. 269)

			GDxPr	GD Pr	130-126 130-126	126-126	126-130 130-130	Juvenile years 2&3 (292 vs. 303 and 302)
8	CH01c06	IVB8.1	AnxSE	An	160-158	160	158	Adults (231 vs. 253)
9	Hi05e07	IVB9.1	AnxSE	An	194-213	194	213	Adults (224 vs. 252) Juveniles (244 vs. 260)
	CH01f03b	IVB9.1?	GDxPr	GD	138-171	138	171	Adults (280 vs. 286)
	A334			GD	209-225	209	225	Adults (280 vs. 287)
10	Hi02d04	IVB10.1	AnxSE	SE	223-null	223	Null	Adults (227 vs. 257) Juveniles (242 vs. 265)
	Hi08h12	IVB 10.2	AnxSE	SE	156-171	156	171	Adults (227 vs. 254)
11	CH04a12	IVB11.1	AnxGD	An GD	186-null 176-182	186-182	null-182	Year1, 3 & 4 (244 vs. 264)
	CH02d12	IVB11.1?	GDxPr	Pr	197-190	197	190	Adults (277 vs. 285)
14	CH05g11	IVB14.1	AnxSE	An	252-214	214	252	Adults (231 vs. 251)
	CH04c07				Null-96	96	null	Adults (233 vs. 252)
16	A267	IVB16.1	AnxGD	An	484-466	466	484	Year 1, 2 & 4 (258 vs. 271)
17	CH04c06y	IVB17.1	AnxSE	An	187-177	187	177	Adults (231 vs. 255) Juveniles (249 vs. 262)
			AnxGD	An	187-177	187	177	Year 1 & 4 (251 vs. 263)
	CH01h01	IVB17.1	AnxGD	An	102-118	102	118	4 years (256 vs. 267)
	CH01h01	IVB17.2	AnxSE	An	102-118	102	118	Juveniles (249 vs. 259)
	CH05g03	IVB17.2	AnxGD	An GD	132-160 132-164	132-164	160-164	Year 1, 3 & 4 (248 vs. 266)
	Hi02f12	IVB17.3	AnxSE	SE	149-136	136	149	Adults (231 vs. 252)
	A236	IVB17.3	AnxSE	An SE	241-235 241-220	235-241	220-241	Adults (219 vs. 256) Juveniles (247 vs. 264)

On LG 1 segregating alleles from the parental cultivar 'Sharpe's Early', when implementing the marker KA4b on the mapping pedigree derived from a cross with 'Anna', could be significantly associated with a difference in the time of IVB. In the adult population budbreak occurred on average (calculated over 3 years) on day 234 in individuals who inherited the 135bp allele form 'Sharpe's Early', compared to day 256 in individuals who inherited the 133bp allele.

On LG 2 the marker CH03d01, associated with the QTL IVB2-1 showing significant LOD thresholds in 2 of the 4 years of investigation conducted on the 'Anna' x 'Golden Delicious' population, showed significant levels of association with the trait of interest in all 4 years of investigation. CH03d01 is heterozygous for the same two alleles in the two parental cultivars involved (segregation type <hkxhk>), and interaction between the alleles renders the resulting heterozygote (hk) not desirable when selecting for early IVB. The specific segment of LG 2 containing the QTL and the associated marker CH03d01, is not represented on the genetic linkage maps derived for the other two mapping populations (Figure 47), and for this reason the association of CH03d01, or any other marker in the region of the identified QTL, with the trait of interest could not be tested in a different genetic background. Similarly both QTLs identified on LG 3 are on segments not represented on the genetic linkage maps of 'Anna' x 'Golden Delicious', or that of 'Anna' x 'Sharpe's Early' and 'Prima' x 'Golden Delicious' in the case of the first and second QTL respectively. The association of markers with the trait of interest in these regions could thus be tested in one specific genetic

background only. Segregating alleles from the parental cultivar 'Golden Delicious' for both markers flanking the QTL, IVB3.1, A209 and A310 could be associated with differences in time of IVB in the 'Prima' x 'Golden Delicious' adult mapping pedigree while segregation results from A310 could be associated with the phenotype in 2 of the 3 years during which the study was conducted on juvenile trees. Interaction between the parental alleles obtained from the two markers flanking the QTL, IVB3.2, Hi04c10? and CH03g12y, could be associated with differences in time of IVB in the juvenile 'Anna' x 'Sharpe's Early' mapping pedigree.

The LG 4 marker Hi07b12, segregating in the parental cultivar 'Prima', could be significantly associated with time of IVB in adult trees, but the association was only evident during the first year of study on juvenile trees in the mapping pedigree derived from a cross with 'Golden Delicious'. On LG 5 the marker CH03a09 segregating from 'Golden Delicious' in the mapping pedigree 'Anna' x 'Golden Delicious' could be associated with a difference observed in the mean time of IVB. Although no QTL exceeding significance thresholds was identified on LG 5 in the 'Prima' x 'Golden Delicious' mapping population, the interaction between segregating alleles from both parental cultivars could be associated with early IVB during 2 of the 3 years of study conducted on juvenile trees.

The segregating alleles from the parental cultivar 'Anna', when implementing the LG 8 marker CH01c06 on 'Anna' x 'Sharpe's Early', showed significant levels of association with time of IVB in adult trees.

Segregation results obtained for the LG 9 marker Hi05e07 when studying the parental cultivar 'Anna' in the mapping pedigree 'Anna' x

'Sharpe's Early' showed highly significant association with the phenotypic trait under investigation in both adult and juvenile trees. The fact that marker Hi05e07 is heterozygous for the same two alleles in the cultivars 'Golden Delicious' and 'Prima' may be the reason why no significant association between genotype and phenotype could be found. The fact that both flanking markers (CH01f03b and A334) show significant levels of association with the phenotype when studying the segregation of alleles from the parental cultivar 'Golden Delicious', results in uncertainty regarding the true number of QTLs. Although likelihood peaks lead to the assumption of two putative QTLs on this LG, the possibility of one single QTL should not be ignored.

LG 10 markers Hi02d04 and Hi08h12, close to likelihood peaks of both putative QTLs identified on the parental genetic linkage map of 'Sharpe's Early', showed significant levels of association with the phenotype under investigation in adult trees, but only the second showed a similar association in juvenile trees.

