Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Comparative Genome Mapping of the Rosaceae

A thesis presented in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in

Plant Molecular Genetics and Genomics

Institute of Molecular BioSciences
Massey University
Palmerston North
New Zealand

Jill M. Bushakra

2012

ABSTRACT

Comparative genome mapping uses genetic map and DNA sequence alignment to assess genome conservation between two or more organisms. This study makes use of the recent genome sequence availability of four Rosaceae genera, and the development of new, and the expansion of existing, linkage maps to: 1) explore overall genome synteny between apple and strawberry; 2) assess homology between, and the degree of ancestral genome rearrangement among, four genera; and 3) compare genome synteny with respect to the production of anthocyanins between raspberry and strawberry.

The inter-tribal comparison of the genomes of apple and diploid strawberry, conducted by adding 56 newly developed orthologous markers to existing linkage maps, identified 21 regions of genomic synteny between the linkage groups of apple and strawberry. In addition, this work identified two each of potential translocations, inversions and insertions, and provided a set of orthologous markers that will be useful for orienting and anchoring other Rosaceae genome sequences.

Orthologous- and other DNA sequence-based markers were used in the construction of new linkage maps for *Rubus occidentalis* 96395S1 and *R. idaeus* 'Latham'. The sequences from which the *Rubus* markers were designed were compared with the draft genome sequences of *Malus* × *domestica* 'Golden Delicious', *Fragaria vesca* 'Hawaii 4', and *Prunus persica* 'Lovell' to identify regions of orthology. This first comparison of *Rubus* linkage maps with other members of the Rosaceae identified a nearly 1:1 homology between the linkage groups of *Rubus* and *F. vesca*, as well as family-wide conservation among some linkage groups.

The F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' was used to conduct a quantitative trait locus (QTL) analysis to explore the presence of associations between genotype and the variation in concentrations of anthocyanins in the fruit. Seven associations of traits with markers designed from the sequences of transcription factors and anthocyanin biosynthetic pathway genes were identified, providing opportunities for further fine-scale mapping, as well as cloning and expression analyses. The comparison of QTL maps of *Rubus* and *Fragaria* × *ananassa* suggests that homologous genomic regions may be important in the expression of various fruit quality traits.

ACKNOWLEDGEMENTS

It's hard to believe that four years ago a PhD was furthest from my mind and yet, thanks to a brief comment by David Chagné at PAG 2008, here I am. I could not have completed this journey without the help and support of so many, but primarily I wish to thank my husband, John, for his faith, patience, support and commitment to me and my thirst for learning. We've been down another long road together.

Next, I would like to thank my advisors, especially David Chagné and Susan Gardiner for giving me this opportunity and always following through with their promises. Through this venture I have relied heavily on you both and you have always held me up. Thanks also to V. Vaughan Symonds and Emily Buck. Thank you also to Plant and Food Research (PFR) for funding from the Excellence Programme and to Tony Conner for the extension. The people of PFR have been very welcoming, encouraging and fun.

I would like to extend my appreciation to Michael McManus, Paul Dijkwell, and Brent Barrett for seeing my potential and providing helpful and constructive comments during my first two years. A special thanks to Nadia Ilyin for making it possible for John to work from NZ, and to Mom and Janis for taking care of the cats and for being okay with us being so far away.

My sanity was maintained with the help of my friends, especially Amy Watson, Toshi Foster, Claudia Wiedow, Cindy Skema, Fronny Plume, Mariam Sharland, Erica Prier and my family and friends back in California, especially Ed Show, Jennifer Wong-Izzo, and Albert Herrera for keeping me in the loop.

Finally, I would like to dedicate this thesis to my parents, who instilled in me a sense of adventure and willingness to try new things. I miss you both.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

A	Adenosine
a/a/p	*Amygdalus*/*Armeniaca*/*Prunocerasus*
ACY	Anthocyanins
AFLP	Amplified fragment length polymorphism
Amplicon	Amplified product of PCR
BLAST	Basic local alignment search tool
bp	Base pair
C	Cytosine
c/l/p	*Cerasus*/*Laurocerasus*/*Padua*
C3G	Cyanindin 3-*O*-glucoside
C3GR	Cyanindin 3-*O*-2^{G}-glucosylrutinoside
C3R	Cyanindin 3-*O*-rutinoside
C3S	Cyanindin 3-*O*-sophoroside
C3Sb	Cyanindin 3-*O*-sambubioside
C3XR	Cyanindin 3-*O*-2^{G}-xylosylrutinoside
cDNA	Complementary DNA
cM	Centimorgans
Contig	Contiguous sequence
COS	Conserved orthologous set
DNA	Deoxyribonucleic acid
dsDNA	Double stranded DNA
ESI	Electrospray interface
EST	Expressed sequence tag
FISH	Fluorescence *in situ* hybridization
FV×FB	*Fragaria vesca* × *F. bucharica*
FW	Fresh weight
g	Gram
G	Guanine
GDD	Growing degree days
gDNA	Genomic DNA
gSSR	Genomic SSR
HG	Homeologous group
HPLC	High performance liquid chromotography
HRM	High-resolution melting
IM	Interval mapping
indel	Insertion or deletion
kb	Kilobase
K-S DMax	Kolmogorov-Smirnov Dmax test
LG	Linkage group
LOD	Logarithm of odds
M.9×R5	‘Malling 9’ × Robusta 5

MAB	Marker assisted breeding
Mb	Megabase
µg	Microgram
mL	Milliliter
MQM	Multiple-QTL model
mRNA	Messenger RNA
MS	Mass spectrometry
my	Million years
nm	Nanometers
P3R	Pelargonidin 3-*O*-rutinoside
PCR	Polymerase chain reaction
pg	Picogram
QTL	Quantitative trait locus
R gene	Resistance gene
RAPD	Random amplification of polymorphic DNA
Rf	Recombination frequency
RFLP	Restriction fragment length polymorphism
RLG	*Rubus* linkage group
SCAR	Sequence characterized amplified region
SD	Standard deviation
s.e.	Standard error
SEM	Standard error of the mean
SNP	Single nucleotide polymorphism
sp.	Species
SSR	Simple sequence repeat
STS	Sequence tagged site
T	Thymine
T×E	‘Texas’ × ‘Earlygold’
TS	Transgressive segregant
UHPLC	Ultra high performance liquid chromotography
unigene	Unique gene
x	Base chromosome number

LIST OF COMMONLY REFERRED TO ROSACEAE SPECIES

Scientific name	Chromosome number	Common name
Cydonia oblonga	2*n*=34	quince
Eriobotyra japonica	2*n*=34	loquat
Fragaria bucharica	2*n*=14	
Fragaria vesca	2*n*=14	alpine strawberry
Fragaria × *ananassa*	8*n*=56	cultivated strawberry
Malus × *domestica*	2*n*=34	apple
Prunus armeniaca	2*n*=16	apricot
Prunus avium	2*n*=16	sweet cherry
Prunus ceracus	2*n*=16	sour cherry
Prunus domestica	2*n*=16	European plum
Prunus dulcis	2*n*=16	almond
Prunus persica	2*n*=16	peach
Pyrus communis	2*n*=34	European pear
Pyrus pyrifolia	2*n*=34	Asian pear
Rosa sp.	2*n*=14	rose
Rubus idaeus	2*n*=14	red raspberry
Rubus occidentalis	2*n*=14	black raspberry
Rubus sp.	4*n*=28	blackberry/hybrid berry

CHAPTER ONE

1 INTRODUCTION

1.1 Comparative genome mapping

Can we use the information gained from the study of one organism to expand our knowledge of another organism? Would the comparison of closely related organisms allow the discovery of lineage-specific differences? Is it possible to reconstruct an ancestral genome by comparing extant genomes? These questions can be addressed using comparative genome mapping, which is the use of molecular markers and DNA sequence to assess genome conservation between two or more organisms (Hardison 2003; Feuillet and Keller 2002). Comparative genome mapping can be used to detect and align orthologous regions of the genome which have been conserved through evolutionary time from a common ancestor. However, as similarity does not necessarily equate to homology, that is, a common origin (Hardison 2003), genome alignment of multiple taxa with differing phylogenetic relationships is required to identify the extent of similarity accurately (Doyle and Davis 1998). Furthermore, lineage-specific differences can be identified by comparing genomes of multiple closely related taxa (Hardison 2003).

Comparative genome mapping encompasses the alignment of genomes using molecular markers or DNA sequence. I will begin this chapter with an introduction of the principles of comparative genome mapping, followed by a discussion of genetic map construction, a prerequisite for genome comparisons in the absence of whole genome sequence. Next, I will describe molecular marker development for genetic mapping, including a discussion on marker development from orthologous sequence sources. I will then discuss an application for genetic mapping, the determination of genome regions involved in the expression of complex traits, and what is required for quantitative trait locus (QTL) mapping. Next, I will present examples of how genetic and QTL maps are used for comparative genome studies in plants. I will then introduce the biology and economic importance of the Rosaceae (rose family), followed by the current state of genetic and comparative genome mapping in the family. Finally, I will present the objectives of my project and demonstrate how they relate to the current knowledge of comparative genome and genetic mapping in the Rosaceae.

1.2 Principles of comparative genome mapping

The principle of comparative genome mapping is the alignment of genomes using molecular markers or DNA sequence in common. Comparative genome mapping requires the understanding of some basic terms. Synteny is defined as markers or genes on the same chromosome (Hartl and Jones 2009 p. 152). For example, the chromosomes illustrated in Fig. 1.1, have five markers in common (thin blue lines), indicating that they are homologous. The order of the markers along the chromosome compared with a chromosome in another genome defines the degree of genome colinearity (Paterson and Bennetzen 2001), as is illustrated by the difference in marker order, such that the chromosome from Genome B has one region of colinearity compared with the chromosome from Genome A. Markers used for comparative genome mapping must be derived from a common ancestor and be orthologous, though difficulties can arise when attempting to distinguish orthology from paralogy, those regions that have duplicated within a genome (Fitch 1970). The relationship between orthologous and paralogous genes is illustrated in Fig. 1.2. Regions of the genome that are inherited by vertical descent are orthologous, while those regions that undergo duplication within the genome are paralogs. Homology is used to define those regions that are not inherited by vertical descent nor duplicated within the same genome, as may occur when genes are lost through evolution (Fitch 1970; Doyle and Gaut 2000).

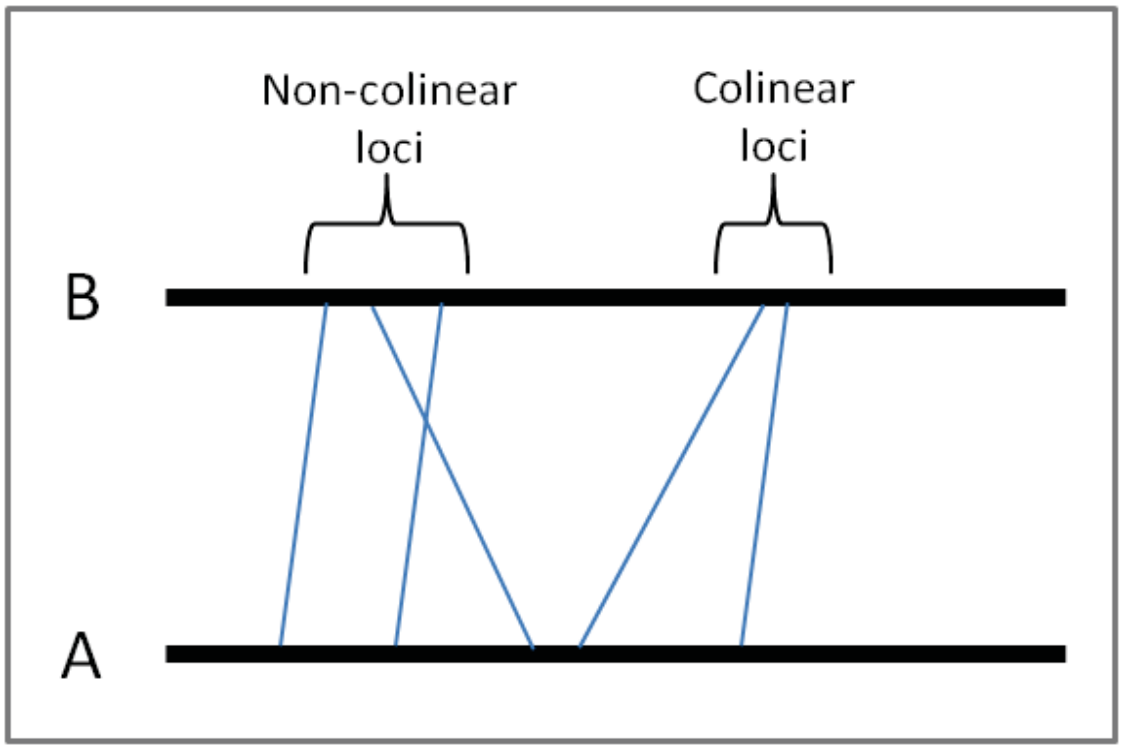


Fig. 1.1: Simple illustration of comparative genome mapping, synteny and colinearity. Markers on the chromosome from Genome A constitute a syntenic group, as do markers on the chromosome from Genome B (thick black lines). Colinearity is demonstrated by loci in the same order on both chromosomes.

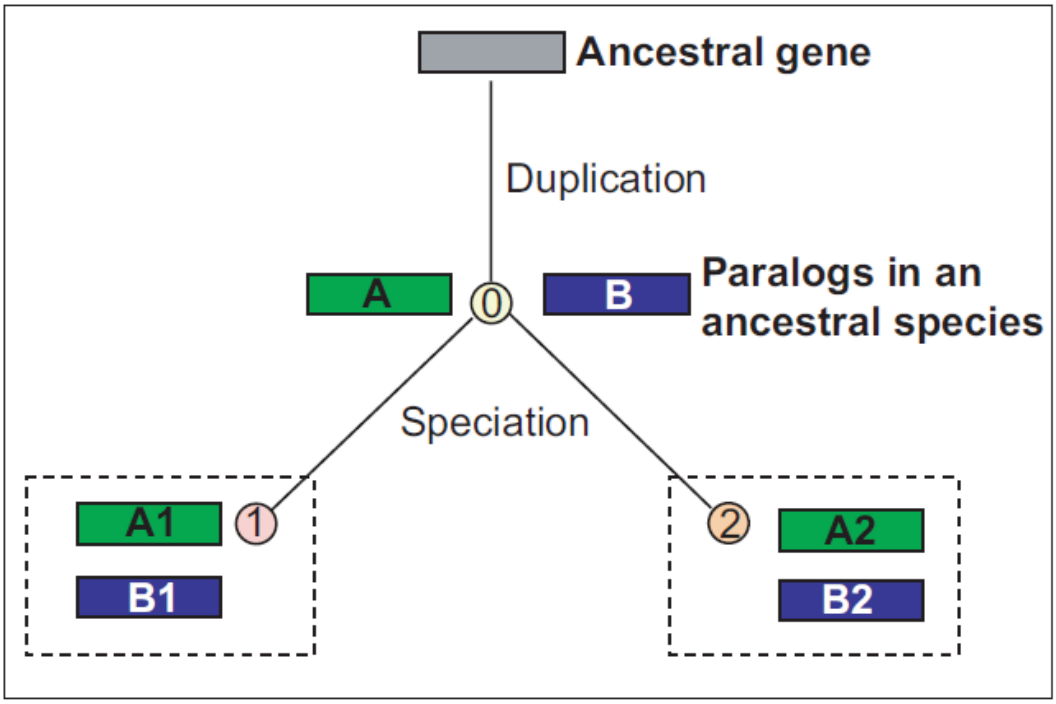


Fig. 1.2: Relationship between orthologs and paralogs. An ancestral gene (grey) undergoes duplication in Species 0 to form paralogous genes A & B. A speciation event leads to Species 1 and 2. Genes A1 & A2 are orthologs, as are genes B1 & B2. A1 & B1 are paralogs, as are A2 & B2. However, A1 & B2 (and A2 & B1) are considered homologous, as they are not inherited by vertical descent (orthologs), nor are they duplicated within the same genome (paralogs). Figure copied from Koonin (2001) Figure 1.

Before a comparative genome mapping project can be attempted, genetic linkage maps are needed for the taxa to be compared.

1.2.1 Genetic linkage map construction

A genetic linkage map is a visual representation of the genome of an individual or a taxon. Each linkage group (LG) represents a whole or partial single chromosome and a complete, or saturated, map will have as many LG as there are chromosomes, with genetic markers spaced evenly along the LG. The relative genetic distances between markers along a chromosome are generally proportional to the meiotic recombination between markers; that is, a greater genetic distance is indicative of more frequent recombination between two markers (Sturtevant 1913) (Fig. 1.3). However, recombination frequency and genetic distance are not completely proportional nor additive since double recombination events can occur within a region. Two equations,

Haldane's or Kosambi's, are used to compensate for the skewing in genetic distance caused by double recombination. The Haldane mapping function is used to calculate the proportion of double recombination events over map distance, without accounting for interference from recombination events in other regions (Crow 1990; Huehn 2011). The more commonly used Kosambi mapping function is based on empirical data regarding the proportion of double recombination events as the physical distance varies. The Kosambi mapping function takes into account the interference on one chromosome region because of a recombination event in another region and adjusts the genetic map distance (Crow 1990; Huehn 2011).

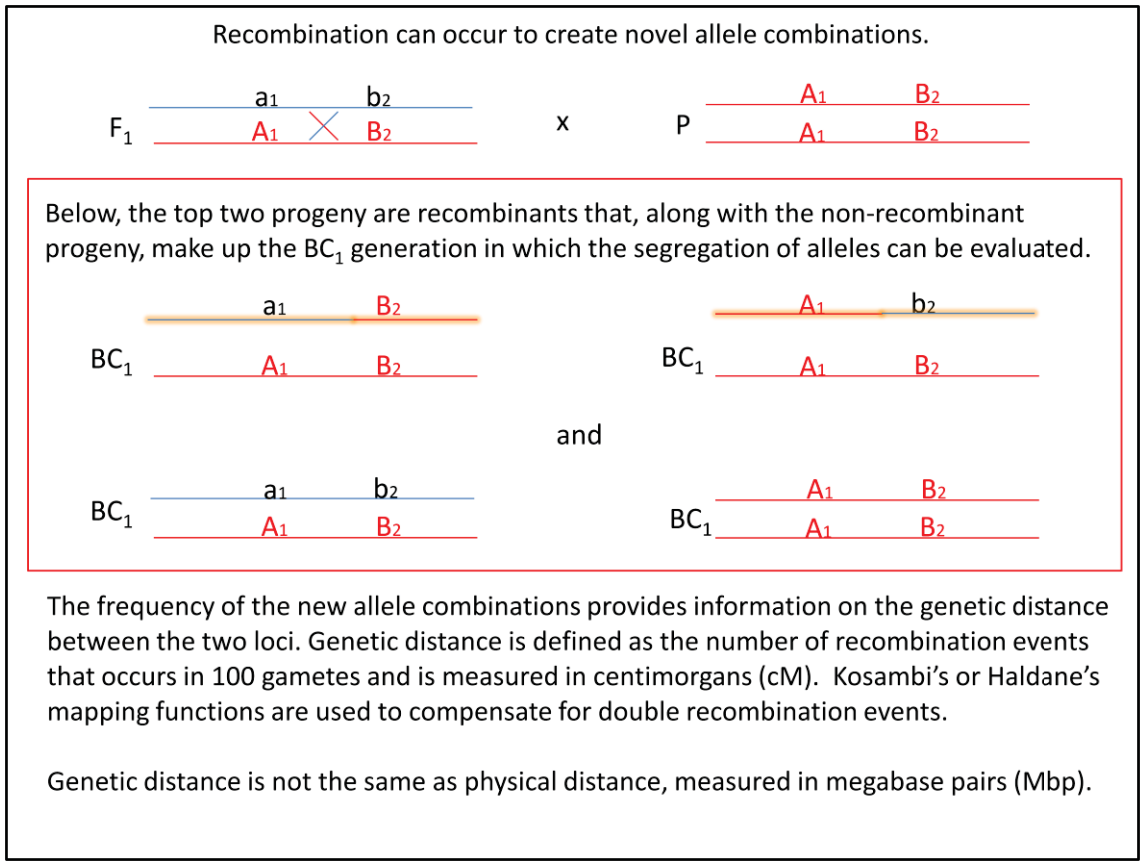


Fig. 1.3: Illustration and explanation of meiotic recombination in a backcross 1 (BC1) progeny. F_1: first filial generation, the result of a cross between two parents; P: parent; A_1, a_1, B_2, b_2: loci; line color indicates parental descent; shading represents recombinant genotypes.

Three main steps are needed to create a linkage map. First, a mapping population of appropriate size must be established to ensure a sufficient number of recombination

events occur to calculate genetic distance; second is the identification of polymorphism at the DNA level by screening markers over the population; and third is the linkage analysis of polymorphic markers (Botstein et al. 1980). A presentation of the steps needed to construct genetic linkage maps follows.

1.2.1.1 Parental selection and population size for genetic mapping

Population establishment requires careful selection and crossing of parents; they must be different enough so that polymorphisms can be passed to the progeny, but similar enough for their genetic makeup to encompass the diversity of the wider taxon (Anderson et al. 1993). The biology of the species with regard to reproductive strategy (inbreeding or outcrossing) and length of generation will influence the population type. The breeding strategy will also determine the expected degree of locus heterozygosity in the population (Young 2001). Table 1-1 summarizes different types of populations commonly used for genetic mapping in plants (Young 2001; Tanksley 1993; Grattapaglia and Sederoff 1994).

Table 1-1: Population types commonly used for genetic mapping in plants.

Genetic Population	Derived from	Advantages	Disadvantages	Selected examples
F_1	Hybrid between two parents	Useful for long-lived species Best for highly-heterozygous species Can use existing populations Knowledge of parental genotypes not required	Two separate maps are constructed, one for each parent Large population size may be needed to integrate parental maps	Strawberry (Sargent et al. 2009a); Pear (Yamamoto et al. 2007; Pierantoni et al. 2007); Almond (Wu et al. 2009c; Viruel et al. 1995); Rose (Crespel et al. 2002); Sweet cherry (Olmstead et al. 2008); Raspberry (Graham et al. 2004); Apple (Liebhard et al. 2003a; Maliepaard et al. 1998) Pseudo-testcross analytical strategy: *Eucalyptus* (Grattapaglia and Sederoff 1994); Apple (Liebhard et al. 2003a; Maliepaard et al. 1998); Sour cherry (Wang et al. 1998)
F_2	Selfing or crossing of F_1	Easy to construct Can be used in outcrossing species	Not a permanent source of mapping material Seed derived from selfing will not breed true	Eggplant (Wu et al. 2009b); Pepper (Wu et al. 2009a; Livingstone et al. 1999); *Physocarpus* (Sutherland et al. 2008); Strawberry (Sargent et al. 2004); *Medicago* (Kaló et al. 2004); *Capsella* (Acarkan et al. 2000)
Recombinant inbred (RI)	Single seed descent from sibling F_2 plants (lines) for 5 or more generations	Permanent mapping resource-all lines breed true Each line contains a different combination of parental linkage blocks Reproducible	Time-consuming to generate Not feasible for obligate outcrossing species	Rice (Xu et al. 1997), Maize (Pejic et al. 1998; Ahn and Tanksley 1993)
Backcross (BC)	F_1 backcross to parent(s)	Easy to construct Can be used in outcrossing species	F_1 sterility may limit parental backcrosses Seed derived from selfing will not breed true	Peach (Verde et al. 2005); *Brassica* (Lagercrantz 1998)

One of the most important considerations for population development for mapping studies is population size. Many genetic linkage maps have been constructed using populations of 55-200 progeny (Sargent et al. 2009a; Grattapaglia and Sederoff 1994;

Ellwood et al. 2008; Foolad et al. 1995; Gupta et al. 2002a; Crespel et al. 2002; Wu et al. 2009a; Yamamoto et al. 2007). When possible, larger populations (more than 100) are preferred over smaller populations since more progeny equals a greater number of meiotic recombination events and this determines the resolution and accuracy of the linkage map (Paterson 1996 pg. 32). However, large populations are not always possible to attain nor economical to analyze, hence the wide range in population sizes as shown above.

1.2.1.2 Molecular marker development

Construction of genetic linkage maps provides information on allelic combinations and their frequency of being inherited together within a population. This relationship can be determined with molecular markers. The term "molecular marker" refers to a region of polymorphic DNA. Markers that are inherited together are said to be linked. The degree of linkage gives an estimate of the distance between markers and is used to order the markers along a linkage group. Other types of genetic markers are morphological (physical), and biochemical such as protein variants (isozymes); however I will be focusing on molecular markers. Molecular markers represent genetic differences between individuals, the variation of which can be classed as point mutations or replication errors. A discussion follows of genetic polymorphisms and the types of markers commonly used to detect them.

1.2.1.3 Point mutations

A point mutation or single nucleotide polymorphism (SNP) is a difference at a nucleotide position and is typically bi-allelic in nature (Picoult-Newberg et al. 1999). In plant species, SNP frequency varies widely with an estimated 1 in 40 base pairs (bp) in maize simple sequence repeat (SSR)-flanking regions (Edwards and Mogg 2001), to 1 in 3.3 kilobase (kb) in the genomes of *Arabidopsis thaliana* strains Landsberg *erecta* and Columbia (Gupta et al. 2001). Following is a discussion of some of the different methods used to detect these variants.

1.2.1.3.1 Restriction fragment length polymorphism

One of the first DNA-based marker systems used for genetic mapping, the RFLP technique uses restriction enzymes to digest genomic DNA into fragments based on sequence recognition by the enzymes. Typically two different enzymes, a frequent cutter (recognizing a sequence of four nucleotides) and a rare cutter (recognizing a

sequence of six nucleotides), are used to digest the DNA (Fig. 1.4). The random nature of the restriction cut sites of the fragments provide variation that can be used to develop markers without prior knowledge of DNA sequence (Botstein et al. 1980). Sources of fragment variation that can be detected using agarose gel electrophoresis are SNP at the homologous enzyme cut sites, resulting in the loss of the cleavage site or the formation of new cleavage sites, and insertion/deletions (indels) within the fragment (Botstein et al. 1980). Additionally, RFLP fragments can be developed into probes that can detect regions of homology in other taxa. While the reproducibility of this "first generation" marker type is high, inheritance is co-dominant, and 1-5 loci can be detected per enzyme pair, the relatively high cost of marker development, probing with radioisotopes, and low-throughput for genotyping (Gupta et al. 2002b) has led to the development of other methods for detecting polymorphism.

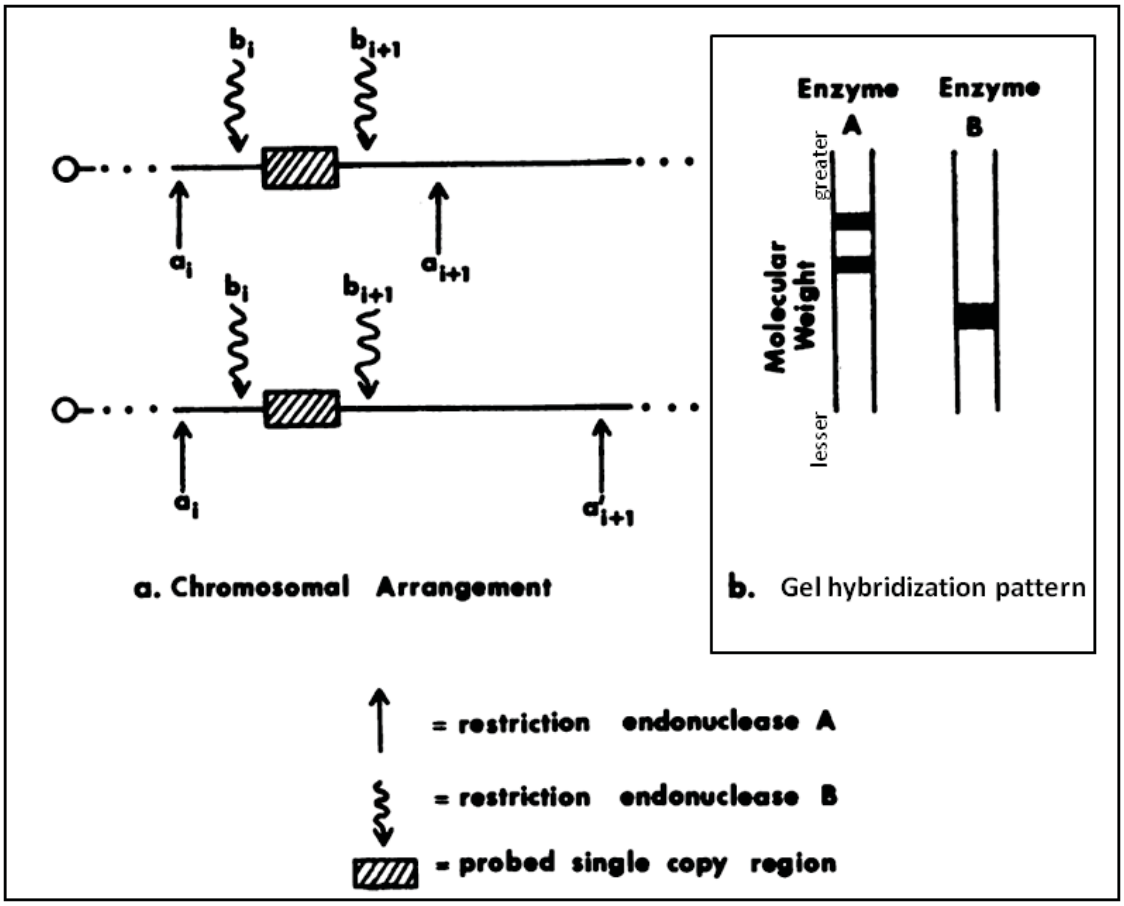


Fig. 1.4: Illustration of restriction fragment length polymorphism (RFLP) technique. (a) Two restriction endonucleases are used to digest DNA into fragments of various lengths. Horizontal lines represent homologous chromosomes. The lower chromosome has a difference in nucleotide sequence relative to the upper chromosome at a_{1+1}, thereby creating a fragment of a different length. (b) The hybridization pattern of the fragments is visualized on the gel. The two different sized fragments generated by enzyme A are shown as two bands of different molecular weight (the band of greater molecular weight is higher on the gel). This indicates heterozygosity at this locus. Enzyme B fragments are the same size and are visualized as a single band. This indicates homozygosity at this locus. The probed single copy region can also be used to detect homologous regions in other taxa. Figure modified from Botstein et al. (1980) Figure 1.

1.2.1.3.2 Amplified fragment length polymorphism

Like RFLP, AFLP begins with the enzymatic digestion of DNA into fragments. Following digestion, ligation of specific adapters with three selector nucleotides determines a subset of fragments that can be amplified by PCR (Fig. 1.5). SNP at the cut sites create variation in fragment length that is detected using either capillary or gel electrophoresis (Vos et al. 1995). Furthermore, AFLP bands can be developed into

probes and used to detect homologous regions in other taxa. The change to PCR-based amplification method improved the efficiency and sensitivity of detection over RFLP (Lin et al. 1996). The dominant inheritance, high reproducibility and large range of locus detection (from 50-100 per enzyme pair) make this a good method for DNA fingerprinting (Vos et al. 1995). However, the adapters and three selector nucleotides are not likely to amplify a homologous product in other taxa, therefore limiting the direct use of primer pairs for comparative genome mapping.

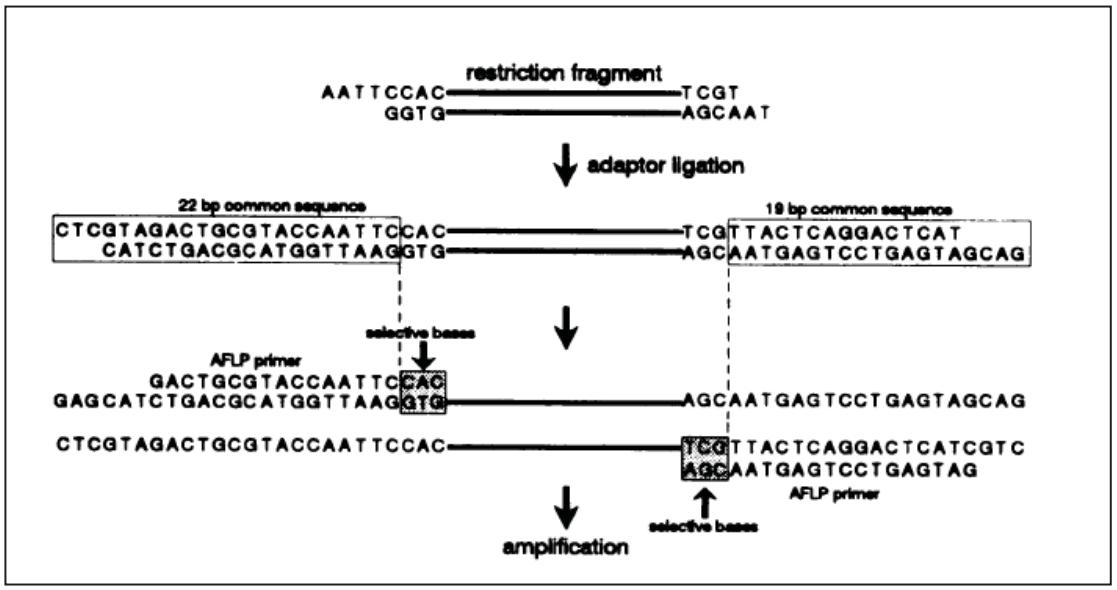


Fig. 1.5: Illustration of amplified fragment length polymorphism (AFLP) technique. Top: restriction fragment formed by the combination of enzymes *Eco*RI-*Mse*I with 5' protruding ends. Center: the same fragment with ligated adapters for the respective enzymes. Bottom: both strands of the fragment with their corresponding AFLP primers. The 3' end of the primers and their recognition sequence in the *Eco*RI-*Mse*I fragment are highlighted. Figure copied from Vos et al. (1995) Figure 1.

1.2.1.3.3 Random amplification of polymorphic DNA

The RAPD technique identifies polymorphisms by the presence or absence of bands amplified using short (usually 10 nucleotide), random sequence primer pairs and visualized by agarose gel electrophoresis (Williams et al. 1990). Sequence variations, such as SNP, in the primer binding sites are detected by the absence of bands or a change in banding pattern (Fig. 1.6). This PCR-based technique provides a rapid, inexpensive method of detecting dominantly inherited, genetic polymorphisms. Although locus detection ranges from 1 to 10, the random nature of the primers limits

reproducibility and therefore transferability between taxa, as amplification of a homologous sequence is unlikely (Gupta et al. 2002b).

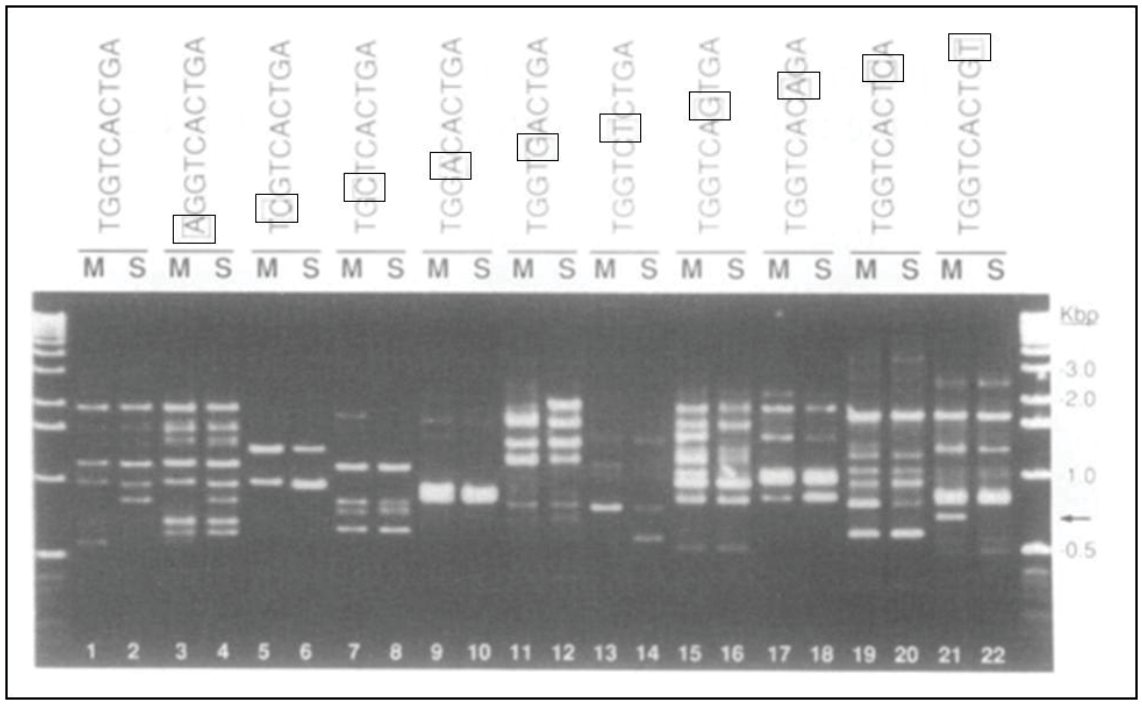


Fig. 1.6: Gel demonstrating random amplification of polymorphic DNA (RAPD) technique. Amplification was performed on soybean *Glycine max* (M) and *G. soja* (S). The original random primer is shown in lanes 1-2. Each subsequent pair of lanes shows how the banding pattern changes with one change of nucleotide in the primer. Nucleotide change is boxed. The arrow on the right points to a polymorphic 0.65 kb band in the final pair of lanes (21-22). This experiment demonstrates that 10 nucleotide-long oligos can be used as PCR primers and detect polymorphism. It also illustrates that nucleotide changes will determine fragment amplification. Figure copied from Williams et al. (1990) Figure 2.

1.2.1.3.4 Sequence characterized amplified region

SCAR markers can be developed from PCR amplification products (amplicons) of other detection methods. The band of interest is extracted from the gel, sequenced, and new primer pairs specific for the fragment of interest are developed (Paran and Michelmore 1993). Sequence variations can be detected by comparing SCAR DNA sequences or as an absence of product during amplicon gel visualization, if SNP in the PCR primer binding sites prohibit primer annealing.

Each of the preceding methods was designed for gel-based analysis methods. The strictly PCR-based, "second generation marker" (Gupta et al. 2002b) methods can be

used to detect SNP if the variation leads to a disruption in primer annealing or changes in banding pattern. All these methods can also be used to detect variation in fragment length that may be caused by an indel. Next I will discuss non-gel based methods of analysis.

1.2.1.4 High-resolution melting

HRM is a fast, PCR-based, high-throughput, closed-tube melting curve analysis that uses a high-affinity double-stranded DNA (dsDNA)-binding fluorescent dye to discriminate between genotypes relative to the presence or absence of a SNP (Wittwer et al. 2003; Reed and Wittwer 2004) or small indel (Montgomery et al. 2007). Marker development usually requires DNA sequence knowledge to design primers that will target a putative SNP and produce an amplicon of 50 - 300 bp. Amplicon size is an important consideration for accurate analysis, as larger fragments hamper the melting process (www.roche-applied-science.com). The small amplicon size also provides a degree of detection accuracy that is not attainable with gel-based electrophoretic analysis methods or denaturing high performance liquid chromatography (dHPLC) (Chou et al. 2005; Montgomery et al. 2007). If an individual is heterozygous at a locus (*e.g.* A/T), during the PCR amplification four different duplexes will be formed: two homoduplexes (A/A and T/T) and two heteroduplexes (both A/T) (Montgomery et al. 2007). Based on the affinity of the nucleotide hydrogen binding, the heteroduplex strands will separate (melt) at a lower temperature than either of the two homoduplexes (Reed and Wittwer 2004). Background fluorescence is subtracted from the analysis and normalized by subtracting the minimum fluorescence from all values, dividing by the resultant maximum fluorescence, and multiplying by 100 (Montgomery et al. 2007). The shape of the melting curve (Seipp et al. 2007), and the temperature at which the PCR product melts (Chou et al. 2005), which is visualized by the rate at which dsDNA-binding dye is released from the melting amplicon, provide information on the genotype of the individuals (Fig. 1.7). Heterozygote scanning is improved by correcting for minor temperature differences. This is accomplished by overlaying the melting curves at an interval where all curves are similar (Montgomery et al. 2007). A randomly selected standard curve is selected and the average temperature difference of each curve is subtracted from the standard thereby equalizing the average temperature of all melting curves (Montgomery et al. 2007). The heteroduplex individuals are identified by the change in melting temperature compared with that of the homoduplex, although the

specific SNP is not identified (Montgomery et al. 2007; Vandersteen et al. 2007). However, sequencing of the PCR product could lead to the discovery of a novel SNP (Picoult-Newberg et al. 1999). Discrimination between the two homoduplexes may or may not be evident, since the dye has high affinity for both genotypes. Advantages of the HRM technique are results that are highly reproducible; the position of the SNP need not be fixed (Reed and Wittwer 2004; Picoult-Newberg et al. 1999); and the melting process is non-destructive, so that the PCR product can be sequenced (Montgomery et al. 2007).

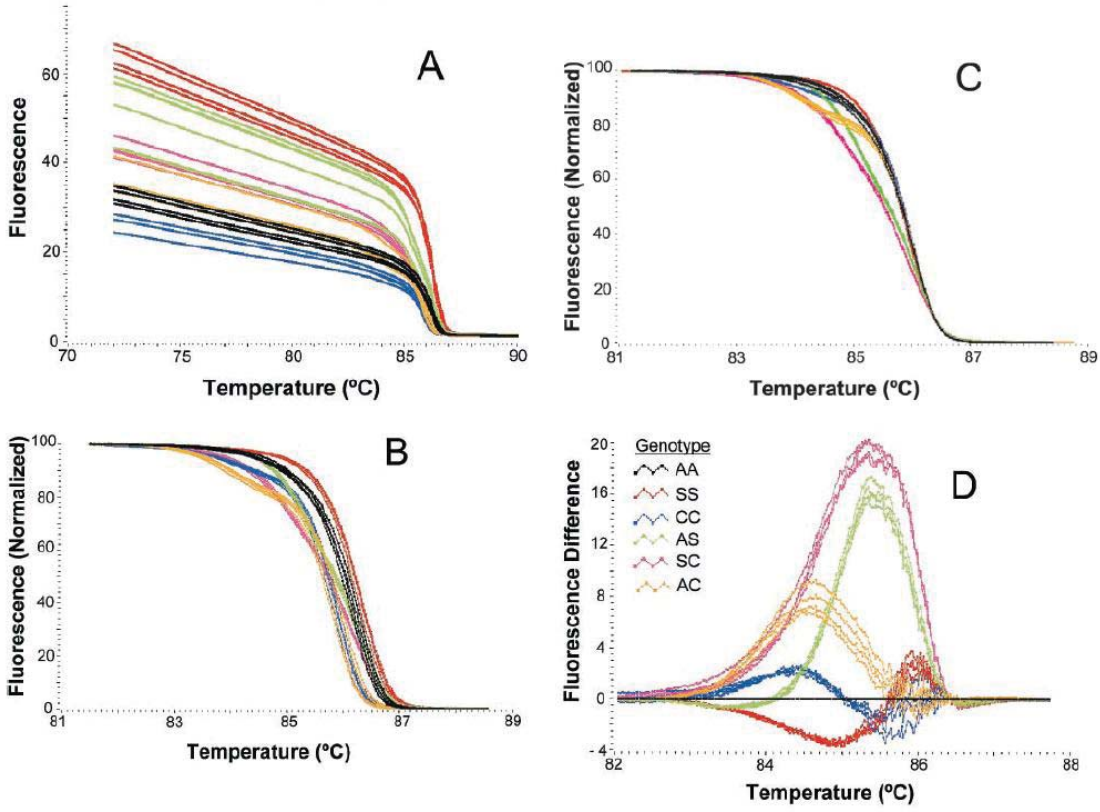


Fig. 1.7: Illustration of high-resolution melting (HRM) technique. An example of genotyping two adjacent single nucleotide polymorphisms (SNP) within a 110 base pair (bp) fragment. (A) Original HRM curves showing two individuals each (replicated) for six genotypes (each color). (B) The same curves as shown in (A), after normalization. (C) Normalized curves after temperature shifting of each curve to overlay them on one of the AA (wildtype) curves between 5% and 10% fluorescence. (D) Fluorescence difference curves obtained by subtracting each curve in (C) from one wildtype curve. The temperature axis in (C) & (D) now reflects the difference in temperature relative to the wildtype, rather than an absolute temperature. The curves shown in (D) demonstrate the difference in shapes seen between homoduplexes (single, high peak shown in green and pink) and heteroduplexes (various shapes shown in yellow, blue and red). The black horizontal line is the wildtype. Genotype abbreviations (panel D): A: adenine; C:

cytosine; S: either guanine or cytosine. Figure copied from Wittwer et al. (2003) Figure 3.