Failure of the marker CH02d08, homozygous in both 'Anna' and 'Prima' but segregating in 'Golden Delicious', to show any significant level of association with the phenotype, leads to the conclusion that the QTL identified on LG 11 in the 'Anna' x 'Golden Delicious' population are much closer to the marker CH04a12 than concluded from likelihood peaks. The closer association observed between the marker CH02d12 and the phenotypic trait when studying the parental genetic linkage map of 'Prima' in the mapping pedigree derived from 'Prima' x 'Golden Delicious', leads to the assumption that the QTL identified on this LG in the two mapping pedigrees mentioned, might be two different QTLs. Interaction between parental alleles obtained

from 'Anna' and 'Golden Delicious' when studying the marker CH04a12, could be significantly associated with earlier IVB, while alleles segregating from the parental cultivar 'Prima' when studying the marker CH02d12, could be associated with differences in time of IVB observed in the 'Prima' x 'Golden Delicious' mapping pedigree.

The two markers flanking the QTL identified on LG 14 of the parental genetic linkage map of 'Anna' in the mapping pedigree 'Anna' x 'Sharpe's Early', CH05g11 and CH04c07, are both associated with the trait of interest and can thus be used to successfully select those individuals that have no recombination between these markers. Although the use of two markers flanking the gene of interest are ideal for MAS, the fact that the association between marker genotype and phenotype was identified in adults only are probably due to the gene not being expressed during all developmental stages.

On LG 16 the segregation ratios observed for the marker A267, when studying the parental cultivar 'Anna' in the mapping population 'Anna' x 'Golden Delicious', could be associated with time of IVB in all 4 years.

The LG 17 marker CH04c06y could be associated with time of IVB in adult and juvenile trees when studying the parental cultivar 'Anna' in the mapping pedigree 'Anna' x 'Sharpe's Early'. This association was confirmed in the mapping pedigree 'Anna' x 'Golden Delicious' during 2 of the 4 years of study. Segregating alleles from the parental cultivar 'Anna' when studying the marker CH01h01, were associated with variation in time of IVB when studying the 'Anna' x 'Sharpe's Early' juvenile mapping pedigree as well as the 'Anna' x 'Golden Delicious' mapping pedigree. Different marker orders in the integrated

genetic linkage maps derived for these populations leads to uncertainty whether the QTL of association is the same in the two respective mapping pedigrees. Interaction between the parental alleles obtained when implementing the marker CH05g03 on the 'Anna' x 'Golden Delicious' mapping pedigree leads to a specific allele combination which can be associated with the trait of interest. Similarly the interaction between parental alleles led to an association between the genotypic code obtained when implementing the marker A236 on seedlings derived from the 'Anna' x 'Sharpe's Early' mapping pedigree and the time of IVB.

6.5 CONCLUSION

This study has led to the successful identification of 18 putative QTLs associated with time of IVB in apple. All of these QTLs have been associated with the trait of interest over more than one year of investigation. At least six of these QTLs have been identified in both juvenile and adult populations and three have been identified in different genetic backgrounds.

The significant levels of association found between marker genotype and the phenotype of interest, enabled the identification of markers that could be used in MAS for the purpose of breeding apples that are better adapted to local climatic conditions. The association between a marker and the phenotype of interest should be viewed in the genetic background/s in which this study has been conducted and in which the association has been established. Further testing in order to verify the associations concluded in this study and to test associations in different genetic backgrounds, need to

be conducted before these markers can be implemented in the breeding program.

7 CONCLUDING REMARKS

The identification of SSRs from ESTs and the subsequent amplification of these regions using primers designed from highly conserved flanking regions, have proven to be very successful. More than 80 % of primers yielded the expected amplification product and from these the level of polymorphism detection was found to be above 80 %. These EST derived SSRs provided an ideal marker for the construction of genetic linkage maps due to the high level of heterozygosity. The development of multiplexes greatly enhances the efficiency with which markers can be implemented on newly developed mapping pedigrees within the breeding programme. The highly efficient database that was developed as part of this investigation proved to be a valuable contribution towards data management and has contributed significantly towards the accessibility of obtained results.

Thirty-four newly developed markers have been implemented and mapped on at least one of the genetic linkage maps constructed for each of the three mapping pedigrees studied during this investigation (27 in Chapter 4 and another 7 in Chapter 5). Implementation of bin mapping enabled prediction of the possible position of markers on the genetic linkage map. This will enable the identification of markers that could result in the expansion of the map through the expansion of the telomeric ends of LGs, markers that will fill the large gaps found between adjacent SSR markers and markers that are located in the region of an identified QTL. The latter will facilitate the identification of markers that closely flank a QTL of interest enabling the selection of those individuals that has no recombination event occurring

between the QTL and the markers used. The greater the density of markers in the region of such a QTL, the more likely that flanking markers could be identified with a minimal distance between the marker and the actual QTL itself.

Eighteen putative QTLs associated with time of IVB budbreak have been identified. All reported QTLs have been identified in more than one year of investigation, of which 6 were identified during different developmental stages and 3 were identified in different genetic backgrounds. One of these QTLs, identified on LG 8 on the integrated map of 'Anna' x 'Sharpe's Early', corresponds to a QTL identified for budbreak by Segura *et al.* (2007) and a QTL identified on the corresponding LG 7 of a previously published map (Conner, *et al.*, 1998, Maliepaard, *et al.*, 1998) for which a different LG numbering system was used. The QTL identified on LG 6 by Segura *et al.* could however not be confirmed during this investigation. The QTL identified on LG 9 in all three mapping populations under investigation, might corresponds to a QTL identified on the corresponding LG 3 (Maliepaard, *et al.*, 1998) by Conner *et al.* (1998). The 2 QTLs identified on LG 10 might include the genomic region/s identified on the corresponding LG 6 (Kenis and Keulemans, 2005) by Lawson *et al.* (1995) to contain two markers of which the segregation could be associated with time of vegetative budbreak.

Sixteen QTLs identified during this investigation could not be aligned to QTLs identified by Conner *et al.* due to a lack in transferable markers on the previously published map (Conner, *et al.*, 1998) hampering the correct alignment of LGs. This investigation thus resulted in the identification of up to 16 newly defined QTLs for time of IVB. The QTLs identified on LG 14 and LG

17 during this investigation were not identified during earlier studies (Maliepaard, *et al.*, 1998) on the corresponding LG 2 and LG 1 of the previously published map.

Markers of which the segregating alleles showed significant levels of association with differences in time of IVB can also be associated with the putative QTLs identified. These markers can potentially be used in marker-assisted selection (MAS) of individuals with an earlier time of IVB in local breeding programs, aimed at breeding cultivars with a lower CR that are better adapted to the mild winter conditions experienced in the Western Cape region of South Africa. Markers can be evaluated in terms of the observed difference between associated means in adult and juvenile populations and in terms of different genetic backgrounds in which these associations were made. These criteria can be used to select markers that will be used during validation studies before implementation in breeding programs.

A study, conducted based on the results obtained during this investigation, involves a second mapping population derived from a cross between 'Anna' and 'Golden Delicious' that is currently being assessed for fruit quality traits. Seedlings will be screened using a set of proposed markers in order to predict time of IVB. These trees will then be phenotyped for the trait in order to determine the success with which markers could be used to select the phenotype of interest. The following markers have been selected for this purpose based on results obtained in Chapter 6.

1. CH04a12 (LG 11). Interaction between parental alleles resulted in a difference of 20 days in mean time of IVB. This marker can

thus be used in crosses involving 'Anna' and 'Golden Delicious' only as involves intra-locus interactions.

2. CH04c06y (LG17). Segregation of alleles from the parental cultivar 'Anna' have been shown to be associated with differences in time of IVB in different genetic backgrounds and in different developmental stages. The difference between time of IVB ranged between 12 days in 'Anna' x 'Golden Delicious' trees to 24 days in adult trees from the 'Anna' x 'Sharpe's Early' mapping population.
3. CH01h01 (LG17). Segregation of alleles from the parental cultivar 'Anna' have been shown to be associated with differences in time of IVB in different genetic backgrounds, with differences in mean time of IVB of 10 days.
4. A267 (LG16). Segregation of alleles from the parental cultivar 'Anna' was shown to be associated with averaged differences in time of IVB of 13 days.

The implementation of 4 markers could theoretically lead to the selection of 6 % of the population expected to have budbreak occuring relatively earlier than other individuals. Looking at the 'Anna' x 'Golden Delicious' population of 88 individuals used during this investigation, 3 individuals (24, 48 and 327), having the desired allelic combinations, could be selected, representing 3.4 % of the population. The average time of IVB observed over the whole 4 year period for these individuals were significantly lower than the overall average time of IVB. The fact that there are individuals for which the average time of IVB are still earlier than for the 3 individuals

identified through the implementation of the above markers is an indication that there are more genes influencing the trait of interest than the 4 linked to the markers listed above.

Markers that are strong candidates for MAS and that could be considered when studying other mapping pedigrees include:

1. markers KA4b (LG 1), Hi02d04 (LG 10), Hi08h12 (LG 10) and Hi02f12 (LG 17) in crosses involving 'Sharpe's Early',
2. markers Hi04c10 (LG 3), CH03g12y (LG 3), CH01c06 (LG 8), Hi05e07 (LG 9), CH05g11 (LG 14), CH04c07 (LG 14) and A236 (LG 17) in 'Anna' x 'Sharpe's Early',
3. A209 and/or A310 (LG 3) and CH03a09 (LG 5) in crosses involving 'Golden Delicious'.

Results obtained during this investigation enhance our understanding of the genetic complexity of the "time of initial vegetative budbreak" trait, one of the many complex characteristics associated with dormancy. It has been proven that the trait has a strong genetic component controlled by several to many genes. This study reports different genomic regions containing genes with a possible involvement in determining the time of IVB. The identification of markers closely linked to such genes will, after verification studies, enable the implementation of MAS in local breeding programs for the selection of cultivars that are better adapted to local climatic conditions. This will have a positive impact as far as health and environmental conditions are concerned because the reduced need for application of dormancy breaking chemicals, together with labour costs, will greatly reduce producer costs. Increasing

production in the face of global warming and sub-optimal growing conditions will ensure that South Africa remains a major global exporter of apples.

8 BIBLIOGRAPHY

1. **Arora, R., L. J. Rowland, and K. Tanino.** 2003. Induction and release of bud dormancy in woody perennials: A science comes of age. HortScience **38:**911-921.
2. **Arus, P., T. Yamamoto, E. Dirlewanger, and A. Abbott.** 2006. Synteny in the Rosaceae, p. 175-211. *In* J. Janick (ed.), Plant breeding reviews, vol. 27. John Wiley & Sons, Inc.
3. **Asins, M. J.** 2002. Present and future of quantitative trait locus analysis in plant breeding. Plant Breeding **121:**281-291.
4. **Benson, G.** 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research **27:**573-580.
5. **Bradshaw, B. D. J., and R. F. Stettler.** 1995. Molecular genetics of growth and development in Populus. IV. Mapping QTLs with large effects on growth, form and phenology traits in a forest tree. Genetics **139:**963-973.
6. **Brownstein, M. J., J. D. Carpten, and J. R. Smith.** 1996. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques **20:**1004-1010.
7. **Burn, J. E., D. J. Bagnall, J. D. Metzger, E. S. Dennis, and W. J. Peacock.** 1993. DNA methylation, vernalization and the initiation of flowering. Proceedings of the National Academy of Sciences **90:**287-291.
8. **Bus, V., W. E. Van de Weg, C. E. Durel, C. Gessler, F. Calenge, L. Parisi, E. Rikkerink, S. Gardiner, A. Patocchi, M. Meulenbroek, H. Schouten, and F. Laurens.** 2004. Delineation of a scab resistance gene cluster on linkage group 2 of apple, p. 57-62, Eucarpia symposium on fruit breeding and genetics, vol. 663. Acta Horticulturae **663:**57-62, Angers, France.
9. **Bus, V. G. M., F. N. D. Laurens, W. E. Van de Weg, R. L. Rusholme, E. H. A. Rikkerink, S. E. Gardiner, H. C. M. Bassett, L. P. Kodde, and K. M. Plummer.** 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. New Phytol **166:**1035-1049.
10. **Bus, V. G. M., E. H. A. Rikkerink, W. E. Van de Weg, R. L. Rusholme, S. E. Gardiner, H. C. M. Bassett, L. P. Kodde, L. Parisi, F. N. D. Laurens, E. J. Meulenbroek, and K. M. Plummer.** 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.
11. **Butler, J. M., C. M. Ruitberg, and P. M. Vallone.** 2001. Capillary electrophoresis as a tool for optimization of multiplex reactions. Fresenius Journal for Analytical Chemistry **369:**200-205.
12. **Calenge, F., and C. E. Durel.** 2006. Both stable an unstable QTL's for resistance to powdery mildew are detected in apple after four years of field assessment. Molecular Breeding **17:**329-339.