1.2.1.5 SNP chips and arrays

Genome-wide SNP discovery is becoming more common as more whole genomes are sequenced and expressed sequence tag (EST) libraries are developed. Various medium- and high-throughput methods have been developed by commercial interests. One protocol is to amplify the DNA sample to create amplicons. The amplicons are digested to create fragments of appropriate size; fragments are hybridized to DNA that is covalently bonded to beads attached to the chip surface; allelic specificity is determined by enzymatic base extension and then the products are fluorescently stained; the fluorescence intensity is measured and analyzed by computer software. SNP chips have the potential to identify tens of thousands to hundreds of thousands SNP loci per genome (Illumina 2006) and the technology is being applied to such crops as apple (Chagné et al. in press), cacao (Livingstone III et al. 2012), maize (Lu et al. 2011), and rice (Thomson et al. 2010).

These examples of "third generation" (Gupta et al. 2002b), non-gel-based techniques, use fluorescence intensity to differentiate individual genotypes by the presence or absence of SNP.

1.2.1.6 Genotyping-by-sequencing

The emergence of next-generation sequencing (NGS) technologies has rapidly increased the amount of sequence data available for polymorphism detection. One method for high-throughput SNP detection is genotype-by-sequencing (GBS) (Elshire et al. 2011). This method, like the restriction-site associated DNA (RAD) method developed by Baird et al. (2008) and the multiplex shotgun genotyping (MSG) developed by Andolfatto et al. (2011), combines NGS with the judicious use of restriction enzymes to fragment genomes. Genome fragmentation is followed by the ligation of adapters individualized by barcode nucleotide sequences. While the three methods differ in complexity, the outcome of the relatively low-cost generation of large numbers of SNP loci and the ability to map these polymorphisms is similar.

1.2.1.7 Replication errors: simple sequence repeat

Simple sequence repeat analysis is an example of a technique used to detect polymorphisms due to replication errors. SSR, or microsatellites, are stretches of mono-, di-, tri-, tetra-, or penta-nucleotide motifs repeated in tandem throughout prokaryotic and eukaryotic genomes (Powell et al. 1996). Motif repetitiveness promotes polymerase slippage during DNA replication, resulting in variation in repeat number, while conservation in the flanking regions tends to be maintained (Schlötterer and Tautz 1992). Microsatellite regions were first developed by the digestion of gDNA with restriction enzymes and the ligation of the fragments into a vector, usually a bacterium, to create a library (Condit and Hubbell 1991). The library was then probed with oligonucleotides of the repeat of interest ($[CA]_n$ or $[AG]_n$, for example) to identify the complementary regions of the digested genome (Condit and Hubbell 1991). Markers developed from microsatellite sequences have been extensively used for diversity analyses and fingerprinting in many crops: the Solanaceae (nightshade family) (McGregor et al. 2000; Milbourne et al. 1997; Tam et al. 2005), the Poaceae (grass family) (Warburton et al. 2002; Eujayl et al. 2002; Pejic et al. 1998), the Fabaceae (bean family) (Fig. 1.8) (Cregan et al. 1994), the genus *Vitus* (grape) (Lin and Walker 1998), and the Rosaceae (Gasic et al. 2009b; Gianfranceschi et al. 1998; Hagen et al. 2004; Hokanson et al. 1998; Igarashi et al. 2008; Lewers et al. 2008; Aranzana et al. 2003a; Aranzana et al. 2003b; Dirlewanger et al. 2006). SSR markers have also been used to saturate existing maps composed of other marker types (Krutovsky et al. 2004; Hackauf et al. 2009). Once developed, their multi-allelic nature, co-dominant inheritance, and PCR-based amplification of these "second generation" markers make them versatile and easy to use for linkage mapping (Gupta et al. 2002b), and with the availability of EST and genomic databases, their development has become simplified. However, primer pairs designed to amplify SSR markers can be limited in their transferability between genera (Lewers et al. 2005; Decroocq et al. 2003).

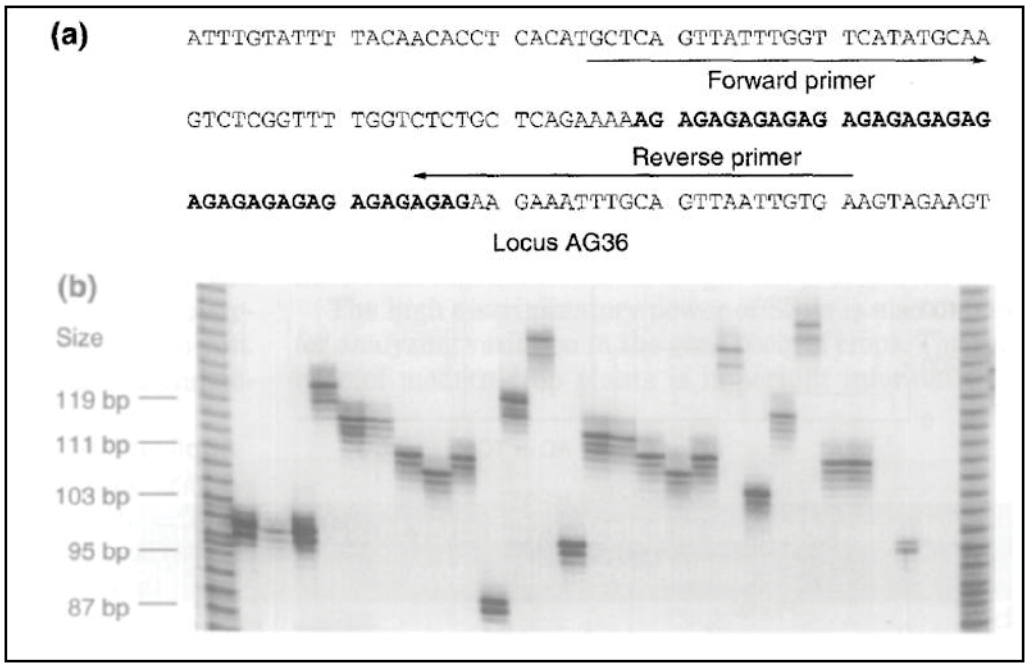


Fig. 1.8: Illustration of simple sequence repeat (SSR) technique. (a) DNA sequence from soybean with forward and reverse PCR primers shown with arrows; repeat motif is in bold. (b) Autoradiogram of a polyacrylamide gel separation of ^{33}P-labelled PCR products for a range of soybean DNA templates (one in each lane); outer-most lanes are size standards. The banding patterns shown in (b) illustrate the number of alleles per individual (bands per lane) and the number of alleles throughout the genus (variation across the gel). Figure after Powell et al. (1996) Figure 1.

The preceding marker types can be developed from random or non-coding DNA sequences and these are not necessarily derived from a common ancestor. The following section discusses the development of molecular markers from orthologous sequence sources.

1.2.1.8 Orthologous marker development

Orthology indicates shared ancestry. Some orthologous regions, such as key regulatory elements and core components of the translation and transcription machineries, are conserved through long evolutionary time scales (Koonin et al. 2000; Riechmann et al. 2000), while other orthologous regions are derived from more recent ancestry. Orthologous sequences, including EST, conserved orthologous set (COS), and gene sequences, can be used to design PCR primers to identify and amplify polymorphisms (usually in the form of a SNP, SSR, or indel) that can be used to distinguish individuals and provide anchor loci between maps of different taxa. However, even if the target

fragment is highly similar, the primer pair that flanks the target may not be, and new primers specific for the species of interest may need to be designed.

1.2.1.8.1 Expressed sequence tags

EST are short segments of protein-coding DNA derived from sequencing either one or both ends of an expressed gene, and the gene function may be unknown. Development of EST requires the extraction of messenger RNA (mRNA), and the subsequent enzymatic conversion by reverse transcriptase into complementary DNA (cDNA) (www.ncbi.nlm.nih.gov/About/primer/est.html). EST information is collected into searchable databases, such as NCBI's dbEST (http://www.ncbi.nlm.nih.gov/sites/entrez), which can be used to identify sequence variation, such as SNP (Picoult-Newberg et al. 1999) and SSR (Gupta and Rustgi 2004), for PCR-based molecular marker development. The sequenced cDNA fragments can also be labeled with a fluorescent or radioactive tag and used as RFLP probes of genomic DNA (gDNA) or cDNA to identify complementary regions (Bonierbale et al. 1988). Overlapping EST fragments can be assembled into longer contiguous sequences (contigs) and used for *in silico* gene identification and detection of variation in the form of SNP or SSR (Gupta and Rustgi 2004; Gupta et al. 2002b). The association of EST fragments with gene expression makes them more likely to be conserved among species (Cato et al. 2001; Van Deynze et al. 1995) and PCR primers designed from these fragments have the potential to amplify homologous products across species (Cato et al. 2001; Gupta and Rustgi 2004).

1.2.1.8.2 Conserved orthologous set

COS sequences are derived from regions of single-copy DNA, including genes. The emphasis on single-copy reduces the difficulties of distinguishing orthology from paralogy as found in multi-gene families (Fulton et al. 2002; Devos and Gale 2000). Variation within a COS sequence, such as the presence of a SNP, SSR, or indel, can be used as the basis for designing PCR-based markers. The orthologous nature of the markers increases the potential for them to be useful for comparative genome mapping. Some of the first COS were single-copy cDNA fragments developed in rice that were used to probe the genomes of other grass genera, and to build a grass consensus map with wheat as the reference (Van Deynze et al. 1995). This consensus map

demonstrated high degrees of colinearity in gene order and chromosome position among the analyzed genera.

The utility of COS markers in grasses led Fulton et al. (2002) to use bioinformatic tools to try to identify a set of markers that would be conserved in both sequence and copy number in all higher plants. They started by comparing the *Arabidopsis thaliana* genomic sequence with the tomato EST database and searching for single-copy tomato genes that aligned to a single region in *A. thaliana*. Tomato and *A. thaliana* belong to different families (the Solanaceae and Brassicaceae, respectively) that diverged approximately 125 million years (my) ago (Fulton et al. 2002). The single-copy regions identified by comparing the two genomes suggest that the COS selected are orthologous rather than paralogous. The stringent selection criteria established by the researchers resulted in the production of a set of highly conserved markers for use in comparative mapping across divergent plant species. Their work has led to the development of additional COS in Solanaceae (Wu et al. 2009b; Van Deynze et al. 2007; Wu et al. 2009a), Pinaceae (Liewlaksaneeyanawin et al. 2009), Poaceae (Hackauf et al. 2009; Bolot et al. 2009), and Rosaceae (Cabrera et al. 2009).

1.2.1.8.3 Gene-based markers

Gene-based markers may be developed from a number of sources including EST, COS, transcriptomes, or gDNA and are derived from genes of known function. Variation can often be found in the introns and may take the form of a SNP, SSR, or indel. If the polymorphism influences the phenotype, then the term "functional marker" is applied (Andersen and Lübberstedt 2003). Primer pairs are often designed to anneal to the exons flanking an intron, as introns are less constrained by selection and are more likely to accumulate differences (Castle 2011; Wolfe et al. 1989). SNP markers associated with gene function can be identified using EST and genomic databases (Picoult-Newberg et al. 1999). PCR primer pairs designed to anneal to exons and amplify the intervening intron are predicted to have a high degree of between-genera transportability (Gupta et al. 2002b) because of the sequence stability of the exons (Picoult-Newberg et al. 1999). While the specific SNP is usually not transferable between species, the amplified intron region may have accumulated other novel SNP (Picoult-Newberg et al. 1999).

Each of the marker types discussed above have their own set of strengths and weaknesses. Newly-constructed genetic linkage maps are often first constructed using "first generation" markers because no prior knowledge of the genome is required for their application. As more is learned about the taxon's genome, other markers types are employed to improve the linkage map and make the map more useful for comparative mapping studies.

1.2.1.9 Linkage analysis of polymorphic markers

Once the segregating population has been established, gDNA from all individuals has been extracted, and molecular marker type(s) selected, screening of the primer pairs on the individuals can begin. As data on polymorphic marker genotypes of the individuals in the mapping population are accumulated, linkage can be calculated using statistical software such as JoinMap® v3.0 (van Ooijen and Voorrips 2001). Programs such as these calculate the recombination frequencies (Rf) between the alleles of polymorphic markers in a pairwise fashion, perform a numerical search for the best fitting linear marker arrangement, and order the markers into LG (Stam 1993). The Rf is used to calculate genetic distance which is defined as the number of recombination events that occur in 100 gametes and is measured in centimorgans (cM) (Stam 1993). The logarithm of odds (LOD) score, a measurement of likelihood of linkage and calculated as the probability that two loci are linked over the probability that they are not linked, is used as a threshold of linkage significance (Stam 1993). The LOD and Rf are inversely related; therefore, a high LOD and a low Rf value indicate tight linkage of markers. The LOD threshold is one of the criteria used to group markers, with higher LOD establishing a more stringent threshold and generally a greater number of shorter LG. Map distances for a given order of markers are calculated using a least squares procedure (Stam 1993).

1.2.1.9.1 Selective mapping for linkage analysis

The construction of a linkage map depends on identifying recombination events within a population. One method to reduce the number of individuals necessary for genotyping to place new markers on a genetic map is called selective (or bin) mapping. Two processes are needed to implement selective mapping: 1) construction of a high-quality framework map, and 2) addition of more markers to the framework (Vision et al. 2000). Genotyping the entire population allows the identification of specific recombination

events (breakpoints) within each individual in the mapping population. Further genotyping is then conducted using a subset of individual progeny that among them have breakpoints that are evenly distributed along the genome (Vision et al. 2000). Additional markers added to the framework map are placed in "bins" that are delimited by framework markers and within which no recombination occurs, hence the phrase "bin mapping" (Vision et al. 2000; Howad et al. 2005). The order of the framework markers is fixed, however the order of additional markers placed in a bin is not fixed and the order can only be determined by mapping on the full population. The advantage of selective or bin mapping is the cost-effective and rapid placement of additional genetic markers, once the initial genetic map has been constructed.

1.2.2 Summary

In this section I have presented the steps necessary to begin a comparative genome mapping study. First is the construction of genetic linkage maps for the taxa to be compared, which begins with the selection and crossing of parents and establishment of a population of appropriate size to ensure sufficient meiotic recombination events. Next, DNA-level polymorphism is detected in the progeny using molecular markers, and the results analyzed to determine marker placement and order on the representative linkage groups. The choice of marker type employed often is based on the prior knowledge of the genome and available analysis tools. I have presented information on different types of polymorphisms exploited for detection with molecular markers, and the development of orthologous markers. Finally, I have summarized the linkage analysis for map construction using analytical software. Next, I will present an application for genetic linkage maps, the mapping of quantitative trait loci.

1.2.3 Quantitative trait locus mapping in plants

Comparative genome mapping studies can be conducted using aligned genetic maps on which QTL have been located. QTL identify the regions of the genome involved in the expression of traits with continuous variation, as these traits are controlled by the expression of more than one gene. Along with the genetic component, traits with continuous variation can also be influenced by non-genetic factors, as well as the interactions between them (Thoday 1961; Sax 1923). Most economically important traits in crop genera, such as fruit quality (Kouassi et al. 2008; Torre and Barritt 1977; Etienne et al. 2002; Lerceteau-Kohler et al. 2004), flowering time (Crespel et al. 2002;

Hibrand-Saint Oyant et al. 2008; Weebadde et al. 2008), agronomic characteristics (Blanc et al. 2006; McCouch and Doerge 1995; Chagné et al. 2003; Crepieux et al. 2005), and disease resistance (Liebhard et al. 2003b; Pierantoni et al. 2007) are quantitative; therefore an understanding of the underlying genetic control is important in a number of biological contexts, including crop improvement (Mackay 2001) (Fig. 1.9). The genetic model used to explain the genetic architecture of a complex trait assumes a two-allele system segregating at each QTL and affecting variation in trait expression (Mackay 2001). The allelic state of the locus (homozygous v. heterozygous), the mode of inheritance (additive or dominant), and how the loci interact (epistasis) means that the relationship between genotype and phenotype is not proportional, and thus the QTL genotype cannot be determined simply based on the segregation of phenotypes of controlled crosses or pedigrees (Thoday 1961). Since multiple genes are responsible for the expression of a complex trait, the different QTL will have various levels of effect. Four major steps are needed to conduct a QTL analysis: 1) establishment of a segregating population; 2) measurement of variation in the trait(s) of interest for each individual in the population; 3) construction of a linkage map; and 4) identification of QTL. QTL maps can also be used for comparative analyses to predict regions of similar function in other taxa.

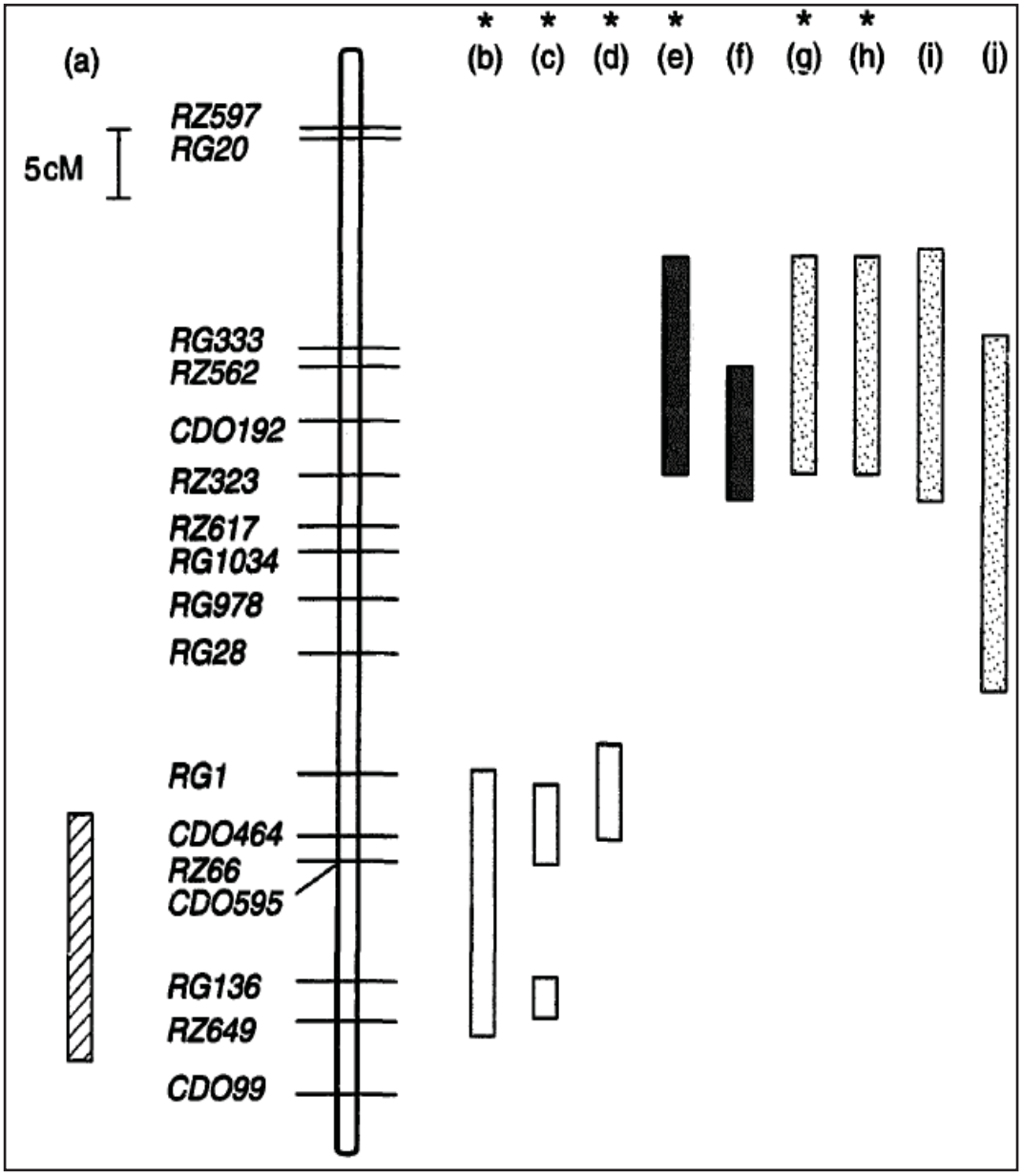


Fig. 1.9: An example of quantitative trait locus (QTL) mapping in rice. Chromosome 8 of rice is represented as a vertical open bar with markers (names on left) distributed along its length. (a) The size and location of a homeologous segment of the wheat chromosome 7A based on markers in common; (b-j) various QTL for agronomic, disease resistance, flowering, and grain ripening characteristics. Asterisk (b-e, g, h) indicates traits mapped in a single population. No asterisk (f, i, j) indicates that the QTL were mapped in multiple populations using a common set of RFLP markers. This figure is an example of using comparative genome mapping to identify regions of similar function in two species (rice and wheat), and locates areas of interest for focused fine-scale mapping to detect underlying genes controlling traits of interest. Figure copied from McCouch & Doerge (1995) Figure 2.

1.2.3.1 Segregating populations

Establishing segregating populations for QTL analysis follows the same procedure as that for genetic linkage as presented in section "1.2.1 *Genetic linkage map construction*", except that for fine-scale or QTL mapping, populations of 200-500 are often necessary to achieve unambiguous resolution of less than 1 cM (Young 2001; Paterson et al. 2000).

1.2.3.2 Phenotyping trait variation

As QTL analysis links a measurable trait with a genetic component, accurate phenotyping is the key to QTL mapping. Replications of measurement, when possible, improve QTL detection as this allows the separation of genetic from environmental effects on the trait. Replicates may be year-to-year, under different growing conditions, or in different locations, and will provide information on the influence of the non-genetic component of the trait, as well as the stability of the QTL in that specific population with respect to the variations detected in the replicates (Lander and Kruglyak 1995).

1.2.3.3 Linkage map construction

Linkage map construction for QTL analysis is the same as presented in section "1.2.1.9 *Linkage analysis of polymorphic markers*", with the added feature of markers spaced evenly 15-20 cM apart (Tanksley 1993). In practice, markers may not be evenly spaced, as the distance between markers reflects the amount of recombination within a linkage group (Hartl and Jones 2009 pg. 169). A linkage map for QTL analysis may be derived by removing markers from a more densely populated genetic linkage map to arrive at the desired marker spacing, although marker removal is not required.

1.2.3.4 QTL identification

QTL are identified by linkage disequilibrium between the phenotypic trait and molecular markers that segregate in a Mendelian fashion. The presence of a QTL is indicated if there is a difference in the mean values of the quantitative trait between individuals with one genotype relative to the alternative genotype (Lander and Botstein 1989) (Fig. 1.10). Strategies have been developed to detect this difference, including single-marker or point analysis (Soller et al. 1976; Tanksley et al. 1982) and simple interval mapping (Lander and Botstein 1989; Jansen and Stam 1994). Single-marker analysis uses statistical methods such as *t*-tests, analysis of variance, and linear

regression (R^2) to detect associations between traits and single markers. Advantages of this method are that a saturated genetic map is not required and analysis can be performed with standard statistical software; however, the further the QTL is from the marker, the less likely it will be detected, as recombination during meiosis can separate the marker from the trait (Lander and Botstein 1989). Simple interval mapping overcomes the disadvantage of distance found in single-marker analysis by analyzing two adjacent markers simultaneously. This compensates for recombination, as the linkage between two markers inherited together will be tighter than if both are considered independently (Jansen and Stam 1994). Simple interval mapping is often followed by multiple-QTL model (MQM) mapping, during which the marker(s) with the greatest QTL effect (cofactor) is/are removed from the calculation to allow the detection of other QTL with lesser effect (Jansen and Stam 1994).

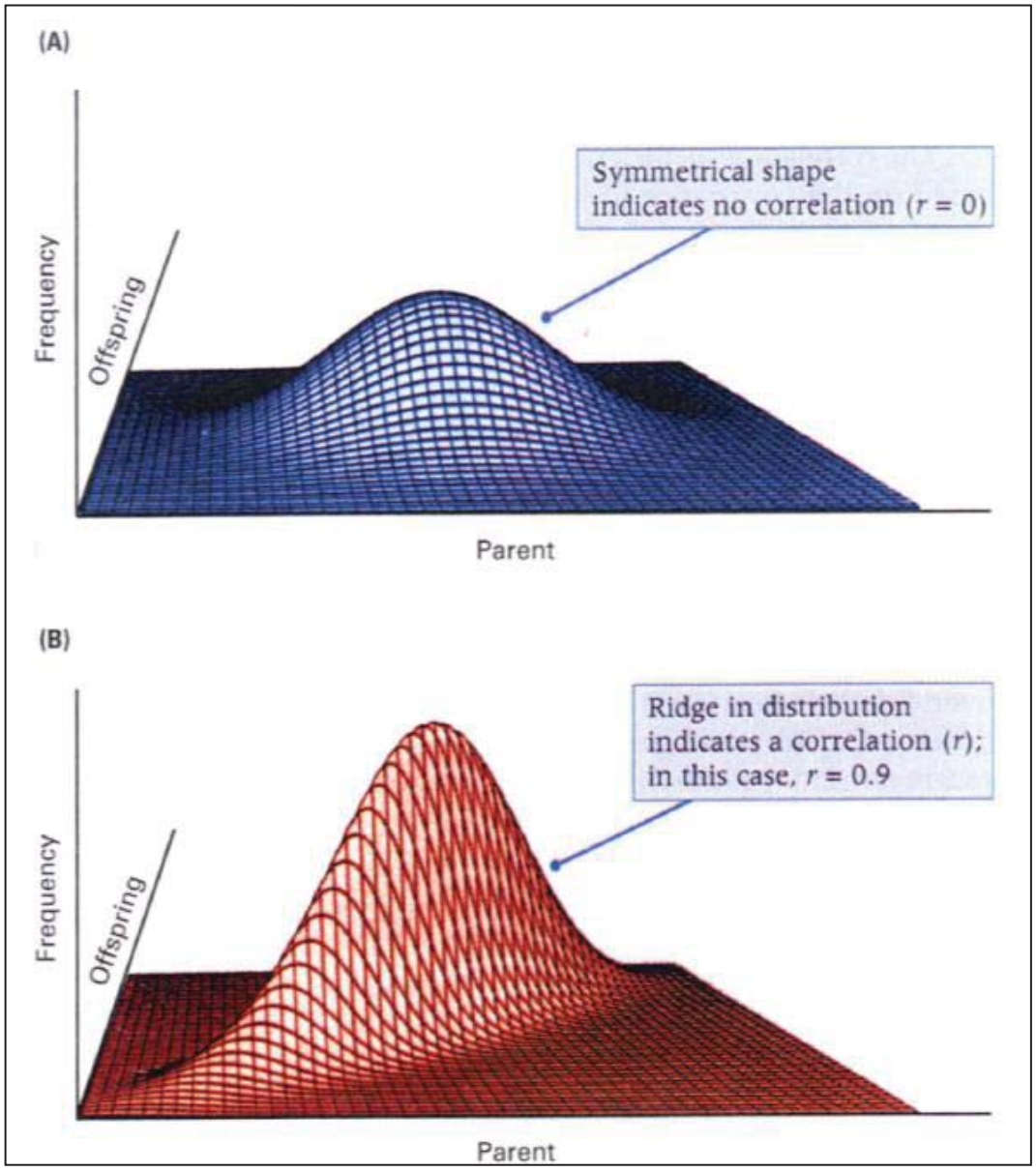


Fig. 1.10: Distribution of a quantitative trait in parents and offspring. (A) The symmetrical shape of the data when the distribution of the trait is independent of the marker. (B) The asymmetrical shape of the data when the distribution of the trait varies according to a particular marker. The correlation coefficient (r) represents the tendency for pairs of numbers (in this case the trait measurement and the marker position) to vary together. Figure copied from Hartl & Jones (2009) Figure 18.13.

1.2.3.5 QTL significance

As with linkage map construction, the LOD score is used as a threshold of significance when a QTL is detected using simple interval and MQM mapping. In general, QTL intervals are delimited by the LOD±2 method based on an increase and decrease of 1 LOD unit representing a 95% confidence interval, and an increase and decrease of 2

LOD units representing a 99% confidence interval (van Ooijen 1992) (Fig. 1.11). Permutation is used to determine the significance of a LOD score and is calculated by sampling the marker data without replacement, to determine the frequency distribution of the maximum LOD (Churchill and Doerge 1994). Typically 1000 permutations are performed to determine the significance at 95% and 99% over the entire genome.

It is important to note that the region of the genome covered by the QTL may contain many genes, some involved in the expression of the trait and some not. The QTL provides a starting point for fine-scale mapping to dissect the genes underlying the locus (Mauricio 2001).

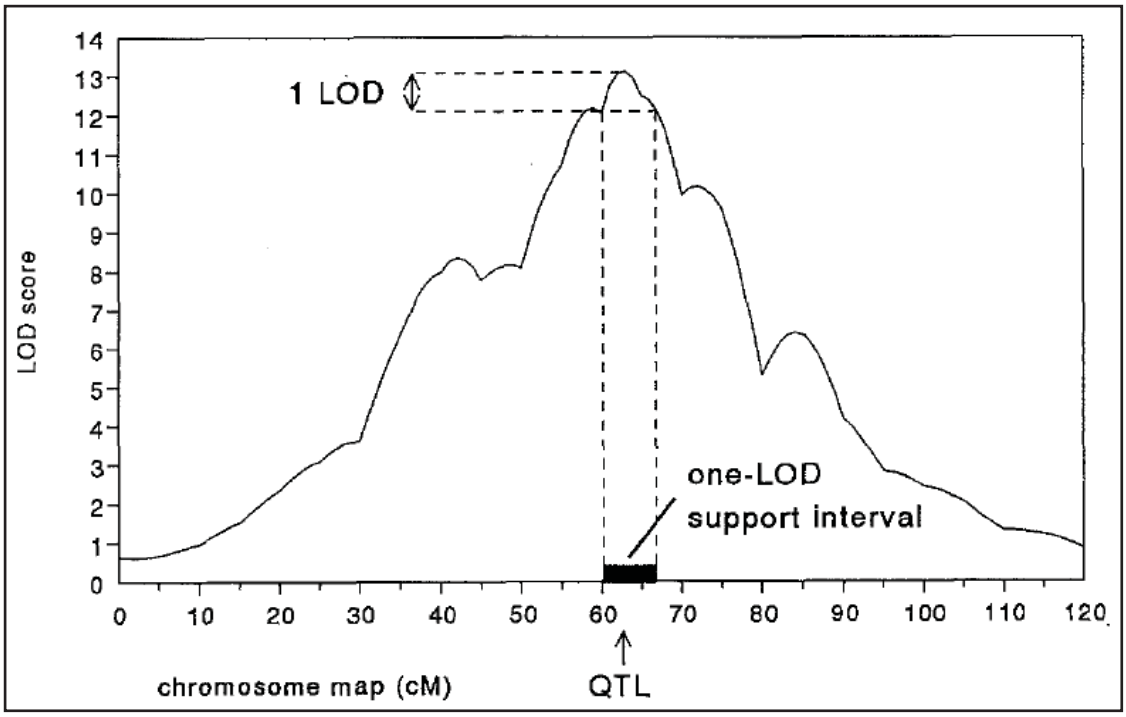


Fig. 1.11: Example of a QTL likelihood map and the construction of a one-LOD support interval. The peak of the QTL is located at 62.5 centimorgans (cM) with a confidence interval that spans from approximately 60 to 66 cM (dark bar and dashed lines), which equals the drop of 1 LOD. Figure copied from van Ooijen (1992) Figure 1.

1.2.4 Association mapping in plants

Another method used to analyze the linkage between traits and genome regions is called association or linkage disequilibrium mapping (reviewed by Oraguzie et al. 2007). In contrast to QTL mapping, which typically uses closely related individuals, usually from

a bi-parental cross, to detect linkage between a trait of interest and a region of the genome, association mapping makes use of groups of individuals that are from a wide range of genetic backgrounds to take advantage of associations detected among dissimilar gene pools (Chakraborty and Weiss 1988). One advantage of this method is that the number of recombination events that have taken place since the populations diverged from a common ancestor has reduced the linkage disequilibrium between the causative variant and the segregating marker and therefore the genetic interval encompassing the trait of interest is narrowed relative to that detected by QTL mapping (Oraguzie et al. 2007). However, genetic map marker density must be much higher for association mapping, relative to map density for QTL mapping (Oraguzie et al. 2007).

1.2.5 Summary

In this section I have introduced the necessary steps for conducting a QTL analysis. Two additional requirements for QTL mapping, beyond those needed for genetic linkage mapping, are accurate phenotyping of the trait of interest and establishment of a large population (where possible). Accuracy of the phenotypic measurements for each individual in the population is crucial for accurate placement of the locus, and larger populations improve the map resolution. QTL analysis in biparental populations can be accomplished using existing genetic linkage maps from which markers have been removed to create LG with markers spaced 15-20 cM apart. This spacing is useful for simple interval mapping to detect linkage between markers and quantitative traits. In the next section I will provide examples of comparative genome mapping in plants.

1.2.6 Comparative genome mapping in plants

Comparative genome mapping requires the construction of genetic linkage maps populated with orthologous markers, the order of which can be used to estimate genome similarity, detect lineage-specific divergence, and determine phylogenetic relationships (Hardison 2003).

1.2.6.1 Previous comparative genome mapping studies in plants

Some of the first genome-wide comparative studies were conducted in the Poaceae. The colinearity of the genomes of rice (subfamily Ehrhartoideae) and maize (subfamily Panicoideae) were compared using RFLP to reveal that colinearity in 62% and 70% of the maize and rice maps, respectively, was conserved (Ahn and Tanksley 1993) (Fig. 1.12). In this example, the organisms of interest have different complements of

chromosomes (rice 2*n* = 24; maize 2*n* = 20), different DNA content (maize has about six times the amount of DNA as rice), and different ploidy origins. Polyploidy is defined as "having more than two complete sets of chromosomes" (Comai 2005). Rice is a diploid, possessing one set each of maternal and paternal chromosomes, and maize is presumed to be an ancient polyploid (Ahn and Tanksley 1993; Bolot et al. 2009) that has undergone "rediploidization" (Paterson et al. 2004). The majority (72%) of the RFLP probes developed and selected to be single copy in rice were located at two loci in maize, supporting the maize polyploid hypothesis. The remaining 28% of loci were single copy, which could indicate a loss of a copy in the maize genome, either through genome rearrangement or high mutation accumulation (Ahn and Tanksley 1993). This scenario was illustrated in Fig. 1.2, as the ancestry of the retained gene is uncertain, and therefore orthology cannot be confirmed.

More extensive comparisons of genomes of Poaceae species have been conducted (Sorrells et al. 2003; Van Deynze et al. 1995). Initially, RFLP probes from barley, oat and rice were used to assess genome colinearity in seven different grass genera from three different subfamilies (Table 1-2) (Van Deynze et al. 1995). The results of this study indicate the conservation of specific chromosomal segments among all tested species and a maintenance of chromosome structure regardless of chromosome number (Van Deynze et al. 1995) (Fig. 1.13). When a combination of bioinformatic alignment and deletion mapping were used to compare the whole genomes of wheat and rice (Sorrells et al. 2003) the findings of previous studies relative to gene order and content were confirmed. However, the higher resolution afforded by an increased number of markers and use of the genome sequences uncovered numerous discontinuities in gene order between the species (Sorrells et al. 2003). This work in Poaceae demonstrates how improvements in technology can be used to validate previous findings, as well as discover new areas for further exploration.

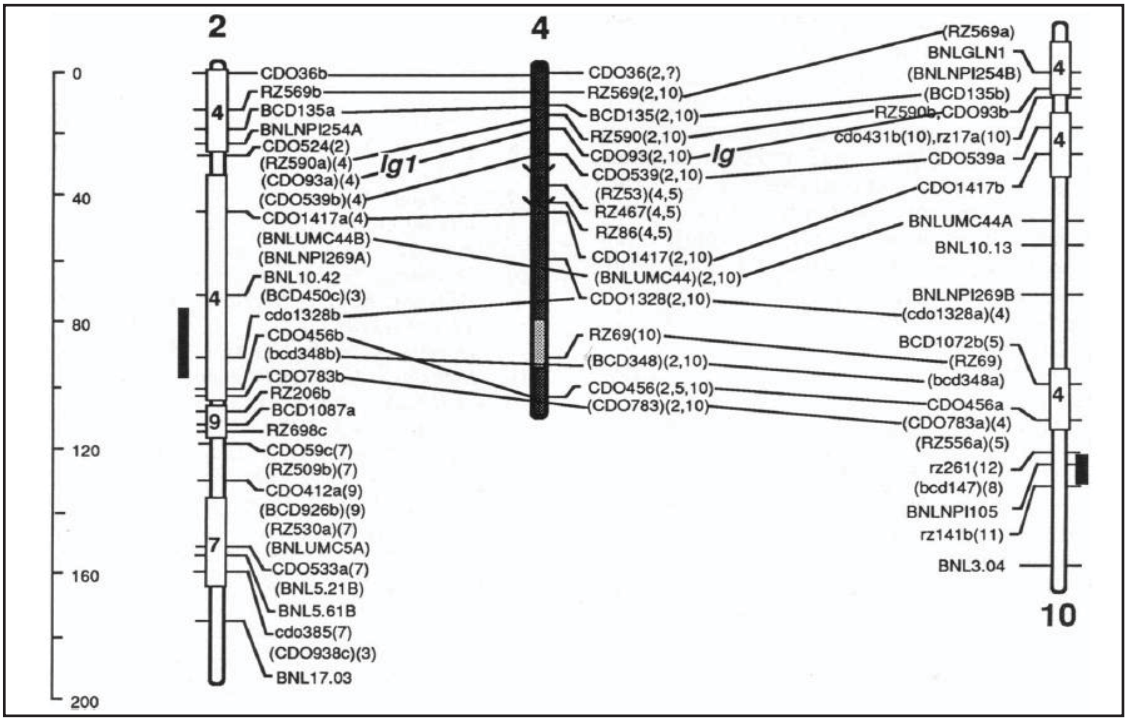


Fig. 1.12: Conserved linkage between rice chromosome 4 (black) and maize chromosomes 2 and 10 (white). Maize chromosome 10 is inverted to clarify the relationship of it with the other chromosomes. Connecting lines show colinearity; solid bars indicate the approximate positions of centromeres; numbers in brackets following marker name or in white bars on chromosomes indicate homologous linkage group association. This figure illustrates the relationship of rice chromosome 4 to specific arms of the maize chromosomes (arms delimited by centromere). Colinearity is not completely conserved as three markers in the middle of rice 4 are homologous to maize chromosomes 4 and 5. Figure copied from Ahn & Tanksley (1993) Figure 3.

Table 1-2: Summary of the percentage hybridization success of restriction fragment length polymorphism (RFLP) probes derived from barley, oat and rice to identify homologous loci on seven grass genera (Van Deynze et al. 1995). Wheat, barley and oats belong to subfamily Pooideae; rice is from Ehrhartoideae; and maize, sorghum and sugarcane are from Panicoideae.

	Approximate % hybridization						
Clone Source	**Wheat**	**Barley**	**Oats**	**Rice**	**Maize**	**Sorghum**	**Sugarcane**
Barley cDNA	100	100	95	87	100	71	44
Oat cDNA	86	92	95	97	70	89	75
Rice cDNA	52	72	46	98	59	80	74

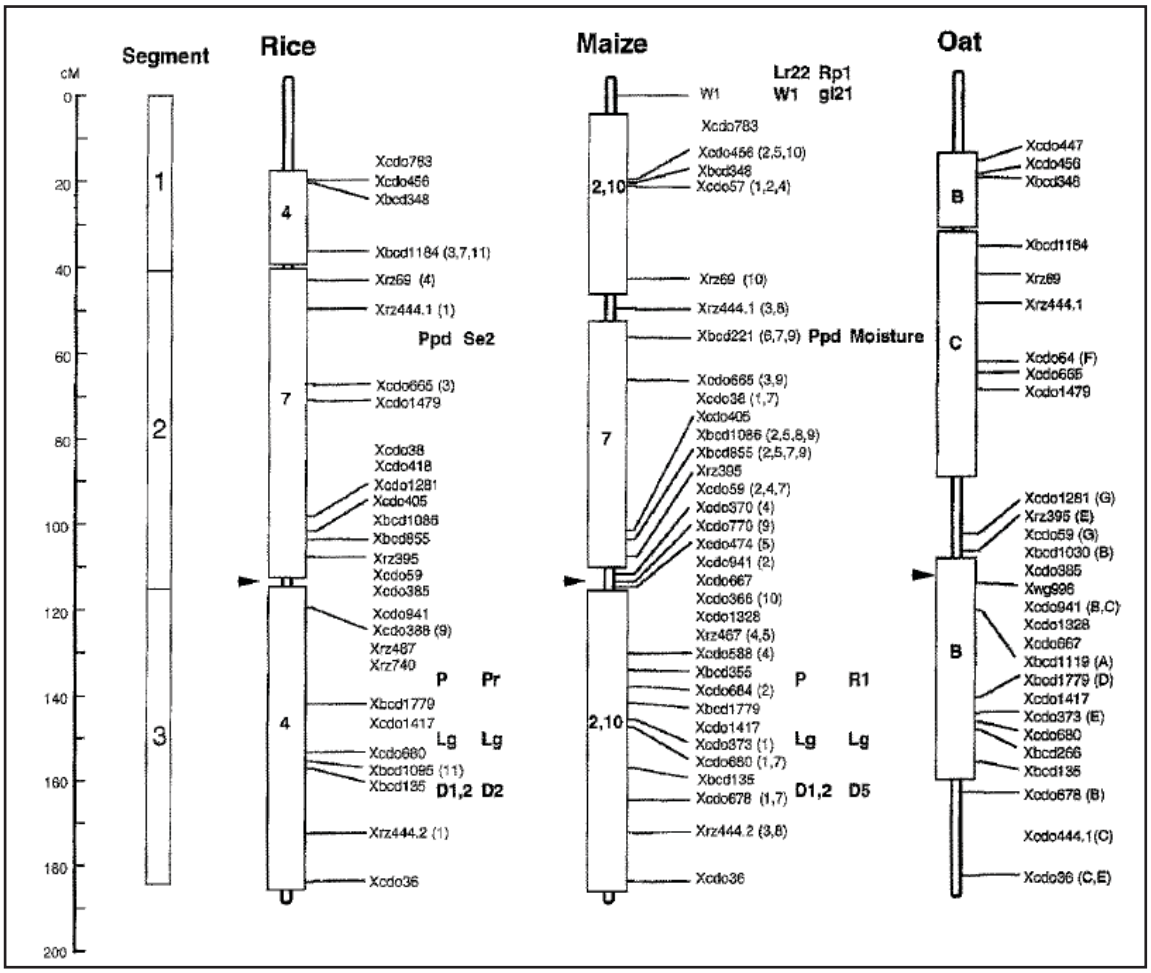


Fig. 1.13: Consensus maps for group 2 chromosomes of three Triticeae species. Conserved segments from homeologous rice, maize, and oat chromosomes are superimposed on the consensus map. Segments 1, 2, and 3 are conserved in all species compared. Symbols in bold type at the right of the linkage maps represent putative gene loci in Triticeae from two sources. Black arrows indicate centromeres. This figure demonstrates high degrees of colinearity across the grass subfamilies Ehrhartoideae (rice), Panicoideae (maize), and Pooideae (oats). Figure copied from Van Deynze et al. (1995) Figure 5.

Other early comparative genome mapping studies were conducted on members of the Brassicaceae (mustard family). A finer-scale genome study compared genome colinearity between the diploid model species *A. thaliana* and the crop species *Brassica nigra* (black mustard) that identified three colinear segments of a 1.5 megabase (Mb) *A. thaliana* contig in *B. nigra*, one of which was interrupted by a large chromosomal inversion (Lagercrantz et al. 1996). In addition, two QTL controlling flowering time in *B. nigra* were identified using probes designed from *A. thaliana* contigs surrounding the *CONSTANS* locus (Fig. 1.14).