13. **Calenge, F., A. Faure, M. Goerre, C. Gebhardt, W. E. Vande Weg, L. Parisi, and C. E. Durel.** 2004. Quantitative trait loci (QTL) analysis reveals both broad-spectrum and isolate-specific QTL for scab resistance in an apple progeny challenged with eight isolates of *Venturia inaequalis*. Phytopathology **94:**370-379.
14. **Calenge, F., C. G. Van der Linden, E. Van de Weg, H. J. Schouten, G. Van Arkel, C. Denance, and C. E. Durel.** 2005. Resistance gene analogues identified through the NBS-profiling method map close to major genes and QTL for disease resistance in apple. Theoretical and Applied Genetics **110:**660-668.
15. **Campbell, C. S., M. J. Donoghue, B. G. Baldwin, and M. F. Wojciechowski.** 1995. Phylogenetic relationships in Maloideae (Rosaceae): evidence from sequences of the internal transcribed spacers of nuclear ribosomal DNA and ITS congruence with morphology. American Journal of Botany **82:**903-918.
16. **Celton, J.-M., B. Ambrose, S. Tustin, I. Ferguson, T. Yamamoto, and S. Gardiner.** 2007. Presented at the Plant & Animal Genome XV conference, San Diego, CA, USA, January 2007.
17. **Celton, J.-M., R. Rusholme, S. Tustin, S. Ward, B. Ambrose, I. Ferguson, and S. Gardiner.** 2006. Presented at the 3rd International Rosaceae genomics conference, Napier, New Zealand, 19-22 March 2006.
18. **Cevik, V., and G. J. King.** 2002. High-resolution genetic analysis of the *Sd-1* aphid resistance locus in *Malus* spp. Theoretical and Applied Genetics **105:**346-354.
19. **Chagne, D., P. Chaumeil, A. Ramboer, C. Collada, A. Guevara, M. T. Cervera, G. G. Vendramin, V. Garcia, J.-M. Frigeria, C. Echt, T. Richardson, and C. Plomion.** 2004. Cross-species transferability and mapping of genomic and cDNA SSR's in pines. Theoretical and Applied Genetics **109:**1204-1214.
20. **Chee, P. W., J. Rong, D. Williams-Coplin, S. R. Schulze, and A. H. Paterson.** 2004. EST derived PCR-based markers for functional gene homologues in cotton. Genome **47:**449-462.
21. **Chen, T. H. H., J. Davis, B. E. Frewen, G. T. Howe, and H. D. Bradshaw.** 2000. Molecular Genetic Analysis of Bud Dormancy-related Traits in Populus, p. 319-329. *In* G. A. Lang (ed.), Dormancy in plants: from whole plant behaviour to cellular control. CABI Publishing, CAB International, Wallingford, Oxon, UK.
22. **Coart, E., X. Vekemans, M. J. M. Smulders, I. Wagner, J. Van Huylenbroeck, E. Van Bockstaele, and I. Roldan-Ruiz.** 2003. Genetic variation in the endangered wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as revealed by amplified fragment length polymorphism and microsatellite markers. Molecular Ecology **12:**845-857.
23. **Conley, E. J., V. Nduati, J. L. Gonzalez-Hernandez, A. Mesfin, M. Trudeau-Spanjers, S. Chao, G. R. Lazo, D. D. Hummel, O. D. Anderson, L. L. Qi, B. S. Gill, B. Echalier, A. M. Linkiewicz, J. Dubcovsky, E. D. Akhunov, J. Dvorak, J. H. Peng, N. L. V. Lapitan, M. S. Pathan, H. T. Nguyen, X.-F. Ma, Miftahudin, J. P. Gustafson, R. A. Greene, M. E. Sorrells, K. G. Hossain, V. Kalavacharla, S. F.**

Kianian, D. Sidhu, M. Dilbirgi, K. S. Gill, D. W. Choi, R. D. Fenton, T. J. Close, P. E. McGuire, C. O. Qualset, and J. A. Anderson. 2004. A 2600-locus chromosome bin map of wheat homoelogous group 2 reveals interstitial gene-rich islands and colinearity with rice. Genetics **104**:625-638.

24. **Conner, P. J., S. K. Brown, and N. F. Weeden.** 1998. Molecular-marker analysis of quantitative traits for growth and development in juvenile apple trees. Theoretical and Applied Genetics **96**:1027-1035.
25. **Conner, P. J., S. K. Brown, and N. F. Weeden.** 1997. Randomly amplified polymorphic DNA-based genetic linkage maps of three apple cultivars. Journal of the American Society for Horticultural Science **122**:350-359.
26. **Cook, N. C., A. Bellen, P. J. R. Cronje, I. De Wit, W. Keulemans, A. Van den Putte, and W. Steyn.** 2005. Freezing temperature treatment induces bud dormancy in 'Granny Smith' apple shoots. Scientia Horticulturae **106**:170-176.
27. **Cook, N. C., and D. U. Bellstedt.** 2001. Chilling response of 'Granny Smith' apple lateral buds inhibited y distal shoot tissues. Scientia Horticulturae **89**:299-308.
28. **Cook, N. C., D. U. Bellstedt, and G. Jacobs.** 2001. Endogenous cytokinin distribution patterns at budburst in 'Granny Smith' and 'Braeburn' apple shoots in relation to bud growth. Scientia Horticulturae **87**:53-63.
29. **Cook, N. C., and G. Jacobs.** 2000. Progression of apple (*Malus* x *domestica* Borkh.) bud dormancy in two mild winter climates. The Jounal of Horticultural Science and Biotechnology **75**:233-236.
30. **Costa, F., S. Stella, W. E. Van de Weg, W. Guerra, M. Cecchinel, J. Dallavia, B. Koller, and S. Sansavini.** 2005. Role of the genes *Md-ACO1* and *Md-ACS1* in ethylene production and shelf life of apple (*Malus domestica* Borkh). Euphytica **141**:181-190.
31. **Da Silva, J. A. G.** 2001. Preliminary analysis of microsatellite markers derived from sugarcane expressed sequence tags (EST's). Genetics and Molecular Biology **24**:155-159.
32. **Dakin, E. E., and J. C. Avise.** 2004. Microsatellite null alleles in parentage analysis. Heredity **93**:504-509.
33. **Doyle, J. J., and J. L. Doyle.** 1990. Isolation of plant DNA from fresh tissue. Focus **12**:13-15.
34. **Dunemann, F., G. Bracker, T. Markussen, and P. Roche.** 1999. Identification of molecular markers for the major mildew resistance gen *Pl2* in apple, p. 411-416, Eucarpia Symposium on fruit breeding and genetics, vol. 484. Acta Horticulturae **484**:411-416, Dresden, Germany.
35. **Durel, C. E., F. Calenge, L. Parisi, W. E. Van de Weg, L. Kodde, R. Liebhard, C. Gessler, M. Thiermann, F. Dunemann, F. Gennari, S. Tartarini, and Y. Lespinasse.** 2004. Overview on position and robustness of scab resistance QTL and major genes by allignment of genetic maps in five apple progenies, Eucarpia symposium on fruit breeding and genetics. Acta Horticulturae **663**:135-140, Angers, France.
36. **Durel, C. E., V. Fresion, C. Denance, F. Laurens, Y. Lespinasse, E. Rat, L. Parisi, V. Bus, E. Depena de la Fuente, M. Minarro, M. D.**

Blaquez, and W. E. Van de Weg. 2006. Presented at the 3rd International Rosaceae genomics conference, Napier, New Zealand, 19-22 March 2006.

37. **Durel, C. E., L. Parisi, F. Laurens, W. E. Van de Weg, R. Liebhard, and M. F. Jourjon.** 2003. Genetic dissection of partial resistance to race 6 of *Venturia inaequalis* in apple. Genome **46:**224-234.
38. **Echt, C., S. Knapp, and B.-H. Liu.** 1992. Genome mapping with non-inbred crosses using GMendel 2.0. Maize Genetics Cooperation Newsletter **66:**27-29.
39. **Erdin, N., S. Tartarini, G. A. L. Broggini, F. Gennari, S. Sansavini, C. Gessler, and A. Patocchi.** 2006. Mapping of the apple scab-resistance gene *Vb*. Genome **49:**1238-1245.
40. **Evans, K. M., and C. M. James.** 2003. Identification of SCAR markers linked to *Pl-w* mildew resistance in apple. Theoretical and Applied Genetics **106:**1178-1183.
41. **Falconer, D. S., and T. F. C. Mackay.** 1996. Introduction to quantitative genetics, 4th ed. Pearson, England.
42. **Frewen, B. E., T. H. H. Chen, G. T. Howe, J. Davis, A. Rodhe, W. Boerjan, and H. D. Bradshaw.** 2000. Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. Genetics **154:**837-845.
43. **Fujimori, S., T. Washio, K. Higo, Y. Ohtomo, K. Murakami, K. Matsubara, J. Kawai, P. Carninci, Y. Hayashizaki, S. Kikuchi, and M. Tomita.** 2003. A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription. FEBS Letters **554:**17-22.
44. **Gao, Z. S., W. E. Van de Weg, J. G. Schaart, H. J. Schouten, D. H. Tran, L. P. Kodde, I. M. Van der Meer, A. H. M. Van der Geest, J. Kodde, H. Breiteneder, K. Hoffmann-Sommergruber, D. Bosch, and L. J. W. J. Gilissen.** 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.
45. **Gao, Z. S., W. E. Van de Weg, J. G. Schaart, I. M. Van der Meer, L. Kodde, M. Laimer, H. Breiteneder, K. Hoffmann-Sommergruber, and G. L. J. W. J.** 2005b. 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 **110:**479-491.
46. **Gardiner, S., V. G. M. Bus, R. L. Rusholme, D. Chagne, and E. H. A. Rikkerink.** 2007. Apple, p. 1-62. *In* C. Kole (ed.), Genome mapping and molecular breeding in plants. Fruits and Nuts, vol. 4. Springer-Verlag, Berlin Heidelberg.
47. **Gianfranceschi, L., and V. Soglio.** 2004. The European project HIDRAS: Innovative Multidisciplinary approaches to breeding high quality disease resistant apples, Eucarpia Symposium on fruit breeding and genetics. Acta Horticulturae **663:**327-330, Angers, France.
48. **Goldstein, D. B., G. W. Roemer, D. A. Smith, D. E. Reich, A. Bergman, and R. K. Wayne.** 1999. The use microsatellite variation to infer population structure and demographic history in a natural model system. Genetics **151:**797-801.

49. **Guilford, P., S. Prakash, J. M. Zhu, E. Rikkerink, S. Gardiner, H. Bassett, and R. Forster.** 1997. Microsatellites in *Malus* x *domestica* (apple): abundance, polymorphism and cultivar identification. Theoretical and Applied Genetics **94:**249-254.

50. **Gupta, P. K., S. Rustgi, S. Sharma, R. Singh, N. Kumar, and H. S. Balyan.** 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. Theoretical and Applied Genetics **270:**315-323.

51. **Gygax, M., L. Gianfranceschi, R. Liebhard, M. Kellerhals, C. Gessler, and A. Patocchi.** 2004. Molecular markers linked to the apple scab resistance gene *Vbj* derived from *Malus baccata jackii.* Theoretical and Applied Genetics **109:**1702-1709.

52. **Hauagge, R., and J. N. Cummins.** 1991. Genetics of length of dormancy period in *Malus* vegetative buds. Journal of the American Society for Horticultural Science **116:**121-126.

53. **Hearne, C. M., S. Ghosh, and J. A. Todd.** 1992. Microsatellites for linkage analysis of genetic traits. Trends in Genetics **8:**288-294.

54. **Hemmat, M., B. S. K., and N. F. Weeden.** 2002. Tagging and mapping scab-resistance genes from R12740-7A apple. Journal of the American Society for Horticultural Science **127:**365-370.

55. **Hemmat, M., N. F. Weeden, A. G. Manganaris, and D. M. Lawson.** 1994. Molecular marker linkage map for apple. Journal of Heredity **85:**4-11.

56. **Hokanson, S. C., A. K. Szewe-McFadden, W. F. Lamboy, and M. 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.