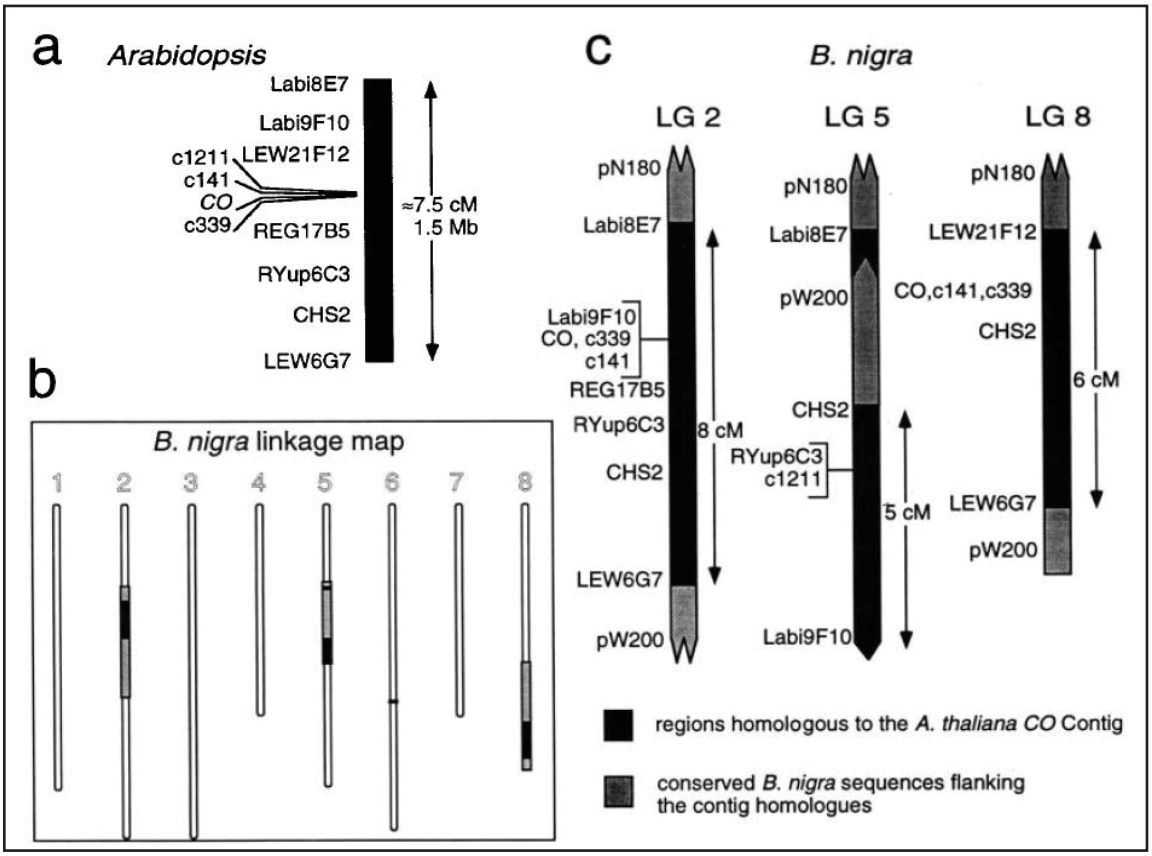


Fig. 1.14: Comparative mapping of a 1.5 megabase (Mb) *Arabidopsis thaliana* contig and its homeologs in *Brassica nigra*. (a) Physical map of a 1.5 Mb *A. thaliana* contig and the RFLP probes used on *B. nigra*. (b) Schematic genetic linkage map of the *B. nigra* linkage groups (LG) 2, 5, and 8 that are homologous to the *A. thaliana* contig (black boxes); grey boxes indicate homeology among the *B. nigra* LG. (c) The three *B. nigra* linkage groups in context. Figure copied from Lagercrantz et al. (1996) Figure 1.

Comparative genome mapping studies in the Solanaceae (nightshade family) are represented by comparisons between potato and tomato (Bonierbale et al. 1988; Iovene et al. 2008), and between tomato and pepper (Tanksley et al. 1988; Wu et al. 2009a).

Potato and tomato both have $x = 12$ chromosomes comprised of approximately 0.7 pg (picograms) of DNA, and nearly identical karyotypes (Bonierbale et al. 1988). The genome comparison between genetic maps of potato and tomato revealed four chromosomal inversions between LG 5, 9 and 10 (Bonierbale et al. 1988) (Fig. 1.15).

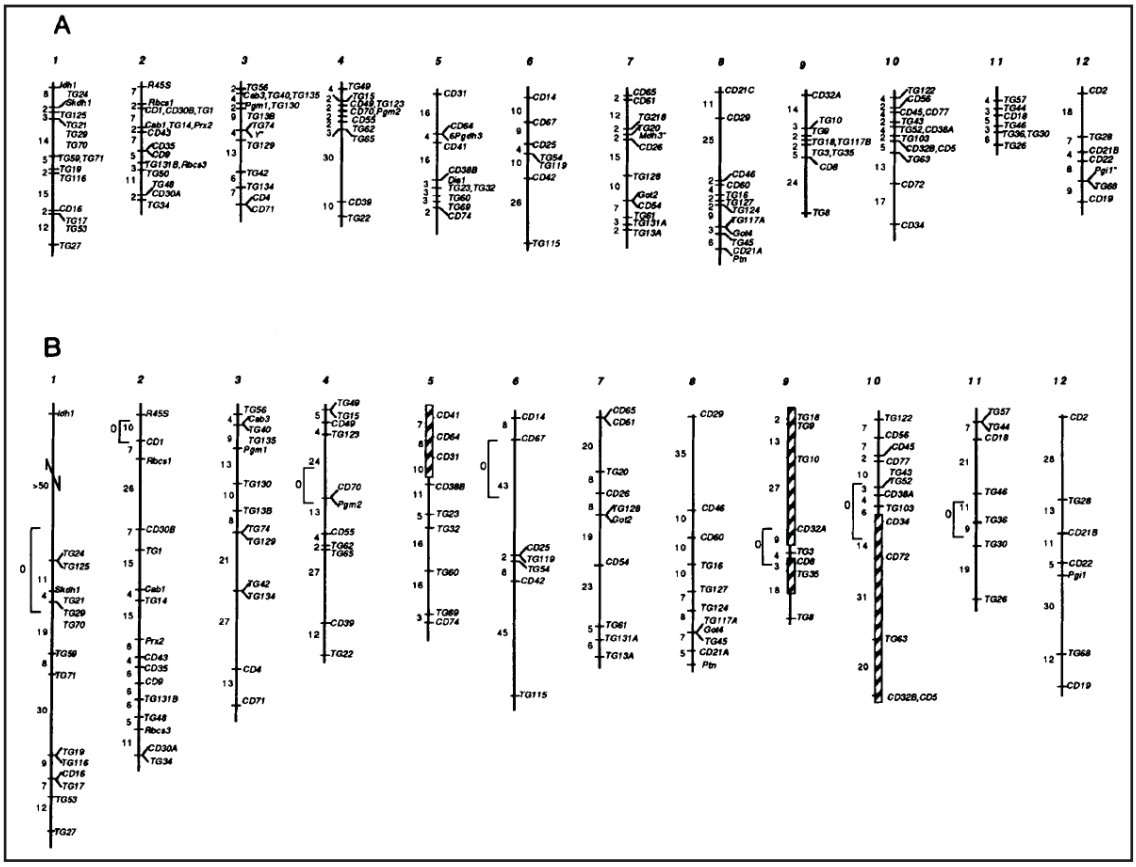


Fig. 1.15: Comparative genetic maps of potato and tomato based on RFLP. (A) Potato linkage map. (B) Tomato: hatched boxes represent intervals in which marker order is inverted between the two genomes. Centromeres are indicated by brackets. Figure copied from Bonierbale et al. (1988) Figure 1.

A more recent genome comparison of potato and tomato using fluorescence *in situ* hybridization (FISH) also revealed a chromosomal rearrangement on LG 6 and narrowed the distance spanned by the inversions identified with RFLP (Iovene et al. 2008). The discovery of an additional rearrangement between these two species suggests that even more structural differences may be identified with additional fine-scale mapping and comparative studies (Iovene et al. 2008).

The tomato genome has also been compared with the genome of pepper. Like other Solanaceae, pepper has $x = 12$ chromosomes, but with nearly four times the amount of DNA (2.76 pg in pepper vs. 0.74 pg in tomato) perhaps as a result of gene duplications (Tanksley et al. 1988). The comparison of the tomato and pepper genomes using maps constructed with RFLP revealed a high degree of genome rearrangement between the species and suggested 32 breakpoints necessary to align the segments (Tanksley et al. 1988). A subsequent genome comparison between tomato and pepper was performed using a newly developed pepper genetic map densely populated with markers derived from solanaceous conserved orthologous set (COSII) sequences (Wu et al. 2009a). The density of orthologous markers allowed the researchers to identify 19 inversions, six translocations, and numerous putative single gene transpositions that differentiate the two genomes, as well as 35 conserved segments (Wu et al. 2009a). The COSII-based pepper map constructed by Wu et al. (2009a) reveals a higher degree of conserved colinearity than that identified by Tanksley et al. (1988) with RFLP. The examples of comparative genome mapping in grasses, mustards, and nightshades presented here show the utility of well-characterized genomes for identifying genomic features in related, but less-studied, taxa.

1.2.6.2 Previous comparative genome QTL mapping studies in plants

QTL for disease resistance in the Solanaceae were used to compare the genomic organization of resistance genes (R genes) and to assess the conservation of structure and function of these loci in the genomes of pepper, tomato and potato (Grube et al. 2000). Positions of four R gene clusters were similar between tomato and potato; two clusters were similar between tomato and pepper; one cluster was similar between potato and pepper; and five clusters were similar among all three species. Even though R gene cluster positions were conserved among the species, resistance to the same type of pathogen was found only at two loci (Grube et al. 2000) (Fig. 1.16). The results of this work demonstrate that in the case of R genes, gene similarity is not enough to define putative QTL, and that functional comparisons are necessary to confirm homology (Grube et al. 2000).

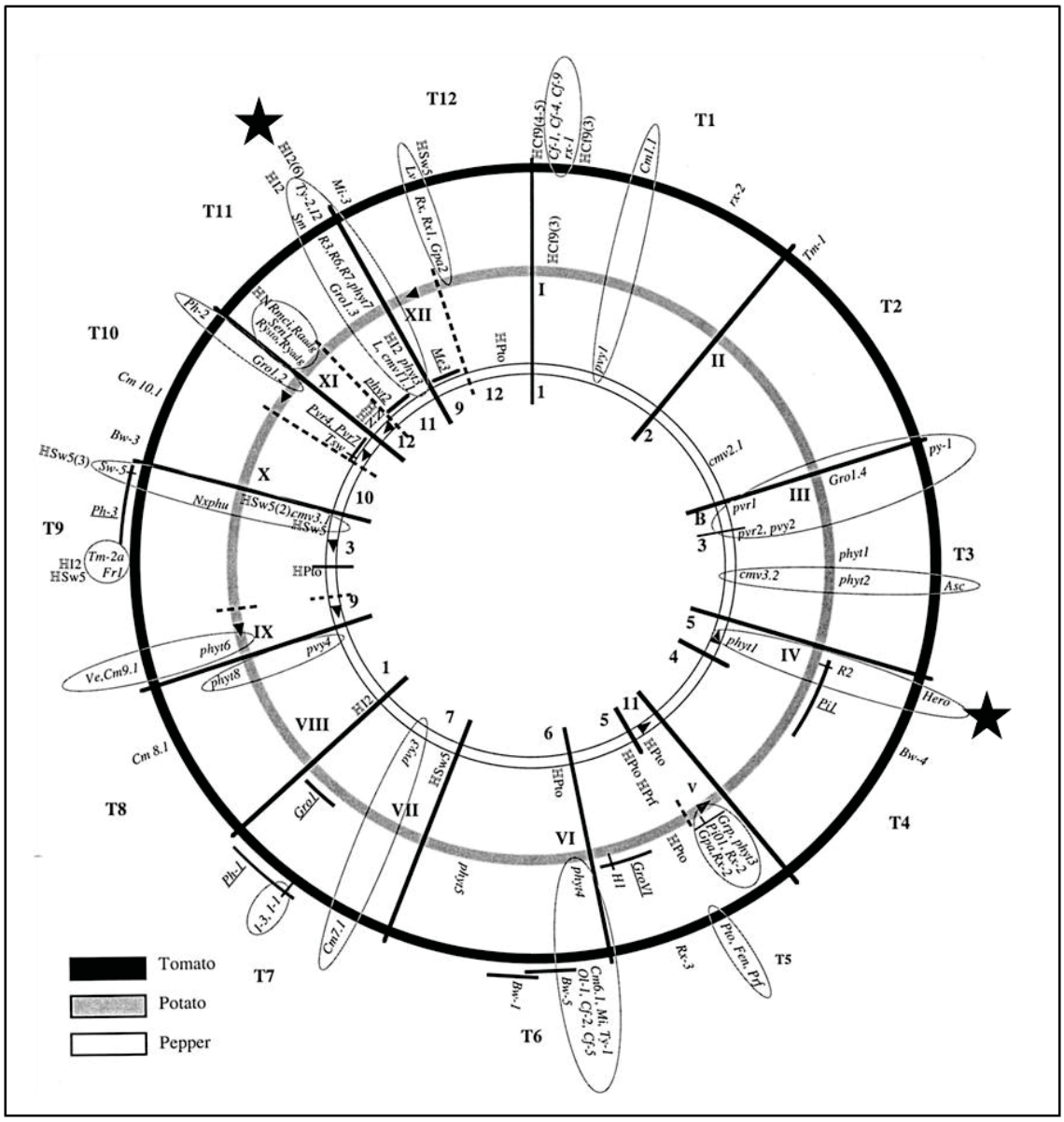


Fig. 1.16: Resistance gene organization in the Solanaceae. Black ring = tomato genome (T1-T12); grey ring = potato genome (I-XII); white ring = pepper genome (1-12). Solid lines perpendicular to the rings indicate chromosome ends; dashed perpendicular lines indicate borders of rearranged regions; black arrow heads indicate regions of inversion; ovals are drawn around R gene clusters. Black stars are added to show the two loci at which QTL for resistance to *Phytophthora* sp. in pepper (*phyt3*, *phyt1*) and potato (*phyt7*, *Pi1*) are located. Figure copied from Grube et al. (2000) Figure 1.

The alignment of the genetic maps of two species of pine from different phylogenetic subsections was used to compare the positions of three QTL for wood quality traits (Chagné et al. 2003). Two linkage groups were found to be involved and the QTL

positions were associated with candidate genes belonging to wood formation physiological pathways (Chagné et al. 2003).

A genetic linkage map constructed from an intraspecific grape cross was aligned to the linkage map constructed using an interspecific cross and used to compare QTL of four traits for fruit quality and one for disease resistance (Costantini et al. 2005). Two regions on one linkage group were found to be involved in the expression of the traits measured: the top of LG5 for disease resistance, and the bottom of LG5 for fruit quality. In addition, the fruit quality traits were found in association with a functional marker (Costantini et al. 2005).

1.2.7 Summary

In this section I have given examples of some of the ways genetic linkage maps populated with orthologous markers can be used for comparative genome mapping studies within plant families. These examples demonstrate that within plant families, differences in genome size and ploidy level do not hinder genome-wide comparisons, although gene duplications can make distinguishing ancestry difficult. The comparative QTL analyses presented here are a sample of how QTL on maps from multiple taxa can be used to predict regions of homology, and also illustrate that prediction of function is not equal to actual function. Next is an introduction to the rose family.

1.3 A brief history of the Rosaceae

The taxonomic history of the Rosaceae is convoluted, with many changes to nomenclature and species ranking over the past 60 years (Potter et al. 2007). I will use the nomenclature of Potter et al. (2007), and their treatment of the family, throughout my thesis. Fig. 1.17 illustrates this revision of the family, which is monophyletic, into three subfamilies based on chromosome number, as well as analyses of molecular and morphological markers (Potter et al. 2007). As subfamily Dryadoideae is comprised of all the nitrogen-fixing symbiotic species and no crop species, I will not discuss it further. Subfamily Rosoideae is comprised of the single supertribe Rosodae, which includes the genera *Rosa* and *Rubus* and the three tribes Colurieae, Sanguisorbeae and Potentilleae. Tribe Potentilleae includes the subtribe Fragariinae, which includes the genus *Fragaria* (strawberry). The third subfamily is Spiraeoideae. This subfamily has been re-assessed to include those genera that were formerly in the tribe Maloideae. Spiraeoideae is divided into two supertribes, Kerriodae and Pyrodae, and the tribes Amygdaleae [which includes the genus *Prunus* (almond, apricot, cherry, nectarine, peach, and plum)], Neillieae, Sorbarieae, Spiraeeae, and the genera *Gillenia* and *Lyonothamnus*. This reclassification recognizes the contribution of a spiraeoid ancestor (*Gillenia* or an ancestor are suggested) to the origin of the Pyrodae (Evans and Campbell 2002; Campbell et al. 2007; Potter et al. 2007). Within the supertribe Pyrodae are located the tribe Pyreae and subtribe Pyrinae. This subtribe is the best-studied of the family and includes such well-known horticultural crops as *Malus* (apple), *Pyrus* (pear), *Cydonia* (quince), and *Eriobotrya* (loquat). Of the members of this subtribe, *Malus* is the most important economically and the most studied.

Within the Rosaceae, base chromosome number (x) seems to be a better indicator of relationship than fruit type, which was used in previous taxonomic treatments, such as the 1964 treatment by Shulze-Menz cited by Potter et al. (2007). Potter and colleagues' treatment groups the subfamilies as follows: Dryadoideae $x = 9$; Rosoideae $x = 7$ (8); Spiraeoideae $x = 8$, 9, 15 or 17. The origin of $x = 17$ in Pyrodae has strong support from two recent hypotheses. The first hypothesis suggests hybridization of two sister taxa of *Gillenia* ($x = 9$) and allopolyploidization (Campbell et al. 2007; Evans and Campbell 2002; Potter et al. 2007), while the second hypothesis suggests autopolyploidization of an $x = 9$ ancestor of *Gillenia* (Velasco et al. 2010). Both hypotheses suggest subsequent diploidization and loss of a chromosome. Fig. 1.18 provides an illustration of the

current genome alignment of *Malus* homeologous chromosomes and demonstrates how part of the ancestral spiraeoid "chromosome 18" could have been lost along with part of chromosome 1, with the remainder incorporated into present-day chromosome 15 with part of present-day chromosome 2 (Velasco et al. 2010).

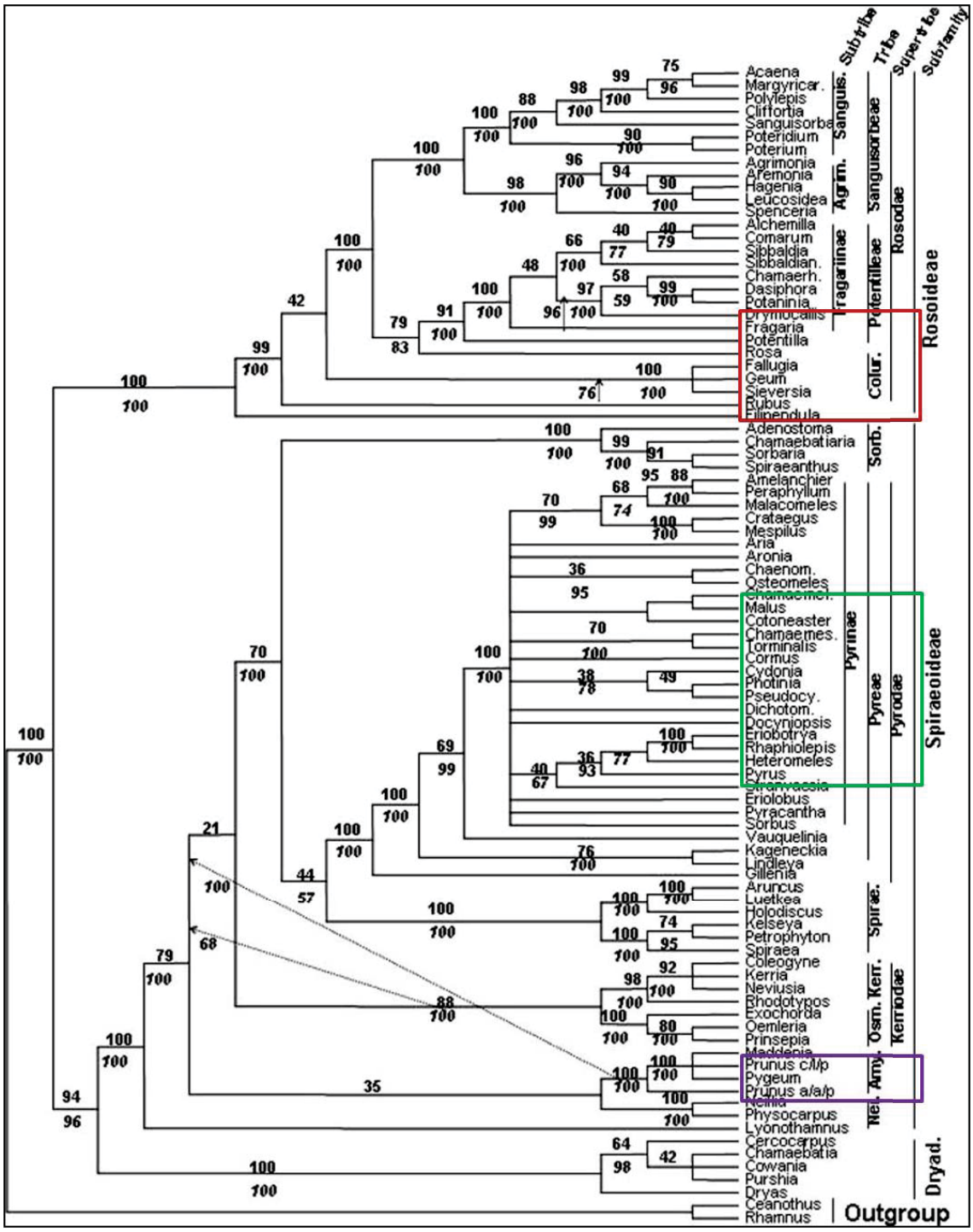


Fig. 1.17: The complete phylogeny of the Rosaceae divided into three subfamilies. Colored boxes highlight the genera addressed in this thesis. Phylogeny established by strict consensus of 226 most parsimonious trees constructed by analyzing ten genes or genomics regions (six nuclear and four chloroplast loci) in 91 taxa and and two cloned sequences in 54 taxa; bootstrap values above branches; Bayesian clade credibility below branches; arrows indicate groups that were supported by the Bayesian analysis but were not recovered in the strict consensus tree. Figure copied from Potter et al. (2007) Figure 1.

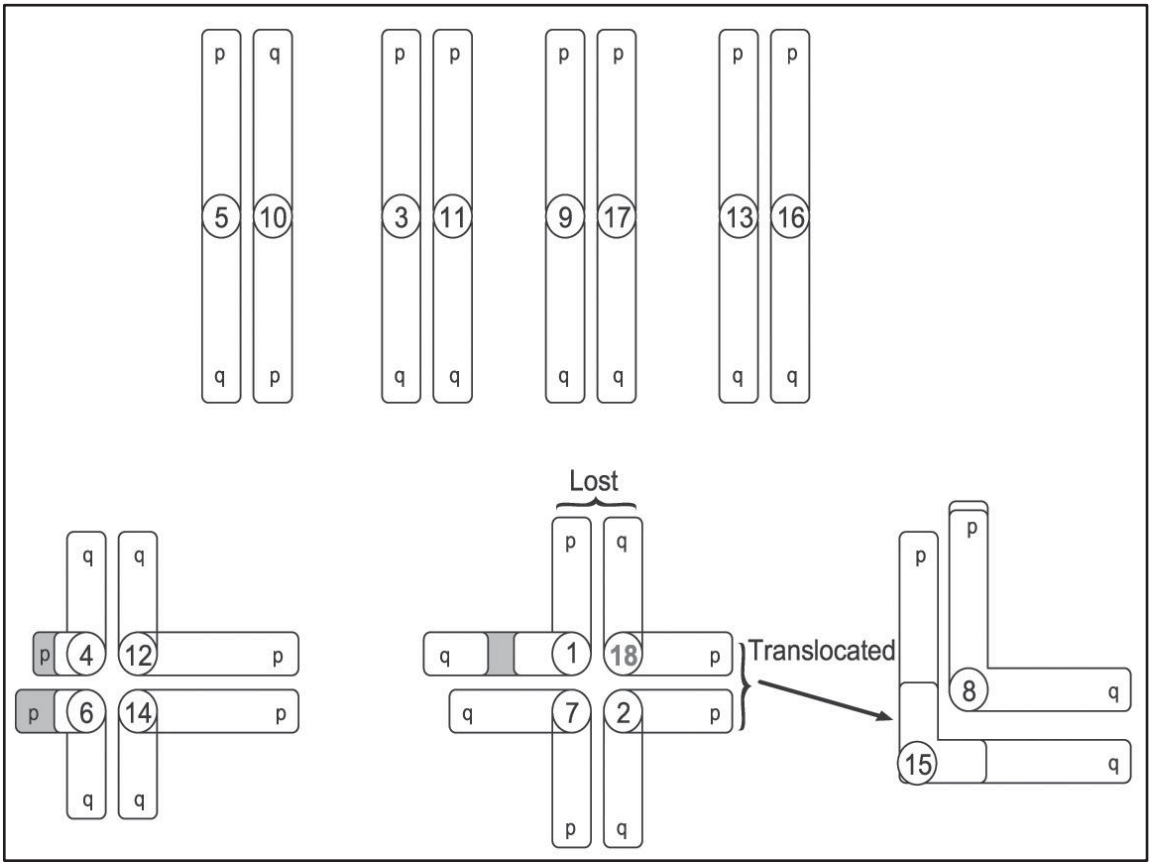


Fig. 1.18: Relationship of apple chromosomes based on a chromosome-by-chromosome comparison as determined by gene homology. Chromosome loss is hypothesized and shown here as the loss of portions of chromosomes 1 and 18, with the remainder of chromosome 18 and part of chromosome 2 translocated to chromosome 15. Grey areas indicate missing duplicate counterparts; letters p & q indicate chromosome orientation. Figure modeled after Velasco et al. (2010).

Polyploidy is common in the family. The genus *Fragaria* includes many polyploids with examples of species ranging from diploid ($2n$ = 14) to octoploid ($8x$ = 56), including one pentaploid (Scott 1951); cultivated strawberry, *Fragaria* × *ananassa*, is an octoploid (Byrne and Jelenkovic 1976). Species and cultivars within the genus *Rubus* span the range of diploid ($2n$ = 14) to decaploid ($10x$ = 70) (although no pentaploid examples are known), with a few cultivars of $12x$ = 84 and $14x$ = 98 (Stafne et al. 2005; Thompson 1995a). Many examples of aneuploids, cultivars which have lost at least one chromosome, have also been identified (Stafne et al. 2005; Thompson 1995a). *Malus* has at least three extant triploid and four extant tetraploid species (Way et al. 1991). Within *Malus*, the cultivated apple 'Jonagold' is triploid as a result of the fertilization of an unreduced 'Golden Delicious' female gamete and 'Jonathan' pollen (Gianfranceschi

et al. 1998). *Pyrus* has 12 triploid cultivars, two tetraploid cultivars, one triploid species, and one tetraploid species (Shenghua and Chengquan 1994). Shenghua and Chengquan (1994) also found that *P. bretschneideri* 'Yali' has different ploidy levels (diploid, triploid or tetraploid) depending on its origin. Within the genus *Prunus*, the species *P. serotina* (black cherry) is tetraploid (Downey and Iezzoni 2000) and *P. domestica* (European plum) is hexaploid (Arús and Gardiner 2007). The level of ploidy can influence mapping efficacy of cross-species markers (Lagercrantz 1998; Bolot et al. 2009) because of duplicated loci (paralogs) on different chromosomes (Doyle and Davis 1998; Comai 2005).

1.4 Economic importance of Rosaceae genera

Rosaceae crop species together rank third in agronomic importance in temperate zones (Dirlewanger et al. 2002). Table 1-3 provides the most recent details of production and value to the producer for eleven common Rosaceae crop species (http://faostat.fao.org/default.aspx). Rosaceae genera have a variety of uses other than fruits (Table 1-4), with many of their members valued for their ornamental characteristics, wood quality, and secondary metabolites (Hummer and Janick 2009).

Table 1-3: World-wide production (2008, in tonnes) and value (2007, in US$) of eleven Rosaceae fruit crops.

	Production (tonnes)*	Producer value† ($US)
Apples	69,603,640	46,016,287,327
Pears	20,998,473	11,492,205,218
Peaches and Nectarines	18,000,853	10,406,982,394
Strawberries	4,068,454	8,419,280,198
Almonds	2,112,815	5,748,393,874
Plums and Sloes	10,340,902	4,118,049,995
Cherries	1,875,618	3,537,278,997
Apricots	3,473,710	3,131,522,953
Raspberries	458,885	1,224,604,837
Sour Cherries	1,153,348	1,012,455,133
Quinces	480,456	309,515,202
Totals	**132,567,154**	**95,416,576,129**

*Production values from FAOSTAT 2008

†Producer value from PRICESTAT 2007

Table 1-4: Economically important species of Rosaceae. Table after Hummer & Janick (2009) Table 2.

Subfamily	*Genus*	*species*	Common name	Uses
Amygyloideae	*Prunus*	*armeniaca*	Apricot	Fresh and processed fruit
		avium	Sweet cherry	Fresh and processed fruit
		cerasus	Tart (sour) cherry	Fresh and processed fruit
		domestica	European plum	Fresh and processed fruit
		dulcis	Almond	Fresh and processed fruit
		mume	Mume	Ornamental
		persica	Peach, nectarine	Fresh and processed fruit
		serotina	Black cherry	Timber species
Maloideae	*Amelanchier*	*alnifolia*	Saskatoon, serviceberry; shadbush	Landscape ornamental
	Aronia	*melanocarpa*	Black chokecherry	Processed fruit for juice, nutriceutical
	Chaenomales	*japonica*	Japanese quince	Landscape ornamental, processed fruit
	Cotoneaster	*spp.*	Cotoneaster	Landscape ornamental
	Crataegus	*spp.*	Hawthorn, thornapple	Landscape ornamental, craft uses for wood
	Cydonia	*oblonga*	European quince	Fresh and processed fruit, dwarfing rootstock for pear and loquat
	Eriobotrya	*mespilus*	Loquat	Fresh and processed fruit
	Malus	x *domestica*	Apple	Fresh and processed fruit
		pumila	Crabapples	Landscape ornamentals
	Pyrus	*calleryana*	Callery pear	Landscape ornamental
		communis	European pear	Fresh and processed fruit
		serotina	Japanese pear (nashi)	Fresh fruit
		ussurienses	Chinese pear	Fresh fruit
	Mespilus	*germanica*	Medlar	Fresh fruit (bletted)
	Photinia	*spp.*	Photinia	Landscape ornamental
	Pyracantha	*spp.*	Firethorn	Landscape ornamental
	Sorbus	*spp.*	Mountain ash, rowan	Landscape ornamental
Rosoideae	*Fragaria*	x *ananassa*	Strawberry	Fresh and processed fruit
	Geum	*spp.*	Avens	Herbaceous perennial
	Kerria	*japonica*	Kerria	Landscape ornamental
	Potentilla	*spp.*	Cinquefoil	Landscape ornamental
	Rosa	*spp.*	Rose	Cut flowers, landscape ornamental, perfume oil, medicinal
	Rubus	*spp. and hybrids*	Blackberry, raspberry hybrid berry	Fresh and processed fruit
Spiraeoideae	*Spirea*	*prunifolia*	Bridal wreath	Landscape ornamental
	Exochorda	*racemosa*	Exochorda	Landscape ornamental
	Physocarpus	*opulitolius*	Ninebark	Landscape ornamental

1.5 Genetic mapping in the Rosaceae

1.5.1 Obstacles to genetic mapping in the Rosaceae

Many of the horticultural Rosaceae species have high heterozygosity, providing useful material for breeding, yet this diversity, along with self-incompatibility, creates difficulties for mapping studies because of the inability to develop inbred lines, such as those commonly used in studies of the Brassicaceae, Solanaceae and Poaceae. The time investment to develop F_2 generations in fruit trees is considerable, as large woody plants often have long juvenile periods; for example, apple trees require at least six years to flower (Jonkers 1971). Consequently, this difficulty leads to a lack of fine-scale mapping compared with genera in the families Brassicaceae, Solanaceae, and Poaceae.

Further obstacles for mapping in the Rosaceae arise from the lack of diversity in the parents of some of the elite cultivar breeding and mapping populations. Sweet cherry has low genetic diversity, with most cultivars having 'Emperor Francis' in their background (Olmstead et al. 2008). Commercial strawberry, in 1987, had only 53 genetic contributors to the 134 cultivars studied by Sjulin and Dale (1987). Similarly, red raspberry cultivars, in 1993, had only 50 founding clones contributing to the 137 cultivars analyzed by Dale et al. (1993); blackberry had 19 founding clones contributing to 32 cultivars (Stafne and Clark 2004). In peach, the genetic background narrowed considerably in the latter part of the 19th century when a few "Chinese Cling"-type varieties became important for the breeding of eating-quality characteristics (Scorza et al. 1985). The creation of the F_2 hybrid almond 'Texas' × peach 'Earlygold' (T×E) progeny for use as a reference mapping population has provided a wealth of information on polymorphism in these two low-diversity species (Joobeur et al. 1998). Apple, on the other hand, is quite amenable to mapping studies, especially compared with the other genera mentioned, and in spite of most of the fruiting cultivars having 'Delicious' and 'Golden Delicious' in their pedigrees (Kellerhals 2009). The interspecific *M.* × *domestica* 'Malling 9' × *M. robusta* 'Robusta 5' (M.9×R5) mapping population (Celton et al. 2009b; Rusholme Pilcher et al. 2008) is, like T×E, highly heterozygous and useful for mapping studies.

Another difficulty affecting the efficiency of mapping studies in the Rosaceae is the localized segregation distortion that is often identified during map construction in specific Rosaceae cultivars. This distortion could be the result of the history of intense

selection for horticultural traits (Igarashi et al. 2008; Joobeur et al. 1998). RFLP and isozyme markers on the *Prunus* T×E reference map exhibit heterozygote-biased segregation distortion on four linkage groups (Joobeur et al. 1998). Segregation distortion leading to the loss of certain allele combinations can arise in interspecific crosses as a result of incompatible chromosome pairing (Comai 2005). Polyploidy, as found in many Rosaceae genera, can be another source of segregation distortion because of inter-genome reorganization, silencing, and gene loss, as "homoeologs" can become silenced immediately or soon after the doubling event (Adams and Wendel 2005) returning the level of gene expression in the polyploidy to the same as that found in the diploid (Mittelsten Scheid et al. 1996).

1.5.2 Genetic map construction in the Rosaceae

While F_2 progeny can be generated in some Rosaceae genera, such as *Prunus* and *Fragaria*, the high heterozygosity of many Rosaceae genera and the difficulty of developing inbred lines have led to the adoption of the "pseudo-testcross" mapping strategy (Grattapaglia and Sederoff 1994; Maliepaard et al. 1998). This strategy was originally developed for genetic mapping in *Eucalyptus*, another highly heterozygous genus. The concept of this mapping strategy is that in an F_1 population derived from heterozygous parents (Table 1-1), some loci will be heterozygous in one parent and homozygous in the other, resulting in a segregation ratio of 1:1, as is found in backcross progeny. Since polymorphic markers can only be mapped in the heterozygous parent, two independent parental linkage maps are constructed, with the *a posteriori* inference of parental genotypes from marker segregation in the progeny (Grattapaglia and Sederoff 1994). The term "double" or "two-way" pseudo-testcross indicates the use of both the progeny and parents to determine parental origin of the alleles; and markers that are heterozygous in both parents will segregate 1:2:1 or 3:1 in the progeny, providing markers in common between the two parental maps, which can be integrated to generate a consensus map (Grattapaglia and Sederoff 1994). Using the F_1 progeny limits the amount of genetic information gained from the cross, since only those loci that are heterozygous in at least one parent will be detected in the progeny, and alleles homozygous in both parents (aa × bb) will produce only heterozygous offspring.

1.5.3 Early linkage map construction in Rosaceae

The first genetic linkage maps for the Rosaceae were constructed of "first generation" markers (Gupta et al. 2002b) for the genera *Prunus* and *Malus* and were developed in peach (*Prunus persica*) and apple (*Malus* × *domestica*). These early maps were comprised of isozymes, RAPD, (Chaparro et al. 1994) and RFLP markers (Hemmat et al. 1994; Rajapakse et al. 1995), along with morphological markers (Table 1-5). For peach, Chaparro et al. (1994) assigned their markers to 15 LG covering approximately 396 cM, while Rajapakse et al. (1995) assigned their markers to eight LG covering 332 cM. Subsequently, in 1998, Joobeur and colleagues used the F_2 progeny of an inter-specific cross between almond 'Texas' and peach 'Earlygold' to develop a map with 248 markers distributed over eight LG and covering 491 cM (Joobeur et al. 1998). The eight LG corresponded to the eight chromosomes of *Prunus*. The map was constructed of isozymes and RFLP markers derived from genomic and cDNA libraries of different *Prunus* species and apple. This was the first map to be considered saturated, with an average density of one marker every 2.0 cM, and it is this map that became the *Prunus dulcis* 'Texas' (almond) × *P. persica* 'Earlygold' (peach) (T×E) reference map (Joobeur et al. 1998). Other early maps constructed for *Prunus* species include RFLP-based linkage maps for almond (*P. dulcis*) (Viruel et al. 1995), peach × almond (Foolad et al. 1995), and tetraploid sour cherry (*P. cerasus*) (Wang et al. 1998) (Table 1-5).

Table 1-5: Genetic maps constructed in Rosaceae genera using predominantly first generation molecular markers. Examples of first generation markers are restriction fragment length polymorphism, isozyme and morphological.

Taxon	Marker types	Linkage groups	Map length (cM)*	Reference
Prunus persica	isozyme, RAPD, morphological	15	396	Chaparro et al. (1994)
Prunus persica	isozyme, RAPD, RFLP, morphological	8	332	Rajapakse et al. (1995)
Almond × Peach	isozyme, RFLP	8	491	Joobeur et al. (1998)†
	RFLP, isozyme	7	474	Jáuregui et al. (2001)
Prunus dulcis	isozyme, RFLP	8	394	Viruel et al. (1995)
	RFLP, RAPD	8	415	Joobeur et al. (2000)
	RFLP, RAPD	8	386	Ballester et al. (2001)
Peach × Almond	isozyme, RFLP	9	800	Foolad et al. (1995)
Prunus cerasus	RFLP	19	461	Wang et al. (1998)
Malus × *domestica*	isozyme, RAPD	24	950	Hemmat et al. (1994)
	isozyme, RAPD, RFLP, (AFLP, SCAR, SSR)	17	984	Maliepaard et al. (1998)
Fragaria vesca	RAPD	7	445	Davis & Yu (1997)
Rosa sp.	RAPD, (AFLP)	7	370	Debener & Mattiesch (1999)

*If more than one map is presented, the longer length is listed.
†First *Prunus* map considered saturated, became the 'Texas' × 'Earlygold' reference map.
Marker types in parentheses are second generation markers and make up a small proportion of the total markers used for mapping.
Abbreviations: AFLP: amplified fragment length polymorphism; cM: centimorgan; RAPD: random amplification of polymorphic DNA; RFLP: restriction fragment length polymorphism; SCAR: sequence characterized amplified region; SSR: simple sequence repeat; sp.: species.

Hemmat et al. (1994) were the first to publish a linkage map of apple. Their map was constructed of 253 RAPD, RFLP, and isozyme markers in a progeny derived from 'Rome Beauty' × 'White Angel' and spanned 950 cM over 24 LG. The first alignment of apple maps constructed of male- and female-specific "first generation" isozyme and RFLP and "second generation" (Gupta et al. 2002b) AFLP, RAPD, SCAR, SSR markers was accomplished using 152 F_1 progeny of 'Prima' × 'Fiesta' (Maliepaard et al. 1998). The markers from this work were distributed over 17 LG corresponding to the 17 chromosomes of apple and covered 842 cM for 'Prima' and 984 cM for 'Fiesta'. Both of these early apple maps were developed using the "pseudo-testcross" mapping strategy.

The first genetic linkage maps for members of subfamily Rosoideae were developed in *Fragaria* (Davis and Yu 1997), *Rosa* (Debener and Mattiesch 1999), and *Rubus*

(Graham et al. 2004). The *Fragaria* map was constructed of 80 markers, predominantly RAPD, and consisted of seven LG covering 445 cM. It was constructed using 80 F_2 progeny resulting from the cross of *F. vesca* 'Baron Solemacher' × *F. vesca* WC6. The *Rosa* map was constructed of 305 RAPD and AFLP markers analyzed in a progeny of 60 F_1 diploid individuals derived from rose genotypes 93/1-117 and 93/1-119 using the "double pseudo-testcross" mapping strategy. The resulting maps consisted of seven LG for each parent and covered 326 cM with a marker every 2.4 cM for 93/1-117, and 370 cM with a marker every 2.6 cM for 93/1-119. The *Rubus* map was constructed of 432 loci derived from genomic SSR, EST-SSR and AFLP using 300 F_1 progeny of *R. idaeus* 'Glen Moy' × 'Latham'. The resulting consensus map consisted of nine LG spanning 789 cM.

These early linkage maps set the stage for genetic mapping in the Rosaceae. The maps provided tools for identifying the genomic regions responsible for the expression of quantitative traits, specific markers for use in marker assisted selection, and for comparative mapping.

1.5.4 The rise of simple sequence repeat molecular markers for Rosaceae studies

"Second generation" (Gupta et al. 2002b) SSR marker use by the Rosaceae scientific community became prevalent in the late 1990s and these markers have been used for linkage map construction and cultivar identification in most of the agronomic crops in the family (Table 1-6). Rosaceae-specific SSR markers were first developed in apple by Guilford et al. (1997) and Hokanson et al. (1998) for cultivar identification, and Gianfranceschi et al. (1998) for germplasm diversity assessment and linkage mapping, with additional SSR markers being developed by Liebhard et al. (2002). The first whole-genome reference map of apple was constructed of 840 AFLP, RAPD, SCAR, and SSR molecular markers using 267 F_1 progeny of 'Fiesta' × 'Discovery' (Liebhard et al. 2003a) and the "pseudo-testcross" mapping strategy (Grattapaglia and Sederoff 1994). The resulting map covered 17 linkage groups of 1,144 cM in 'Fiesta' and 1,454 cM in 'Discovery'. SSR markers from *Prunus*, *Malus*, and *Fragaria* have been used in the construction of a rose linkage map (Hibrand-Saint Oyant et al. 2008); and SSR markers from apple and pear have been used in the construction of a loquat (*Eriobotrya japonica*) linkage map (Gisbert et al. 2009), and for cultivar identification

in both loquat (Soriano et al. 2005) and quince (*Cydonia oblonga*) (Yamamoto et al. 2004b).

Table 1-6: Summary of the most prevalent marker types used for cultivar identification in the Rosaceae.

Genus	Major Marker Type	Reference
Pyrus	SSR	Kimura et al. (2002)
Malus & *Pyrus*	SSR	Yamamoto et al. (2001a)
Cydonia, *Malus* & *Pyrus*	SSR	Yamamoto et al. (2004b)
Eriobotrya & *Malus*	SSR	Soriano et al. (2005)
Prunus	SSR	Cipriani et al. (1999)
	SSR	Sosinski et al. (2000)
	SSR	Testolin et al. (2000)
	SSR	Cantini et al. (2001)
	SSR	Dirlewanger et al. (2002)
	SSR	Aranzana et al. (2003a)
	SSR	Romero et al. (2003)
	EST-SSR	Hagen et al. (2004)
Rubus	SSR	Stafne et al. (2005)
	EST-SSR	Lewers et al. (2008)
	SSR	Woodhead et al. (2008)
Rubus & *Fragaria*	SSR	Lewers et al. (2005)

Abbreviations: SSR: simple sequence repeat; EST: expressed sequence tag

SSR markers have been developed for peach (Cantini et al. 2001; Cipriani et al. 1999; Dirlewanger et al. 2002; Sosinski et al. 2000; Testolin et al. 2000), Asian pear (*Pyrus pyrifolia*) (Kimura et al. 2002; Yamamoto et al. 2002a), rose (*Rosa* sp.) (Hibrand-Saint Oyant et al. 2008), strawberry (*Fragaria* sp.) (James et al. 2003; Sargent et al. 2003; Bassil et al. 2006; Gil-Ariza et al. 2006; Keniry et al. 2006), and raspberry and blackberry (*Rubus* sp.) (Amsellem et al. 2001; Graham et al. 2002; Lopes et al. 2006; Lewers et al. 2008) and used to enhance existing maps, or for *de novo* map construction (Table 1-7).

Table 1-7: Summary of the most prevalent marker types used for genetic map construction or augmentation of existing maps in the Rosaceae.