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

58. **Iqbal, M., A. Navabi, R.-C. Yang, D. F. Salmon, and D. Spaner.** 2007. Molecular characterization of vernalization response genes in Canadian spring wheat. Genome **50:**511-516.

59. **Isidore, E., H. Van Os, A. Andrzejewski, J. Bakker, I. Barrena, G. J. Bryan, B. Caromel, H. van Eck, B. Ghareeb, W. de Jong, P. Van Koert, V. Lefebvre, D. Milbourne, E. Ritter, J. R. van der Voort, F. Rouselle-Bourgeois, J. Van Vliet, and R. Waugh.** 2003. Towards a marker-dense meiotic map of the potato genome: lessons from linkage group 1. Genetics **165:**2107-2116.

60. **James, C. M., J. B. Clarke, and K. M. Evans.** 2004. Identification of molecular markers linked to the mildew resistance gene *Pl-d* in apple. Theoretical and Applied Genetics **110:**175-181.

61. **James, C. M., and K. M. Evans.** 2004. Identification of molecular markers linked to the powdery mildew resistance genes *Pl-d* and *Pl-w* in apple, p. 123-127, Eucarpia Symposium on fruit breeding and genetics, vol. 663. Acta Horticulturae **663:**123-127, Angers, France.

62. **Janick, J.** 2005. The Origins of Fruit, Fruit Growing, and Fruit Breeding. John Wiley & Sons, Inc.

63. **Janick, J., J. N. Cummins, S. K. Brown, and M. Hemmat.** 1996. Apples, p. 3-37. *In* J. Janick and J. N. Moore (ed.), Fruit breeding, Volume 1: Tree and tropical fruits. John Wiley & sons, Inc.
64. **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.
65. **Jansen, J., A. G. De Jong, and J. W. Van Ooijen.** 2001. Constructing dense genetic linkage maps. Theoretical and Applied Genetics **102:**1113-1122.
66. **Jansen, R. C.** 1994. Controlling the Type I and Type II errors in mapping quantitative trait loci. Genetics **138:**871-881.
67. **Kenis, K., and J. Keulemans.** 2005. Genetic linkage maps of two apple cultivars (*Malus* x *domestica* Borkh.) based on AFLP and microsatellite markers. Molecular Breeding **15:**205-219.
68. **Kenis, K., and J. Keulemans.** 2004. QTL Analysis of growth characterisitcs in apple, p. 369-374, Eucarpia symposium on fruit breeding and genetics, vol. 663. Acta Horticulturae **663:**369-374, Angers, France.
69. **Khan, M. A., B. Duffy, C. Gessler, and A. Patocchi.** 2006. QTL mapping of fire blight resistance in apple. Molecular Breeding **17:**299-306.
70. **Kim, C. S., C. H. Lee, J. S. Shin, Y. S. Chung, and N. I. Hyung.** 1997. A simple and rapid method for isolation of high quality genomic DNA from fruit trees and conifers using PVP. Nucleic Acids Research **25:**1085-1086.
71. **King, G. J., J. R. Lynn, C. J. Dover, K. M. Evans, and G. B. Seymour.** 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.:1227-1235.
72. **King, G. J., C. Maliepaard, J. R. Lynn, F. H. Alston, C. E. Durel, K. M. Evans, B. Griffon, F. Laurens, A. G. Manganaris, E. Schrevens, S. Tartarini, and J. Verhaegh.** 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.
73. **Korban, S. S., and R. M. Skirvin.** 1994. Nomenclature of the cultivated apple. HortScience **19:**177-180.
74. **Korban, S. S., L. O. Vodkin, L. Liu, H. S. Aldwinckle, K. Gasic, D. O. Gonzalez, M. Malnoy, J. Thimmapuram, H. J. Carroll, P. Goldsbrough, K. Orvis, S. Clifton, D. Pape, M. Martin, and R. Meyer.** 2005. Presented at the Plant & Animal Genomes XIII Conference, SanDiego, CA., January 2005.
75. **Kosambi, D. D.** 1994. The estimation of map distances from recombination values. Ann. Eugen. **12:**172-175.
76. **Kotoda, N., M. Wada, S. Kusaba, Y. Kano-Marukami., T. Masuda, and J. Soejima.** 2002. Overexpression of *MdMADS5*, an *APETALA1*-like gene of apple, causes early flowering in transgenic *Arabidopsis*. Plant Science **162:**679-687.

77. **Kuleung, C., P. S. Baenzinger, and I. Dweikat.** 2004. Transferability of SSR markers among wheat, rye and triticale. Theoretical and Applied Genetics. **108:**1147-1150.
78. **Labuschagné, I. F.** 2004. Presented at the Eleventh Eucarpia symposium on fruit breeding and genetics., Angers, France., 1 - 3 Sept 2003.
79. **Labuschagné, I. F., J. H. Louw, K. Schmidt, and A. Sadie.** 2002a. Genetic variation in chilling requirement in apple progeny. Journal of the American Society for Horticultural Science **127:**663-672.
80. **Labuschagné, I. F., J. H. Louw, K. Schmidt, and A. Sadie.** 2002b. Genotypic variation in prolonged dormancy symptoms in apple families. HortScience **37:**157-163.
81. **Labuschagné, I. F., J. H. Louw, K. Schmidt, and A. Sadie.** 2003. Selection for increased budbreak in apple. Journal of the American Society for Horticultural Science **128:**363-374.
82. **Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly, S. E. Lincoln, and L. Newburg.** 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:**174-181.
83. **Lawson, D. M., M. Hemmat, and N. F. Weeden.** 1995. The use of molecular markers to analyze the inheritance of morphological and developmental traits in apple. Journal of the American Society for Horticultural Science **120:**532-537.
84. **Lespinasse, Y., and C. E. Durel.** 1999. Presented at the Plant & Animal Genome VII conference, San Diego, CA, USA, January 1999.
85. **Liebhard, R., L. Gianfranceschi, B. Koller, C. D. Ryder, R. Tarchini, E. Van de Weg, and C. Gessler.** 2002. Development and characterisation of 140 new microsatellites in apple (*Malus* x *domestica* Borkh.). Molecular Breeding **10:**217-241.
86. **Liebhard, R., M. Kellerhals, W. Plammatter, M. Jertmni, and C. Gessler.** 2003a. Mapping quantitative physiological traits in apple (*Malus* x *domestica* Borkh.). Plant Molecular Biology **52:**511-526.
87. **Liebhard, R., B. Koller, L. Gianfranceschi, and C. Gessler.** 2003b. Creating a saturated reference map for the apple (*Malus* x *domestica* Borkh.) genome. Theoretical and Applied Genetics **106:**1497-1508.
88. **Lin, Z., X. Cui, and H. Li.** 1996. Multiplex genotype determination at a large number of gene loci. Proceedings of the National Academy of Sciences **93:**2582-2587.
89. **Linsley-Noakes, G. C., P. Allan, and G. Matthee.** 1994. Modification of rest completion prediction models for improved accuracy in South African stone fruit orchards. Journal of South African Horticultural Science **4:**13-15.
90. **Liu, Z. W., R. M. Biyashev, and M. A. Saghai Maroof.** 1996. Development of simple sequence repeat DNA markers and their integration into a barley linkage map. Theoretical and Applied Genetics **93:**869-876.
91. **Maliepaard, C., F. H. Alston, G. Van Arkel, L. M. Brown, E. Chevreau, F. Dunemann, K. M. Evans, S. Gardiner, P. Guilford, A. W. Vand Heusden, J. Janse, F. Laurens, J. R. Lynn, A. G. Manganaris, A. P. M. Den Nijs, N. Periam, R. E., P. Roche, C.**

Ryder, S. Sansavini, H. Schmidt, S. Tartarini, J. J. Verhaegh, M. Vrielink-van Ginkel, and G. J. King. 1998. Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theoretical and Applied Genetics **97**:60-73.

92. **Maliepaard, C., J. Jansen, and J. W. Van Ooijen.** 1997. Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences of applications. Genetical Research **70**:237-250.
93. **Michaels, S. D., and R. M. Amasino.** 1999. *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell **11**:949-956.
94. **Naik, S., C. Hampson, K. Gasic, G. Bakkeren, and S. S. Korban.** 2006. Development and linkage mapping of E-STS and RGA markers for functional gene homologues in apple. Genome **49**:959-968.
95. **Newcomb, R. D., R. N. Crowhurst, A. P. Gleave, E. H. A. Rikkerink, A. C. Allan, L. L. Beuning, J. H. Bowen, E. Gera, K. R. Jamieson, B. J. Janssen, W. A. Laing, S. McArtney, B. Nain, G. C. Ross, S. K. C., E. J. F. Souleyre, E. F. Walton, and Y.-K. Yauk.** 2006. Analysis of expressed sequence tags from apple (*Malus* x *domestica*). Plant Physiology **141**:147-166.
96. **Nicot, N., V. Chiquet, B. Gandon, L. Amilhat, F. Legeai, P. Leroy, M. Bernard, and P. Sourdille.** 2004. Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (EST's). Theoretical and Applied Genetics **109**:800-805.
97. **Okubo, H.** 1996. Growth Cycle and Dormancy in Plants, p. 1-22. *In* G. A. Lang (ed.), Dormancy in plants: from whole plant behaviour to cellular control. CABI Publishing, CAB International, Wallingford, Oxon, UK.
98. **Oraguzie, N. C., T. Yamamoto, J. Soejima, T. Suzuki, and H. N. De Silva.** 2005. DNA fingerprinting of apple (*Malus* spp.) rootstocks using simple sequence repeats. Plant Breeding **124**:197-202.
99. **Patocchi, A., B. Bigler, B. Koller, M. Kellerhals, and C. Gessler.** 2004. *Vr(2)*: a new apple scab resistance gene. Theoretical and Applied Genetics **109**:1087-1092.
100. **Patocchi, A., M. Walser, S. Tartarini, G. A. L. Broggini, F. Gennari, S. Sansavini, and C. Gessler.** 2005. Identification by genome scanning approach (GSA) of a microsatellite tightly associated with the apple scab resistance gene *Vm*. Genome **48**:630-636.
101. **Peil, A., F. Dunemann, T. Garcia, K. Richter, B. Trognitz, V. Hanke, and H. Flachowsky.** 2006. Presented at the The 3rd International Rosaceae Genomics Conference, 19-22 March 2006.
102. **Pierantoni, L., K.-M. Cho, I.-S. Shin, R. Chiodini, S. Tartarini, L. Dondini, S.-J. Kang, and S. Sansavini.** 2004. Characterisation and transferability of apple SSRs to two European pear F1 populations. Theoretical and Applied Genetics **109**:1519-1524.
103. **Qi, L. L., B. Echalier, S. Chao, G. R. Lazo, J. M. Butler, J. R. Andersen, E. D. Akhunov, J. Dvorak, A. M. Linkiewicz, A. Ratnasiri, J. Dubcovsky, C. E. Bermudez-Kandianis, R. A. Greene, R. Kantety, C. M. La Rota, J. D. Munkvold, S. F. Sorrells, M. E. Sorrells, M. Dilbirgi, D. Sidhu, M. Erayman, H. S. Randhawa, D. Sandhu, S. N. Bondareva, K. S. Gill, A. A. Mahmoud, X.-F. Ma,**

Miftahudin, J. P. Gustafson, E. J. Conley, V. Nduati, J. L. Gonzalez-Hernandez, J. A. Anderson, J. H. Peng, N. L. V. Lapitan, K. G. Hossain, V. Kalavacharla, S. F. Kianian, M. S. Pathan, D. S. Zhang, H. T. Nguyen, D. W. Choi, R. D. Fenton, T. J. Close, P. E. McGuire, C. O. Qualset, and B. S. Gill. 2004. A chromosome bin map of 16,000 espressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. Genetics **104:**701-712.