Genus	Major marker type	New or existing map	Reference
Prunus	AFLP	New	Lu et al. (1998)
	AFLP, RAPD, SSR	New	Yamamoto et al. (2001b)
	EST-SSR, RFLP	Existing	Etienne et al. (2002)
	SSR	Existing	Aranzana et al. (2003b)
	SSR	Existing	Verde et al. (2005)
	SSR	New	Dirlewanger et al.(2004a)
	SSR, EST-SSR	Existing	Dirlewanger et al. (2006)
	SSR, AFLP	New	Olmstead et al. (2008)
Malus	AFLP, RAPD, SSR	Existing	Liebhard et al. (2003a)
	SSR	Existing	Silfverberg-Dilworth et al. (2006)
	SSR	New	Fernández-Fernández et al. (2008)
	AFLP	New	Igarashi et al. (2008)
	SSR	New	Celton et al. (2009b)
Pyrus	AFLP, SSR	New	Yamamoto et al. (2002b)
Malus & *Pyrus*	SSR	Existing	Pierantoni et al. (2004)
	AFLP	New	Yamamoto et al. (2004a)
	SSR	New	Celton et al. (2009a)
Prunus & *Fragaria*	RFLP	Existing	Vilanova et al. (2008)
Rubus	SSR	New	Graham et al. (2004)
	SSR	Existing	Stafne et al. (2005)
	AFLP	Existing	Sargent et al. (2007b)
	EST-SSR	Existing	Graham et al. (2004)
	EST-SSR	Existing	Woodhead et al. (2008)
	EST-SSR	Existing	Woodhead et al. (2010)
Eriobotrya	SSR	New	Gisbert et al. (2009)
Rosa	RAPD, AFLP	New	Debener and Mattiesch (1999)
	EST-SSR, SSR	New	Hibrand-Saint Oyant, et al. (2008)

Abbreviations: AFLP: amplified fragment length polymorphism; EST: expressed sequence tag; RAPD: random amplification of polymorphic DNA; RFLP: restriction fragment length polymorphism; SSR: simple sequence repeat.

1.5.5 Gene expression analyses and mapping in Rosaceae

One of the first large-scale gene expression analyses to occur in the Rosaceae was performed on the tissues of the receptacle and achenes of cultivated strawberry using DNA microarray technology (Aharoni and O'Connell 2002). This study compared gene expression during tissue maturation at three stages for both tissue types. The researchers clustered the 1700 expressed genes into 21 putative function categories based on basic local alignment search tool (BLAST) (Altschul et al. 1990) search output. The next major study generated three EST libraries from cDNA of ripening apricot (*Prunus armeniaca*) fruit, which were compared to sequences in public databases using

BLASTX (Grimplet et al. 2005). The researchers assigned their 3425 unique genes (unigenes) to 21 categories similar to those of strawberry.

The first EST sequencing project for apple was performed in 43 different tissues of 'Royal Gala' during which over 150,000 EST were collected (Newcomb et al. 2006). Additionally, the authors identified more than 4,000 SSR, as well as a SNP locus every 706 bp, and estimated that approximately half the expressed genes of apple were captured during their analysis. Subsequently, additional EST libraries were developed from different tissues from nine apple cultivars (Gasic et al. 2009a). The first EST sequencing project for cultivated blackberry (*Rubus* sp.) utilized leaf tissue, which was used to generate a cDNA library of 3,000 expressed genes and to identify SSR (Lewers et al. 2008).

These EST libraries are sources of sequences that can be used to design molecular markers based on expressed genes. The categorization of the sequences into putative functional groups allows the sequences from the four genera to be compared. Furthermore, the sequence information can be used to design genus-specific primer pairs for identification of putative orthologous regions to link genetic maps.

Even with the availability of large EST data sets representing each of the major agronomic genera, small-scale EST libraries are often developed for specific tissues or species and used for a range of purposes. Small numbers of EST-SSR markers have been generated *de novo* from: leaves of apricot and used to analyze genetic diversity in *Prunus* (Hagen et al. 2004); roots, leaves, and fruit (Graham et al. 2004), and buds and roots (Woodhead et al. 2008) of red raspberry (*Rubus idaeus*) to enhance an existing linkage map; vegetative and floral buds of *Rosa* to enhance an existing map (Hibrand-Saint Oyant et al. 2008); and mature shoots and mature fruits of apple to construct a new genetic map (Igarashi et al. 2008). Previously developed *Prunus*-derived EST-SSR sequences have been used to design primers to amplify new polymorphic loci and augment existing *Prunus* linkage maps (Dirlewanger et al. 2006; Ogundiwin et al. 2009). The apple EST database was used to design markers to develop a new apple linkage map (Celton et al. 2009b). EST-SSR have been useful for intrageneric comparisons such as within *Prunus* (Decroocq et al. 2003; Dirlewanger et al. 2002), *Pyrus* (Yamamoto et al. 2002b), and *Rubus* (Stafne et al. 2005); and between genera such as *Malus* and *Pyrus* (Celton et al. 2009b; Pierantoni et al. 2004; Yamamoto et al.

2004a; Yamamoto et al. 2001a), *Malus*, *Pyrus* and *Cydonia* (Yamamoto et al. 2004b), *Malus* and *Prunus* (Illa et al. 2011), between *Prunus* and *Fragaria* (Vilanova et al. 2008), and with limited success between *Rubus* and *Fragaria* (Lewers et al. 2005).

A significant alternative resource used for marker development is the gene-based conserved orthologous set (COS) sequences developed by Cabrera et al. (2009). The Rosaceae COS (RosCOS) markers were produced by comparing 412,827 Rosaceae-specific EST sequences (mainly from *Prunus*, *Malus* and *Fragaria*) from GenBank by BLAST analysis against *A. thaliana* single copy genes. The total of 30,801 Rosaceae-specific EST identified were assembled into contigs to create 7,247 Rosaceae-specific unigenes. The researchers designed PCR primer pairs to span a putative intron for 857 sequences selected from 1,039 contigs constructed from a minimum of two EST from at least two of the Rosaceae genera, or from *Prunus* alone. Of the 857 PCR primer pairs, 91% amplified a product in peach 'Earlygold', and this was sequenced. The interspecific *Prunus* T×E progeny were genotyped using the final set of 613 polymorphic RosCOS markers which were then located on the reference bin map. These orthologous sequences have the potential to be used as markers for comparative mapping within the Rosaceae. When the authors evaluated the transferability of primer pairs among the three genera by selecting a subset of RosCOS derived from different combinations of Rosaceae-specific EST sources, they found that PCR amplification across the genera was most successful if the RosCOS was derived from EST of at least two Rosaceae genera, although primer design was the most difficult when all three genera were represented. The authors do not discuss whether the amplicons were polymorphic, nor do they discuss what DNA was used for evaluation of primer pair transferability.

RosCOS have yet to be widely tested on the Rosaceae. Two studies have been published to date: an inter-tribal pairwise comparative genome study in which a subset of 126 RosCOS that were located on the T×E reference map were also located on the interspecific *Fragaria vesca* × *F. bucharica* (FVxFB) reference bin map (Illa et al. 2011); and the use of RosCOS to augment an existing sweet cherry (*Prunus avium*) linkage map (Cabrera et al. 2011).

1.5.6 Summary

Both the large number of taxa and high morphological diversity has made phylogenetic studies of the Rosaceae a challenge. The most recent treatment has moved from morphological characters to grouping genera according to base chromosome number and molecular characters, resulting in the current classification of three subfamilies, three supertribes, 10 tribes, and four subtribes.

Linkage map construction in *Prunus* is the most complete, with maps representing the major economic crops (peach, almond, sweet and sour cherry). Mapping in Pyrinae follows, with maps available for apple, pear and loquat. Within these two groups, many markers are shared among the maps but few markers are shared between the clades. Linkage map construction in *Fragaria*, *Rosa* and *Rubus* lags behind, with only a few maps having been developed for each genus.

In the past 20 years, genome mapping in the Rosaceae has progressed rapidly as a result of advances in molecular marker type and higher-throughput methods of genotyping. As more populations are mapped genetically, and the genomic regions involved in the control of important agronomic traits are identified, trait- or gene-specific markers can be developed to aid plant breeders who are using marker assisted breeding methods (for examples, see Baldi et al. 2004; Bus et al. 2010; Bus et al. 2002; Cao et al. 2011; Costa et al. 2009; Graham et al. 2011; Kumar et al. 2012). More advances are on the horizon as entire genomes are sequenced, and these will be discussed in section "1.7.3 *Recent advances in genomics and technology for the Rosaceae*".

1.6 Quantitative trait locus mapping in Rosaceae

As the development of genetic maps and genomic resources has progressed, so has the ability to examine the expression of complex traits and associate these traits with specific genomic regions. Traits of interest for Rosaceae breeders include disease resistance, fruit quality (e.g. firmness, size, flavor, content of volatiles, polyphenolics), developmental characteristics (e.g. branching, flowering time, fruiting time) and responses to stress-producing environmental conditions (e.g. drought, cold, heat, salt). I will present specific examples to illustrate how QTL mapping is being utilized in the Rosaceae.

1.6.1 Apple

Of the economically important Rosaceae crop species, QTL studies have been conducted most often in apple. A number of segregating populations have been employed to detect a variety of different trait/locus associations. Resistance to the causative bacterial pathogen (*Erwinia amylovora*) for fire blight has received much attention from researchers. Major loci associated with resistance to this disease, which also affects pear, have been mapped to the top of LG3 in three different apple populations derived from 'Robusta 5' of 133, 132, and 83 F_1 progeny (Gardiner et al. 2012), to the bottom of LG7 in 'Fiesta' (Khan et al. 2007), and to the bottom of LG12 in two interspecific apple populations (Durel et al. 2009). Several of these markers are currently being used for marker-assisted selection by apple breeders.

The production and accumulation of polyphenolics in apple fruit has been mapped to several loci, with loci on LG16 involved in the production of epicatechins, catechins, and procyanidins in a population of 170 'Royal Gala' × 'Braeburn' F_1 progeny (Chagné et al. 2012) and a population of 251 'Prima' × 'Fiesta' F_1 progeny and 171 'Milwa' × '1217' F_1 progeny (Khan et al. 2012). These two studies independently identified *leucoanthocyanidin reductase* (*LAR1*) as a candidate gene for the control of these traits. Additional QTL for other polyphenolics were identified on other LG, with a point of difference relating to the production of quercetins. Khan et al. (2012) identified loci on LG1, 8 and 13, whereas Chagné et al. (2012) identified a QTL on LG17.

Flower development is an important trait especially in crops where fruit mature simultaneously. Biennial bearing in apple, where number of flowers and subsequent fruit yield varies from year to year, was investigated over a five year span in a population of 122 'Starkrimson' × 'Granny Smith' F_1 progeny replicated twice (Guitton et al. 2012). Three large-effect QTL for biennial bearing were identified on LG4, 8 and 10. The candidate gene *MdEFL3* (*Early Flowering 3*) mapped within the QTL interval on LG8 and genome alignment with the 'Golden Delicious' draft whole genome sequence (Velasco et al. 2010) identified five additional candidate genes involved in plant hormone synthesis or control. Two candidate genes involved in plant hormone synthesis or control were mapped within the QTL interval on LG10. No candidate genes were identified associated with the QTL on LG4.

1.6.2 Strawberry

The most comprehensive QTL analysis in octoploid commercial strawberry to date combines the analyses of agronomic and fruit quality traits (Zorrilla-Fontanesi et al. 2011a). This study was conducted using a population of 95 F_1 progeny of '232' × '1392', which segregated for a number of important traits, such as yield, fruit firmness, and quantities of acids, sugars, anthocyanins (ACY), and L-ascorbic acid. The researchers identified several candidate genes underlying or within the confidence intervals of some QTL. For example, a putative R2R3 MYB transcription factor was identified underlying a QTL for total anthoycanin production on LGII. This class of MYB transcription factor encodes a component of the MYB/bHLH/WD protein complex involved in the regulation of the expression of ACY biosynthetic pathway genes and the subsequent accumulation of ACY in plant tissues (Allan et al. 2008; Baudry et al. 2004). Another locus/candidate gene association was detected for fruit firmness and its co-location with a gene encoding a fruit-specific auxin-independent expansin.

1.6.3 *Prunus*

A population of 73 F_1 progeny of interspecific *Prunus persica* 'Summergrand' × *P. davidiana* was used to map QTL for resistance of this peach cross to the green peach aphid [*Myrus persicae* (Sulzer)], as well as aphid feeding behavior (Sauge et al. 2012). QTL for resistance and feeding behavior occurred in close proximity in seven of nine instances, with the strongest effect QTL for both traits occurring within approximately 2 cM on LG3. The co-occurrence of these traits suggest a strong genetic basis for the relationship (Sauge et al. 2012).

Self-incompatibility (SI) in *Prunus* has been considered a qualitative trait controlled by the *S*-locus (Crane and Lawrence 1929) and which as been mapped to LG6 (Ballester et al. 1998). Those genotypes that are self-compatible possess an inactive form (S_{fi}). However, a recent study in 77 F_1 progeny of almond 'Vivot' × 'Blanquerna' has identified two additional loci, one on LG6 approximately 10 cM from the *S*-locus, and one on LG8, which may modulate the *S*-locus (Fernández i Martí et al. 2011).

A recent QTL study was conducted to determine how chilling and heat units influence bloom date in peach (Fan et al. 2010). The goal of this study was to try to understand the genetic control of dormancy break in temperate fruit tree species and attempt to

identify loci that can be used for marker assisted selection. The researchers used a population of 378 F_2 peach progeny and conducted the study over two years. They identified 20 QTL on all *Prunus* LG except LG3. The strongest-effect QTL encompassing all three traits was mapped on LG1 in an interval that includes the *Evergrowing* locus. In most cases the three traits collocated. The authors suggest that the genes regulating the different traits are tightly linked, or alternatively, the presence of a single temperature sensing and response system.

Each of the examples presented above illustrate some of the different complex traits that can be examined using QTL analysis. Population sizes range from 73 F_1 individuals to 378 F_2 individuals, and in each example small effect QTL are not accurately mapped. However, the studies identify the major genomic regions that are important for the expression of specific traits. The recent increase in genomic information available for apple (Velasco et al. 2010), peach (Sosinski et al. 2010), and strawberry (Shulaev et al. 2011) provides the opportunity to use gene mining for the identification of the putative controlling genes.

I will now present the status of comparative genome mapping in the family, including some of the factors that limit progress in genetic mapping across the Rosaceae.

1.7 Comparative genome mapping in the Rosaceae

A few genetic map comparisons within and between Rosaceae genera have been performed using SSR markers. The reference linkage map for T×E has been used to develop and align linkage maps of other *Prunus* species. Lambert et al. (2006) constructed a linkage map of apricot using the T×E genetic map as a source of markers. Their results revealed three order inversions of markers mapped to apricot LG 1 and 2, while colinearity is otherwise maintained among apricot, peach and almond. The sweet cherry genetic map constructed by Olmstead et al. (2008), was also compared with the T×E linkage map and supports colinearity between the two species (Fig. 1.19); however, the researchers had difficulty with the transferability of markers, which will be discussed in section "1.7.2 *Primer pair transferability within subtribes*".

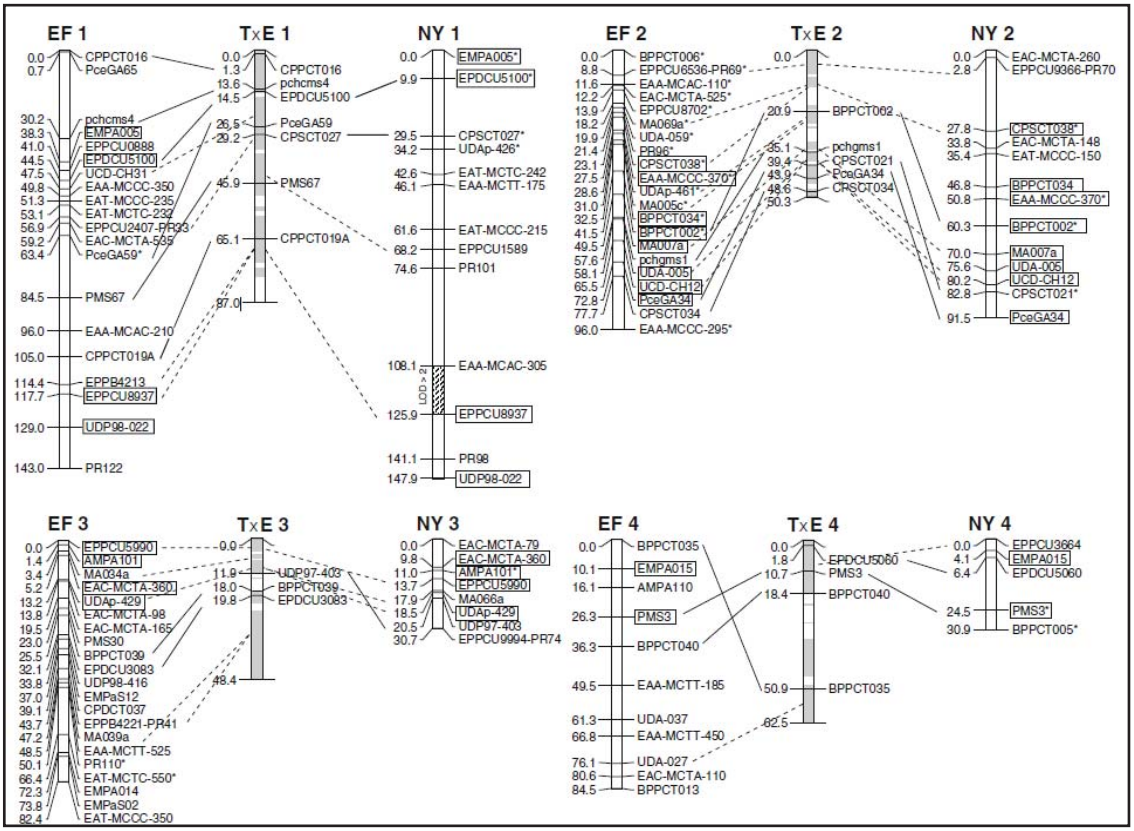


Fig. 1.19: Alignment of 'Emperor Francis' (EF) and 'New York 54 (NY) sweet cherry parental maps with the *Prunus* reference bin map ['Texas' almond × 'Earlygold' peach (T×E)]. Only markers in common between T×E and both EF and NY are shown on the T×E map. Solid connecting lines represent homology between the mapped markers; dashed connecting lines represent markers common between T×E and one sweet cherry parent. Shaded regions on the T×E linkage groups delimit the mapping bins. Boxed markers are anchors between EF and NY. Asterisks indicate deviation (P<0.05) from the expected chi-square segregation value. This figure illustrates the overall colinearity found in the genus *Prunus* as well as some rearrangements that may be unique to sweet cherry. Figure copied from Olmstead et al. (2008) Figure 1.

Within subtribe Pyrinae, markers and maps developed from apple have been used to construct and align linkage maps in pear (Yamamoto et al. 2004a; Pierantoni et al. 2004). Alignments of genetic maps of apple with maps constructed from two *Pyrus communis* (European pear) populations (Pierantoni et al. 2004), and apple with maps constructed from *P. pyrifolia* (Japanese pear) and *P. communis* (Celton et al. 2009a; Yamamoto et al. 2004a), revealed high marker transferability and genomic colinearity among the species, while also detecting differences in marker order over short genetic distances. Linkage maps of *Eriobotrya japonica* (loquat) were constructed using markers derived from *Malus*, *Pyrus*, *Prunus* and *Eriobotyra* (Gisbert et al. 2009) (Fig. 1.20). Markers developed from *Malus* and *Pyrus* had higher rates of transferability than

markers developed from *Prunus* and this will be discussed in section "1.7.2 *Primer pair transferability within subtribes*".

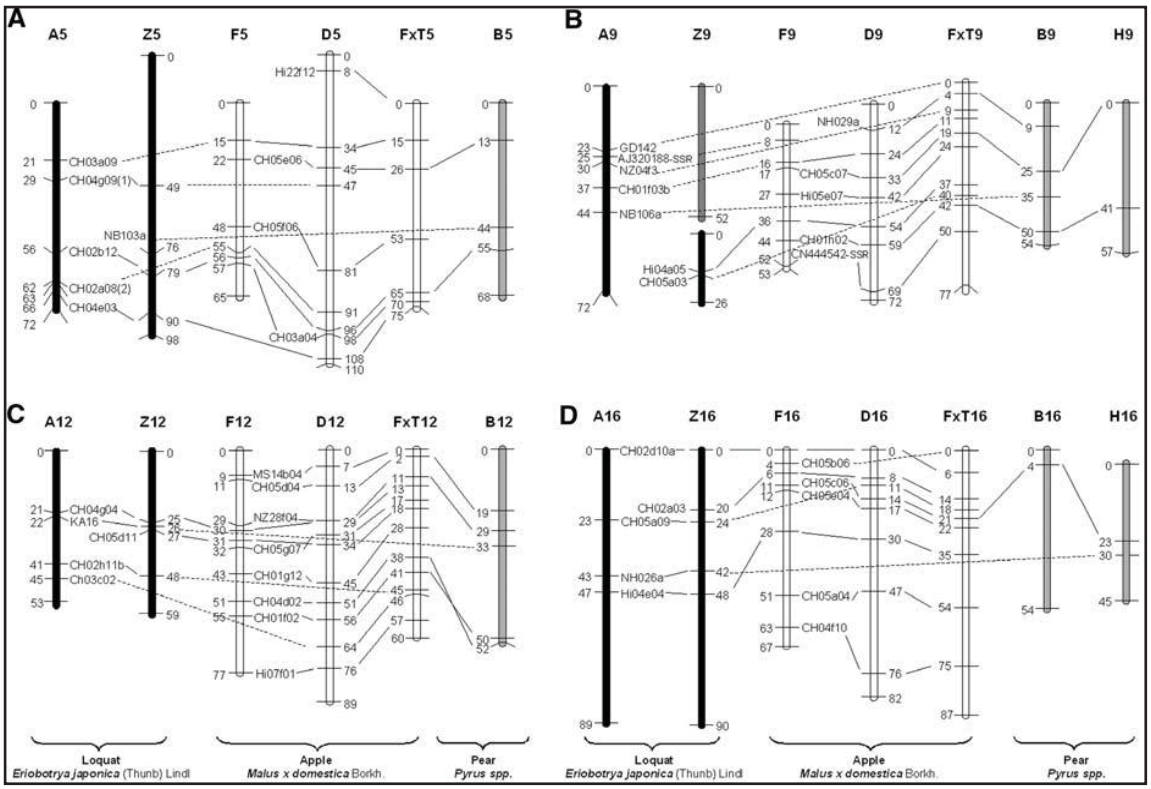


Fig. 1.20: Alignment of *Eriobotyra, Malus* and *Pyrus* linkage groups (LG) sharing at least five linked SSR from different SSR-based maps. (A) LG5; (B) LG9; (C) LG12; (D) LG16; Apple maps represented by: F: 'Fiesta'; D: 'Discovery'; F×T: 'Fiesta' × 'Totem'; Pear maps represented by B: 'Bartlett'; H: 'Housui'; Loquat represented by A: 'Algerie' and Z: 'Zaozhong-6'. This figure illustrates the overall colinearity among the genera, with some rearrangements that may be specific to loquat. Figure copied from Gisbert et al. (2009) Figure 2.

Three wider phylogenetic comparisons have been conducted in the family. The first was a partial comparative analysis between linkage maps of *Malus* × *domestica* ($2n = 2x = 34$) and *Prunus* T×E ($2n = 2x = 16$) aligned by 24 RFLP and 6 isozyme markers, which suggested genome homology (Dirlewanger et al. 2004b). However, a current lack of transferable PCR-based markers has prevented the performance of a comprehensive comparison between the two genera. A comparison between maps of *Prunus* T×E and *Fragaria vesca* (alpine strawberry; $2n = 2x = 14$) using RFLP, EST, and only two SSR markers was used to reconstruct a hypothetical ancestral genome and demonstrated that

there is homology between the eight LG of *Prunus* and the seven LG of *Fragaria*, and suggests that 27 inversions and nine fusion/fission or translocation events have occurred between the two genera (Vilanova et al. 2008). Most recently, a study was conducted that combined genetic mapping of orthologous markers on the T×E and FV×FB bin mapping sets with the bioinformatic location of the marker sequences on the 'Golden Delicious' draft genome (Illa et al. 2011). A total of 129 markers in common among the three genera allowed the alignment of syntenic blocks, which were used to model a hypothetical ancestral genome with nine chromosomes (Illa et al. 2011).

The examples presented above demonstrate that comparative mapping analyses within the genus *Prunus* and within the subtribe Pyrinae have been most successful, with SSR primer pairs being highly transferable within these two clades. However, wider phylogenetic comparisons outside these groups are lacking, as the transferability of SSR primer pairs declines with phylogenetic distance, which requires the design of taxon-specific primer pairs. These limitations are being overcome with the release of whole genome sequences, the development of PCR-based orthologous markers, and the use of existing linkage maps. A review of primer pair transferability among different taxonomic levels of the family follows.

1.7.1 Primer pair transferability between subtribes

Primer pair transferability within the Rosaceae has shown varying degrees of success depending on the phylogenetic relationship and type of sequence used to design the primer pair. It is important to note that the amplification of a PCR product does not necessarily make the marker useful for genetic mapping. A study of apple-derived EST-SSR primer pair transferability by Gasic and colleagues (2009b) illustrates this caveat. The researchers selected apple-derived EST-SSR markers to explore the transferability of primer pairs across 14 species in three Rosaceae genera from three different subtribes (Pyrinae, Amygdaleae, and Fragariinae) (Fig. 1.17). The primer pairs designed to amplify SSR were screened over the species, resulting in an amplification success rate of 75%. However, depending on the marker, 25% to 59% of the amplicons were monomorphic.

The gel shown in Fig. 1.21 illustrates the difference between successful amplification and utility of a marker for mapping. Of the six apple-derived EST-SSR markers shown, only CN918509 and CN857658 are polymorphic in a single representative of the eight

apple cultivars tested. These two markers have the potential to be useful for mapping in apple, yet the remaining markers are all monomorphic in apple and are not useful for mapping. Additionally, when marker CN918509 was screened over individuals from the other rosaceous species, amplification was successful only in pear (one allele amplified in five accessions) and strawberry (one allele in one accession), and failed to amplify a product in the other 12 species. Marker CN918509 was not determined by the researchers to have applicability across species because of its low intergeneric amplification success.

One marker that the researchers determined to be widely useful, CO414802, amplified two alleles in apple, and while amplifying a fragment of expected size in most of the tested species, appears monomorphic (Fig. 1.22); therefore its applicability for mapping is limited, although this application was not addressed by the authors. The marker may be monomorphic in the other genera because of a disruption in the SSR motif, perhaps caused by the presence of a SNP, which has led to the loss of variability at this locus. The gel-based method of fragment visualization used by the authors is not precise enough to detect differences at the SNP level. However, the conservation of primer sites and amplicon size of this marker do make it a good candidate for sequencing for SNP identification and subsequent marker development.

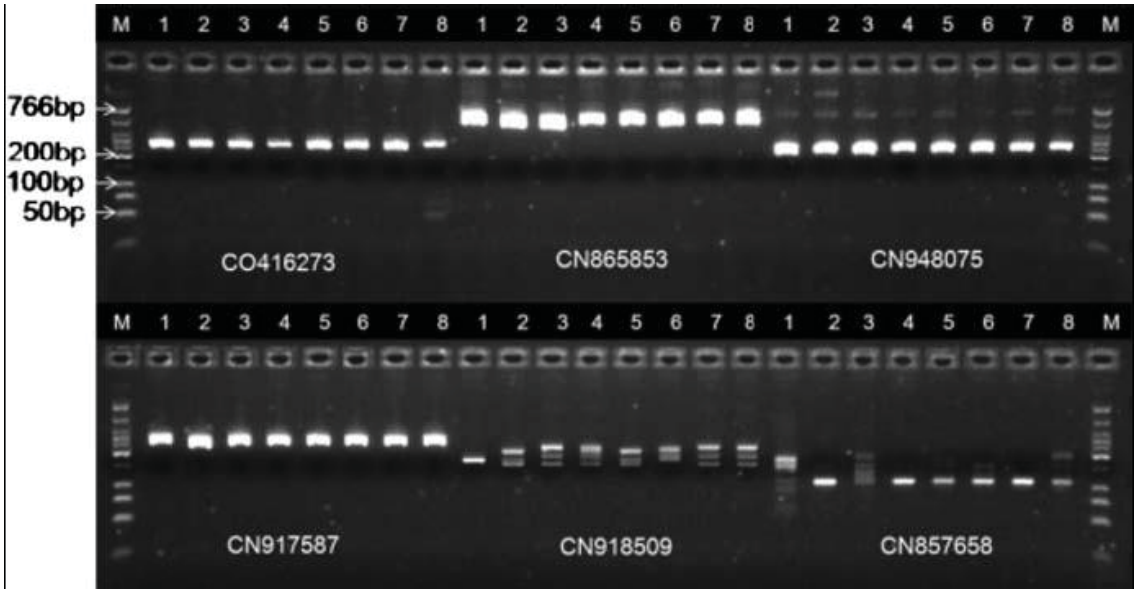


Fig. 1.21: Detection of polymorphism by six apple-derived expressed sequence tag-simple sequence repeat (EST-SSR) markers across eight apple cultivars (lanes 1-8 for each marker). Only markers CN918509 and CN857658 exhibit polymorphism in apple; all other markers appear monomorphic. Gel photograph copied from Gasic et al. (2009b) Figure 1.

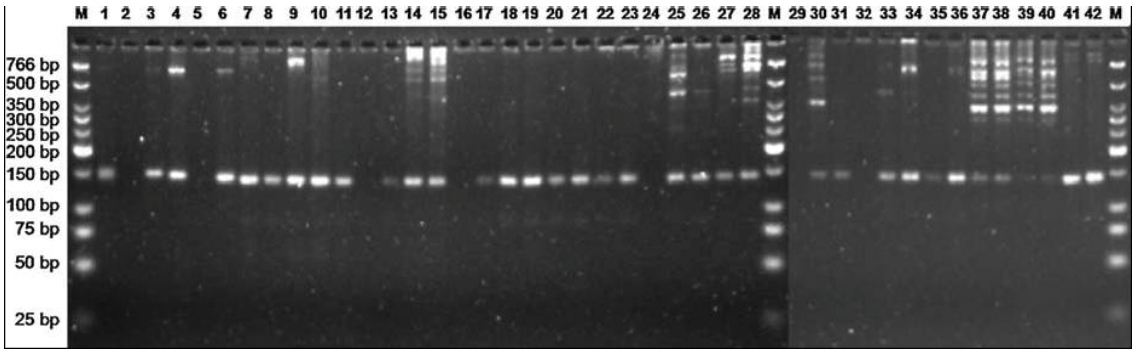


Fig. 1.22: Gel demonstrating the amplification of apple-derived expressed sequence tag-simple sequence repeat (EST-SSR) marker CO414802 across different Rosaceae species. Repeat type $(GGA)_7$; size predicted from apple EST database is 142 base pairs (bp); M, 1kb DNA standard; *lanes 1-6* pear; *7-9* rose; *10-17* strawberry; *18-21* apricot; *22-26* European plum; *27-31* almond; *32-36* Japanese plum; *37-40* peach; and *41-42* apple. This gel demonstrates that this marker, while able to amplify a band in many species, is uninformative for mapping. Gel photograph copied from Gasic et al. (2009b) Figure 2.

In another example, Dirlewanger et al. (2002) evaluated genomic SSR (gSSR) markers developed from the peach 'Merrill O'Henry' on other *Prunus* species, apple and strawberry. They found PCR product amplification within other *Prunus* species to be very high (80-95%), with 56% polymorphism detected in sweet cherry (the only species discussed). Amplification success was lower in *Malus* (19.5%) and *Fragaria* (29%); the degree of polymorphism detected in these two genera is not addressed by the authors.

The transferability of primer pairs designed from peach, apple and strawberry gSSR were evaluated during the construction of a *Rosa* linkage map (Hibrand-Saint Oyant et al. 2008). In this study, 17.5% of primer pairs from peach, and 4% of primer pairs from apple amplified a polymorphic product in *Rosa*; whereas 40% of the primer pairs from strawberry gSSR were successful in amplifying a polymorphic product. *Rosa* and *Fragaria* are the most closely related of the horticultural crops in subfamily Rosoideae (Fig. 1.17).

In the study of transferability of EST-SSR conducted by Decroocq et al. (2003), the necessity of polymorphism detection as a true measure of marker usefulness is addressed. cDNA libraries enriched for SSR from grape (*Vitis vinifera*) roots and apricot (*Prunus armeniaca*) leaves were constructed; then the ability of primer pairs to detect fragment length polymorphism and variation in the flanking regions, assessed by the efficiency of primer annealing, was evaluated. Within Vitaceae, the flanking regions of all eight SSR were highly conserved, with high degrees of length polymorphism detected within the genus *Vitis*, with less polymorphism and more amplification failures detected outside the genus, although many primer pairs were still successful. Within the Rosaceae tribes studied (Amygdaleae and Pyreae), only one EST-SSR marker, out of the ten markers designed, successfully detected polymorphism in all four genera tested (*Prunus*, *Pyrus*, *Malus*, and *Cydonia*), with the remaining nine primer pairs either failing to amplify a product or producing uninterpretable results. Transferability was most successful within the subfamily *Prunoideae* [Potter and colleagues' (2007) tribe Amygdaleae], with transferability decreasing as phylogenetic distance increased. The authors conclude that this change of transferability occurred over a shorter phylogenetic distance in Rosaceae than was seen in Vitaceae (Decroocq et al. 2003).

Lewers et al. (2005) conducted an experiment to assess the functional taxonomic distance for transferability of primer pairs designed to amplify SSR in *Fragaria* and

Rubus. In brief, using BLAST, they searched GenBank for genomic and EST sequences containing CT repeats that were found in *A. thaliana*, the most distantly related tested taxon from *Fragaria* and *Rubus*; *Glycine max* (soybean), *Medicago truncatula* (a model legume), and *Gossypium* sp. (cotton) to represent the next taxonomic level; *Prunus persica*, *Pyrus communis*, and *Malus* × *domestica* to represent the sister clade to *Fragaria* and *Rubus*; and *Fragaria* × *ananassa* and *Rosa* to represent the closest taxonomic level. Primer pairs to amplify the SSR were designed for all nine species. In addition, they searched the *Fragaria* GenBank genomic and EST sequences for all possible di- and tri-nucleotide combinations and designed primer pairs to amplify the SSR. Furthermore, they developed gSSR from *F.* × *ananassa* 'Earliglow' and designed primer pairs to amplify the SSR. All primer pairs were evaluated on genomic DNA of diploid and octoploid *Fragaria* species, and diploid and tetraploid *Rubus* species.

The results of this intricate experiment suggest that taxonomic distance is an important consideration for transferability of primer pairs to amplify SSR in *Fragaria* and *Rubus.* None of the primer pairs designed from the seven species with wide phylogenetic relationships amplified a product in either of the target genera. Of the primer pairs designed from the GenBank search and from published sources, 27% and 19% of the *Fragaria*-derived primer pairs amplified a product in tetraploid blackberry and/or diploid raspberry, respectively; of the *Rubus alceifolius*-derived primer pairs (the only source evaluated) none amplified a product in *Fragaria* and only two amplified a product in the target *Rubus* species; of the *Rosa*-derived primer pairs, none amplified a product in either *Fragaria* or *Rubus*.

In addition to the above evaluations on cross-clade amplification success, the authors also assessed the ability to detect polymorphisms of primer pairs designed from GenBank-derived *Fragaria*-specific EST-SSR to be 32% in blackberry, and 20% in raspberry. They also found the ability of primer pairs designed from GenBank-derived *Fragaria*-specific gSSR to detect polymorphisms was lower, at 26% in blackberry, and 18% in raspberry. Overall, the authors conclude that the ability of primer pairs designed to amplify SSR in Rosoideae is genus-specific, with very little cross-over. These findings are in agreement with the work by Decroocq (2003).

1.7.2 Primer pair transferability within subtribes

While the transferability of SSR primer pairs between subfamilies and subtribes has relatively low success, intergeneric amplification of SSR within Pyrinae is quite successful, especially between *Malus* and *Pyrus* (Yamamoto et al. 2004a; Yamamoto et al. 2004b). For example, 66-72% of primer pairs designed from apple gSSR were useful for genetic mapping in European pear, depending on the target population (Pierantoni et al. 2004). Celton et al. (2009b) also used primer pairs designed from apple- and pear-derived SSR to construct a linkage map of apple rootstocks. They found that of the 117 primer pairs designed from pear gSSR, 42% were useful for genetic mapping in apple (Celton et al. 2009b).

As briefly described in section "1.7 *Comparative genome mapping in the Rosaceae*", primer pairs for SSR from four genera, *Eriobotrya*, *Malus*, *Prunus*, and *Pyrus*, were used along with AFLP markers to produce a linkage map of *Eriobotrya japonica* (loquat) that was then compared with linkage maps of *Malus* and *Pyrus* (Gisbert et al. 2009). The polymorphic markers were distributed over 17 linkage groups corresponding to the 17 chromosomes of loquat with coverage of 900 cM for mapping population parents 'Algerie' and 870 cM for 'Zaozhong-6'. Markers developed from loquat were most efficient (81%) at detecting polymorphisms (Table 1-8). Of the other species used for marker development, apple and pear were about equal (27% and 22%, respectively) in the ability to detect polymorphisms in loquat, although they were much less efficient that the loquat-derived markers. Markers from *Prunus* had the lowest rate of success, with an overall average of 9% (Gisbert et al. 2009).

Table 1-8: An example of cross-generic simple sequence repeat (SSR) primer pair amplification and detection of a polymorphic product during construction of an *Eriobotrya japonica* (loquat) linkage map. Primer pairs developed from *Malus* (apple), *Pyrus* (pear) and *Prunus* spp. were assessed.

Source	Total SSR tested	Number amplified	Number polymorphic	Percent polymorphic of total tested
Malus (apple)	249	184	68	27%
Pyrus (pear)	36	21	8	22%
Eriobotrya (loquat)	21	21	17	81%
Prunus (peach)	96	70	9	9%
Prunus (apricot)	21	11	2	10%
Prunus (plum)	6	4	1	17%
Prunus (almond)	5	1	0	0%
Prunus (sweet cherry)	6	1	0	0%
Total *Prunus*	134	87	12	9%

Within the genus *Prunus*, the transfer of primer pairs and their amplification of polymorphic products are species- and marker-type dependent. As presented in section "1.7 *Comparative genome mapping in the Rosaceae*", during the construction of a linkage map for sweet cherry (*Prunus avium*), 433 published primer pairs designed from *Prunus*-derived gSSR or EST sequences were evaluated on the parents of the mapping population 'Emperor Francis' and 'New York 54' (Olmstead et al. 2008). The authors found SSR primer pairs derived from other sweet cherry sources had the highest amplification success and detection of polymorphism (87% for primer pairs designed from gDNA and 89% for primer pairs designed from EST), and only 26% of primer pairs from other *Prunus* sources (mainly peach) amplified a polymorphic product. The authors attribute the poor mapping results to the low heterozygosity of the mapping parents and the phylogenetic distance between cherry and peach (Olmstead et al. 2008). Cherries and laurel-cherries belong to the clade *Cerasus*/*Laurocerasus*/*Padua* (*Prunus* c/l/p), while peach, almond, plum and apricot belong to the *Amygdalus*/*Armeniaca*/*Prunocerasus* (*Prunus* a/a/p) based on chloroplast DNA analysis (Badenes and Parfitt 1995; Shaw and Small 2004) (Fig. 1.17).

Limits in the transferability of primer pairs designed from SSR are also demonstrated by a recent study on the transferability of *Fragaria*-derived primer pairs to *Rubus* and *Rosa*, three members of subfamily Rosoideae, yet of relatively wide phylogenetic relationship within the clade (Fig. 1.17). Zorrilla-Fontanesi (2011b) reported that 20% of primer pairs developed from *Fragaria*-specific EST-SSR successfully amplified a polymorphic product in *Rubus*, and 29% in *Rosa*. In the same study, 16% of primer pairs developed from *Fragaria*-specific gSSR successfully amplified a polymorphic product in *Rubus*, and 19% in *Rosa*. In general, and in agreement with Lewers et al. (2005), these results indicate that the transferability of primer pairs for EST-SSR markers is higher than that of primer pairs for gSSR markers, though neither type is especially efficient for inter-generic comparisons in Rosoideae.

1.7.3 Recent advances in genomics and technology for the Rosaceae

The recent release of draft genome sequences for three Rosaceae genera ['Golden Delicious' apple (*Malus* × *domestica* Borkh.) (Velasco et al. 2010), peach doubled haploid 'Lovell' (*Prunus persica*) (Sosinski et al. 2010), and woodland strawberry

'Hawaii 4' (*F. vesca* ssp. *vesca*) (Shulaev et al. 2011)], along with the availability of reference maps for *Prunus* and *Fragaria* as well as RosCOS markers and EST databases, have provided tools that have allowed the development of genus-specific markers to identify regions of homology. Additionally, the sequencing of the red raspberry genome is underway (J. Udall, pers. comm.) and SNP chips for apple and pear (Chagné et al. in press), and peach (Martinez Garcia et al. 2012; Micheletti et al. 2012) have been developed. These new developments add to the molecular toolbox available to Rosaceae researchers by providing searchable databases for genome alignment and primer design, and high-throughput genotyping of Pyrinae genera and *Prunus* species using the new SNP chips.

1.7.4 Summary

Primer pair transferability among the Rosaceae shows varying degrees of success depending on sequence source and the phylogenetic relationship of targets compared with the source. In some cases, as with the Pyrinae, transferability, especially between apple and pear, seems straightforward, with high degrees of success in both amplification and detection of polymorphism in the target species. The comparative genome mapping in this clade supports strong conservation of colinearity. Colinearity is also maintained in the genus *Prunus*, yet marker transferability success varies between the c/l/p and a/a/p clades, although, the decline in transferability may be because of low genetic diversity found in the sweet cherry parents used for map construction and comparative genome analysis. The success of transferability and detection of polymorphism within Rosoidae also seems to be dependent on phylogenetic distance, as primer pairs developed in *Fragaria* seem to have greater transferability to *Rosa* than to *Rubus*, although few studies of transferability have been conducted. In general, the data suggest a short phylogenetic distance is necessary for SSR primer pair transferability in the Rosaceae.

To date, only three inter-tribal genome comparisons have been conducted in the family as insufficient numbers of transferable PCR-based primer pairs have hindered comparisons outside *Prunus* and Pyrinae. The examples presented demonstrate the limited potential of primer pairs derived from a single genus to amplify a polymorphic product within the wider Rosaceae and emphasize the need to develop transferable markers that can be used for comparative mapping studies.

1.8 THESIS AIM AND OBJECTIVES

In the Introduction, I have demonstrated that while genetic linkage mapping in Rosaceae genera has proceeded relatively rapidly, only in the last several years has molecular and genomic information for different Rosaceae genera developed to a stage that allows in-depth comparative analyses. A lack of reliably transferable primer pairs has constrained wide phylogenetic, genome-level comparisons; therefore, the construction of genetic linkage maps populated with orthologous markers is needed for genetic map alignment. Comparative genome mapping studies conducted by aligning the genetic maps or genome sequences of multiple genera within the Rosaceae will identify genomic regions that are conserved throughout the family, will shed light on the genomic regions that are unique to each genus, and might uncover lineage-specific segmental duplications, deletions, or repetitive regions. Additionally, comparative genome mapping studies that align genetic maps or genome sequences of Rosaceae genera with other members of the order Rosidae will provide insight into the evolution of the core Eudicots (APG 2003). Eventually, wide-scale comparisons will allow estimates of rates and types of selection on different regions of each genome. Of immediate importance, both genetic and comparative genome mapping assist in the assembly of genome sequences, and can be used to identify regions of coding DNA, as well as predict the function of those regions.

The difficulties encountered with the lack of primer pair transferability have led me to re-evaluate the importance of phylogenetic distance and sequence source during primer design. For this reason, in this study I have evaluated the utility of RosCOS and the sequences of genes of known function as sources of orthologous sequences for primer pair design. The primer pairs were evaluated for amplification efficiency and polymorphism detection within apple, strawberry and red raspberry, with the goals of constructing linkage maps for homolog identification, genome alignment, and comparative genome analysis.

The aim of this thesis was to develop orthologous, transferable molecular markers in order to align genetic maps of different Rosaceae genera to provide insight into lineage-specific differences and the evolutionary history of the family. To accomplish this aim, I 1) designed and assessed new types of molecular markers and technologies for genotyping; 2) evaluated the effectiveness and efficiency of these new tools by adding markers to existing genetic maps and comparing widely divergent members of the

Rosaceae; 3) constructed a new linkage map of *Rubus* using these tools; and 4) conducted a QTL analysis on anthocyanin production in *Rubus* and predicted the locations of orthologous loci on the existing *Fragaria vesca* linkage map.