104. **Reeves, P. H., and G. Coupland.** 2000. Response of plant development to environment: control of flowering by daylength and temperature. Current Opinion in Plant Biology **3:**37-42.
105. **Schoske, R., P. M. Vallone, C. M. Ruitberg, and J. M. Butler.** 2003. Multiplex PCR design strategy used for the simultaneous amplification of 10Y chromosome short tandem repeat (STR) loci. Analytical and Bioanalytical Chemistry **375:**333-343.
106. **Seglias, N. P., and C. Gessler.** 1997. Genetics of apple powdery mildew resistance from *Malus zumi* (*Pl2*). IOBC/WPRS Bull **20:**195-208.
107. **Segura, V., C. Denance, C. E. Durel, and E. Costes.** 2007. Wide range QTL analysis for complex architectural traits in a 1-year-old apple progeny. Genome **50:**159-171.
108. **Seymour, G. B., K. Manning, E. M. Eriksson, A. H. Popovich, and G. J. King.** 2002. Genetic identification and genomic organization of factors affecting fruit texture. Journal of Experimental Botany **53:**2065-2071.
109. **Sheldon, C. C., E. J. Finnegan, D. T. Rouse, M. Tadege, D. J. Bagnall, C. A. Helliwell, W. J. Peacock, and E. S. Dennis.** 2000. The control of flowering by vernalization. Current Opinion in Plant Biology **3:**418-422.
110. **Silfverberg-Dilworth, E., S. Besse, R. Paris, E. Belfanti, S. Tartarini, S. Sansavini, A. Patocchi, and C. Gessler.** 2005. Identification of functional apple scab resistance gene promoters. Theoretical and Applied Genetics **110:**1119-1126.
111. **Silfverberg-Dilworth, E., C. L. Matasci, W. E. Van de Weg, M. P. W. Van Kaauwen, M. Walser, L. P. Kodde, V. Soglio, L. Gianfranceschi, C. E. Durel, F. Costa, T. Yamamoto, B. Koller, C. Gessler, and A. Patocchi.** 2006. Microsatellite markers spanning the apple (*Malus* x *domestica* Borkh.) genome. Tree Genetics and Genomes **2:**202-224.
112. **Stam, P.** 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant Journal **3:**739-744.
113. **Stankiewicz-Kosyl, M., E. Pitera, and S. W. Gawronski.** 2005. Mapping QTL involved in powdery mildew resistance of the apple clone U 211. Plant Breeding **124:**63-66.
114. **Sung, S., and R. M. Amisino.** 2004. Vernalization and epigenetics: how plants remember winter. Current Opinion in Plant Biology **7:**4-10.
115. **Tartarini, S., F. Gennari, D. Pratesi, C. Palazzetti, S. Sansavini, L. Parisi, A. Fouillet, V. Fouillet, and C. E. Durel.** 2004. Characterisation and genetic mapping of a major scab resistance gene from the old Italian apple cultivar 'Durello di Forli', Eucarpia symposium

on fruit greeding and genetics. Acta Horticulturae **663:**129-133, Angers, France.

116. **Van de Weg, W. E., R. E. Voorrips, R. Finkers, L. P. Kodde, J. Jansen, and M. C. A. M. Bink.** 2003. Pedigree genotyping: A new pedigree-based approach of QTL identification and allele mining, Eucarpia symposium on fruit breeding and genetics, vol. 663. Acta horticulturae **663:**45-50, Angers, France.
117. **Van Ooijen, J. W.** 1992. Accuracy of mapping quantitative trati loci in autogamous species. Theoretical and Applied Genetics **84:**803-811.
118. **Van Ooijen, J. W.** 2006. JoinMap® 4. Software for the calculation of genetic linkage maps in experimental populations. Kyazma B. V., Wageningen, Netherlands.
119. **Van Ooijen, J. W.** 1999. LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity **83:**613-624.
120. **Van Ooijen, J. W.** 2004. MapQTL® 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma B. V., Wageningen, Netherlands.
121. **Varshney, R. K., A. Graner, and M. E. Sorrells.** 2005. Genic microsatellite markers in plants: features and applications. Trends in Biotechnology **23:**48-55.
122. **Vision, T. J., D. G. Brown, D. B. Shmoys, R. T. Durrett, and S. D. Tanksley.** 2000. Selective mapping: a strategy for optimizing the construction of high density linkage maps. Genetics **155:**407-420.
123. **Voorrips, R. E.** 2002. MapChart: Software for the graphical representation of linkage maps and QTL's. The Journal of Heredity **93:**77-78.
124. **Wand, S. J. E.** 2007. Is Global climate change becoming a regional issue? Perspectives for fruit production in the Western Cape. SA Fruit Journal:24-26.
125. **Weeden, N. F., M. Hemmat, D. M. Lawson, M. Lodhi, R. L. Bell, A. G. Manganaris, B. I. Reisch, S. K. Bown, and G.-N. Ye.** 1994. Development and application of molecular marker linkage maps in woody fruit crops. Euphytica **77:**71-75.
126. **Yakovlev, I., C.-G. Fossdal, O. Johnsen, O. Junttila, and T. Skroppa.** 2006. Analysis of gene expression during bud burst initiation in Norway spruce via ESTs from subtracted cDNA libraries. Tree Genetics and Genomes **2:**39-52.
127. **Yamamoto, T., T. Kimura, T. Saito, K. Kotobuki, N. Matsuia, R. Liebhard, C. Gessler, W. E. Van de Weg, and T. Hayashi.** 2004. Genetic linkage maps of Japanese and European pears aligned to apple consensus map, Eucarpia symposium on fruit breeding and genetics, vol. 663. Acta Horticulturae **663:**51-56, Angers, France.
128. **Yamamoto, T., T. Kimura, Y. Sawamura, K. Kotobuki, Y. Ban, T. Hayashi, and N. Matsuta.** 2001. SSR's isolated from apple can identify polymorphism and genetic diversity in pear. Theoretical and Applied Genetics **102:**865-870.
129. **Yamamoto, T., T. Kimura, Y. Sawamura, T. Manabe, K. Kotobuki, T. Hayashi, Y. Ban, and N. Matsuta.** 2002a. Simple sequence repeats for genetic analysis in pear. Euphytica **124:**129-137.

130. **Yamamoto, T., T. Kimura, M. Shoda, Y. Ban, T. Hayashi, and N. Matsuta.** 2002b. Development of microsatellite markers in the Japanese pear (*Pyrus pyrifolie* Nakai). Molecular Ecology Notes **2:**14-16.
131. **Yamamoto, T., T. Kimura, M. Shoda, T. Imai, T. Saito, Y. Sawamura, K. Kotobuki, T. Hayashi, and N. Matsuta.** 2002c. Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. Theoretical and Applied Genetics **106:**9-18.
132. **Yin, X., P. Stam, M. J. Kropff, and A. H. C. M. Schapendonk.** 2003. Crop modeling, QTL mapping and their complementary role in plant breeding. Agronomy Journal **95:**90-98.
133. **Young, N. D.** 1996. QTL mapping and quantitative disease resistance in plants. Annual Review of Phytopathology **34:**479-501.