REFERENCES

Acarkan A, Roßberg M, Koch M, Schmidt R (2000) Comparative genome analysis reveals extensive conservation of genome organisation for *Arabidopsis thaliana* and *Capsella rubella*. Plant Journal 23 (1):55-62

Adams KL, Wendel JF (2005) Novel patterns of gene expression in polyploid plants. Trends in Genetics 21 (10):539-543

Aharoni A, O'Connell AP (2002) Gene expression analysis of strawberry achene and receptacle maturation using DNA microarrays. Journal of Experimental Botany 53 (377):2073-2087. doi:10.1093/jxb/erf026

Ahn S, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. Proceedings of the National Academy of Sciences of the United States of America 90 (17):7980-7984

Allan AC, Hellens RP, Laing WA (2008) MYB transcription factors that colour our fruit. Trends in Plant Science 13 (3):99-102

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215 (3):403-410

Amsellem L, Dutech C, Billotte N (2001) Isolation and characterization of polymorphic microsatellite loci in *Rubus alceifolius* Poir. (Rosaceae), an invasive weed in La Réunion island. Molecular Ecology Notes 1 (1-2):33-35. doi:10.1046/j.1471-8278.2000.00013.x

Andersen JR, Lübberstedt T (2003) Functional markers in plants. Trends in Plant Science 8 (11):554-560

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36 (1):181-186. doi:10.1139/g93-024

Andolfatto P, Davison D, Erezyilmaz D, Hu TT, Mast J, Sunayama-Morita T, Stern DL (2011) Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Research 21 (4):610-617. doi:10.1101/gr.115402.110

APG (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Botanical Journal of the Linnean Society 141 (4):399-436. doi:10.1046/j.1095-8339.2003.t01-1-00158.x

Aranzana M, Carbó J, Arús P (2003a) Microsatellite variability in peach [*Prunus persica* (L.) Batsch]: cultivar identification, marker mutation, pedigree inferences and population structure. Theoretical and Applied Genetics 106 (8):1341-1352

Aranzana MJ, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder C, Testolin R, Abbott A, King G, Iezzoni A, Arús P (2003b) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. Theoretical and Applied Genetics 106 (5):819-825

Arús P, Gardiner S (2007) Genomics for improvement of Rosaceae temperate tree fruit. In: Varshney RK, Tuberosa R (eds) Genomics-Assisted Crop Improvement: Genomics Applications in Crops, vol 2. Springer, Dordrecht, The Netherlands,

Badenes ML, Parfitt DE (1995) Phylogenetic relationships of cultivated *Prunus* species from an analysis of chloroplast DNA variation. Theoretical and Applied Genetics 90 (7):1035-1041

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3 (10):e3376. doi:10.1371/journal.pone.0003376

Baldi P, Patocchi A, Zini E, Toller C, Velasco R, Komjanc M (2004) Cloning and linkage mapping of resistance gene homologues in apple. Theoretical and Applied Genetics 109 (1):231-239

Ballester J, Bošković R, Batlle I, Arús P, Vargas F, de Vicente MC (1998) Location of the self-incompatibility gene on the almond linkage map. Plant Breeding 117 (1):69-72. doi:10.1111/j.1439-0523.1998.tb01450.x

Ballester J, Socias i Company R, Arús P, De Vicente MC (2001) Genetic mapping of a major gene delaying blooming time in almond. Plant Breeding 120 (3):268-270. doi:10.1046/j.1439-0523.2001.00604.x

Bassil NV, Gunn M, Folta K, Lewers K (2006) Microsatellite markers for *Fragaria* from 'Strawberry Festival' expressed sequence tags. Molecular Ecology Notes 6 (2):473-476. doi:10.1111/j.1471-8286.2006.01278.x

Baudry A, Heim M, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L (2004) TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant Journal 39 (3):366 - 380

Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L (2006) Connected populations for detecting quantitative trait loci and testing for epistasis: An application in maize Theoretical and Applied Genetics 113 (2):206-224

Bolot S, Abrouk M, Masood-Quraishi U, Stein N, Messing J, Feuillet C, Salse J (2009) The 'inner circle' of the cereal genomes. Curr Opin Plant Biol 12 (2):119-125

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120 (4):1095-1103

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in Man using restriction fragment length polymorphisms. American Journal of Human Genetics 32 (3):314-331

Bus V, Bassett H, Bowatte D, Chagné D, Ranatunga C, Ulluwishewa D, Wiedow C, Gardiner S (2010) Genome mapping of an apple scab, a powdery mildew and a woolly apple aphid resistance gene from open-pollinated Mildew Immune Selection. Tree Genetics & Genomes 6 (3):477-487

Bus V, White A, Gardiner S, Weskett R, Ranatunga C, Samy A, Cook M, Rikkerink E (2002) An update on apple scab resistance breeding in New Zealand. Acta Horticulturae (ISHS) 595:43-47

Byrne D, Jelenkovic G (1976) Cytological diploidization in the cultivated octoploid strawberry *Fragaria* x *ananassa*. Canadian Journal of Genetics and Cytology 18 (4):653-659

Cabrera A, Kozik A, Howad W, Arús P, Iezzoni A, van der Knaap E (2009) Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. BMC Genomics 10 (1):562

Cabrera A, Rosyara U, De Franceschi P, Sebolt A, Sooriyapathirana S, Dirlewanger E, Quero-Garcia J, Schuster M, Iezzoni A, van der Knaap E (2011) Rosaceae conserved orthologous sequences marker polymorphism in sweet cherry germplasm and construction of a SNP-based map. Tree Genetics & Genomes:1-11. doi:10.1007/s11295-011-0436-9

Campbell CS, Evans RC, Morgan DR, Dickinson TA, Arsenault MP (2007) Phylogeny of subtribe Pyrinae (formerly the Maloideae, Rosaceae): Limited resolution of a complex evolutionary history. Plant Systematics and Evolution 266 (1):119-145

Cantini C, Iezzoni AF, Lamboy WF, Boritzki M, Struss D (2001) DNA fingerprinting of tetraploid cherry germplasm using simple sequence repeats. Journal of the American Society for Horticultural Science 126 (2):205-209

Cao K, Wang LR, Zhu GR, Fang WC, Chen CW, Zhao P (2011) Construction of a linkage map and identification of resistance gene analog markers for root-knot nematodes in wild peach, *Prunus kansuensis*. Journal of the American Society for Horticultural Science 136 (3):190-197

Castle JC (2011) SNPs occur in regions with less genomic sequence conservation. Public Library of Science One 6 (6):e20660. doi:10.1371/journal.pone.0020660

Cato SA, Gardner RC, Kent J, Richardson TE (2001) A rapid PCR-based method for genetically mapping ESTs Theoretical and Applied Genetics 102 (2-3):296-306

Celton J-M, Chagné D, Tustin SD, Terakami S, Nishitani C, Yamamoto T, Gardiner SE (2009a) Update on comparative genome mapping between *Malus* and *Pyrus*. BMC Research Notes 2 (182)

Celton J-M, Tustin DS, Chagné D, Gardiner SE (2009b) Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. Tree Genetics & Genomes 5:93-107

Chagné D, Brown G, Lalanne C, Madur D, Pot D, Neale D, Plomion C (2003) Comparative genome and QTL mapping between maritime and loblolly pines. Molecular Breeding 12 (3):185-195

Chagné D, Crowhurst RN, Troggio M, Davey MW, Gilmore B, Lawley C, Vanderzande S, Hellens RP, Kumar S, Cestaro A, Velasco R, Main D, Rees J, Iezzoni A, Mockler TC, Wilhelm LJ, Van de Weg E, Gardiner SE, Bassil NV, Peace C (in press) Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. Public Library of Science One

Chagné D, Krieger C, Rassam M, Sullivan M, Fraser J, Andre C, Pindo M, Troggio M, Gardiner S, Henry R, Allan A, McGhie T, Laing W (2012) QTL and candidate gene mapping for polyphenolic composition in apple fruit. BMC Plant Biology 12 (1):12

Chakraborty R, Weiss KM (1988) Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. Proceedings of the National Academy of Sciences 85 (23):9119-9123

Chaparro JX, Werner DJ, O'Malley D, Sederoff RR (1994) Targeted mapping and linkage analysis of morphological, isozyme, and RAPD markers in peach. Theoretical and Applied Genetics 87 (7):805-815

Chou L-S, Lyon E, Wittwer CT (2005) A comparison of High-Resolution Melting analysis with Denaturing High-Performance Liquid Chromatography for mutation scanning: Cystic fibrosis transmembrane conductance regulator gene as a model. American Journal of Clinical Pathology 124 (3):330-338

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138 (3):963-971

Cipriani G, Lot G, Huang WG, Marrazzo MT, Peterlunger E, Testolin R (1999) AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L.) Batsch]: isolation, characterisation and cross-species amplification in *Prunus*. Theoretical and Applied Genetics 99 (1):65-72

Comai L (2005) The advantages and disadvantages of being polyploid. Nature Reviews Genetics 6:836-846

Condit R, Hubbell SP (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome 34 (1):66-71. doi:10.1139/g91-011

Costa F, Costa G, Sansavini S, Soglio V, Gianfranceschi L, Schouten HJ, Alba R, Giovannoni J (2009) Heterologous comparative genomics to identify candidate genes impacting fruit quality in apple (*Malus* x *domestica* Borkh.). Acta Horticulturae (ISHS) 814:517-522

Costantini L, Madini A, Battilana J, Grando MS (2005) Comparative mapping of quantitative traits in grape. In: Williams LE (ed) Proceedings of the Seventh International Symposium on Grapevine Physiology and Biotechnology. Acta Horticulturae, vol 689. pp 409-415

Crane MB, Lawrence WJC (1929) Genetical and cytological aspects of incompatibility and sterility in cultivated fruits. Journal of Pomology and Horticultural Science 7:276-301

Cregan PB, Bhagwat AA, Akkaya MS, Rongwen J (1994) Microsatellite fingerprinting and mapping of soybean. Methods in Molecular and Cellular Biology 5 (1):49-61

Crepieux S, Lebreton C, Flament P, Charmet G (2005) Application of a new IBD-based QTL mapping method to common wheat breeding population: Analysis of kernel hardness and dough strength Theoretical and Applied Genetics 111 (7):1409-1419

Crespel LC, Chirollet MC, Durel CD, Zhang DZ, Meynet JM, Gudin SG (2002) Mapping of qualitative and quantitative phenotypic traits in *Rosa* using AFLP markers. Theoretical and Applied Genetics 105 (8):1207-1214. doi:10.1007/s00122-002-1102-2

Crow JF (1990) Mapping Functions. Genetics 125 (4):669-671

Dale A, Moore PP, McNicol RJ, Sjulin TM, Burmistrov LA (1993) Genetic diversity of red raspberry varieties throughout the world. Journal of the American Society for Horticultural Science 118 (1):119-129

Davis TM, Yu H (1997) A linkage map of the diploid strawberry, *Fragaria vesca*. Journal of Heredity 88 (3):215-221

Debener T, Mattiesch L (1999) Construction of a genetic linkage map for roses using RAPD and AFLP markers. Theoretical and Applied Genetics 99 (5):891-899. doi:10.1007/s001220051310

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theoretical and Applied Genetics 106 (5):912-922

Devos KM, Gale MD (2000) Genome relationships: The grass model in current research. Plant Cell 12 (5):637-646. doi:10.1105/tpc.12.5.637

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

Dirlewanger E, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Voisin R, Poizat C, Lafargue B, Baron O, Laigret F, Kleinhentz M, Arús P, Esmenjaud D (2004a) Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid—location of root-knot nematode resistance genes. Theoretical and Applied Genetics 109 (4):827-838

Dirlewanger E, Cosson P, Tavaud M, Aranzana M, Poizat C, Zanetto A, Arús P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theoretical and Applied Genetics 105 (1):127-138

Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldera F, Cosson P, Howad W, Arús P (2004b) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proceedings of the National Academy of Sciences of the United States of America 101 (26):9891-9896. doi:10.1073/pnas.0307937101

Downey SL, Iezzoni AF (2000) Polymorphic DNA markers in black cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach, and sour cherry. Journal of the American Society for Horticultural Science 125 (1):76-80

Doyle JJ, Davis JI (1998) Homology in molecular phylogenetics: A parsimony perspective. In: Soltis DE, Soltis PS, Doyle JJ (eds) Molecular Systematics of Plants II: DNA Sequencing. Klewer Academic Publishers, Norwell, MA,

Doyle JJ, Gaut BS (2000) Evolution of genes and taxa: a primer. Plant Molecular Biology 42 (1):1-23. doi:10.1023/a:1006349518932

Durel CE, Denancé C, Brisset MN (2009) Two distinct major QTL for resistance to fire blight co-localize on linkage group 12 in apple genotypes 'Evereste' and *Malus floribunda* clone 821. Genome 52 (2):139-147. doi:10.1139/g08-111

Edwards KJ, Mogg R (2001) Plant genotyping by analysis of Single Nucleotide Polymorphisms. In: Henry RJ (ed) Plant Genotyping: The DNA Fingerprinting of Plants. CABI, Oxon and New York,

Ellwood S, Phan H, Jordan M, Hane J, Torres A, Avila C, Cruz-Izquierdo S, Oliver R (2008) Construction of a comparative genetic map in faba bean (*Vicia faba* L.); conservation of genome structure with *Lens culinaris*. BMC Genomics 9 (1):380

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6 (5):e19379. doi:10.1371/journal.pone.0019379

Etienne C, Rothan C, Moing A, Plomion C, Bodénès C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E (2002) Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]. Theoretical and Applied Genetics 105 (1):145-159. doi:10.1007/s00122-001-0841-9

Eujayl I, Sorrells ME, Baum M, Wolters P, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theoretical and Applied Genetics 104 (2):399-407

Evans RC, Campbell CS (2002) The origin of the apple subfamily (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicated *GBSSI* genes. American Journal of Botany 89 (9):1478-1484. doi:10.3732/ajb.89.9.1478

Fan S, Bielenberg DG, Zhebentyayeva TN, Reighard GL, Okie WR, Holland D, Abbott AG (2010) Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). New Phytologist 185 (4):917-930. doi:10.1111/j.1469-8137.2009.03119.x

Fernández-Fernández F, Evans K, Clarke J, Govan C, James C, Marič S, Tobutt K (2008) Development of an STS map of an interspecific progeny of *Malus*. Tree Genetics & Genomes 4 (3):469-479

Fernández i Martí À, Howad W, Tao R, Alonso Segura J, Arús P, Socias i Company R (2011) Identification of quantitative trait loci associated with self-compatibility in a *Prunus* species. Tree Genetics & Genomes 7 (3):629-639. doi:10.1007/s11295-010-0362-2

Feuillet C, Keller B (2002) Comparative genomics in the grass family: Molecular characterization of grass genome structure and evolution. Annals of Botany 89 (1):3-10. doi:10.1093/aob/mcf008

Fitch WM (1970) Distinguishing homologous from analogous proteins. Systematic Zoology 19 (2):99-113

Foolad MR, Arulsekar S, Becerra V, Bliss FA (1995) A genetic map of *Prunus* based on an interspecific cross between peach and almond. Theoretical and Applied Genetics 91 (2):262-269

Fulton TM, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of Conserved Ortholog Set markers for comparative genomics in higher plants. Plant Cell 14 (7):1457-1467. doi:10.1105/tpc.010479

Gardiner S, Norelli J, de Silva N, Fazio G, Peil A, Malnoy M, Horner M, Bowatte D, Carlisle C, Wiedow C, Wan Y, Bassett C, Baldo A, Celton J-M, Richter K, Aldwinckle H, Bus V (2012) Putative resistance gene markers associated with quantitative trait loci for fire blight resistance in Malus 'Robusta 5' accessions. BMC Genetics 13 (1):25

Gasic K, Gonzalez DO, Thimmapuram J, Liu L, Malnoy M, Gong G, Han Y, Vodkin LO, Aldwinckle HS, Carroll NJ, Orvis KS, Goldsbrough P, Clifton S, Pape D, Fulton L, Martin J, Theising B, Wisniewski ME, Fazio G, Feltus FA, Korban SS (2009a) Comparative analysis and functional annotation of a large expressed sequence tag collection of apple. Plant Genetics 2 (1):23-38. doi:10.3835/plantgenome2008.11.0014

Gasic K, Han Y, Kertbundit S, Shulaev V, Iezzoni A, Stover E, Bell R, Wisniewski M, Korban S (2009b) Characteristics and transferability of new apple EST-derived SSRs to other Rosaceae species. Molecular Breeding 23 (3):397-411

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998) Simple sequence repeats for the genetic analysis of apple. Theoretical and Applied Genetics 96 (8):1069-1076

Gil-Ariza DJ, Amaya I, Botella MA, Blanco JM, Caballero JL, López-Aranda JM, Valpuesta V, Sánchez-Sevilla JF (2006) EST-derived polymorphic microsatellites from cultivated strawberry (*Fragaria* × *ananassa*) are useful for diversity studies and varietal identification among *Fragaria* species. Molecular Ecology Notes 6 (4):1195-1197. doi:10.1111/j.1471-8286.2006.01489.x

Gisbert A, Martínez-Calvo J, Llácer G, Badenes M, Romero C (2009) Development of two loquat [*Eriobotrya japonica* (Thunb.) Lindl.] linkage maps based on AFLPs and SSR markers from different Rosaceae species. Molecular Breeding 23 (3):523-538

Graham J, Hackett C, Smith K, Woodhead M, MacKenzie K, Tierney I, Cooke D, Bayer M, Jennings N (2011) Towards an understanding of the nature of resistance to *Phytophthora* root rot in red raspberry. Theoretical and Applied Genetics 123 (4):585-601. doi:10.1007/s00122-011-1609-5

Graham J, Smith K, MacKenzie K, Jorgenson L, Hackett C, Powell W (2004) The construction of a genetic linkage map of red raspberry (*Rubus idaeus* subsp. *idaeus*) based on AFLPs, genomic-SSR and EST-SSR markers. Theoretical and Applied Genetics 109 (4):740-749

Graham J, Smith K, Woodhead M, Russell J (2002) Development and use of simple sequence repeat SSR markers in *Rubus* species. Molecular Ecology Notes 2 (3):250-252. doi:10.1046/j.1471-8286.2002.00203.x

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

Grimplet J, Romieu C, Audergon J-M, Marty I, Albagnac G, Lambert P, Bouchet J-P, Terrier N (2005) Transcriptomic study of apricot fruit (*Prunus armeniaca*) ripening among 13 006 expressed sequence tags. Physiologia Plantarum 125 (3):281-292. doi:10.1111/j.1399-3054.2005.00563.x

Grube RC, Radwanski ER, Jahn M (2000) Comparative genetics of disease resistance within the Solanaceae. Genetics 155 (2):873-887

Guilford P, Prakash S, Zhu JM, Rikkerink E, Gardiner S, Bassett H, Forster R (1997) Microsatellites in *Malus* x *domestica* (apple): Abundance, polymorphism and cultivar identification. Theoretical and Applied Genetics 94 (2):249-254

Guitton B, Kelner J-J, Velasco R, Gardiner SE, Chagné D, Costes E (2012) Genetic control of biennial bearing in apple. Journal of Experimental Botany 63 (1):131-149. doi:10.1093/jxb/err261

Gupta P, Balyan H, Edwards K, Isaac P, Korzun V, Röder M, Gautier MF, Joudrier P, Schlatter A, Dubcovsky J, De la Pena R, Khairallah M, Penner G, Hayden M, Sharp P, Keller B, Wang R, Hardouin J, Jack P, Leroy P (2002a) Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. Theoretical and Applied Genetics 105 (2):413-422

Gupta P, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. Functional & Integrative Genomics 4 (3):139-162

Gupta PK, Roy JK, Prasad M (2001) Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Current Science 80 (4):524-535

Gupta PK, Varshney RK, Prasad M (2002b) Molecular Markers: Principles and Methodology. In: Jain SM, Brar DS, Ahloowalia BS (eds) Molecular Techniques in Crop Improvement. Kluwer Academic Publishers, Dordrecht and Boston, pp 9-54

Hackauf B, Rudd S, van der Voort J, Miedaner T, Wehling P (2009) Comparative mapping of DNA sequences in rye (*Secale cereale* L.) in relation to the rice genome. Theoretical and Applied Genetics 118 (2):371-384

Hagen LS, Chaib J, Fady B, Decroocq V, Bouchet JP, Lambert P, Audergon JM (2004) Genomic and cDNA microsatellites from apricot (*Prunus armeniaca* L.). Molecular Ecology Notes 4 (4):742-745

Hardison RC (2003) Comparative Genomics. Public Library of Science, Biology 1 (2):e58

Hartl DL, Jones EW (2009) Genetics. 7th edn. Jones and Bartlett, Sudbury

Hemmat M, Weedon NF, Manganaris AG, Lawson DM (1994) Molecular marker linkage map for apple. Journal of Heredity 85 (1):4-11

Hibrand-Saint Oyant L, Crespel L, Rajapakse S, Zhang L, Foucher F (2008) Genetic linkage maps of rose constructed with new microsatellite markers and locating QTL controlling flowering traits. Tree Genetics & Genomes 4:11-23

Hokanson SC, Szewc-McFadden AK, Lamboy WF, McFerson JR (1998) Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus* × *domestica* Borkh. core subset collection. Theoretical and Applied Genetics 97 (5):671-683

Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P (2005) Mapping with a few plants: Using selective mapping for microsatellite saturation of the *Prunus* reference map. Genetics 171 (3):1305-1309. doi:10.1534/genetics.105.043661

Huehn M (2011) On the bias of recombination fractions, Kosambi's and Haldane's distances based on frequencies of gametes. Genome 54 (3):196-201. doi:10.1139/g10-109

Hummer KE, Janick J (2009) Rosaceae: Taxonomy, economic importance, genomics. In: Folta K, Gardiner S (eds) Genetics and genomics of Rosaceae, vol 6. Springer, New York, pp 1-17

Igarashi M, Abe Y, Hatsuyama Y, Ueda T, Fukasawa-Akada T, Kon T, Kudo T, Sato T, Suzuki M (2008) Linkage maps of the apple (*Malus* × *domestica* Borkh.) cvs 'Ralls Janet' and 'Delicious' include newly developed EST markers. Molecular Breeding 22 (1):95-118

Illa E, Sargent D, Lopez Girona E, Bushakra J, Cestaro A, Crowhurst R, Pindo M, Cabrera A, van der Knaap E, Iezzoni A, Gardiner S, Velasco R, Arús P, Chagne D, Troggio M (2011) Comparative analysis of rosaceous genomes and the reconstruction of a putative ancestral genome for the family. BMC Evolutionary Biology 11 (1):9

Illumina I (2006) Illumina (R) SNP genotyping.

Iovene M, Wielgus SM, Simon PW, Buell CR, Jiang J (2008) Chromatin structure and physical mapping of chromosome 6 of potato and comparative analyses with tomato. Genetics 180 (3):1307-1317. doi:10.1534/genetics.108.093179

James CM, Wilson F, Hadonou AM, Tobutt KR (2003) Isolation and characterization of polymorphic microsatellites in diploid strawberry (*Fragaria vesca* L.) for mapping, diversity studies and clone identification. Molecular Ecology Notes 3 (2):171-173. doi:10.1046/j.1471-8286.2003.00365.x

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136 (4):1447-1455

Jáuregui B, de Vicente MC, Messeguer R, Felipe A, Bonnet A, Salesses G, Arús P (2001) A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. Theoretical and Applied Genetics 102 (8):1169-1176

Jonkers H (1971) An international experiment on juvenility in apple. Euphytica 20 (1):57-59

Joobeur T, Periam N, Vicente MCd, King GJ, Arús P (2000) Development of a second generation linkage map for almond using RAPD and SSR markers. Genome 43 (4):649-655. doi:10.1139/g00-040

Joobeur T, Viruel MA, de Vicente MC, Jáuregui B, Ballester J, Dettori MT, Verde I, Truco MJ, Messeguer R, Batlle I, Quarta R, Dirlewanger E, Arús P (1998) Construction of a saturated linkage map for *Prunus* using an almond × peach F_2 progeny. Theoretical and Applied Genetics 97 (7):1034-1041

Kaló P, Seres A, Taylor S, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis T, Kiss G (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. Molecular Genetics and Genomics 272 (3):235-246

Kellerhals M (2009) Introduction to Apple (*Malus* x *domestica*). In: Folta KM, Gardiner SE (eds) Plant Genetics and Genomics: Crops and Models, vol 6. Plant Genetics and Genomics: Crops and Models. Springer, New York, pp 73-84

Keniry A, Hopkins CJ, Jewell E, Morrison B, Spangenberg GC, Edwards D, Batley J (2006) Identification and characterization of simple sequence repeat (SSR) markers from Fragaria×ananassa expressed sequences. Molecular Ecology Notes 6 (2):319-322. doi:10.1111/j.1471-8286.2005.01215.x

Khan MA, Durel C-E, Duffy B, Drouet D, Kellerhals M, Gessler C, Patocchi A (2007) Development of molecular markers linked to the 'Fiesta' linkage group 7 major QTL for fire blight resistance and their application for marker-assisted selection. Genome 50 (6):568-577. doi:10.1139/g07-033

Khan SA, Chibon P-Y, de Vos RCH, Schipper BA, Walraven E, Beekwilder J, van Dijk T, Finkers R, Visser RGF, van de Weg EW, Bovy A, Cestaro A, Velasco R, Jacobsen E, Schouten HJ (2012) Genetic analysis of metabolites in apple fruits indicates an mQTL hotspot for phenolic compounds on linkage group 16. Journal of Experimental Botany 63 (8):2895-2908. doi:10.1093/jxb/err464

Kimura T, Shi YZ, Shoda M, Kotobuki K, Matsuta N, Hayashi T, Ban Y, Yamamoto T (2002) Identification of Asian pear varieties by SSR analysis. Breeding Science 52 (2):115-121

Koonin E (2001) An apology for orthologs - or brave new memes. Genome Biology 2 (4):comment1005.1001 - comment1005.1002

Koonin EV, Aravind L, Kondrashov AS (2000) The impact of comparative genomics on our understanding of evolution. Cell 101 (6):573-576

Kouassi AB, Durel CE, Mathis F, Gianfranceschi L, Komjanc M, Mott D, Cova V, Patocchi A, Gobbin D, Rezzonico F, Evans K, Fernandez-Fernandez F, Dunemann F, Boudichevskaja A, Stankiewicz-Kosyl M, Antofie A, Weg Evd, Bink M, Laurens F, de Weg Ev, van de Weg E (2008) QTL mapping for fruit firmness in apple using a pedigree-based Bayesian approach. Modern variety breeding for present and future needs. Proceedings of the 18th EUCARPIA general congress, Valencia, Spain, 9-12 September, 2008.

Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB (2004) Comparative mapping in the Pinaceae. Genetics 168 (1):447-461. doi:10.1534/genetics.104.028381

Kumar S, Bink MCAM, Volz RK, Bus VGM, Chagné D (2012) Towards genomic selection in apple (*Malus* x *domestica* Borkh.) breeding programmes: Prospects, challenges and strategies. Tree Genetics & Genomes 8 (1):1-14. doi:10.1007/s11295-011-0425-z

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150 (3):1217-1228

Lagercrantz U, Putterill J, Coupland G, Lydiate D (1996) Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. Plant Journal 9 (1):13-20

Lambert P, Hagen LS, Audergon JM, Arús A (2006) Evaluation of a mapping strategy based on synteny: Using the *Prunus* reference map for apricot. Acta Horticulturae (ISHS) 701:295-301

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121 (1):185-199

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits:guidelines for interpreting and reporting linkage results. Nature Genetics 11:241-247

Lerceteau-Kohler E, Moing A, Guerin G, Renaud C, Courlit S, Camy D, Praud K, Parisy V, Bellec F, Rolin D, Roudeillac P, Denoyes-Rothan B (2004) QTL analysis for fruit quality traits in octoploid strawberry (*Fragaria* x *ananassa*). Acta Horticulturae (ISHS) 663:331-335

Lewers K, Saski C, Cuthbertson B, Henry D, Staton M, Main D, Dhanaraj A, Rowland L, Tomkins J (2008) A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. BMC Plant Biology 8 (1):69-76

Lewers KS, Styan SMN, Hokanson SC, Bassil NV (2005) Strawberry GenBank-derived and genomic simple sequence repeat (SSR) markers and their utility with strawberry, blackberry, and red and black raspberry. Journal of the American Society for Horticultural Science 130 (1):102-115

Liebhard R, Gianfranceschi L, Koller B, Ryder CD, Tarchini R, Van De Weg E, Gessler C (2002) Development and characterisation of 140 new microsatellites in apple (*Malus* x *domestica* Borkh.). Molecular Breeding 10 (4):217-241

Liebhard R, Koller B, Gianfranceschi L, Gessler C (2003a) Creating a saturated reference map for the apple (*Malus* x *domestica* Borkh.) genome. Theoretical and Applied Genetics 106:1497 - 1508

Liebhard R, Koller B, Patocchi A, Kellerhals M, Pfammatter W, Jermini M, Gessler C (2003b) Mapping quantitative field resistance against apple scab in a 'Fiesta' x 'Discovery' progeny. Phytopathology 93:493 - 501

Liewlaksaneeyanawin C, Zhuang J, Tang M, Farzaneh N, Lueng G, Cullis C, Findlay S, Ritland C, Bohlmann J, Ritland K (2009) Identification of COS markers in the Pinaceae. Tree Genetics & Genomes 5 (1):247-255

Lin H, Walker MA (1998) Identifying grape rootstocks with simple sequence repeat (SSR) DNA markers. American Journal of Enology and Viticulture 49 (4):403-407

Lin J-J, Kuo J, Ma J, Saunders J, Beard H, MacDonald M, Kenworthy W, Ude G, Matthews B (1996) Identification of molecular markers in soybean comparing RFLP, RAPD and AFLP DNA mapping techniques. Plant Molecular Biology Reporter 14 (2):156-169. doi:10.1007/bf02684905

Livingstone III D, Royaert S, Mockaitis K, May GD, Farmer A, Saski C, Schnell RJ, Motamayor JC, Kuhn D (2012) Identification, analysis and mapping of SNP markers in *Theobroma cacao*. Paper presented at the Plant & Animal Genome XX, San Diego, CA, 14-18 January 2012

Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the Solanaceae. Genetics 152 (3):1183-1202

Lopes MS, Maciel GB, Mendonça D, Gil FS, Da Câmara Machado A (2006) Isolation and characterization of simple sequence repeat loci in *Rubus hochstetterorum* and their use in other species from the Rosaceae family. Molecular Ecology Notes 6 (3):750-752. doi:10.1111/j.1471-8286.2006.01329.x

Lu Y, Shah T, Hao Z, Taba S, Zhang S, Gao S, Liu J, Cao M, Wang J, Prakash AB, Rong T, Xu Y (2011) Comparative SNP and haplotype analysis reveals a higher genetic diversity and rapider LD decay in tropical than temperate germplasm in maize. Public Library of Science One 6 (9):e24861. doi:10.1371/journal.pone.0024861

Lu Z-X, Sosinski B, Reighard GL, Baird WV, Abbott AG (1998) Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks. Genome 41 (2):199-207. doi:10.1139/g98-008

Mackay TFC (2001) The genetic architecture of quantitative traits. Annual Review of Genetics 35:303 - 339

Maliepaard C, Alston FH, Van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, Van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, Den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrielink-van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theoretical and Applied Genetics 97:60 - 73

Martinez Garcia PJ, Parfitt D, Ogundiwin E, Fass J, Chan HM, Gradziel T, Crisosto CH (2012) High-density SNP mapping on peach (*Prunus persica* L. Batsch) populations. Paper presented at the Plant and Animal Genome XX, San Diego, CA, 14-18 Jan 2012

Mauricio R (2001) Mapping quantitative trait loci in plants: Uses and caveats for evolutionary biology. Nature Genetics 2:370-381

McCouch SR, Doerge RW (1995) QTL mapping in rice. Trends in Genetics 11 (12):482-487. doi:10.1016/s0168-9525(00)89157-x

McGregor CE, Lambert CA, Greyling MM, Louw JH, Warnich L (2000) A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. Euphytica 113 (2):135-144

Micheletti D, Aramini V, Arús P, Banchi E, Barreneche T, Bassi D, Dirlewanger E, Gao Z, Gazza L, Lambert P, Laurens F, Li X, Pascal T, Quilot-Turion B, Rossini L, Troggio M, Van De Weg E, Verde I, Aranzana MJ (2012) Application of high throughput genotyping techniques in peach germplasm and breeding lines within the FruitBreedomics project. Paper presented at the Plant and Animal Genome XX, San Diego, CA, 14-18 Jan, 2012

Milbourne D, Meyer R, Bradshaw JE, Baird E, Bonar N, Provan J, Powell W, Waugh R (1997) Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. Molecular Breeding 3 (2):127-136

Mittelsten Scheid O, Jakovleva L, Afsar K, Maluszynska J, Paszkowski J (1996) A change of ploidy can modify epigenetic silencing. Proceedings of the National Academy of Sciences 93 (14):7114-7119

Montgomery J, Wittwer CT, Palais R, Zhou L (2007) Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. Nature Protocols 2 (1):59-66. doi:10.1038/nprot.2007.10

Newcomb RD, Crowhurst RN, Gleave AP, Rikkerink EHA, Allan AC, Beuning LL, Bowen JH, Gera E, Jamieson KR, Janssen BJ, Laing WA, McArtney S, Nain B, Ross GS, Snowden KC, Souleyre EJF, Wallon EF, Yauk YK (2006) Analyses of expressed sequence tags from apple. Plant Physiology 141:147 - 166

Ogundiwin E, Peace C, Gradziel T, Parfitt D, Bliss F, Crisosto C (2009) A fruit quality gene map of *Prunus*. BMC Genomics 10 (1):587

Olmstead J, Sebolt A, Cabrera A, Sooriyapathirana S, Hammar S, Iriarte G, Wang D, Chen C, van der Knaap E, Iezzoni A (2008) Construction of an intra-specific sweet cherry (*Prunus avium* L.) genetic linkage map and synteny analysis with the *Prunus* reference map. Tree Genetics & Genomes 4 (4):897-910

Oraguzie NC, Rikkerink EHA, Gardiner SE, de Silva HN (2007) Association mapping in plants. Springer-Verlag GmbH, Heidelberg

Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theoretical and Applied Genetics 85 (8):985-993. doi:10.1007/bf00215038

Paterson AH (1996) Genome mapping in Plants. Academic Press,

Paterson AH, Bennetzen JL (2001) Comparative mapping of plant chromosomes. In: Phillips RL, Vasil IK (eds) DNA-Based Markers in Plants, vol 6. Advances in Cellular and Molecular Biology of Plants, 2nd edn. Kluwer Academic Publishers, Dordrecht, pp 101-114

Paterson AH, Bowers JE, Burow MD, Draye X, Elsik CG, Jiang C-X, Katsar CS, Lan T-H, Lin Y-R, Ming R, Wright RJ (2000) Comparative genomics of plant chromosomes. Plant Cell 12 (9):1523-1540. doi:10.1105/tpc.12.9.1523

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proceedings of the National Academy of Sciences of the United States of America 101 (26):9903-9908. doi:10.1073/pnas.0307901101

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M (1998) Comparative analysis of genetic similarity among maize inbred

lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theoretical and Applied Genetics 97 (8):1248-1255

Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce-Jacino M (1999) Mining SNPs from EST databases. Genome Research 9:167-174. doi:PMC310719

Pierantoni L, Cho KH, Shin IS, Chiodini R, Tartarini S, Dondini L, Kang SJ, Sansavini S (2004) Characterisation and transferability of apple SSRs to two European pear F_1 populations. Theoretical and Applied Genetics 109 (7):1519-1524

Pierantoni L, Dondini L, Cho KH, Shin IS, Gennari F, Chiodini R, Tartarini S, Kang SJ, Sansavini S (2007) Pear scab resistance QTLs via a European pear (*Pyrus communis*) linkage map. Tree Genetics & Genomes 3 (4):311-317

Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, Campbell CS (2007) Phylogeny and classification of Rosaceae. Plant Systematics and Evolution 266 (1):5-43

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends in Plant Science 1 (7):215-222

Rajapakse S, Belthoff LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R, Abbott AG (1995) Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. Theoretical and Applied Genetics 90:503-510

Reed GH, Wittwer CT (2004) Sensitivity and specificity of single-nucleotide polymorphism scanning by High-Resolution Melting analysis. Clinical Chemistry 50 (10):1748-1754. doi:10.1373/clinchem.2003.029751

Riechmann J, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe O, Samaha R (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290:2105 - 2110

Romero C, Pedryc A, Munoz V, Llácer G, Badenes M (2003) Genetic diversity of different apricot geographical groups determined by SSR markers. Genome 46 (2):244-252

Rusholme Pilcher RL, Celton JM, Gardiner SE, Tustin DS (2008) Genetic markers linked to the dwarfing trait of apple rootstock 'Malling 9'. Journal of the American Society for Horticultural Science 133 (1):100-106

Sargent D, Fernandéz-Fernandéz F, Ruiz-Roja J, Sutherland B, Passey A, Whitehouse A, Simpson D (2009a) A genetic linkage map of the cultivated strawberry (*Fragaria* x *ananassa*) and its comparison to the diploid *Fragaria* reference map. Molecular Breeding 24 (3):293-303. doi:10.1007/s11032-009-9292-9

Sargent DJ, Davis TM, Tobutt KR, Wilkinson MJ, Battey NH, Simpson DW (2004) A genetic linkage map of microsatellite, gene-specific and morphological markers in diploid *Fragaria*. Theoretical and Applied Genetics 109 (7):1385-1391

Sargent DJ, Fernández-Fernández F, Rys A, Knight VH, Simpson DW, Tobutt KR (2007b) Mapping of A_1 conferring resistance to the aphid *Amphorophora idaei* and *dw* (dwarfing habit) in red raspberry (*Rubus idaeus* L.) using AFLP and microsatellite markers. BMC Plant Biology 7 (15)

Sargent DJ, Hadonou AM, Simpson DW (2003) Development and characterization of polymorphic microsatellite markers from *Fragaria viridis*, a wild diploid strawberry. Molecular Ecology Notes 3 (4):550-552. doi:10.1046/j.1471-8286.2003.00507.x

Sauge MH, Lambert P, Pascal T (2012) Co-localisation of host plant resistance QTLs affecting the performance and feeding behaviour of the aphid Myzus persicae in the peach tree. Heredity 108 (3):292-301

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8 (6):552-560

Schlötterer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Research 20 (2):211-215

Scorza R, Mehlenbacher SA, Lightner GW (1985) Inbreeding and coancestry of freestone peach cultivars of the eastern United States and implications for peach germplasm improvement. Journal of the American Society for Horticultural Science 110:547-552

Scott DH (1951) Cytological studies on polyploids derived from tetraploid *Fragaria vesca* and cultivated strawberries. Genetics 36 (4):311-331

Seipp MT, Durtschi JD, Liew MA, Williams J, Damjanovich K, Pont-Kingdon G, Lyon E, Voelkerding KV, Wittwer C (2007) Unlabeled oligonuclotides as internal temperature controls for genotyping by amplicon melting. Journal of Molecular Diagnostics 9 (3):284-289

Shaw J, Small RL (2004) Addressing the "hardest puzzle in American pomology:" Phylogeny of *Prunus* sect. Prunocerasus (Rosaceae) based on seven noncoding chloroplast DNA regions. American Journal of Botany 91 (6):985-996. doi:10.3732/ajb.91.6.985

Shenghua L, Chengquan F (1994) Studies on chromosome of *Pyrus* in China. Acta Horticulturae (ISHS) 367:27-32

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton J-M, Rees DJG, Williams KP, Holt SH, Rojas JJR, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman T-L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Girona EL, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arús P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). Nature Genetics 43 (2):109-116. doi:http://www.nature.com/ng/journal/v43/n2/abs/ng.740.html#supplementary-information

Silfverberg-Dilworth E, Matasci CL, van de Weg WE, Van Kaauwen MPW, Walser M, Kodde LP, Soglio V, Gianfranceschi L, Durel CE, Costa F, Yamamoto T, Koller B, Gessler C, Patocchi A (2006) Microsatellite markers spanning the apple (*Malus* x *domestica* Borkh.) genome. Tree Genetics & Genomes 2:202 - 224

Sjulin TM, Dale A (1987) Genetic diversity of North American strawberry cultivars. Journal of the American Society for Horticultural Science 112 (2):375-385

Soller M, Brody T, Genizi A (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theoretical and Applied Genetics 47 (1):35-39. doi:10.1007/bf00277402

Soriano JM, Romero C, Vilanova S, Llácer G, Badenes M (2005) Genetic diversity of loquat germplasm (*Eriobotrya japonica* (Thunb.) Lindl) assessed by SSR markers. Genome 48 (1):108-114

Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD, Miftahudin, Mahmoud A, Ma X, Gustafson PJ, Qi LL, Echalier B, Gill BS, Matthews DE, Lazo GR, Chao S, Anderson OD, Edwards H, Linkiewicz AM,

Dubcovsky J, Akhunov ED, Dvorak J, Zhang D, Nguyen HT, Peng J, Lapitan NLV, Gonzalez-Hernandez JL, Anderson JA, Hossain K, Kalavacharla V, Kianian SF, Choi D-W, Close TJ, Dilbirligi M, Gill KS, Steber C, Walker-Simmons MK, McGuire PE, Qualset CO (2003) Comparative DNA sequence analysis of wheat and rice genomes. Genome Research 13 (8):1818-1827. doi:10.1101/gr.1113003

Sosinski B, Gannavarapu M, Hager LD, Beck LE, King GJ, Ryder CD, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000) Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. Theoretical and Applied Genetics 101 (3):421-428

Sosinski B, Verde I, Rokhsar DS (2010) International Peach Genome Initiative Peach Genome v1.0. http://www.rosaceae.org/peach/genome.

Stafne ET, Clark JR (2004) Genetic relatedness among eastern North American blackberry cultivars based on pedigree analysis. Euphytica 139 (2):95-104. doi:10.1007/s10681-004-2436-4

Stafne ET, Clark JR, Weber CA, Graham J, Lewers KS (2005) Simple sequence repeat (SSR) markers for genetic mapping of raspberry and blackberry. Journal of the American Society for Horticultural Science 130 (5):722-728

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: Join Map. Plant Journal 3 (5):739-744. doi:10.1111/j.1365-313X.1993.00739.x

Sturtevant AH (1913) The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of inheritance. Journal of Experimental Zoology 14:43-59

Sutherland BG, Tobutt KR, Marchese A, Paternoster G, Simpson DW, Sargent DJ (2008) A genetic linkage map of *Physocarpus*, a member of the Spiraeoideae (Rosaceae), based on RAPD, AFLP, RGA, SSR and gene specific markers. Plant Breeding 127 (5):527-532

Tam SM, Mhiri C, Vogelaar A, Kerkveld M, Pearce SR, Grandbastien M-A (2005) Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. Theoretical and Applied Genetics 110 (5):819-831

Tanksley SD (1993) Mapping polygenes. Annual Review of Genetics 27:205-233. doi:10.1146/annurev.ge.27.120193.001225

Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988) Conservation of gene repertoire but not gene order in pepper and tomato. Proceedings of the National Academy of Sciences of the United States of America 85 (17):6419-6423

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49 (1):11-25

Testolin R, Marrazzo M, Cipriani G, Quarta R (2000) Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerpinting and testing the genetic origin of cultivars. Genome 43 (3):512-520

Thoday JM (1961) Location of polygenes. Nature 191:368-370. doi:10.1038/191368a0

Thompson MM (1995a) Chromosome numbers of *Rubus* cultivars at the National Clonal Germplasm Repository. HortScience 30 (7):1453-1456

Thomson MJ, Zhao K, Wright M, McNally KL, Leung H, McCouch SR (2010) Development and application of 96- and 384-plex single nucleotide polymorphism (SNP) marker sets for diversity analysis, mapping and marker-assisted selection in rice. Paper presented at the Second Africa Rice Congress, Bamako, Mali, 22-26 March 2010

Torre LC, Barritt BH (1977) Quantitative evaluation of *Rubus* fruit anthocyanin pigments. Journal of Food Science 42 (2):488-490. doi:10.1111/j.1365-2621.1977.tb01528.x

Van Deynze A, Stoffel K, Buell CR, Kozik A, Liu J, van der Knaap E, Francis D (2007) Diversity in conserved genes in tomato. BMC Genomics 8:465

Van Deynze AE, Nelson JC, Yglesias ES, Harrington SE, Braga DP, McCouch SR, Sorrells ME (1995) Comparative mapping in grasses. Wheat relationships. Molecular and General Genetics MGG 248 (6):744-754

van Ooijen JW (1992) Accuracy of mapping quantitative trait loci in autogamous species. Theoretical and Applied Genetics 84 (7):803-811. doi:10.1007/bf00227388

van Ooijen JW, Voorrips RE (2001) JoinMap(R) 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands

Vandersteen JG, Bayrak-Toydemir P, Palais RA, Wittwer CT (2007) Identifying common genetic variants by High-Resolution Melting. Clinical Chemistry 53 (7):1191-1198. doi:10.1373/clinchem.2007.085407

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nature Genetics 42 (10):833-839. doi:http://www.nature.com/ng/journal/v42/n10/abs/ng.654.html#supplementary-information

Verde I, Lauria M, Dettori M, Vendramin E, Balconi C, Micali S, Wang Y, Marrazzo M, Cipriani G, Hartings H, Testolin R, Abbott A, Motto M, Quarta R (2005) Microsatellite and AFLP markers in the *Prunus persica* [L. (Batsch)] × *P. ferganensis* BC_1 linkage map: Saturation and coverage improvement. Theoretical and Applied Genetics 111 (6):1013-1021

Vilanova S, Sargent D, Arús P, Monfort A (2008) Synteny conservation between two distantly-related Rosaceae genomes: *Prunus* (the stone fruits) and *Fragaria* (the strawberry). BMC Plant Biology 8 (1):67

Viruel MA, Messeguer R, De Vicente MC, Garcia-Mas J, Puigdomenech P, Vargas F, Arús P (1995) A linkage map with RFLP and isozyme markers for almond. Theoretical and Applied Genetics 91:964-971

Vision TJ, Brown DG, Shmoys DB, Durrett RT, Tanksley SD (2000) Selective mapping: A strategy for optimizing the construction of high-density linkage maps. Genetics 155 (1):407-420

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23 (21):4407-4414

Wang D, Karle R, Brettin TS, Iezzoni AF (1998) Genetic linkage map in sour cherry using RFLP markers. Theoretical and Applied Genetics 97 (8):1217-1224

Warburton ML, Xianchun X, Crossa J, Franco J, Melchinger AE, Frisch M, Bohn M, Hoisington D (2002) Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. Crop Science 42 (6):1832-1840

Way RD, Aldwinckle HS, Lamb RC, Rejman A, Sansavini S, Shen T, Watkins R, Westwood MN, Yoshida Y (1991) Apples (*Malus*). Acta Horticulturae (ISHS) 209:3-46

Weebadde CK, Wang D, Finn CE, Lewers KS, Luby JJ, Bushakra J, Sjulin TM, Hancock JF (2008) Using a linkage mapping approach to identify QTL for day-neutrality in the octoploid strawberry. Plant Breeding 127 (1):94-101. doi:10.1111/j.1439-0523.2007.01430.x

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18 (22):6531-6535

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-Resolution Genotyping by Amplicon Melting Analysis using LCGreen. Clinical Chemistry 49 (6):853-860. doi:10.1373/49.6.853

Wolfe KH, Sharp PM, Li WH (1989) Mutation rates differ among regions of the mammalian genome. Nature 337:283-285

Woodhead M, McCallum S, Smith K, Cardle L, Mazzitelli L, Graham J (2008) Identification, characterisation and mapping of simple sequence repeat (SSR) markers from raspberry root and bud ESTs. Molecular Breeding 22 (4):555-563

Woodhead M, Weir A, Smith K, McCallum S, MacKenzie K, Graham J (2010) Functional markers for red raspberry. Journal of the American Society for Horticultural Science 135 (5):418-427

Wu FN, Eannetta NT, Xu YM, Durrett R, Mazourek M, Jahn MM, Tanksley SD (2009a) A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus Capsicum. Theoretical and Applied Genetics 118 (7):1279-1293. doi:10.1007/s00122-009-0980-y

Wu FN, Eannetta NT, Xu YM, Tanksley SD (2009b) A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theoretical and Applied Genetics 118 (5):927-935. doi:10.1007/s00122-008-0950-9

Wu S-B, Tavassolian I, Rabiei G, Hunt P, Wirthensohn M, Gibson J, Ford C, Sedgley M (2009c) Mapping SNP-anchored genes using high-resolution melting analysis in almond. Molecular Genetics and Genomics 282 (3):273-281

Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F_2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). Molecular and General Genetics 253 (5):535-545. doi:10.1007/s004380050355

Yamamoto T, Kimura T, Saito T, Kotobuki K, Matsuta N, Liebhard R, Gessler C, van de Weg WE, Hayashi T (2004a) Genetic linkage maps of Japanese and European pears aligned to the apple consensus map. Acta Horticulturae (ISHS) 663:51-56

Yamamoto T, Kimura T, Sawamura Y, Kotobuki K, Ban Y, Hayashi T, Matsuta N (2001a) SSRs isolated from apple can identify polymorphism and genetic diversity in pear. Theoretical and Applied Genetics 102 (6):865-870

Yamamoto T, Kimura T, Shoda M, Ban Y, Hayashi T, Matsuta N (2002a) Development of microsatellite markers in the Japanese pear (*Pyrus pyrifolia* Nakai). Molecular Ecology Notes 2 (1):14-16

Yamamoto T, Kimura T, Shoda M, Imai T, Saito T, Sawamura Y, Kotobuki K, Hayashi T, Matsuta N (2002b) Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. Theoretical and Applied Genetics 106 (1):9-18

Yamamoto T, Kimura T, Soejima J, Sanada T, Ban Y, Hayashi T (2004b) Identification of quince varieties using SSR markers developed from pear and apple. Breeding Science 54 (3):239-244

Yamamoto T, Kimura T, Terakami S, Nishitani C, Sawamura Y, Saito T, Kotobuki K, Hayashi T (2007) Integrated reference genetic linkage maps of pear based on SSR and AFLP markers. Breeding Science 57 (4):321-329

Yamamoto T, Shimada T, Imai T, Yaegaki H, Haji T, Matsuta N, Yamaguchi M, Hayashi T (2001b) Characterization of morphological traits based on a genetic linkage map in peach Breeding Science 51 (4):271-278

Young ND (2001) Constructing a plant genetic linkage map with DNA markers. In: Phillips RL, Vasil IK (eds) DNA-Based Markers in Plants vol 6. Advances in Cellular and Molecular Biology of Plants, 2nd edn. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 31-47

Zorrilla-Fontanesi Y, Cabeza A, Domínguez P, Medina J, Valpuesta V, Denoyes-Rothan B, Sánchez-Sevilla J, Amaya I (2011a) Quantitative trait loci and underlying candidate genes controlling agronomical and fruit quality traits in octoploid strawberry (*Fragaria* × *ananassa*). Theoretical and Applied Genetics 123 (5):755-778. doi:10.1007/s00122-011-1624-6

Zorrilla-Fontanesi Y, Cabeza A, Torres A, Botella M, Valpuesta V, Monfort A, Sánchez-Sevilla J, Amaya I (2011b) Development and bin mapping of strawberry genic-SSRs in diploid *Fragaria* and their transferability across the Rosoideae subfamily. Molecular Breeding 27 (2):137-156. doi:10.1007/s11032-010-9417-1

CHAPTER TWO

This chapter published as:
***Bushakra JM**, Sargent DJ, Cabrera A, Crowhurst RN, Lopez Girona E, Velasco R, Symonds VV, van der Knaap E, Troggio M, Gardiner SE, Chagné D (2011) Rosaceae conserved orthologous set (RosCOS) markers as a tool to assess genome synteny between Malus and Fragaria. Tree Genetics & Genomes:1-16. DOI 10.1007/s11295-011-0450-y. See Appendix for reprint.*

2 Rosaceae conserved orthologous set (RosCOS) markers as a tool to assess genome synteny between *Malus* and *Fragaria*

2.1 Abstract

Inter-tribal comparisons of genome synteny between phylogenetically distant genera in Rosaceae, such as *Malus* (apple) and *Fragaria* (strawberry), have previously been hampered by a lack of transferable markers that can be used as anchor points between genetic maps. The availability of conserved orthologous set (COS) markers recently developed for this family, coupled with the release of the *Malus* × *domestica* and *Fragaria vesca* draft genome sequences, provide new tools for comprehensive pair-wise comparisons. The genetic mapping of 56 Rosaceae COS (RosCOS) markers revealed 21 regions of genomic synteny between apple and strawberry. Information concerning the location of RosCOS markers on 15 of 17 apple linkage groups (LG) and all seven LG of strawberry was used to assess the ancestral relationships between the two genera. Four differences in orientation of ancestral chromosome fragments on extant LG were identified in comparison with previous studies, as well as two potential insertions, two potential translocations, and two potential inversions. The set of orthologous markers developed for use in genetic mapping in Rosaceae, in combination with high-throughput analysis, will allow the exploration of chromosome evolution and refinement of ancestral relationships within the family, orientation and anchoring of genome sequences as they become available, and provide resources to develop markers for non-sequenced genomes within the family.

2.2 Introduction

In organisms descended from a common ancestor, the information found in the DNA is conserved at several levels, including in the sequences of orthologous genes and the order of the orthologous genes over large parts of chromosomes. Comparative genome mapping is the study of the conservation of gene content within chromosomal regions (synteny) between organisms utilizing comparative sequence analyses and genetic mapping of orthologous markers. In plant families, such as Solanaceae (Bonierbale et al. 1988; Tanksley et al. 1988; Wu et al. 2009a; Wu et al. 2009b), Brassicaceae (Acarkan et al. 2000; Lagercrantz 1998; Lagercrantz et al. 1996) and Poaceae (Ahn and Tanksley 1993; Moore et al. 1995; Van Deynze et al. 1995; Ventelon et al. 2001),

comparative genome mapping studies have identified conserved genomic regions, chromosomal rearrangements, and genome duplication providing information on the evolution of these families, and allowing the genomic resources to be transferred to less well-studied species.

The Rosaceae is a large, complex, diverse plant family consisting of approximately 90 genera and 3,000 species (Potter et al. 2007). Economically important fruit crops include apple, pear, apricot, peach, raspberry, and strawberry. Many Rosaceae crop species are trees with long juvenile periods. They are highly heterozygous and often self-incompatible, making genetic mapping studies more challenging than in annual selfing species where recombinant inbred lines can be generated. An understanding of genome homology and evolution among the wider family is required for the development of tools for genomic studies in Rosaceae.

Recently, the genomes of apple (*Malus* × *domestica*; $2n = 2x = 34$) (Velasco et al. 2010) and diploid alpine strawberry (*Fragaria vesca*; $2n = 2x = 14$) (Shulaev et al. 2011) have been fully sequenced. Apple and strawberry are phylogenetically distant genera (Potter et al. 2007) of the Rosaceae (Fig. 2.1). In addition to their obvious morphological differences (one is a pome-bearing tree, the other is an achene-producing herbaceous perennial), these genera differ greatly in their chromosome number and genome size [*Malus* spp., $n = 17$ chromosomes, genome size ~750 Mb (Arumuganathan and Earle 1991); *Fragaria* spp., $n = 7$ chromosomes, genome size ~240 Mb (Shulaev et al. 2011)]. The draft whole genome sequence (WGS) of strawberry accession Hawaii 4 was assembled from a sequence-by-synthesis approach, and anchored to the *Fragaria* reference linkage map and associated bin map (*Fragaria vesca* × *F. bucharica*, FVxFB) (Sargent et al. 2008; Sargent et al. 2009c), using 320 genetic markers (Shulaev et al. 2011). The 'Golden Delicious' (GD) apple draft WGS was assembled using a combination of Sanger and sequence-by-synthesis reads and anchored to an integrated genetic map with 1,643 simple sequence repeat (SSR) and single nucleotide polymorphic (SNP) markers (Velasco et al. 2010). The origin of $x = 17$ in Pyrodae has strong support from two recent hypotheses, the first hypothesis suggests hybridization of two sister taxa of *Gillenia* ($x = 9$) and allopolyploidization (Campbell et al. 2007; Evans and Campbell 2002; Potter et al. 2007), alternatively the second hypothesis suggests autopolyploidization of an $x = 9$ ancestor of *Gillenia* (Velasco et al. 2010). Both hypotheses suggest diploidization and loss of a chromosome. A chromosome-by-

chromosome comparison of the genome revealed paralogous (duplication within a genome) regions resulting from the *Gillenia*-derived autopolyploidization (Fig. 2.2) that gave rise to the genera in the subtribe Pyrinae (Velasco et al. 2010).

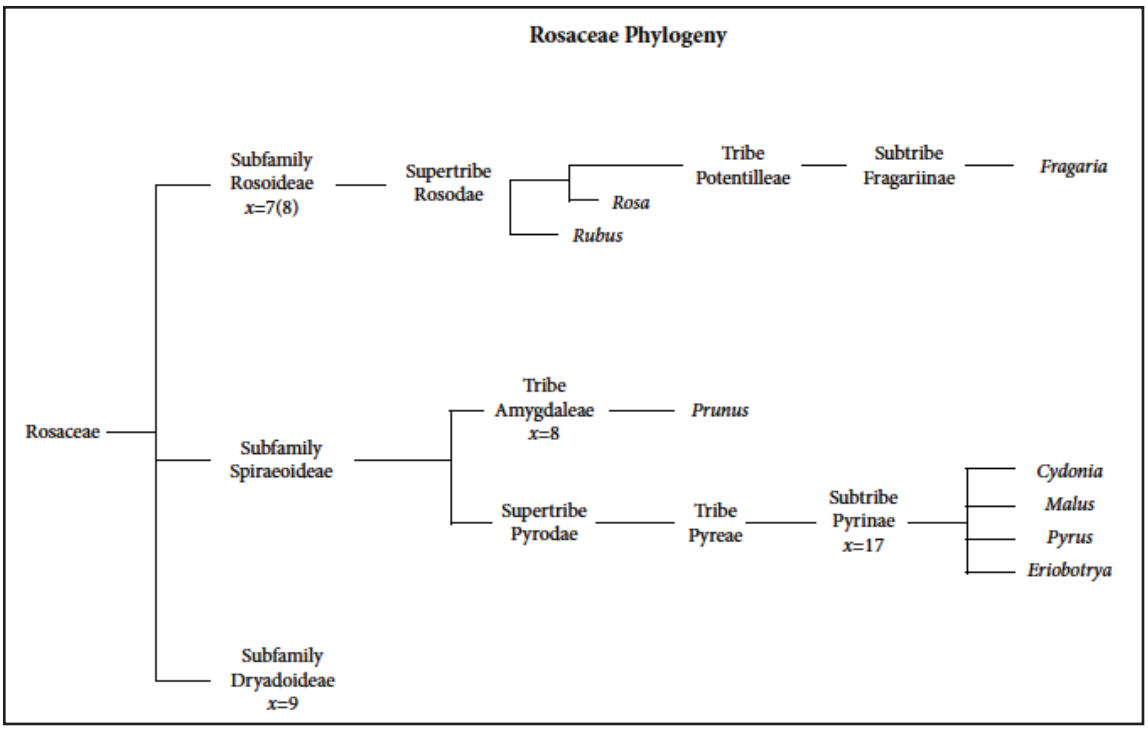


Fig. 2.1: Simplified and abbreviated Rosaceae phylogeny, highlighting economically important genera (modeled after Potter et al. 2007). Branches approximate shared ancestry and are not to scale nor all-inclusive. Base chromosome numbers (*x*) are indicated. Phylogeny illustrates the relative distance between *Malus* (apple) and *Fragaria* (strawberry).

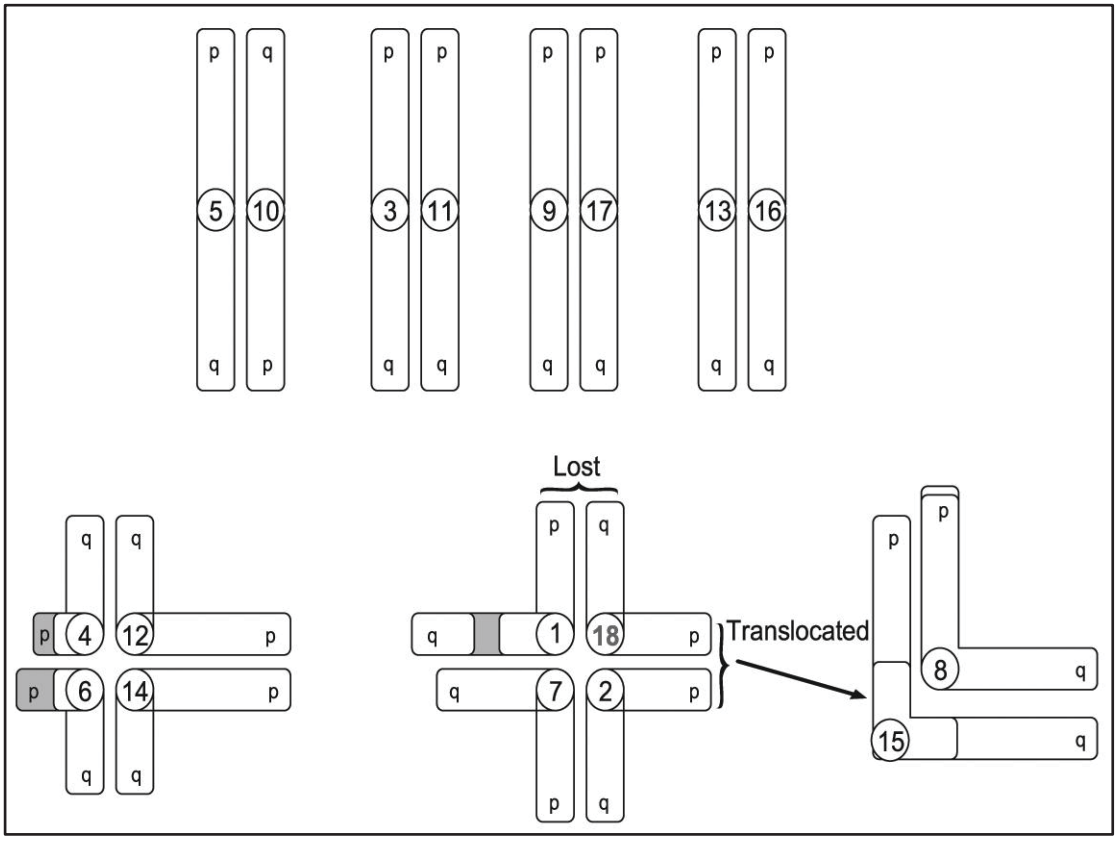


Fig. 2.2: Relationship of apple chromosomes based on a chromosome-by-chromosome comparison as determined by gene homology (modeled after Velasco et al. 2010). Grey areas indicated missing duplicate counterparts. Letters p & q indicate chromosome orientation.

Few inter-generic genetic map comparisons have been performed in Rosaceae, mainly within subtribes (Dirlewanger et al. 2004b; Lambert et al. 2006; Olmstead et al. 2008; Pierantoni et al. 2004; Yamamoto et al. 2004a; Celton et al. 2009a; Chen et al. 2008), with only one inter-tribal comparison (Vilanova et al. 2008). SSR markers were the main marker type used for these comparisons and their transferability within subtribes was demonstrated by using apple-derived SSR markers to develop and compare linkage maps in pear (Booi et al. 2005; Pierantoni et al. 2004; van Dyk et al. 2005; Celton et al. 2009a; Celton et al. 2009b; Yamamoto et al. 2004a), loquat (Gisbert et al. 2009; Soriano et al. 2005) and quince (Yamamoto et al. 2004b). Although peach-derived SSR markers were used to create linkage maps within the genus *Prunus* (Downey and Iezzoni 2000; Dirlewanger et al. 2004a; Dirlewanger et al. 2002; Dirlewanger et al. 2006; Jáuregui et al. 2001; Lambert et al. 2006; Olmstead et al. 2008), transferability within *Prunus* decreases with subgeneric (Shaw and Small 2004) phylogenetic distance (Decroocq et

al. 2003). SSR markers from subfamily Spiraeoideae transfer poorly to subfamily Rosoideae (Lewers et al. 2005; Hibrand-Saint Oyant et al. 2008). Clearly, the development of orthologous genetic markers will benefit genetic mapping across species in Rosaceae, with the lesser-studied species, such as pear (related to apple within Pyrinae) and raspberry (related to strawberry within Rosoideae) profiting most.

A subfamily-level comparative map between Spiraeoideae (Prunoideae) and Rosoideae using a combination of 71 anchored markers derived from expressed sequence tags (EST) developed from both *Prunus* (8) and *Fragaria* (13), eight sequence tag site (STS) markers developed from *Fragaria*, two SSR markers, and 40 restriction fragment length polymorphism (RFLP) markers previously mapped in *Prunus* identified orthology between the eight linkage groups (LG) of *Prunus* and the seven LG of *Fragaria* (Vilanova et al. 2008). This comparison suggested that 27 inversions and nine fusion/fission or translocation events occurred in the evolution of these two genera from a common ancestor (Vilanova et al. 2008). Additionally, an inter-tribal comparative map based on 24 RFLP markers and six isozyme genes homologous between *Malus* and *Prunus* provided evidence of orthology between these two genera (Dirlewanger et al. 2004b), however the map does not cover all chromosomes. Due to the lack of a sufficient number of transferable markers of any type, no direct comparison has been performed between apple and strawberry genetic linkage maps.

Gene-based conserved orthologous set (COS) markers present an alternative to SSR markers derived from genomic DNA. COS are molecular markers that are designed specifically from single- or low-copy genes. The emphasis on developing molecular markers from genes that are single-copy reduces the difficulty of distinguishing orthology from paralogy found when multigene families are used for marker development (Fulton et al. 2002; Bolot et al. 2009), and the use of coding sequence enhances the sequence conservation of primer binding sites across greater phylogenetic distances. A set of 613 polymorphic COS markers have been developed for Rosaceae (RosCOS) and located on the interspecific *Prunus* 'Texas' (almond) × 'Earlygold' (peach) (TxE) reference bin map (Cabrera et al. 2009). These were proposed as a set of markers for use in comparative mapping within Rosaceae, and a subset of 126 has subsequently been bin mapped on the FVxFB bin mapping set in the course of a recent inter-tribal pair-wise comparative genome study (Illa et al. 2011). The location of RosCOS primer binding sites in the conserved exons increases the transferability among

genera and the potential of amplifying a variable intron. This provides an advantage over non-gene-based SSR, as mutations within the SSR primer binding sites can lead to a decrease in transferability between genera (Powell et al. 1996).

We demonstrate here that RosCOS is an efficient orthologous marker system for the assessment of genome synteny between apple and strawberry. We compare the inter-tribal transfer efficiency of RosCOS markers with that of published SSR markers using *in silico* analysis and show that RosCOS markers have a greater success rate for transferability than SSR. Using this tool, we identify regions of synteny between the two genera that support earlier work based on bioinformatics alignment, and identify new syntenic regions, demonstrating that RosCOS can be used to create a framework of orthologous markers for the alignment of Rosaceae genetic maps. This is the first comprehensive pair-wise linkage map comparison of *Malus* and *Fragaria* genomes that uses orthologous markers and linkage mapping analysis as tools to evaluate the degree of synteny between these two phylogenetically distant genera.

2.3 Materials and methods

2.3.1 Segregating populations used for RosCOS mapping in apple

Two apple pseudo-test cross (Grattapaglia and Sederoff 1994) progenies were used to map RosCOS markers: an inter-specific 'Malling 9' × 'Robusta 5' (M.9xR5) population (Celton et al. 2009b) and an intra-specific 'Royal Gala' × 'Braeburn' (RGxBB) population (Velasco et al. 2010). The selection of the apple populations for RosCOS genetic mapping was based on completeness of their respective maps. Three additional intra-specific populations, 'Starkrimson' × 'Granny Smith' (StKxGS) (Segura et al. 2007), 'Fuji' × MIS OP 93.051 (FxMIS) (Bus et al. 2010), and 'Royal Gala' × A689/24 (RGxA689) (Chagné et al. 2008) were used to map RosCOS markers that were not polymorphic in the first two populations (Table 2-1). Apple LG numbering (1 to 17) follows Maliepaard et al. (1998) and homeologous pair designation follows Velasco et al. (2010).

Table 2-1: Description of apple segregating populations used for Rosaceae conserved orthologous set (RosCOS) genetic mapping.

Population	Type of cross	Abbreviation	Total individuals	Individuals used for mapping	Saturated map?	Primary type of marker	Citation
'Malling 9' × 'Robusta 5'	Interspecific	M.9xR5	218	94	yes	SSR	(Rusholme Pilcher et al. 2008)
'Royal Gala' × 'Braeburn'	Intraspecific elite cultivars	RGxBB	572	91	yes	SNP	(Velasco et al. 2010)
'Starkrimson' × 'Granny Smith'	Intraspecific elite cultivars	StKxGS	125	125	yes	SSR	(Segura et al. 2007)
'Fuji' × MIS OP 93.051	Intraspecific	FxMIS	153	94	no	various	(Bus et al. 2010)
'Royal Gala' × A689/24	Intraspecific	RGxA689	176	94	yes	SNP	(Chagné et al. 2008)

2.3.2 RosCOS marker development for use with HRM in *Malus*

Two methods were used for developing RosCOS markers specifically for use in *Malus* with the high-resolution melting (HRM) analysis technique. For both marker design strategies, polymerase chain reaction (PCR) primers were designed using the GD apple genome sequence and were selected using Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi), using default parameters with the following exceptions: amplicon size of 75-300 base pairs (bp), GC content 40 to 55% (alternative 30-60%). Maximum self-complementarity and maximum 3′ self-complementarity of 4 and 1, respectively, were chosen to reduce the formation of primer dimers. All newly designed primer sequences were compared back to the GD apple draft genome sequence using BLASTN (Altschul et al. 1990) to evaluate amplicon potential and confirm primer specificity. If more than one match was found, primers were redesigned or discarded.

2.3.2.1 Marker Design Method 1: Intron-spanning

PCR primer design focused on a subset of 126 RosCOS markers bin mapped in the TxE (Dirlewanger et al. 2004b) and FVxFB (Sargent et al. 2008) bin mapping sets (Illa et al. 2011). BLASTN local alignment of the RosCOS markers on the GD apple draft WGS was used to identify introns of less than 300 bp and primers were designed to span one intron, and be specific to apple homeologous loci, if more than one significant match (minimum E-value of 10^{-20}) was identified.

2.3.2.2 Marker Design Method 2: SNP

Primer pairs were designed to target computationally predicted SNP (T. Foster, unpublished) between GD, M.9 and R5 located within the regions of homology between the 126 RosCOS marker sequences and the apple WGS. Primer pairs were designed to target the SNP and five bp on either side for a PCR product length of 75-300 bp.

2.3.2.3 PCR and HRM conditions

Polymerase chain reactions were carried out in a Roche LightCycler® 480 in volumes of either 7 µl (384-well plates) or 10 µl (96-well plates). Reaction master mix comprised 1× Roche® master mix, 2.5 mM $MgCl_2$, and 0.2 µM each forward and reverse primer, and 1 ng genomic DNA. T_m was set at 50ºC, 55ºC, or 60ºC for 40-50 cycles depending on the primer pair performance during initial screening. The PCR amplification program started with denaturing at 95ºC for 5 min, followed by 95ºC for 10 sec, T_m for 30 sec and extension at 72ºC for 15 sec. The PCR cycles were followed by HRM of 95ºC for 1 min (ramp rate 4.4ºC/sec), cooling to 40ºC for 1 min (ramp rate 1.5ºC/sec) to form duplexes, an increase to 65ºC (ramp rate 1ºC/sec), and a final melting increasing to 95ºC at 0.2ºC/sec with 25 data acquisitions/ºC for 20 min. Melting curves were analyzed using the Roche LightCycler® 480 *gene scanning* module.

2.3.3 Data collection and linkage analysis

RosCOS primer pairs were first screened for the production of polymorphic PCR products in the M.9xR5 bin set. If the primer pair did not amplify, or the product was monomorphic, or both parents were homozygous, the primer pair was tested in at least one other population subset (either a bin set or 14 individuals). RosCOS primer pairs that amplified a polymorphism were then screened over their respective population sets and genotypes scored based on the differences in HRM melting curves (Wittwer et al. 2003).

JoinMap® v3.0 (van Ooijen and Voorrips 2001) software was used to locate polymorphic RosCOS markers to the existing linkage maps constructed for the apple populations used in this study (Bus et al. 2010; Celton et al. 2009b; Chagné et al. 2008; Segura et al. 2007; Velasco et al. 2010). The LOD threshold for grouping was 5 and the Kosambi mapping function was used to convert recombination units into genetic distances.

2.3.4 PCR fragment sequencing

The PCR fragments for RosCOS markers that mapped to a genetic locus different from their *in silico* location were directly sequenced in the forward direction only using BigDye Terminator v. 3.1 per manufacturer's instructions (BigDye™ Terminator Version 3.1 Ready Reaction Cycle Sequencing Kit, Applied Biosystems, Inc.). The sequences obtained were compared against the GD apple draft WGS using BLASTN.

2.3.5 RosCOS genome coverage in apple

A goal of mapping at least two RosCOS markers per LG for LG orientation and apple genome coverage was established. Construction of integrated LG was necessary to evaluate fully the RosCOS marker coverage over linkage groups and, as the RGxBB genetic map (based mainly on SNP markers developed from the GD SNPlex data) (Velasco et al. 2010) and the M.9xR5 genetic map (comprised of 224 SSR, 18 sequence-characterized amplified regions, 14 SNP, and 42 random amplified polymorphic DNA markers) (Celton et al. 2009b) share very few markers, 95 primer pairs for HRM-based markers were designed from sequence flanking the SNP screened by SNPlex for each RGxBB LG. Markers were selected for placement along the length of the LG, then evaluated for polymorphism in the M.9xR5 bin set, and those that were polymorphic were screened over 94 individuals from this population. Additionally, 12 published SSR markers, six for each LG, were evaluated for polymorphism in 14 RGxBB progeny in an attempt to increase the number of markers available for merging LG 8 and also LG 13. The LG from StKxGS, FxMIS and RGxA689 were merged with LG from either RGxBB or M.9xR5 using existing SSR markers in common using the *Map Integration* function of JoinMap® v3.0. A minimum of three markers in common was needed to create integrated LG successfully.

2.3.6 *In silico* detection of *Malus* SSR sequences in the *Fragaria vesca* WGS

To assess the potential for transferability of *Malus* SSR to *F. vesca*, the flanking and primer sequences of 103 apple SSR markers developed from genomic DNA from 'Florina' (HiDRAS ; Liebhard et al. 2002) were searched for within the *Fragaria* whole genome assembly scaffolds (Shulaev et al. 2011) using BLASTN. The degree of sequence similarity of each *Malus* SSR fragment to *F. vesca* sequence was assessed using the score and E-values of the two best matches. Presence or absence of the SSR

motif in *F. vesca* was noted. Amplification ability of forward and reverse primers was assessed using the score and E-values of the best match for each.

2.3.7 *In silico* detection of strawberry gene-based markers in the GD apple draft whole genome sequence

Sequences for the 30 gene-based markers used to construct the FVxFB linkage map were obtained from GenBank. The sequences were compared to the GD apple draft whole genome sequence using BLASTN and the two best BLASTN matches used to indicate physical position in the apple WGS.

2.3.8 Genetic map alignment and assignment of ancestral syntenic blocks

RosCOS markers mapped on the apple genetic maps and FVxFB bin map were used to assess the degree of ancestral relationship between *Malus* and *F. vesca*. The FVxFB bin mapping set is derived from the F_2 mapping population of 65 seedlings (Sargent et al. 2004) and consists of an F_1 hybrid individual, six F_2 progeny, and the *F. vesca* grandparent (Sargent et al. 2008). The F_1 and *F. vesca* grandparent, which is more homozygous than the *F. bucharica* grandparent, were included to assist in interpretation of the genotypes of the F_2 individuals (Sargent et al. 2008). The FVxFB bin map is based on *Fragaria*-derived SSR, EST and gene sequences, and four ripening-related genes from *Prunus* (Sargent et al. 2008) and consists of seven LG (I to VII), one representing each chromosome of the *Fragaria* haploid genome and following the numbering convention of Sargent *et al.* (Sargent et al. 2004). A subset of 126 RosCOS markers, from the 613 markers mapped in TxE (Cabrera et al. 2009) was independently mapped on the FVxFB bin map (Illa et al. 2011).

The ancestral chromosome numbers are as described in Illa et al. (2011). Additional information for ancestral relationships was achieved from the analysis of sequence similarity of strawberry gene-based markers to the GD apple draft whole genome as described above. Ancestral relationship was supported if two or more RosCOS markers, or one or more RosCOS and one or more *in silico* apple orthologs were located on the same apple LG.

2.4 Results

2.4.1 RosCOS marker development for use with HRM in *Malus*

Of the 126 RosCOS markers bin mapped in FVxFB (Illa et al. 2011), 113 were also identified in the GD apple draft WGS (Velasco et al. 2010) (Table 2-2; Supplemental Table S2-1). Of the 113 markers, 55 were detected using BLASTN on more than one LG, 34 of which were on homeologous LG. The remaining 58 markers were found on one LG only. The 113 markers were used to design primer pairs specifically for use in apple (Table 2-3, Supplemental Table S2-1). A total of 122 pairs of intron-spanning PCR primers were developed that covered 96 unique RosCOS loci, 19 of which were identified on two different LG. Fifty-one of the intron-spanning primer pairs were mapped in at least one of the five apple segregating populations. Additionally, 19 pairs of potential SNP PCR primers were designed that covered 12 unique loci including one locus that was identified on homeologous LG. Five of the SNP primers pairs were mapped in at least one of the five apple segregating populations. In total, of 141 primer pairs covering 108 loci, 56 mapped in at least one of the five apple segregating populations. Three RosCOS markers (RosCOS509, RosCOS1341 and RosCOS1354) were mapped on both of their homeologs using specific primer pairs designed for each homeolog (Supplemental Table S2-2). At least two new RosCOS markers were placed on all apple LG except LG 1 and LG 9.

Table 2-2: Rosaceae conserved orthologous set (RosCOS) marker distribution and density on *Fragaria vesca* × *F. bucharica* (FVxFB I-VII) linkage groups (LG) in centimorgans (cM). Length of FVxFB LG and coverage of RosCOS mapped in apple is included to give approximate equivalent to the genome coverage of apple. Bioinformatics was used to located RosCOS markers on the 'Golden Delicious' (GD) apple genome, and the DNA sequence information was used to design primer pairs suitable for use with high-resolution melting (HRM) analysis.

FVxFB LG	RosCOS/ LG	cM/ RosCOS	LG length in cM	Coverage of RosCOS in common with apple in cM	RosCOS-containing bins	RosCOS-containing bins w/mapped marker	RosCOS found in GD genome database	RosCOS markers suitable for HRM primer design
I	9	10	90	69	4	2	8	6
II	27	2.88	75	67	6	6	27	22
III	19	3.17	57	44	4	3	18	13
IV	10	8	80	42	4	2	9	7
V	30	2.7	81	44	5	5	26	21
VI	20	6.05	115	108	6	6	15	11
VII	11	8.18	90	43	3	2	10	6
Totals	**126**		**588**	**417**	**32**	**26**	**113**	**86**

Table 2-3: Rosaceae conserved orthologous set (RosCOS) marker performance based on method of primer design. Categories were based on marker performance in one or more apple populations.

	Total Primer Pairs Designed	Unique Loci Covered[a]	Loci Identified on Homeologous LG[b]	Loci Identified on Non-homeologous LG[b]	Loci covered by both design methods	Unique Loci[c]	Monomorphic	Monomorphic/ Complex[d]	Complex	Polymorphic Not Mapped[e]	No Amplification[f]	% Mapping Success
Intron-spanning	122	96	14	5		51	19	6	6	3	36	53
SNP	19	12	1	0		5	4	0	3	1	3	42
Total	141	108	15	5	2	56	23	6	9	4	39	

[a]Includes RosCOS markers with sequence homology on two different apple linkage groups (LG)
[b]Subset of Unique Loci; *in silico* identification using BLASTN
[c]Includes three RosCOS markers mapped to homeologous linkage groups
[d]Monomorphic/Complex indicates that the marker was monomorphic in one apple population and had a complex pattern in a second apple population
[e]Polymorphic Unmapped indicates markers with a clear segregation pattern that did not map in any tested apple population
[f]No Amplification indicates that the marker failed to amplify a product in two or more apple populations

2.4.2 Confirmation of the genetic map position of RosCOS compared to the draft apple genome

The genetic locations of 12 out of 56 mapped RosCOS markers (21.4%) were different from their physical locations, based on best BLASTN matches (E-values ranging from 0.004 to E-110), in the GD apple draft whole genome sequence (Supplemental Table S2-3). Six of the markers mapped in a single parent of a single population, three markers mapped in two different populations, and three markers mapped in both parents of the same population. These regions were sequenced to determine if the primer pair amplified the correct target. Alignment of the sequences to the RosCOS markers using BLASTN revealed that the primers amplified their targets, indicating that the genetic mapping of these 12 markers was accurate.

2.4.3 Genome coverage of RosCOS in apple

Five apple segregating populations were used for the synteny comparison of apple and strawberry. In every case, more than one apple population was needed to achieve the goal of mapping at least two RosCOS per LG (Table 2-4). This required the merging of the LG of the different apple populations to create integrated maps of the 15 apple LG on which RosCOS markers were placed. The resulting integrated apple genetic map of 15 LG was 1277.5 cM, of which 629.1 cM (49%) was covered by RosCOS markers (Table 2-5). There were too few markers in common to integrate LG 8 between M.9xR5 and RGxBB consensus maps leaving RosCOS1125 out of consensus LG 8. For the same reason, LG 10 could not be integrated between M.9xR5, StK and RGxBB leaving RosCOS1341 out of integrated LG 10.

Table 2-4: Apple parents from five segregating populations needed to establish a minimum of two Rosaceae conserved orthologous set (RosCOS) markers per linkage group (LG). RosCOS markers were mapped on 15 of the total 17 apple LG.

Apple Parent and LG	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
'Malling 9'		x	x	x	x	x	x	x		x	x	x	x	x	x	x	x
'Robusta 5'		x	x	x	x	x		x		x	x	x	x	x		x	x
'Royal Gala'		x	x	x	x			x				x	x	x		x	
'Braeburn'		x		x	x	x		x			x	x	x	x			
A689/24															x		
'Starkrimson'										x							
'Granny Smith'																	x
'Fuji'							x										

Table 2-5: Rosaceae conserved orthologous set (RosCOS) marker coverage, in centimorgans (cM), of apple consensus linkage groups (LG).

Apple Consensus LG	RosCOS Coverage (cM)	Total LG (cM)	% RosCOS Coverage	No. RosCOS Markers
2	44.2	89	50	3
3	55.7	88.5	63	4
4	68.1	68.1	100	3
5	50.8	87.7	58	6
6	79.3	79.3	100	6
7	22.3	74	30	2
8	30.9	77.9	40	3[a]
10	31.5	63.2	50	4[a]
11	18.9	93.3	20	4
12	20.7	77.2	27	4
13	14.8	106.7	14	5
14	18.7	66	28	4
15	111.2	133.2	83	2
16	20	82.6	24	2
17	42	90.8	46	4
Totals	**629.1**	**1277.5**	**49**	**56[b]**

[a]Indicates markers not on integrated map and therefore not included in cM calculation
[b]Includes three homeologous markers

2.4.4 Genome coverage of RosCOS in FVxFB

Based on the positions of the 56 RosCOS markers in common between FVxFB and apple LG, approximately 71% (417 cM of 588 cM total length) of the FvxFB reference map was covered (Table 2-2). FVxFB LG VII had the least coverage with 48% (43 cM

of 90 cM total length), while FVxFB LG VI had the greatest coverage with 94% (108 cM of 115 cM total length).

2.4.5 *In silico* detection of *Malus* SSR sequences in the *Fragaria vesca* genome

A BLASTN search of *F. vesca* scaffolds, conducted with both primers and amplicon sequences for 103 mapped *Malus* SSR markers developed from genomic DNA from 'Florina' (HiDRAS ; Liebhard et al. 2002), revealed only 13 markers with an E-value of less than $6.00E^{-11}$ and sequence similarity of greater than 85% (Supplemental Table S2-4). In no case were the best primer matches on the same genome assembly scaffold as the best amplicon sequence match and furthermore, the SSR repeat was never present in the *F. vesca* scaffolds.

2.4.6 *In silico* detection of strawberry gene-based markers in the draft apple whole genome sequence

In addition to gene-based RosCOS, other sequences corresponding to genes in strawberry were identified in apple (Deng and Davis 2001; Medina-Escobar et al. 1997; Sargent et al. 2009b; Sargent et al. 2007a). When 28 gene-based markers used to construct the FVxFB linkage map were used for the BLASTN analysis to identify putative apple gene orthologs (Supplemental Table S2-5), three markers (*ADH*, *ANS*, and *CHS*) were identified only once in the apple genome. Of the remaining markers, 12 were identified in duplicate on homeologous apple LG, 12 were identified in duplicate on non-homeologous apple LG, while three copies of marker *CEL-1* were identified on three non-homeologous apple LG. The ancestral derivations of the apple orthologs generally agreed with the adjacent RosCOS markers with the exception of *APB*, which was located on FVxFB LG II:08 and homeologous apple LG 3 and 11 (Supplemental Fig. S2.1a); and *ZIP*, which was located on FVxFB LG VI:72 and non-homeologous apple LG 3 and 8 (Supplemental Fig. S2.1d).

2.4.7 Genetic map synteny between *Malus* and *Fragaria vesca*

Genetic mapping of 56 RosCOS markers identified 21 regions of synteny of *Malus* with *F. vesca* and covered 49% of the apple genome. At least two RosCOS markers were found in common between apple and strawberry on nine of the twelve apple homeologs and each of the seven *Fragaria* LG (Table 2-6, Fig. 2.3: Syntenic genomic regions between *Malus* and *Fragaria vesca* as determined by genetic mapping of 56 Rosaceae

conserved orthologous set (RosCOS) markers in common between the two genera). In total, 12 homeologous blocks created by genome duplication were used for the comparison of apple and strawberry. RosCOS marker density on the FVxFB linkage map ranges from one marker every 2.7 cM on LG V to one marker every 10 cM on LG I (D. Sargent, pers. comm.) (Table 2-2). Each FVxFB LG is composed of regions of at least two homeologous apple LG. Seven apple homeologous LG on which RosCOS markers were located are composed of regions from at least two FVxFB LG. Five apple homeologous LG are composed of regions from single FVxFB LG (Fig. 2.3). No RosCOS markers that demonstrate homeology between the distal parts of LG 1 and 7, and proximal part of LG 2 and central part of LG 15 were mapped (Fig. 2.2). FVxFB LG I, IV, and VII have gaps within which no RosCOS markers were placed (Table 2-7).

Table 2-6: Distribution of RosCOS markers mapped on apple homeologous linkage groups (LG) and their *Fragaria vesca* × *F. bucharica* (FV×FB) orthologous LG.

Apple homeologs	1/7	2/7	2/15	3/11	4/6	4/12	5/10	6/14	8/15	9/17	12/14	13/16
Mapped RosCOS markers	0[a]	5	0[a]	8[b]	1	6[b]	10[bc]	7	5[c]	4[d]	3	7
FVxFB orthologous LG		I, II, VI, VII		III	IV, V	I, VI	II	V	I, II, V	III, VI	I,IV, VI	V

[a]Indicates no markers mapped
[b]Indicates homeologous markers mapped on homeologous LG
[c]Indicates marker not on integrated map
[d]Indicates markers located on only one homeolog

Table 2-7: Distribution of 56 Rosaceae conserved orthologous set (RosCOS) markers common between FVxFB linkage groups (LG) and apple homeologs. FVxFB LG are listed on the y-axis, apple paralogs on listed on the x-axis. Each cell is populated with the number of orthologous RosCOS markers. Bold font indicates contiguous markers.

FVxFB LGI		1				1			1			
FVxFB LGII		1					**10**[abc]		**3**[c]			
FVxFB LGIII				**8**[b]						**2**		
FVxFB LGIV					1						1	
FVxFB LGV						1		**7**	1			**7**
FVxFB LGVI		1				**4**[b]				**2**	2	
FVxFB LGVII		2										
	1/7	2/7	2/15	3/11	4/6	4/12	5/10	6/14	8/15	9/17	12/14	13/16

[a]Indicates markers are divided into two regions on LG II, one region with six contiguous markers, and one region with four contiguous markers
[b]Indicates homeologous markers mapped on homeologous LG
[c]Indicates markers not on integrated map

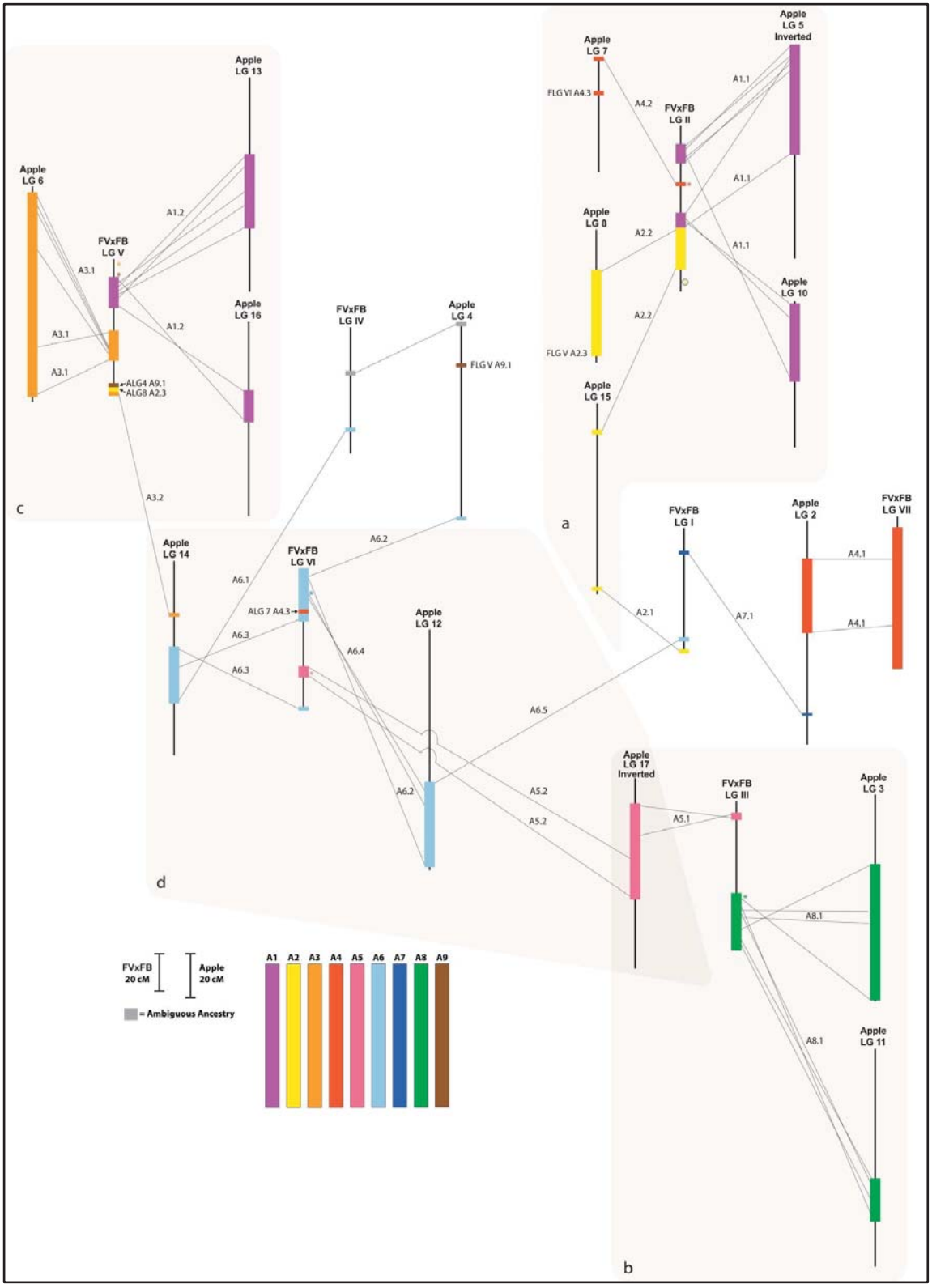

Apple LG 13
Apple LG 6
FVxFB LG V
A1.2
A3.1
A3.1
A1.2
A3.1
Apple LG 16
ALG4 A9.1
ALG8 A2.3
A3.2
c
FVxFB LG IV
Apple LG 4
FLG V A9.1
A6.2
Apple LG 14
FVxFB LG VI
A6.1
ALG 7 A4.3
A6.3
A6.3
A6.4
Apple LG 12
A6.5
A6.2
A5.2
A5.2
d
Apple LG 7
FLG VI A4.3
A4.2
FVxFB LG II
Apple LG 5 Inverted
A1.1
A1.1
A1.1
Apple LG 8
A2.2
A2.2
Apple LG 10
FLG V A2.3
Apple LG 15
a
FVxFB LG I
A2.1
A7.1
Apple LG 2
FVxFB LG VII
A4.1
A4.1
Apple LG 17 Inverted
FVxFB LG III
A5.1
Apple LG 3
A8.1
A8.1
Apple LG 11
b
FVxFB 20 cM
Apple 20 cM
= Ambiguous Ancestry
A1
A2
A3
A4
A5
A6
A7
A8
A9

Fig. 2.3: Syntenic genomic regions between *Malus* and *Fragaria vesca* as determined by genetic mapping of 56 Rosaceae conserved orthologous set (RosCOS) markers in common between the two genera (see also Supplemental Fig. S2.1 and Supplemental Tables S2-1, S2-2 and S2-5). Linkage maps of *Fragaria vesca* × *F. bucharica* (FVxFB) linkage groups (LG) I-VII and apple integrated LG 2-8 and 10-17 are shown. Connecting lines illustrate numbers and locations of RosCOS markers in common. Color blocks and notations above connecting lines indicate the ancestral relationship based on Vilanova et al. (2008) and Illa et al. (2011). For example, A1.1 indicates derivation from Ancestral chromosome 1 (A1) and one instance of occurrence (.1). Solid black vertical lines indicate regions in which RosCOS markers were not mapped in either apple or strawberry. Presence of an * followed by ancestral derivation indicates support from *in silico* location of apple orthologs and their approximate position on the FVxFB LG.

a: FVxFB LG II and apple homeologs LG 5 and 10, LG 8 and 15, and LG 7

Synteny: Genetic mapping of nine RosCOS markers supports synteny between LG II and LG 5 and 10 (pink), as proposed by Illa et al. (2011). Synteny with LG 8 and 15 (yellow) is supported by the genetic mapping of two RosCOS. Synteny with apple LG 7 is suggested by the genetic mapping of one RosCOS marker.

Ancestry: Apple LG 5 is inverted with respect to LG 10 and LG II to demonstrate the common ancestry through A1 (A1.1) among the three LG, as proposed by Illa et al. 2011. A previously unknown ancestral relationship through A2 of LG II with LG 8 (A2.2) and 15 (A2.2) is further supported by the *in silico* location of three apple orthologs on LG 8 (yellow * on LG II). A previously unknown ancestral relationship through A4 of LG II with LG 7 (A4.2) is further supported by the in *silico* location of an apple ortholog on LG 7 (red * on LG II).

b: FVxFB LG III and apple homeologs LG 3 and 11, and LG 17

Synteny: Genetic mapping of eight RosCOS markers supports synteny between LG III and LG 3 and 11 (green), as proposed by Illa et al. (2011). Synteny with LG 17 is supported by the genetic mapping of two RosCOS markers (salmon).

Ancestry: The genetic mapping of eight RosCOS and the *in silico* location of one apple ortholog on LG 3 and 11 (green * on LG III) supports a change in orientation of LG 3 and 11 relative to that proposed by Illa et al. (2011) and suggests the ancestral relationship through A8 (A8.1) of the bottom portions of all three LG. The ancestral relationship through A5 (A5.1) suggests a previously unknown translocation.

c: FVxFB LG V and apple homeologs LG 6 and 14, and LG 13 and 16

Synteny: Genetic mapping of seven RosCOS markers supports synteny between LG V and homeologs LG 6 and 14 (orange), as proposed by Illa et al. (2011). Synteny between LG V and homeologs LG 13 and 16 (pink) is supported by the genetic mapping of seven RosCOS markers. Synteny with apple LG 4 is suggested by the genetic mapping of one RosCOS marker (brown).

Ancestry: The ancestral relationship through A3 of LG V and LG 6 (A3.1) and 14 (A3.2) is further supported by the *in silico* location of two apple orthologs on LG 6 and one apple ortholog on LG 14 (orange * on LG V). The insertion of a segment

of A1 (A1.2) within the region derived from A3, as proposed by Illa et al. (2011), is supported with these data. A previously unknown relationship through A9 (A9.1) is suggested by the genetic mapping of one RosCOS marker on LG 4 (brown) and the *in silico* location of one apple ortholog on LG 5 and 10 (brown * on LG V).

d: FVxFB LG VI and apple homeologs LG 12 and 14, and LG 17

Synteny: Support for synteny between LG VI and LG 12 and 14 comes from the genetic mapping of three RosCOS on LG 12 and two on LG 14 (light blue), as suggested by Illa et al. (2011). The genetic mapping of two RosCOS markers from LG 17 (salmon) supports synteny as suggested by Illa et al. (2011). The two RosCOS mapped to LG 14 suggest an inversion in FVxFB relative to apple, as a single change in the orientation of the lower part of LG VI (bins 68-115) would bring all RosCOS markers from homeologous LG 12 and 14 together.

Ancestry: The ancestral relationship through A6 of LG VI and LG 12 (A6.2) and 14 (A6.3) is further supported by the *in silico* location of two apple orthologs on LG 12 (light blue * on LG VI). The ancestral relationship through A5 of LG VI and LG 17 is further supported by the *in silico* location of two apple orthologs on LG 17 (salmon * on LG VI). The genetic mapping suggests a change in orientation of the ancestral regions, as proposed by Illa et al. (2011), such that the tops of LG VI and LG 12 and 14 are derived from A6, and the lower parts of LG VI and LG 17 are derived from A5.

2.4.7.1 FVxFB LG II and apple homeologs LG 5 and 10, LG 8 and 15, and LG 7

Genetic mapping of nine RosCOS markers supports synteny between FVxFB LG II and apple LG 5 and 10 (Fig. 2.3a), as proposed by others (Illa et al. 2011). Synteny with apple LG 8 and 15 is supported by the genetic mapping of two RosCOS markers. Synteny with apple LG 7 is suggested by the genetic mapping of a single RosCOS marker. Apple LG 5 is inverted with respect to apple LG 10 and FVxFB LG II to demonstrate the common ancestry through A1 (A1.1) among the three LG as proposed by (Illa et al. 2011). A previously unknown ancestral relationship through A2 of FVxFB LG II with apple LG 8 and 15 (A2.2) is further supported by the sequence alignment on LG 8 of three apple orthologs from LG II (yellow * on LG II, Fig. 2.3a). A second previously unknown ancestral relationship through A4 of FVxFB LG II with apple LG 7 (A4.2) is further supported by the *in silico* detection of an apple ortholog on apple LG 7 from LG II (red * on LG II, Fig. 2.3a).

2.4.7.2 FVxFB LG III and apple homeologs LG 3 and 11, and LG 17

Genetic mapping of eight RosCOS markers supports synteny between FVxFB LG III and apple LG 3 and 11 (Fig. 2.3b), as proposed by (Illa et al. 2011). Synteny with apple LG 17 is supported by the genetic mapping of two RosCOS markers. The genetic mapping of eight RosCOS (A8.1) and the *in silico* detection of a single apple ortholog (green * on LG III, Fig. 2.3b) on apple LG 3 and 11 from FVxFB LG III supports a change in orientation of apple LG 3 and 11 relative to that proposed by (Illa et al. 2011). Genetic mapping suggests the ancestral relationship through A8 of the lower sections of all three LG. The ancestral relationship through A5 (A5.1) suggests a previously unknown translocation.

2.4.7.3 FVxFB LG V and apple homeologs LG 6 and 14, and LG 13 and 16

Genetic mapping of seven RosCOS markers supports synteny between FVxFB LG V and apple LG 6 and 14 (Fig. 2.3c), as proposed by (Illa et al. 2011). Synteny between FVxFB LG V and apple LG 13 and 16 is supported by the genetic mapping of seven RosCOS markers. The ancestral relationship through A3 of FVxFB LG V and apple LG 6 (A3.1) and apple LG 14 (A3.2) is further supported by the *in silico* detection of two apple orthologs on apple LG 6 and one apple ortholog on apple LG 14 from FVxFB LG V (orange * on LG V, Fig. 2.3c). The insertion of a segment of A1 (A1.2) within the region derived from A3 as proposed by (Illa et al. 2011) is supported by these data and

identifies a previously unknown ancestral relationship of apple LG 13 and 16 through A1. A previously unknown relationship through A9 (A9.1) is suggested by the genetic mapping of one RosCOS marker and the *in silico* detection of one apple ortholog on apple LG 4 from FVxFB LG V (brown * on LG V, Fig. 2.3c).

2.4.7.4 FVxFB LG VI and apple homeologs LG 12 and 14, and LG 17

Support for synteny between FVxFB LG VI and apple LG 12 and 14 comes from the genetic mapping of three RosCOS on apple LG 12 and two on apple LG 14 (Fig. 2.3: Syntenic genomic regions between *Malus* and *Fragaria vesca* as determined by genetic mapping of 56 Rosaceae conserved orthologous set (RosCOS) markers in common between the two genera d), as suggested by (Illa et al. 2011). The genetic mapping of two RosCOS markers from LG 17 supports synteny, as suggested by (Illa et al. 2011). The two RosCOS mapped to apple LG 14 suggest an inversion in FVxFB relative to apple, as a single change in the orientation of the bottom of FVxFB LG VI (bins 68-115) would bring all RosCOS markers from apple homeologous LG 12 and 14 together. The ancestral relationship through A6 of FVxFB LG VI and apple LG 12 (A6.4) and apple LG 14 (A6.3) is further supported by the *in silico* detection of two apple orthologs on apple LG 12 from FVxFB LG VI (light blue * on LG VI, Fig. 2.3d). The ancestral relationship through A5 of FVxFB LG VI and apple LG 17 (A5.2) is further supported by the *in silico* detection of two apple orthologs on apple LG 17 from FVxFB LG VI (salmon * on LG VI, Fig. 2.3d). The genetic mapping suggests a change in orientation of the ancestral regions as proposed by (Illa et al. 2011), such that the tops of FVxFB LG VI and apple homeologs LG 12 and 14 are derived from A6, and the bottoms of FVxFB LG VI and apple LG 17 are derived from A5.

2.5 Discussion

2.5.1 Orthologous marker development in Rosaceae

The use of molecular markers for comparative genomics in Rosaceae is a relatively new field of study when compared to other plant families such as Solanaceae, Brassicaceae, and Poaceae. COS markers have been used for comparative genomic studies among sister clades in Solanaceae (Wu and Tanksley 2010) and Poaceae (Bolot et al. 2009; Ventelon et al. 2001). This study provides a baseline for the use of RosCOS markers for comparative mapping between apple and strawberry, as well as general information on the design and application of these markers in the large and diverse Rosaceae. The 56

RosCOS markers used in this study comprise a set of transferable, gene-based markers for inter-tribal comparative mapping in the Rosaceae. These markers have the potential to be incorporated into existing genetic maps, and are instrumental for alignment of sequenced scaffolds and anchoring of genomic assemblies to genetic maps. RosCOS markers provide an advantage over both of the two marker types commonly used for comparative mapping, restriction fragment length polymorphism (RFLP) probes (Acarkan et al. 2000; Ahn and Tanksley 1993; Dirlewanger et al. 2004b; Lagercrantz 1998; Lagercrantz et al. 1996; Tanksley et al. 1988; Van Deynze et al. 1995; Ventelon et al. 2001; Krutovsky et al. 2004), and genomic SSR (Dirlewanger et al. 2004a; Dirlewanger et al. 2004b; Lewers et al. 2005; Olmstead et al. 2008; Stafne et al. 2005). While RFLP markers are transferable and co-dominant, they are time consuming and labor intensive to use and their analysis requires relatively large amounts of genomic DNA (Gupta et al. 2002b). The comparison of *Malus* genomic SSR markers with the *Fragaria vesca* whole genome scaffold sequences revealed that none of the sequences was found on the same *F. vesca* scaffolds as the primer sequences, all *Malus* sequences were found fragmented over two or more *F. vesca* scaffolds, and the SSR repeat was absent in all *F. vesca* BLASTN matches. In contrast, 40% of the RosCOS markers designed for this study and analyzed by HRM mapped in both *Malus* and *F. vesca*. HRM has an advantage over the RFLP method because of its suitability for high-throughput analysis. All 15 apple LG on which RosCOS were located, and all seven FVxFB LG could be oriented using this set of RosCOS markers. For these reasons, we propose wider use be made of COS markers for comparative genome mapping across species of relatively distant phylogenetic relationships (Cabrera et al. 2009; Fulton et al. 2002; Liewlaksaneeyanawin et al. 2009), and for species that have yet to have their genomes sequenced.

2.5.2 Genetic map synteny using RosCOS and the HRM technique

A previous comparison between *Fragaria* and *Prunus* using 71 genetically mapped markers in common identified 18 regions of synteny between the genera (Vilanova et al. 2008). Another pairwise comparison was conducted between *Malus* and *Prunus*, *Prunus* and *Fragaria*, and *Malus* and *Fragaria* using 129 markers common to all three genera. These shared markers were genetically mapped using the bin sets of TxE and FVxFB as well as bioinformatically located on the apple WGS, resulting in the identification of 34 regions of synteny between *Malus* and *Prunus*, and 33 regions of synteny between

Malus and *Fragaria* (Illa et al. 2011). Our 56 RosCOS markers, from the same subset of 126 RosCOS markers that were independently genetically mapped in both *Fragaria* and *Prunus* and analyzed with the HRM technique, identified 21 regions of synteny between *Malus* and *F. vesca*. Our choice of the same subset of markers allowed us to develop a set of genetically mapped orthologous markers that, in combination with these other studies, link the three most economically important genera of Rosaceae, *Malus*, *Prunus*, and *Fragaria*. The genetic mapping of the RosCOS markers covered 49% of the apple genome and 71% of the FVxFB bin map, and allowed us to orient 15 of 17 apple LG, with the exception of apple LG 1 and LG 9. The lack of RosCOS markers mapped on apple LG 1 could be due to the relatively high degree of rearrangement which this chromosome has undergone (Velasco et al. 2010) and the relatively low number and density of RosCOS markers on orthologous FVxFB LG VII (11 RosCOS markers over 90 cM). Additionally, we constrained primer design for HRM analysis by limiting amplicon size, as small sequence variations in short amplicons (less than 300 bp) are more likely to be detected (Wittwer et al. 2003; Chou et al. 2005). This reduced the options for designing markers for apple LG 1 and LG 9. The lack of RosCOS markers genetically mapped on LG 9 could also be due to the high similarity between RosCOS sequences in LG 9 and its homeolog LG 17, which hampers the design of HRM markers specific to LG 9. Our newly designed set of RosCOS markers specifically distinguish homeologs in apple and have the potential to assist molecular mapping and linkage map alignment in other members of the subtribe Pyrinae, such as pear, loquat and quince that currently have do not have whole genome assemblies.

2.5.3 Ancestral chromosome contributions to *Malus* and *Fragaria* linkage groups

We have used our results to build on previous studies that assessed the extent of synteny between *Malus* and *F. vesca* (Illa et al. 2011; Vilanova et al. 2008). The set of 126 RosCOS markers that was common to all three genera identify conserved syntenic blocks and a model was developed for the ancestral genome of the progenitor of the Rosaceae (A1-9, Fig. 2.3) (Illa et al. 2011; Vilanova et al. 2008; Velasco et al. 2010). It is hypothesized that the extant *Prunus* genome arose from three translocation events that occurred in an ancestral set of nine similarly sized chromosomes, and the extant *Fragaria* genome arose from four translocation events. Similarly, the extant *Malus* genome arose from seven translocation events, with two of the events preceding the

whole genome duplication in the ancestor of Pyrinae (Illa et al. 2011). The RosCOS marker set developed in this study has yet to be widely tested as markers for genetic mapping and genome alignment across the Rosaceae. However, we found that our 56 genetically mapped RosCOS markers were of sufficient density to validate ancestry of six of seven FVxFB LG with the exception of LG IV, and 10 of 12 apple homeologous blocks (Table 2-6). Our genetic mapping has led to the identification of differences in the proposed ancestral chromosome assignment for apple homeologs LG 3 and 11, homeologs LG 5 and 10, homeologs LG 13 and 16, and LG 12, and for FVxFB LG VI (Fig. 2.3, Supplemental Fig. S2.1) compared with what was previously suggested using *in silico* methods (Illa et al. 2011). Our results demonstrate that the lower parts of apple homeologs LG 3 and 11 and FVxFB LG III are derived from A8; the tops of apple homeologs LG 5 and 10 and FVxFB LG II are derived from A1; the central regions of apple homeologs LG 13 and 16 and the top of FVxFB LG V are derived from A1; and the base of apple LG 12 and the top of FVxFB LG VI are derived from A6, contradicting the relationships assigned by Illa et al. (2011). These specific discrepancies may be due to wrongly oriented scaffolds in the apple genome assembly indicating more research is needed to improve the apple genome assembly. The use of RosCOS that are genetically mapped in segregating populations will ensure the development of a more robust whole genome sequence of apple.

The genetic mapping of RosCOS markers identified an insertion of A5 (A5.1) into FVxFB LG III from apple LG 17, and insertions of A4 (A4.2) and A2 (A2.2) into FVxFB LG II from apple LG 7 and homeologs LG 8 and 15, respectively. In addition, the genetic mapping of RosCOS markers identified two potential insertions from A9, and a possible inversion of A3 on FVxFB LG V. On FVxFB LG VI, the placement of RosCOS markers identified a potential inversion determined by the disruption in FVxFB bin order. These insertions were not detected by Illa et al. (2011) due to the criterion used, i.e. three strawberry markers within the same or adjacent bins and within 3.5 Mb of the apple genome. There are possibly still more chromosomal rearrangements between apple and strawberry than those that have been revealed by *in silico* means and further research is clearly needed to identify the precise breakpoint positions.

With 56 RosCOS markers we identified six previously unknown chromosomal differences between apple and strawberry (Illa et al. 2011). Our strategy was based on genetic mapping and relied on highly significant LOD scores that give strong statistical

support to the genome locations. Additional support was gained from sequencing of PCR products followed by BLASTN analysis. Indeed, an unexpected benefit of the comparison of apple LG to FVxFB was the ability to assess the accuracy of the recent GD apple draft genome assembly (Velasco et al. 2010) by identifying markers whose genetic placement did not agree with their strawberry orthologs. The discrepancies between *in silico* and linkage map position identified in this study emphasize the importance of genetic mapping for verifying draft genome assemblies. For example, 12 RosCOS markers, whose PCR products were subsequently verified by sequencing, did not locate as expected in apple according to analysis of best BLASTN matches to the GD apple draft genome assembly. However, when compared to the linkage map of FVxFB and the ancestral relationships of the orthologous regions, eight of these 12 markers do group with other markers derived from the same ancestral chromosome; hence, only the remaining four of these 12 markers might be placed in error on the genome assembly. Interestingly, four additional markers (RosCOS2105, RosCOS1217, RosCOS1378, and RosCOS1367) whose genetic mapping agreed with their location according to best BLASTN match but which grouped with markers derived from different ancestry were identified and this may also indicate placement errors. These discrepancies could also be due to tensions created during map integration using a linkage mapping program as marker location on linkage maps can be quite variable, especially when creating consensus or integrated maps with only a few markers in common. The *in silico* location of apple orthologs of strawberry gene-based markers on homologous apple linkage groups lends support for ancestral relationships for regions with minimal RosCOS coverage by demonstrating linkages through different types of genetic markers. In addition, we propose that the lack of ancestral support for the four gene markers (*PES*, *APB*, *ZIP*, and *CHS*) may be explained by 1) genomic rearrangement, 2) the loci being paralogs that are not derived from whole genome duplication, or 3) being placed erroneously on the GD apple draft genome assembly.

RosCOS markers were developed from an EST database of Rosaceae unigene sequences and compared against *Arabidopsis* to identify those sequences that were orthologous to single copy genes in *Arabidopsis* (Cabrera et al. 2009). As EST database and transcriptome coverage increases across the Rosaceae, this marker type will become a valuable addition to the molecular toolbox. This study demonstrates the utility of COS

for designing orthologous markers to construct framework maps for comparative genomics in phylogenetically distant family members.

2.6 Conclusions

The goals of this study were to evaluate the genome synteny of two phylogenetically distant members of Rosaceae (*Malus* and *Fragaria*), and to test the applicability of RosCOS markers screened using the HRM technique for the ability to assess orthology between the two genera. This is the first study to combine these elements. These tools allowed us to identify 21 regions of synteny that support previous work, identify four differences in extant LG ancestry relative to the ancestral hypothesis, and identify two each of potential insertions, translocations, and inversions. A set of orthologous markers for use in genetic mapping in Rosaceae has been developed that will allow the exploration of unsequenced genomes, orientation and anchoring of genome sequences as they become available, and determination of ancestral relationships within the family.

We suggest future work in comparative mapping in Rosaceae should expand the number of RosCOS markers shared among the genera, beyond this initial set of 126 markers, to address the gaps on both the FVxFB and apple genetic maps. We recommend that RosCOS sequences should be included in high-throughput genotyping platforms. As more EST datasets and genome sequences become available, RosCOS sequences can be used to identify regions of orthology and primer pairs can be designed for genetic mapping for a range of phylogenetic levels within Rosaceae.

2.7 Acknowledgements

This work was part of a Ph.D. thesis and supported in part by the PFR Excellence Program (JMB) and a subcontract from USDA's National Institute of Food and Agriculture – Speciality Crop Research Initiative RosBREED project (2009-51181-05808). The authors would like to thank Tony Corbett of PFR for graphic design of all figures, and Baptiste Guitton and Dr Evelyne Costes (INRA) for access to the StKxGS apple population.

2.8 Authors' contributions

JMB designed and conducted the research, analyzed the data, and wrote the manuscript. AC and EVK provided sequence information for RosCOS markers and information on genetic mapping in *Prunus* prior to publication; DJS and ELG provided DNA of the

FVxFB bin mapping set and information on genetic mapping of RosCOS on FVxFB prior to publication. RV provided access to the 'Golden Delicious' apple genome sequence prior to release. RC provided technical support and advice for the **'**Golden Delicious' apple whole genome and *F. vesca* scaffolds. MT, VVS, SEG and DC provided manuscript editing, experimental design, project supervision, and leadership.

REFERENCES

Acarkan A, Roßberg M, Koch M, Schmidt R (2000) Comparative genome analysis reveals extensive conservation of genome organisation for *Arabidopsis thaliana* and *Capsella rubella*. Plant Journal 23 (1):55-62

Ahn S, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. Proceedings of the National Academy of Sciences of the United States of America 90 (17):7980-7984

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215 (3):403-410

Arumuganathan K, Earle E (1991) Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter 9 (3):208-218

Bolot S, Abrouk M, Masood-Quraishi U, Stein N, Messing J, Feuillet C, Salse J (2009) The 'inner circle' of the cereal genomes. Curr Opin Plant Biol 12 (2):119-125

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120 (4):1095-1103

Booi S, van Dyk MM, du Preez MG, Rees DJG, Labuschagné I (2005) Molecular typing of red and green phenotypes of 'Bon Rouge' pear trees, with the use of microsatellites. Acta Horticulturae (ISHS) 671:293-297

Bus V, Bassett H, Bowatte D, Chagné D, Ranatunga C, Ulluwishewa D, Wiedow C, Gardiner S (2010) Genome mapping of an apple scab, a powdery mildew and a woolly apple aphid resistance gene from open-pollinated Mildew Immune Selection. Tree Genetics & Genomes 6 (3):477-487

Cabrera A, Kozik A, Howad W, Arús P, Iezzoni A, van der Knaap E (2009) Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. BMC Genomics 10 (1):562

Campbell CS, Evans RC, Morgan DR, Dickinson TA, Arsenault MP (2007) Phylogeny of subtribe Pyrinae (formerly the Maloideae, Rosaceae): Limited resolution of a complex evolutionary history. Plant Systematics and Evolution 266 (1):119-145

Celton J-M, Chagné D, Tustin SD, Terakami S, Nishitani C, Yamamoto T, Gardiner SE (2009a) Update on comparative genome mapping between *Malus* and *Pyrus*. BMC Research Notes 2 (182)

Celton J-M, Tustin DS, Chagné D, Gardiner SE (2009b) Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. Tree Genetics & Genomes 5:93-107

Chagné D, Gasic K, Crowhurst RN, Han Y, Bassett HC, Bowatte DR, Lawrence TJ, Rikkerink EHA, Gardiner SE, Korban SS (2008) Development of a set of SNP markers present in expressed genes of the apple. Genomics 92 (5):353-358

Chen L, Zhang S, Illa E, Song L, Wu S, Howad W, Arús P, van de Weg E, Chen K, Gao Z (2008) Genomic characterization of putative allergen genes in peach/almond and their synteny with apple. BMC Genomics 9 (1):543

Chou L-S, Lyon E, Wittwer CT (2005) A comparison of High-Resolution Melting analysis with Denaturing High-Performance Liquid Chromatography for mutation scanning: Cystic fibrosis transmembrane conductance regulator gene as a model. American Journal of Clinical Pathology 124 (3):330-338

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theoretical and Applied Genetics 106 (5):912-922

Deng C, Davis TM (2001) Molecular identification of the yellow fruit color (*c*) locus in diploid strawberry: A candidate gene approach. Theoretical and Applied Genetics 103 (2):316-322

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

Dirlewanger E, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Voisin R, Poizat C, Lafargue B, Baron O, Laigret F, Kleinhentz M, Arús P, Esmenjaud D (2004a) Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid—location of root-knot nematode resistance genes. Theoretical and Applied Genetics 109 (4):827-838

Dirlewanger E, Cosson P, Tavaud M, Aranzana M, Poizat C, Zanetto A, Arús P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theoretical and Applied Genetics 105 (1):127-138

Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldera F, Cosson P, Howad W, Arús P (2004b) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proceedings of the National Academy of Sciences of the United States of America 101 (26):9891-9896. doi:10.1073/pnas.0307937101

Downey SL, Iezzoni AF (2000) Polymorphic DNA markers in black cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach, and sour cherry. Journal of the American Society for Horticultural Science 125 (1):76-80

Evans RC, Campbell CS (2002) The origin of the apple subfamily (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicated *GBSSI* genes. American Journal of Botany 89 (9):1478-1484. doi:10.3732/ajb.89.9.1478

Fulton TM, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of Conserved Ortholog Set markers for comparative genomics in higher plants. Plant Cell 14 (7):1457-1467. doi:10.1105/tpc.010479

Gisbert A, Martínez-Calvo J, Llácer G, Badenes M, Romero C (2009) Development of two loquat [*Eriobotrya japonica* (Thunb.) Lindl.] linkage maps based on AFLPs and SSR markers from different Rosaceae species. Molecular Breeding 23 (3):523-538

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

Gupta PK, Varshney RK, Prasad M (2002b) Molecular Markers: Principles and Methodology. In: Jain SM, Brar DS, Ahloowalia BS (eds) Molecular Techniques in Crop Improvement. Kluwer Academic Publishers, Dordrecht and Boston, pp 9-54

Hibrand-Saint Oyant L, Crespel L, Rajapakse S, Zhang L, Foucher F (2008) Genetic linkage maps of rose constructed with new microsatellite markers and locating QTL controlling flowering traits. Tree Genetics & Genomes 4:11-23

High-quality Disease Resistant Apples for a Sustainable Agriculture, http://users.unimi.it/hidras/. http://users.unimi.it/hidras/.

Illa E, Sargent D, Lopez Girona E, Bushakra J, Cestaro A, Crowhurst R, Pindo M, Cabrera A, van der Knaap E, Iezzoni A, Gardiner S, Velasco R, Arús P, Chagne D, Troggio M (2011) Comparative analysis of rosaceous genomes and the reconstruction of a putative ancestral genome for the family. BMC Evolutionary Biology 11 (1):9

Jáuregui B, de Vicente MC, Messeguer R, Felipe A, Bonnet A, Salesses G, Arús P (2001) A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. Theoretical and Applied Genetics 102 (8):1169-1176

Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB (2004) Comparative mapping in the Pinaceae. Genetics 168 (1):447-461. doi:10.1534/genetics.104.028381

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150 (3):1217-1228

Lagercrantz U, Putterill J, Coupland G, Lydiate D (1996) Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. Plant Journal 9 (1):13-20

Lambert P, Hagen LS, Audergon JM, Arús A (2006) Evaluation of a mapping strategy based on synteny: Using the *Prunus* reference map for apricot. Acta Horticulturae (ISHS) 701:295-301

Lewers KS, Styan SMN, Hokanson SC, Bassil NV (2005) Strawberry GenBank-derived and genomic simple sequence repeat (SSR) markers and their utility with strawberry, blackberry, and red and black raspberry. Journal of the American Society for Horticultural Science 130 (1):102-115

Liebhard R, Gianfranceschi L, Koller B, Ryder CD, Tarchini R, Van De Weg E, Gessler C (2002) Development and characterisation of 140 new microsatellites in apple (*Malus* x *domestica* Borkh.). Molecular Breeding 10 (4):217-241

Liewlaksaneeyanawin C, Zhuang J, Tang M, Farzaneh N, Lueng G, Cullis C, Findlay S, Ritland C, Bohlmann J, Ritland K (2009) Identification of COS markers in the Pinaceae. Tree Genetics & Genomes 5 (1):247-255

Maliepaard C, Alston FH, Van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, Van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, Den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrielink-van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theoretical and Applied Genetics 97:60 - 73

Medina-Escobar N, Cárdenas J, Moyano E, Caballero JL, Muñoz-Blanco J (1997) Cloning, molecular characterization and expression pattern of a strawberry ripening-specific cDNA with sequence homology to pectate lyase from higher plants. Plant Molecular Biology 34 (6):867-877. doi:10.1023/a:1005847326319

Moore G, Devos KM, Wang Z, Gale MD (1995) Grasses, line up and form a circle. Current Biology 5:737-739

Olmstead J, Sebolt A, Cabrera A, Sooriyapathirana S, Hammar S, Iriarte G, Wang D, Chen C, van der Knaap E, Iezzoni A (2008) Construction of an intra-specific sweet cherry (*Prunus avium* L.) genetic linkage map and synteny analysis with the *Prunus* reference map. Tree Genetics & Genomes 4 (4):897-910

Pierantoni L, Cho KH, Shin IS, Chiodini R, Tartarini S, Dondini L, Kang SJ, Sansavini S (2004) Characterisation and transferability of apple SSRs to two European pear F_1 populations. Theoretical and Applied Genetics 109 (7):1519-1524

Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, Campbell CS (2007) Phylogeny and classification of Rosaceae. Plant Systematics and Evolution 266 (1):5-43

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends in Plant Science 1 (7):215-222

Rusholme Pilcher RL, Celton JM, Gardiner SE, Tustin DS (2008) Genetic markers linked to the dwarfing trait of apple rootstock 'Malling 9'. Journal of the American Society for Horticultural Science 133 (1):100-106

Sargent D, Marchese A, Simpson D, Howad W, Fernández-Fernández F, Monfort A, Arús P, Evans K, Tobutt K (2009b) Development of "universal" gene-specific markers from *Malus* spp. cDNA sequences, their mapping and use in synteny studies within Rosaceae. Tree Genetics & Genomes 5 (1):133-145

Sargent D, Rys A, Nier S, Simpson D, Tobutt K (2007a) The development and mapping of functional markers in *Fragaria* and their transferability and potential for mapping in other genera. Theoretical and Applied Genetics 114 (2):373-384

Sargent DJ, Cipriani G, Vilanova S, Gil-Ariza D, Arús P, Simpson DW, Tobutt KR, Monfort A (2008) The development of a bin mapping population and the selective mapping of 103 markers in the diploid *Fragaria* reference map. Genome 51 (2):120-127

Sargent DJ, Davis TM, Simpson DW (2009c) Strawberry (*Fragaria* spp.) Structural Genomics. In: Folta KM, Gardiner SE (eds) Genetics and Genomics of Rosaceae, vol 6. Plant Genetics and Genomics: Crops and Models. Springer New York, pp 437-456. doi:10.1007/978-0-387-77491-6_21

Sargent DJ, Davis TM, Tobutt KR, Wilkinson MJ, Battey NH, Simpson DW (2004) A genetic linkage map of microsatellite, gene-specific and morphological markers in diploid *Fragaria*. Theoretical and Applied Genetics 109 (7):1385-1391

Segura V, Denance C, Durel C-E, Costes E (2007) Wide range QTL analysis for complex architectural traits in a 1-year-old apple progeny. Genome 50 (2):159-171. doi:10.1139/G07-002

Shaw J, Small RL (2004) Addressing the "hardest puzzle in American pomology:" Phylogeny of *Prunus* sect. Prunocerasus (Rosaceae) based on seven noncoding chloroplast DNA regions. American Journal of Botany 91 (6):985-996. doi:10.3732/ajb.91.6.985

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton J-M, Rees DJG, Williams KP, Holt SH, Rojas JJR, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman T-L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Girona EL, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arús P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). Nature Genetics 43 (2):109-116. doi:http://www.nature.com/ng/journal/v43/n2/abs/ng.740.html#supplementary-information

Soriano JM, Romero C, Vilanova S, Llácer G, Badenes M (2005) Genetic diversity of loquat germplasm (*Eriobotrya japonica* (Thunb.) Lindl) assessed by SSR markers. Genome 48 (1):108-114

Stafne ET, Clark JR, Weber CA, Graham J, Lewers KS (2005) Simple sequence repeat (SSR) markers for genetic mapping of raspberry and blackberry. Journal of the American Society for Horticultural Science 130 (5):722-728

Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988) Conservation of gene repertoire but not gene order in pepper and tomato. Proceedings of the National Academy of Sciences of the United States of America 85 (17):6419-6423

Van Deynze AE, Nelson JC, Yglesias ES, Harrington SE, Braga DP, McCouch SR, Sorrells ME (1995) Comparative mapping in grasses. Wheat relationships. Molecular and General Genetics MGG 248 (6):744-754

van Dyk MM, Koning G, Simayi Z, Booi S, Maharaj R, Selala MC, du Preez MG, Rees DJG, Labuschagné IF, Warnich L (2005) Molecular genetic studies on pears (*Pyrus* spp.) in the Western Cape. Acta Horticulturae (ISHS) 671:259-265

van Ooijen JW, Voorrips RE (2001) JoinMap(R) 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nature Genetics 42 (10):833-839. doi:http://www.nature.com/ng/journal/v42/n10/abs/ng.654.html#supplementary-information

Ventelon M, Deu M, Garsmeur O, Doligez A, Ghesquière A, Lorieux M, Rami JF, Glaszmann JC, Grivet L (2001) A direct comparison between the genetic maps of sorghum and rice. Theoretical and Applied Genetics 102 (2):379-386

Vilanova S, Sargent D, Arús P, Monfort A (2008) Synteny conservation between two distantly-related Rosaceae genomes: *Prunus* (the stone fruits) and *Fragaria* (the strawberry). BMC Plant Biology 8 (1):67

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-Resolution Genotyping by Amplicon Melting Analysis using LCGreen. Clinical Chemistry 49 (6):853-860. doi:10.1373/49.6.853

Wu FN, Eannetta NT, Xu YM, Durrett R, Mazourek M, Jahn MM, Tanksley SD (2009a) A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus Capsicum. Theoretical and Applied Genetics 118 (7):1279-1293. doi:10.1007/s00122-009-0980-y

Wu FN, Eannetta NT, Xu YM, Tanksley SD (2009b) A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theoretical and Applied Genetics 118 (5):927-935. doi:10.1007/s00122-008-0950-9

Wu FN, Tanksley SD (2010) Chromosomal evolution in the plant family Solanaceae. BMC Genomics 11. doi:10.1186/1471-2164-11-182

Yamamoto T, Kimura T, Saito T, Kotobuki K, Matsuta N, Liebhard R, Gessler C, van de Weg WE, Hayashi T (2004a) Genetic linkage maps of Japanese and European pears aligned to the apple consensus map. Acta Horticulturae (ISHS) 663:51-56

Yamamoto T, Kimura T, Soejima J, Sanada T, Ban Y, Hayashi T (2004b) Identification of quince varieties using SSR markers developed from pear and apple. Breeding Science 54 (3):239-244

2.9 Supplemental files

Supplemental Table S2-1: Rosaceae conserved orthologous set (RosCOS) marker design notes for mapped markers. Each *Fragaria vesca* × *F. bucharica* (FVxFB) linkage group (LG) is discussed along with information on the bin (LG:bin) in which the RosCOS marker is located, its orthologous 'Golden Delicious' (GD) apple LG, and additional comments.

LG:bin	RosCOS	Apple LG	Comments	LG:bin	RosCOS	Apple LG	Comments
I:08			1 RosCOS available, none mapped	V:15			No RosCOS available
I:21	2343	2 & 15	Homeologous LG; mapped on LG2 only	V:21			No RosCOS available
I:41			1 RosCOS available, none mapped	V:29			No RosCOS available
I:45			No RosCOS available	V:37	3749	13	Single occurrence in GD genome database
I:47			No RosCOS available	V:37	3055	13	Single occurrence in GD genome database
I:90	1623(759)	12	Designed for GD LG3	V:37	1435	4 & 13	Non-homeologous LG; mapped on LG13 only
I:90	2105	15	Single occurrence in GD genome database	V:37	1629	13	Single occurrence in GD genome database
II:00			No RosCOS available	V:37	2372(2143)	4 & 13	Non-homeologous LG; mapped on LG13 only
II:08	1222(596)	5	Designed for GD LG10	V:37	1271	16	Single occurrence in GD genome database
II:08	530	5	Designed for GD LG7	V:46	3695	16	Single occurrence in GD genome database
II:08	2134	10 & 12	Non-homeologous LG; mapped on LG10 only	V:48	2591	6	Single occurrence in GD genome database
II:14	1338	5	Single occurrence in GD genome database	V:73	1814	6	Single occurrence in GD genome database
II:14	52	5	Single occurrence in GD genome database	V:73	1410	6	Single occurrence in GD genome database
II:14	1272	7	Single occurrence in GD genome database	V:73	1167	6	Single occurrence in GD genome database
II:29	1341	5 & 10	Homeologous LG; mapped to LG5 only; designed for GD LG12	V:73	3782	6	Designed for GD LG15
II:29	3781	5	Designed for GD LG8	V:73	1261	6	Single occurrence in GD genome database
II:29	2054	8 & 10	Non-homeologous LG; mapped on LG10 only	V:73	2411	4	Designed for GD LG16
II:29	2596	5 & 10	Homeologous LG; mapped on LG10 only	V:73	1378	8 & 15	Homeologous LG; mapped on LG8 only
II:29	2118	8	Single occurrence in GD genome database	V:81	510	14	Designed for GD LG13
II:47	1125	8	Designed for GD LG11	VI:07	1354	4 & 12	Homeologous LG; mapped on both LG
II:56	1549(2673)	15	Designed for GD LG1	VI:14	1659	4 & 12	Homeologous LG; mapped on LG12 only
II:75			No RosCOS available	VI:14	1617	12	Single occurrence in GD genome database
III:13	1381	17	Single occurrence in GD genome database	VI:31	1367	4 & 7	Non-homeologous LG; mapped on LG7 only
III:13	645	9 & 17	Homeologous LG; mapped to LG17 only	VI:43			No RosCOS available
III:17			No RosCOS available	VI:68	1360	14	Single occurrence in GD genome database
III:44			2 RosCOS available, none mapped	VI:68	1412	17	Designed for GD LG5
III:53	1248(1920)	3	Single occurrence in GD genome database	VI:70			No RosCOS available
III:53	601	11	Single occurrence in GD genome	VI:72	1097	17	Single occurrence in GD

			database				genome database
III:53	509	3 & 11	Homeologous LG; mapped on both LG	VI:87			No RosCOS available
III:53	1363(3268)	3	Single occurrence in GD genome database	VI:115	2970	12 & 14	Homeologous LG; mapped on LG14 only
III:53	2409	3	Single occurrence in GD genome database	VII:21	1281	2	Designed for GD LG10
III:53	2603	11	Single occurrence in GD genome database	VII:40			2 RosCOS available, none mapped
III:57	1259	3 & 11	Homeologous LG; mapped on LG11 only	VII:45			No RosCOS available
IV:20			4 RosCOS available, none mapped	VII:64	1291	2	Single occurrence in GD genome database
IV:26	3524	4	Designed for GD LG5	VII:70			No RosCOS available
IV:46			No RosCOS available	VII:89			No RosCOS available
IV:68	1217	14	Single occurrence in GD genome database	VII:90			No RosCOS available
IV:80			1 RosCOS available, none mapped				

Supplemental Table S2-2: Forward and reverse primer pair sequences for all mapped Rosaceae conserved orthologous set (RosCOS) markers. Primer names indicate the RosCOS sequence used for primer pair design, the linkage group for which the marker was designed, the exons in which the primers were placed or if the target was a SNP; a and b indicate that more than one primer pair was designed for that locus.

Primer name	Primer sequence 5'-3'	Comments
COS2343_2_ex2_3_F	TCAAGCAAGCTGAGGAACAC	
COS2343_2_ex2_3_R	GAAGCAAATTCCACCGAAAC	
COS1623b(759)_3_SNP_F	CAGCGTTTACTTTGGGAGGA	Mapped to LG12
COS1623b(759)_3_SNP_R	TCGAATCCCCTGTATGAACC	
COS2105_15_ex1_2_F	TCAGTGGATTTTGCTGGTTTC	
COS2105_15_ex1_2_R	CTTCCTTGTGCGGAATTGT	
COS530_7_ex7_8_F	TGCACTAAAATGGGAGTCCTA	Mapped to LG5
COS530_7_ex7_8_R	AATTTCACAAGCAGCCACAA	
COS1222(596)_10_SNP_F	AGCGGAGTGTCAACAGCAAC	Mapped to LG5
COS1222(596)_10_SNP_R	GCATCAGCATCTGGGTAGAA	
COS2134_10_ex6_7_F	CTCCTGCTCCTCTGACCTTT	
COS2134_10_ex6_7_R	CCGCAAACCTTTCACACA	
COS52_5_ex4_5_F	CTGGATCACATGCCGAGTT	
COS52_5_ex4_5_R	TCCTAACAGATCGCAAGCAC	
COS1248a(1920)_3_SNP_F	TTGAGTTGTGGAGTAGACTTGTGA	
COS1248a(1920)_3_SNP_R	CAGTTCTTTGTTCATGTTTCTGG	
COS1272_7_ex3_4_F	TCTGTTGCCCTCATGTTTCA	
COS1272_7_ex3_4_R	GGCCTTTGCAGGTCTTGTT	
COS1338_5_ex1_2_F	TCAGGACGCCAAAGATATGA	
COS1338_5_ex1_2_R	CAACACACCAGCATCCACA	
COS1341a_12_ex1_2_F	TGAGAGAGAGGACCACACCA	Mapped to LG5
COS1341a_12_ex1_2_R	CCGGCAAACAAAACATTCTC	
COS1341_12_ex2_3_F	GGGACGAGAATGTTTTGTTTG	Mapped to LG10
COS1341_12_ex2_3_R	ATTGGTGAAGACTGGCTGCT	
COS2054_10_SNP_F	GGTCTTCTCAAATCCCTCCA	
COS2054_10_SNP_R	CCAACACCGAAACGAATCTAA	
COS2118_8_ex2_3_F	CAAAATTGGGCAAGTCATAG	
COS2118_8_ex2_3_R	GAGGGCTCCACAGTTTCAAG	
COS2596_10_ex6_7_F	AAGGGGATGTTTTGTCGATG	
COS2596_10_ex6_7_R	TTTGGAAGAAGATGGGCAGT	
COS3781_8_ex1_2_F	GAGAAAATTGGCGGAATTGA	Mapped to LG5
COS3781_8_ex1_2_R	ATTCTTGTCGGGGTTCCAG	

COS1125_11_ex1_2_F	TGGCGACTCAAGACGATAACT	Mapped to LG8
COS1125_11_ex1_2_R	CACCATCCCATTTCTTTCCA	
COS601b_11_ex3_4_F	AGAAGCCGAAGAGGAGGAGT	
COS601b_11_ex3_4_R	ATGATTCCCTGAGTGCCTTTC	
COS1549(2673)_1_ex1_2_F	GCCAACTTCTCTCTGCCAAAC	Mapped to LG15
COS1549(2673)_1_ex1_2_R	TCGTCCTCGCCTTATCAATC	
COS645_17_ex7_8_F	TGGCATATCGACACAAATCC	
COS645_17_ex7_8_R	GAACACATTAGAAGCAGGTCGTT	
COS1381_17_ex2_3_F	TGGGCTTTTTCGTCTTCTTG	
COS1381_17_ex2_3_R	CCGGTCCACACTTCTTTTGT	
COS509_11_ex2_3_F	GAAGCTGTTCACACAGACAAGGC	
COS509_11_ex2_3_R	CATTATCCGACACAAACTTCC	
COS509_3_ex1_2_F	TTGCTTAGCCATTTCCACCGAC	
COS509_3_ex1_2_R	ATGCAGCGGGAGCCTTTT	
COS1363c(3268)_3_ex5_6_F	GCAACACGCCTTATACTTTACCA	
COS1363c(3268)_3_ex5_6_R	CCAGCTATCTTCTCTATTGCATC	
COS2409_3_ex2_3_F	TATTGGAAGCCCCAGACAAG	
COS2409_3_ex2_3_R	ACAGGATGGGCATGAAACA	
COS2603_11_ex2_3_F	TCTCCCGTGTCAAGAGGTGT	
COS2603_11_ex2_3_R	ATGGGGCTTCACGTTCATT	
COS1259_11_ex1_2_F	AAGGTGCTGATGGTTCTTGG	
COS1259_11_ex1_2_R	GACTTTTGCACTCCGGGTTT	
COS3524_5_ex2_3_F	AGAAAGCGTCCCCATCAA	Mapped to LG4
COS3524_5_ex2_3_R	ACTATACGTGCTGTGGCAATC	
COS1217_14_ex2_3_F	CGCCAGAAACCTTGACAGA	
COS1217_14_ex2_3_R	CACCTGAAGATTCCCAGTCC	
COS1271_16_ex1_2_F	CGGCTACATCTCCATCTGTG	
COS1271_16_ex1_2_R	AACGTACTCCTCGCGCTTCT	
COS1435_13_ex2_3_F	CATTTTCGCCCGAGTTTTAC	
COS1435_13_ex2_3_R	CATCCCTCGCTGCATATCTT	
COS1629_13_ex2_3_F	AGCGGATCACTGGACAAAAC	
COS1629_13_ex2_3_R	CAATGCACTCCAATTTCCTG	
COS2372_4_ex2_3_F	CCTGCCTACCCCGATACTTT	Mapped to LG13
COS2372_4_ex2_3_R	AAATGCTTGGAGAGCTTGGT	
COS3055_13_ex3_4_F	AGTACGAGGCATTGGGTTCA	
COS3055_13_ex3_4_R	CACGAAGTTTGGTCACAGCA	
COS3749_13_ex4_5_F	TTGGTTAAAGAAGGGGATCGT	
COS3749_13_ex4_5_R	TCAAGATTACAACAGAGAGGGAAG	
COS3695_16_ex2_3_F	CTGGCTTTAGAAGATGCAACAA	
COS3695_16_ex2_3_R	AGGCCGATCAAAGGTCACA	
COS2591_6_ex2_3_F	TCCTCGTCTTTTGACTTGTCCT	
COS2591_6_ex2_3_R	GCATTTTCGCTATGTTCGACT	
COS1167_6_ex3_4_F	CATGGCTGACCAAAGACAAA	
COS1167_6_ex3_4_R	TGGCAACTGAAGAAACCATC	
COS1378_8_ex4_5_F	GGGGGAGATCCTACCACATAA	
COS1378_8_ex4_5_R	TGTTTGGGTCCTTGCTAAGAGT	
COS1410_6_ex1_2_F	GCCCCAAACCCAGAGAAAA	
COS1410_6_ex1_2_R	CCAAAACTTCAGCCAGCTTC	
COS1814_6_ex1_2_F	CCACTTCCTCTGCCTCTCAA	
COS1814_6_ex1_2_R	CTGCCACAAATGCCTTCTTC	
COS2411_16_ex2_3_F	GTGGGCTGACACCAAGAAAG	Mapped to LG4
COS2411_16_ex2_3_R	ACGGGATTATTTTCCACCAA	
COS3782_15_ex1_2_F	AGGATGTTGAGGTTGTGATTG	Mapped to LG6
COS3782_15_ex1_2_R	TTGCAGCTCGCTTGCTCTT	
COS1261_6_ex1_2_F	GTCATCTTGTCACCCCCAAG	
COS1261_6_ex1_2_R	TCCTCCCCAATTTCCTTCAC	
COS510_13_ex3_4_F	AATTGCACCTCATAGTGATTGG	Mapped to LG14
COS510_13_ex3_4_R	GGGTACATGCTTTGCCTCAC	
COS1354_4_ex7_8_F	TTGTACCACCACCAGAAAGGT	

COS1354_4_ex7_8_R	ATGACAATGAGCCCAAGAGG	
COS1354_12_ex2_3_F	CCGAAGAGGAGAAGGAGAAG	
COS1354_12_ex2_3_R	TCCCAGGAGGATTCAAAACA	
COS1617_12_ex1_2_F	TCACCTCACGATTCTTCTGCT	
COS1617_12_ex1_2_R	GAGTAGCACCACTTTGAGCTTT	
COS1659_12_ex1_2_F	CCAAAGAAATCTCCTCTTATTGTGG	
COS1659_12_ex1_2_R	ATACGGTCCATCACGAATGT	
COS1367_7_ex2_3_F	CTGTGTGATCGGAACTTTGG	
COS1367_7_ex2_3_R	GCAATGAATCTGGCAAAACA	
COS1360_14_ex4_5_F	CAATCTACCTGAAAAACCTGGA	
COS1360_14_ex4_5_R	ATTCACAATGGGTGCAAAGG	
COS1412_5_ex4_5_F	CAAGAGGCGATTTGCTAGAGT	Mapped to LG17
COS1412_5_ex4_5_R	TTCACTGTCATAGGCTCCAAG	
COS1097_17_ex4_5_F	GCACAAAGGTTCCACCATCA	
COS1097_17_ex4_5_R	TCGAGTTTCGCTTTCTTCATC	
COS2970_14_ex2_3_F	ACATTGCTACTATGGTGCTTGAGA	
COS2970_14_ex2_3_R	CCACAGACACCCATCTTTCC	
COS1281b_10_SNP_F	GTTTTAAGCAAAGGGCGATG	Mapped to LG2
COS1281b_10_SNP_R	TGTTGCCTCCTGAATGGTTT	
COS1291_2_ex3_4_F	GGAAAATTAGACGATGGCAAAG	
COS1291_2_ex3_4_R	CCACATCCAAGTAATCTCCAA	

Supplemental Table S2-3: Twelve markers placed *in silico* on the 'Golden Delicious' (GD) apple whole genome sequence, but which mapped genetically to a different apple linkage group (LG), and the *Fragaria vesca* x *F. bucharica* (FVxFB) LG and bin (LG:bin) to which the marker is genetically mapped. The two best recombination frequency (Rf) and log of odds (LOD) scores and their associated framework markers show support for the genetic placement. Primer names indicate the RosCOS sequence used to design the primer pair, the LG for which the marker was designed, the exons in which the primers were placed or if the target was a SNP; a and b indicate that more than one primer pair was designed for that locus.

Primer name	FVxFB LG:bin	Expected apple LG	Apple LG	Apple parent	Rf	LOD	Associated frame work marker 1	Rf	LOD	Associated frame work marker 2
COS1623b(759)_3_SNP	I:90	3	12	M.9	0.08	13.3	ch04d02b	0.09	5.2	AY12(300)
COS530_7_ex7_8	II:08	7	5	M.9	0.03	15.7	RosCOS52	0.12	4.5	TsuENH065
COS1341a_12_ex1_2	II:29	12	5	M.9	0.03	8.4	AY08(850)	0.08	16.4	A17_750
COS1125_11_ex1_2	II:47	11	8	RG	0.006	36	GD_SNP01479	0.02	12	GD_SNP00627
COS1125_11_ex1_2	II:47	11	8	BB	0.006	36	GD_SNP01479	0.043	10	GD_SNP02685
COS1549(2673)_1_ex1_2	II:56	1	15	A689	0.024	19.8	32845P7FR_ER3	0.044	14.9	E33M47_06
COS1549(2673)_1_ex1_2	II:56	1	15	RG	0.024	19.8	32845P7FR_ER3	0.044	14.9	E33M47_06
COS3524_5_ex2_3	IV:26	5	4	M.9	0.14	7	ch01d03	0.24	2.3	Hi23d11
COS2372_4_ex2_3	V:37	4	13	M.9	0.03	17.6	RosCOS1435	0.08	11	RosCOS3749
COS3782_15_ex1_2	V:73	15	6	M.9	0.06	18	NZ_23g04	0.12	3	AV05(1000)
COS3782_15_ex1_2	V:73	15	6	RG	0.06	9.6	GD_SNP01716	0.09	15	GD_SNP00472
COS2411a_16_ex2_3	V:73	16	4	M.9	0.06	9.4	NH011b	0.03	9	ch01d03
COS2411a_16_ex2_3	V:73	16	4	R5	0.06	8.5	GD6_b	0.06	9	ch01d03
COS2411a_16_ex2_3	V:73	16	4	RG	0.11	6	GD_SNP00928	0.24	2	GD_SNP00277
COS510_13_ex3_4	V:81	13	14	M.9	0.17	9.4	ch05e05	0.2	6	RosCOS2970
COS1412b_5_ex4_5	VI:68	5	17	M.9	0.1	14	Hi02f12	0.13	5.6	TsuENH041
COS1412b_5_ex4_5	VI:68	5	17	BB	0.09	14	GD_SNP00058	0.34	2	GD_SNP01932
COS1281b_10_SNP	VII:21	10	2	RG	0.045	18.4	GD_SNP02376	0.05	7.6	GD_SNP00992
COS1281b_10_SNP	VII:21	10	2	BB	0.082	17.9	GD_SNP00833	0.16	13.2	GD_SNP02076

Supplemental Table S2-4: Apple simple sequence repeat (SSR) markers compared with *Fragaria vesca* scaffold sequences. The best 13 of 103 markers [56, 57] are indicated by having the lowest E-values using BLASTN analysis. Forward (F) and reverse (R) apple primer sequence BLASTN results are given for each marker. In no case was the SSR motif present in the *Fragaria vesca* scaffold target sequence.

Marker Name	Apple LG	*Fragaria vesca* Scaffold BLASTN Results for Apple SSR	% Identity	*Fragaria vesca* Scaffold BLASTN Results for Apple Primers	BLAST E-value	Identities
CH05g08	1	scf0513190:162672-162817	87%		$1.00E^{-31}$	127/146
				F: scf0513160:3725248-3725264	0.27	17/17
				R: scf0513140:351340-351358	4.2	18/19
CH02c02	2	scf0513105:74398-74463	97%		$1.00E^{-24}$	64/66
				F: scf0513138:555774-555790	0.29	17/17
				R: scf0513152:903212-903226	4.5	15/15
CH02f06	2	scf0513105:1049565-1049664	90%		$9.00E^{-26}$	90/100
				F: scf0513105:1049643-1049662	0.005	20/20
				R: scf0513196:451825-451839	4.4	15/15
CH02b12	5	scf0513119:1181728-1181879	87%		$4.00E^{-34}$	132/152
				F: scf0513160:993581-993596	1.1	16/16
				R: scf0513044:1610388-1610402	4.2	15/15
CH04e03	5	scf0512973:154226-154358	86%		$6.00E^{-27}$	115/133
				F: scf0513095:10789-10804	1.1	16/16
				R: scf0513168:1225005-1225019	4.2	15/15
CH04g09	5&10	scf0513173:145165-145280	91%		$5.00E^{-33}$	105/116
				R: scf0513173:145257-145276	1.1	19/20
CH05a02	8&15	scf0513134:1110194-1110263	94%		$2.00E^{-22}$	66/70
				F: scf0513190:551015-551030	1.2	16/16
				R: scf0513151:190328-190344	0.3	17/17
CH05d11	12	scf0513113:1283262-1283369	92%		$8.00E^{-32}$	99/108
				F: scf0513165:2373021-2373035	4.5	15/15
CH04g04	12	scf0512949:12100-12152	95%		$6.00E^{-11}$	41/43
				F: scf0513194:1194740-1194759	1.1	19/20
				R: scf0513122:637577-637591	4.4	15/15
CH03b06	15	scf0513134:1986547-1986656	93%		$4.00E^{-37}$	102/110
				F: scf0513192:2539848-2539862	4.9	15/15
				R: scf0513134:1986548-1986569	0.079	21/22
CH02a03	16	scf0513158:306976-307143	91%		$2.00E^{-54}$	153/168
				F: scf0513158:307009-307027	4	18/19
				R: scf0513124:1927273-1927288	1	16/16
CH05e04	16	scf0513158:2403845-2403983	85%		$8.00E^{-23}$	118/139
				F: scf0513175:563577-563591	4.9	15/15
				R: scf0513162:667060-667075	1.2	16/16
CH04c06	17	scf0512982:30335-30507	85%		$2.00E^{-32}$	147/173
				F: scf0513130:114523-114538	1.1	16/16
				R: scf0513130:777429-777445	0.27	17/17

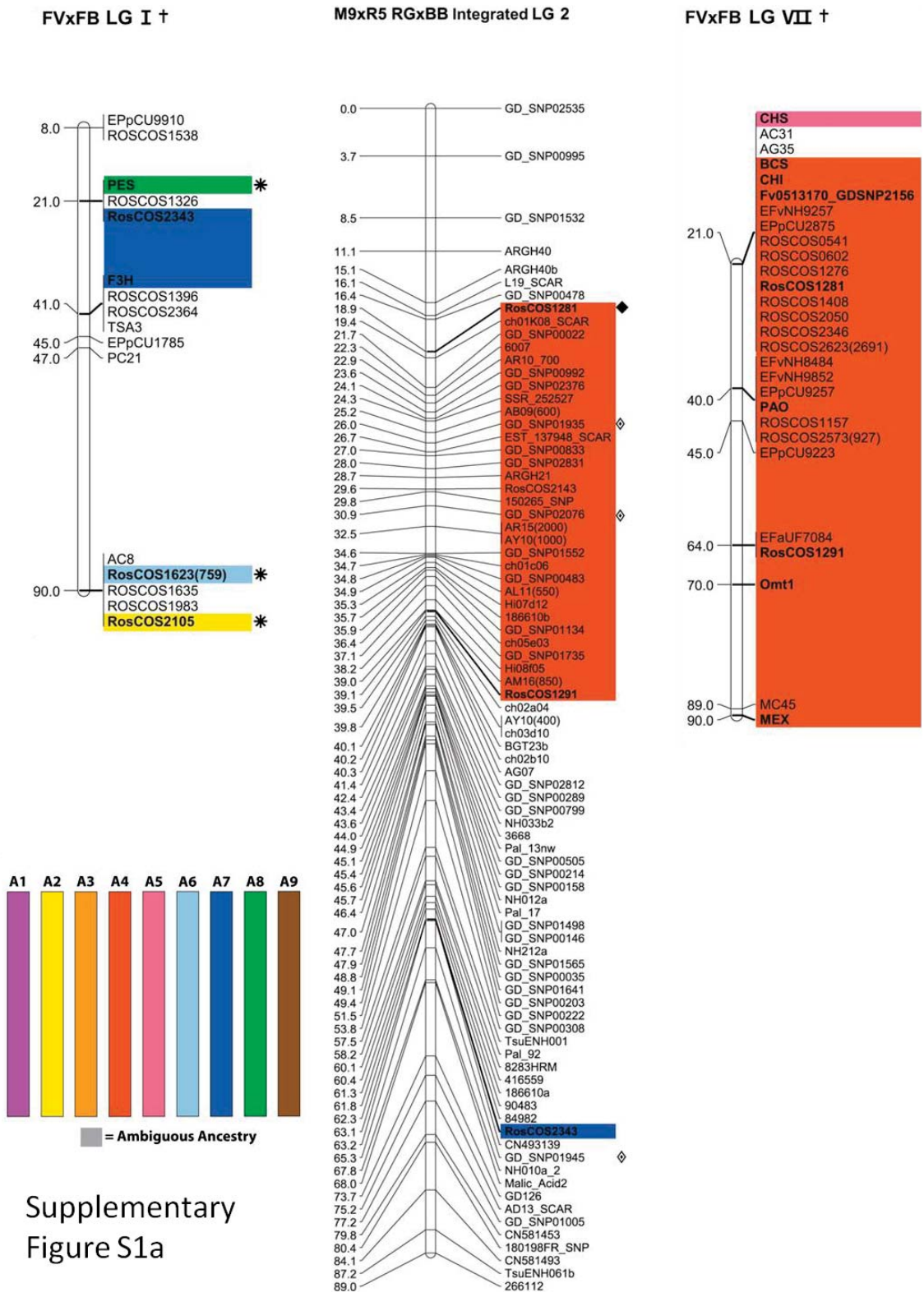
Supplemental Table S2-5: Strawberry gene-based markers genetically mapped *Fragaria vesca* × *F. bucharica* (FVxFB) compared with the 'Golden Delicious' apple whole genome sequence.

FVxFB LG	Bin	Gene Name	GenBank Accession	Apple LG hits	Region start	Region end	Fragment size	BLAST score	BLAST E-value	Identities	% Identity	Strand	Source
I	21	*PES*	CV881712	11	805392	805479	87	174	1.00E^{-41}	88/88	100	+/-	[62]
I	21	*PES*	CV881712	11	804835	804889	54	94	3.00E^{-17}	53/55	96	+/-	[62]
I	21	*PES*	CV881712	4	10299359	10299446	87	170	2.00E^{-40}	87/88	98	+/-	[62]
I	21	*PES*	CV881712	4	10299551	10299616	65	115	9.00E^{-24}	64/66	96	+/-	[62]
I	41	*F3H*	AY017479	5	23800985	23801040	55	88	5.00E^{-15}	53/56	94	+/-	[59]
I	41	*F3H*	AY017479	2	11906211	11906262	51	80	1.00E^{-12}	49/52	94	+/-	[59]
I	41	*F3H*	AY017479	2	11906980	11907043	63	64	7.00E^{-08}	56/64	87	+/-	[59]
II	0	*EKO*	AY462247	8	4356707	4357013	306	236	3.00E^{-59}	260/307	84	+/+	[62]
II	0	*EKO*	AY462247	8	4347838	4348144	306	172	3.00E^{-40}	251/307	81	+/+	[62]
II	0	*EKO*	AY462247	1	32699839	32700145	306	236	3.00E^{-59}	260/307	84	+/+	[62]
II	0	*EKO*	AY462247	1	32699076	32699219	143	119	4.00E^{-24}	123/144	85	+/+	[62]
II	8	*APB*	X91839	3	37186100	37186311	211	242	5.00E^{-61}	189/212	89	+/-	[62]
II	8	*APB*	X91839	3	37186622	37186714	92	113	3.00E^{-22}	84/93	90	+/-	[62]
II	8	*APB*	X91839	11	29358375	29358586	211	214	1.00E^{-52}	186/212	87	+/-	[62]
II	8	*APB*	X91839	11	29358873	29358965	92	129	5.00E^{-27}	86/93	92	+/-	[62]
II	29	*ADH*	AF000216	7	28542855	28542951	96	113	7.00E^{-23}	87/97	89	+/+	Gen-Bank
II	29	*ADH*	AF000216	7	28567712	28567804	92	105	2.00E^{-20}	83/93	89	+/+	Gen-Bank
II	50	*ACL5*	AB204522	15	392406	392667	261	519	E^{-145}	262/262	100	+/+	[61]
II	50	*ACL5*	AB204522	15	395260	395431	171	341	4.00E^{-91}	172/172	100	+/+	[61]
II	50	*ACL5*	AB204522	8	10566673	10566826	153	270	1.00E^{-69}	149/154	96	+/-	[61]
II	50	*ACL5*	AB204522	8	10568647	10568915	268	228	4.00E^{-57}	227/269	84	+/-	[61]
II	56	*AMY*	AF153828	8	505097	505574	477	902	0.00E^{+00}	474/479	98	+/-	[61]
II	56	*AMY*	AF153828	8	507546	507775	229	416	E-^{113}	225/230	97	+/-	[61]
II	56	*AMY*	AF153828	12	22901330	22901559	229	345	4.00E^{-92}	216/230	93	+/-	[61]
II	56	*AMY*	AF153828	12	22899090	22899295	205	281	4.00E^{-73}	190/206	92	+/-	[61]
II	75	*DFR*	AY575057	8	4395430	4395592	162	180	2.00E^{-42}	145/163	88	+/+	[62]
II	75	*DFR*	AY575057	8	4394104	4394270	166	129	8.00E^{-27}	141/164	84	+/+	[62]
II	75	*DFR*	AY575057	12	26023843	26023996	153	155	1.00E^{-34}	135/154	87	+/+	[62]
II	75	*DFR*	AY575057	12	26022247	26022350	103	103	5.00E^{-19}	91/104	87	+/+	[62]
III	13	*CKP*	AY395701	5	1365684	1366246	562	1084	0.00E^{+00}	558/563	99	+/-	[61]
III	13	*CKP*	AY395701	5	1366321	1366638	317	306	E^{-169}	312/318	98	+/-	[61]
III	13	*CKP*	AY395701	10	36474663	36475214	551	864	0.00E^{+00}	523/552	94	+/+	[61]
III	13	*CKP*	AY395701	10	36476741	36476911	170	283	1.00E^{-73}	164/171	95	+/+	[61]
III	13	*FPS*	AY083165	5	171219	171555	336	636	E^{-180}	333/337	98	+/-	[61]
III	13	*FPS*	AY083165	5	175030	175205	175	256	2.00E^{-65}	166/177	93	+/-	[61]
III	13	*FPS*	AY083165	10	37651133	37651377	244	470	E^{-130}	243/245	99	+/+	[61]
III	13	*FPS*	AY083165	10	37654658	37654814	156	307	7.00E^{-81}	156/157	99	+/+	[61]
III	53	*APX*	AF158654	3	10052073	10052205	132	184	2.00E^{-43}	123/133	92	+/+	[62]
III	53	*APX*	AF158654	3	10052354	10052461	107	103	5.00E^{-19}	94/108	97	+/+	[62]
III	53	*APX*	AF158654	11	11340456	11340590	134	172	6.00E^{-40}	123/135	91	+/+	[62]
III	53	*APX*	AF158654	11	11340680	11340860	180	113	5.00E^{-22}	149/181	92	+/+	[62]
IV	46	*LOX*	AJ578035	16	15933865	15934593	728	533	E^{-148}	614/729	84	+/-	[62]
IV	46	*LOX*	AJ578035	16	15935873	15936184	311	345	6.00E^{-92}	277/312	88	+/-	[62]
IV	46	*LOX*	AJ578035	13	20601325	20602053	728	434	E^{-119}	601/729	82	+/+	[62]
IV	46	*LOX*	AJ578035	13	20599750	20600061	311	341	9.00E^{-91}	277/312	88	+/+	[62]
IV	68	*CEL-1*	AF051346	10	21628106	21628882	776	295	3.00E^{-77}	620/777	79	+/+	[62]
IV	68	*CEL-1*	AF051346	7	1719336	1719614	278	172	3.00E^{-40}	232/279	83	+/+	[62]
IV	68	*CEL-1*	AF051346	13	1250179	1250280	101	84	2.00E^{-13}	87/102	85	+/-	[62]
IV	68	*CEL-1*	AF051346	13	1233979	1234056	77	60	3.00E^{-06}	66/78	84	+/-	[62]
V	15	*ANS*	AY695818	6	17352409	17352935	526	430	E^{-118}	449/527	85	+/-	[62]

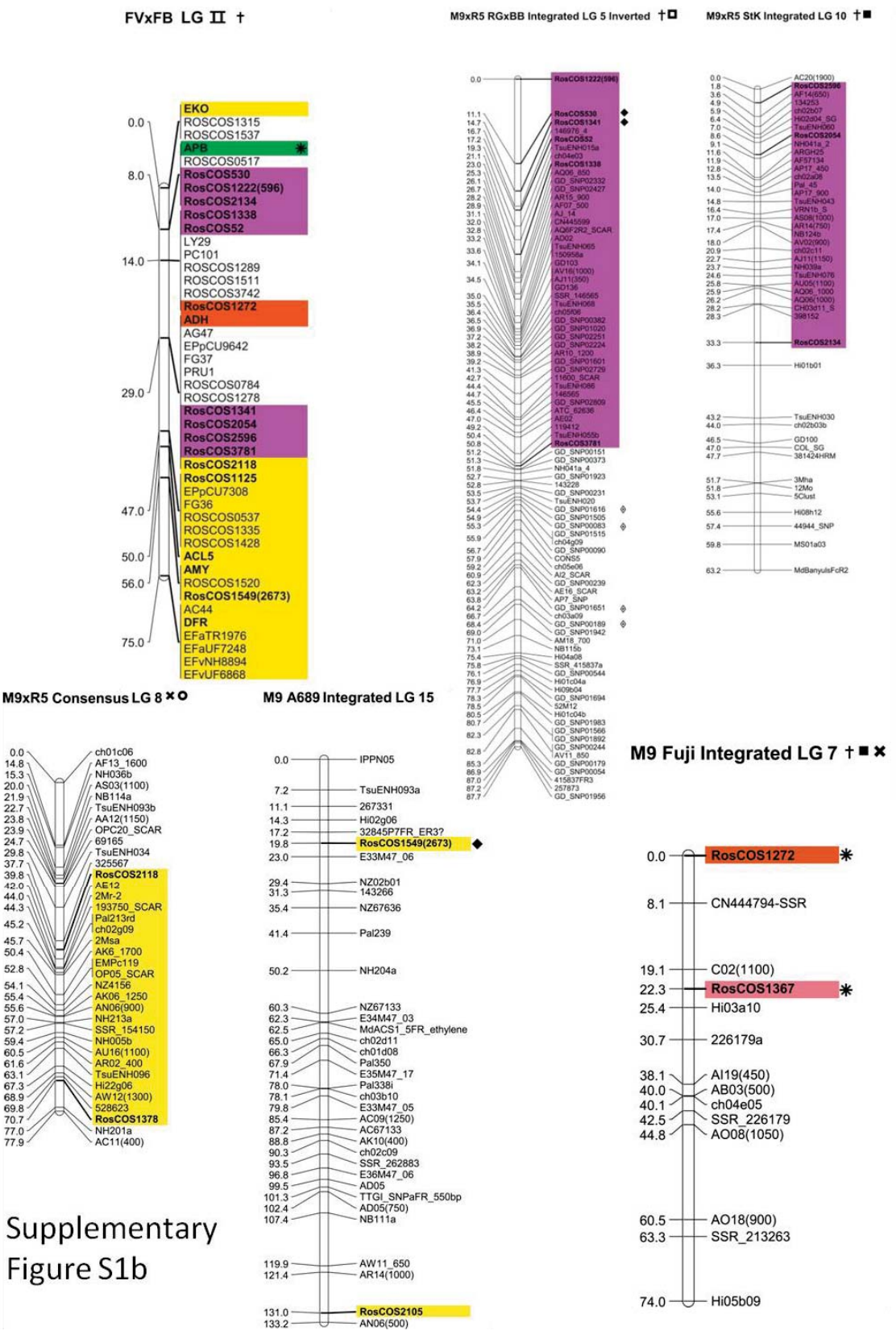
V	15	*ANS*	AY695818	6	17130362	17130853	491	428	E^{-117}	423/492	85	+/-	[62]
V	21	*GP*	AF403707	10	35393054	35393298	244	482	E^{-113}	244/245	99	+/+	[61]
V	21	*GP*	AF403707	10	35395259	35395489	230	458	E^{-126}	231/231	100	+/+	[61]
V	21	*GP*	AF403707	14	20987017	20987185	168	240	$8.00E^{-61}$	157/169	92	+/+	[61]
V	21	*GP*	AF403707	14	20985321	20985557	236	208	$3.00E^{-51}$	204/237	86	+/+	[61]
V	29	*CEL-2*	AF054615	2	20087948	20088257	309	377	E^{-101}	280/310	90	+/+	[62]
V	29	*CEL-2*	AF054615	2	20087630	20087856	226	220	$2.00E^{-54}$	198/227	87	+/+	[62]
V	29	*CEL-2*	AF054615	6	23066709	23067019	310	375	E^{-101}	280/311	90	+/-	[62]
V	29	*CEL-2*	AF054615	6	23068663	23068847	184	212	$5.00E^{-52}$	165/185	89	+/-	[62]
V	37	*PL*	FXU63550	10	19357553	19358175	622	486	E^{-134}	528/623	84	+/+	[60]
V	37	*PL*	FXU63550	10	19358563	19358684	121	155	$6.00E^{-35}$	111/122	90	+/+	[60]
V	37	*PL*	FXU63550	5	18855591	18856189	598	339	$2.00E^{-90}$	491/599	81	+/-	[60]
V	37	*PL*	FXU63550	5	18855090	18855229	139	167	$2.00E^{-38}$	126/140	90	+/-	[60]
VI	31	*PGLM*	AJ004915	12	31225234	31225747	513	936	$0.00E^{+00}$	504/514	98	+/-	[62]
VI	31	*PGLM*	AJ004915	12	31227016	31227442	426	672	$0.00E^{+00}$	405/427	94	+/-	[62]
VI	31	*PGLM*	AJ004915	9	13067576	13068089	513	932	$0.00E^{+00}$	503/514	97	+/+	[62]
VI	31	*PGLM*	AJ004915	9	13065883	13066309	426	672	$0.00E^{+00}$	405/427	94	+/+	[62]
VI	43	*SIP*	AF220202	4	9882601	9882845	244	438	E^{-120}	242/248	97	+/+	[61]
VI	43	*SIP*	AF220202	4	9881990	9882197	207	396	E^{-108}	206/208	99	+/+	[61]
VI	43	*SIP*	AF220202	12	22672972	22673178	206	371	E^{-100}	202/207	97	+/+	[61]
VI	43	*SIP*	AF220202	12	22671529	22671704	175	222	$2.00E^{-55}$	160/176	90	+/+	[61]
VI	72	*ZIP*	AY805422	8	20866933	20867207	274	141	$2.00E^{-30}$	224/275	81	+/-	[62]
VI	72	*ZIP*	AY805422	8	20831673	20831828	155	119	$6.00E^{-24}$	132/356	84	+/-	[62]
VI	72	*ZIP*	AY805422	3	20398071	20398345	274	141	$2.00E^{-30}$	224/275	81	+/+	[62]
VI	72	*ZIP*	AY805422	3	20397774	20397929	155	119	$6.00E^{-24}$	132/156	84	+/+	[62]
VI	87	*ACO*	AJ851828	9	7837757	7838100	343	280	$8.00E^{-73}$	296/347	85	+/+	[62]
VI	87	*ACO*	AJ851828	9	7874522	7874865	343	280	$8.00E^{-73}$	296/347	85	+/+	[62]
VI	87	*ACO*	AJ851828	17	8873389	8873732	343	236	$1.00E^{-59}$	290/347	83	+/+	[62]
VI	87	*ACO*	AJ851828	17	8873842	8874039	197	218	$3.00E^{-54}$	176/198	88	+/+	[62]
VI	115	*QR*	AY158836	9	1873930	1874265	335	283	$4.00E^{-73}$	290/337	86	+/-	[62]
VI	115	*QR*	AY158836	9	1895165	1895500	335	276	$1.00E-^{70}$	289/337	85	+/-	[62]
VI	115	*QR*	AY158836	11	32562534	32562710	176	176	$7.00E^{-41}$	155/177	87	+/+	[62]
VI	115	*QR*	AY158836	11	32562791	32563064	273	137	$6.00E^{-29}$	225/277	81	+/+	[62]
VII	21	*BCS*	DQ471308	15	42988775	42989094	319	634	E^{-179}	320/320	100	+/-	[61]
VII	21	*BCS*	DQ471308	15	42990350	42990616	266	529	E^{-147}	267/267	100	+/-	[61]
VII	21	*BCS*	DQ471308	1	12356902	12357163	261	369	$2.00E^{-99}$	243/262	92	+/+	[61]
VII	21	*BCS*	DQ471308	1	12357272	12357411	139	246	$2.00E^{-62}$	136/140	97	+/+	[61]
VII	21	*CHI*	AY017478	1	20814447	20814645	198	157	$8.00E^{-36}$	169/199	84	+/-	[59]
VII	21	*CHI*	AY017478	1	20831307	20831499	192	149	$2.00E^{-33}$	163/193	84	+/-	[59]
VII	21	*CHI*	AY017478	11	36804716	36804914	198	157	$8.00E^{-36}$	168/199	84	+/-	[59]
VII	21	*CHI*	AY017478	11	19398566	19398758	192	145	$3.00E^{-32}$	163/193	84	+/-	[59]
VII	21	*CHS*	AY017477	9	16052504	16052690	186	153	$7.00E^{-35}$	159/187	85	+/-	[59]
VII	21	*CHS*	AY017477	9	16933290	16933477	187	151	$3.00E^{-34}$	160/188	85	+/+	[59]
VII	40	*PAO*	AB250235	2	38697274	38697826	552	1065	$0.00E^{+00}$	545/553	98	+/-	[61]
VII	40	*PAO*	AB250235	2	38698414	38698671	257	507	E^{-141}	257/258	99	+/-	[61]
VII	40	*PAO*	AB250235	7	1517907	1518157	250	375	E^{-101}	235/251	93	+/-	[61]
VII	40	*PAO*	AB250235	7	1510794	1511165	371	345	$4.00E^{-92}$	326/374	87	+/-	[61]
VII	70	*Omt1*	AF220491	1	35423683	35423988	305	329	$2.00E^{-87}$	270/306	88	+/-	Gen-Bank
VII	70	*Omt1*	AF220491	1	35418709	35419013	304	327	$1.00E^{-86}$	270/305	88	+/-	Gen-Bank
VII	70	*Omt1*	AF220491	7	30464814	30465111	297	281	$5.00E^{-73}$	259/298	86	+/-	Gen-Bank
VII	70	*Omt1*	AF220491	7	30448288	30448585	297	274	$1.00E^{-70}$	258/298	86	+/-	Gen-Bank
VII	90	*MEX*	DQ648082	7	31123433	31123724	291	547	E^{-153}	288/292	98	+/-	[61]
VII	90	*MEX*	DQ648082	7	31124389	31124680	291	462	E^{-127}	273/292	93	+/-	[61]
VII	90	*MEX*	DQ648082	1	36082355	36082642	287	523	E^{-146}	282/288	97	+/+	[61]
VII	90	*MEX*	DQ648082	1	36081916	36082095	179	317	$6.00E^{-84}$	175/180	97	+/+	[61]

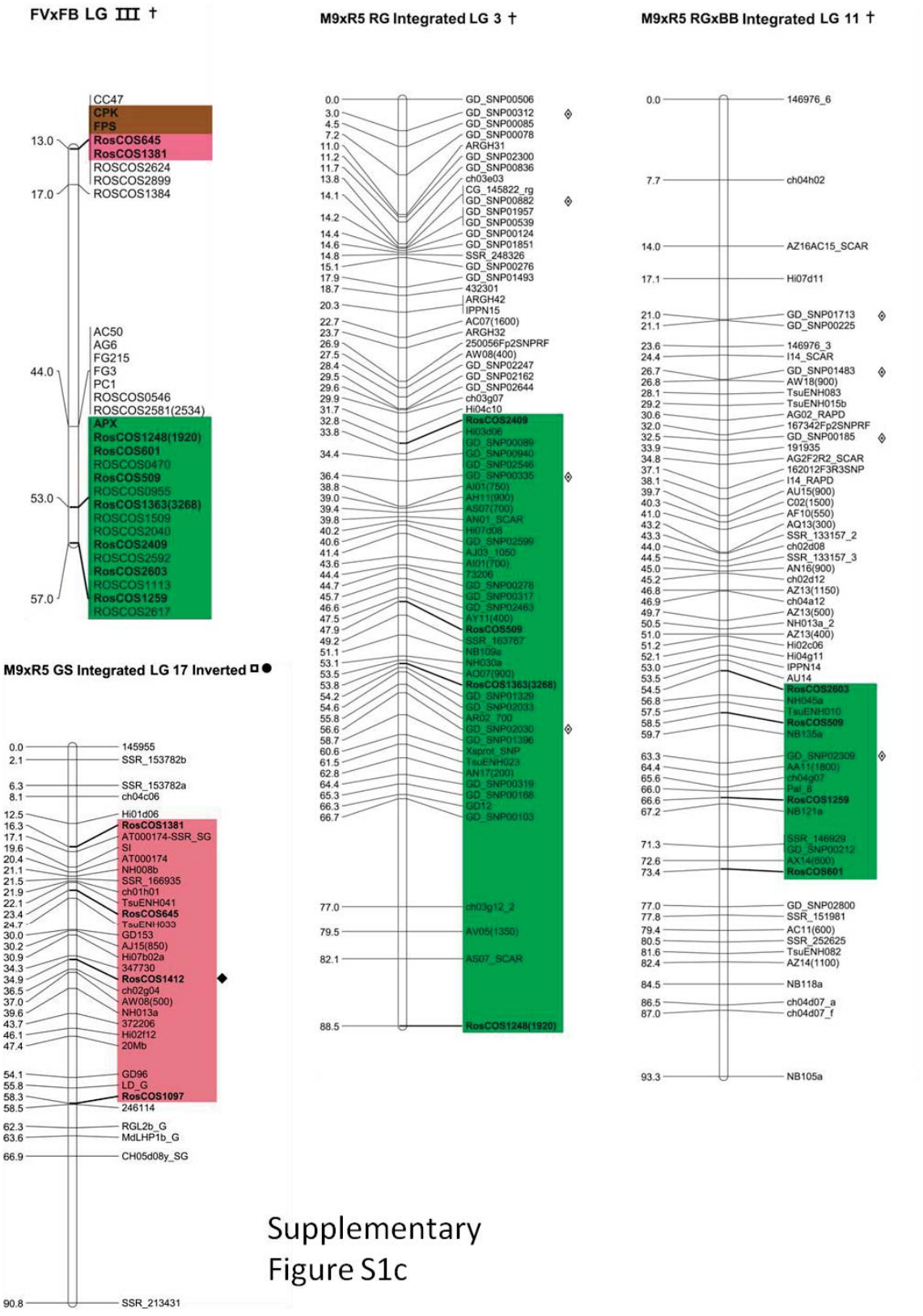
Supplemental Fig. S2.1: Synteny between *Fragaria vesca* × *F. bucharica* (FVxFB) and apple as determined by genetic mapping of 56 Rosaceae conserved orthologous set (RosCOS) markers. Colors indicate ancestral relationships; *in silico* homolog of apple orthologs of strawberry gene-based markers is provided as additional support.

*Indicates markers lacking supporting data
◆Indicates markers with genetic placement different from bioinformatic placement
†Indicates linkage group with ancestry or orientation different from Illa et al. (2011)
▪Indicates linkage group with poor support
◘Indicates linkage group inverted to illustrate orthologous regions better
×Indicates consensus linkage group with missing RosCOS marker
°Indicates incomplete consensus
●Indicates only one homeolog represented
◊ Indicates markers used to merge linkage groups of different apple populations

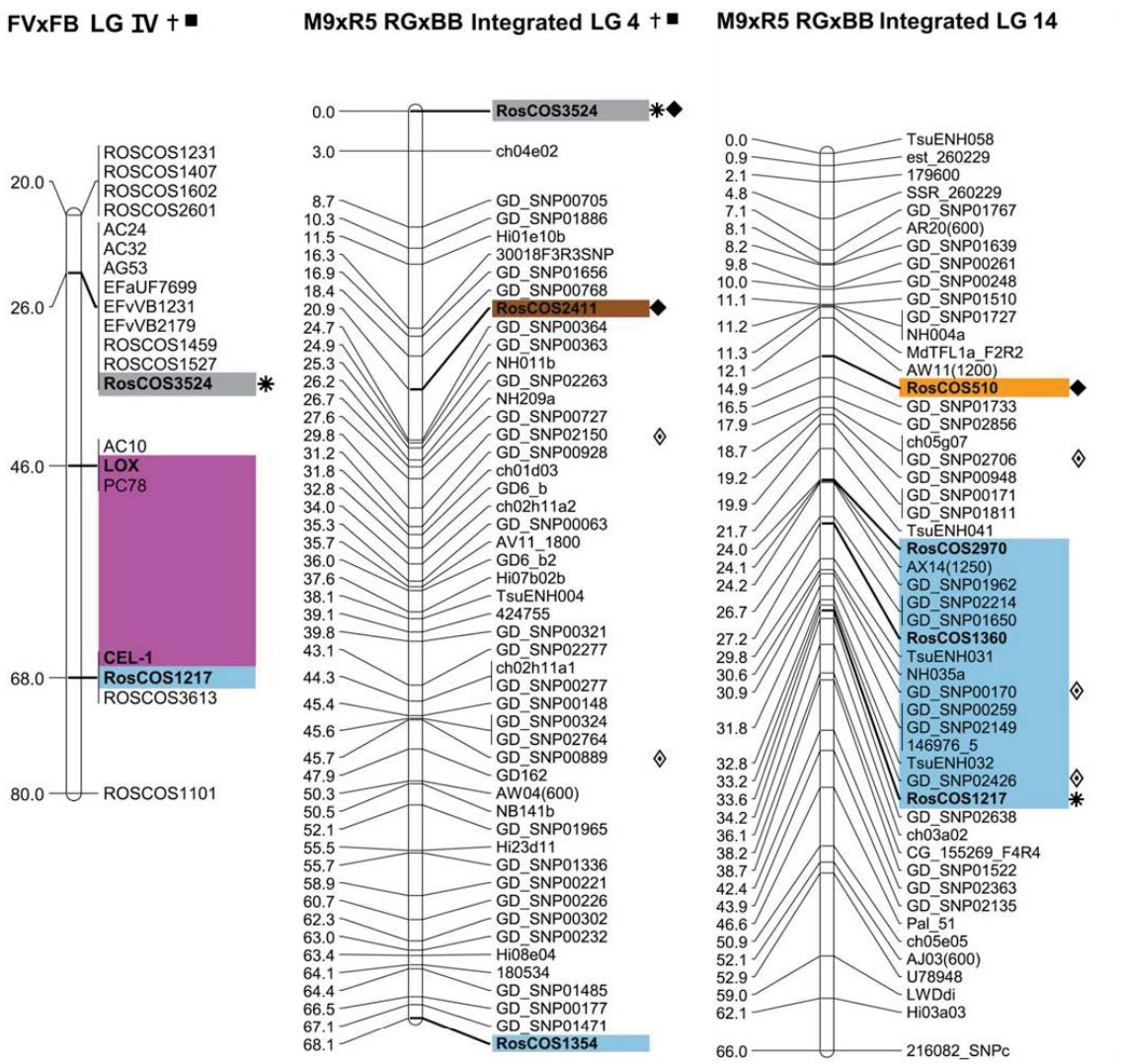


Supplementary Figure S1a

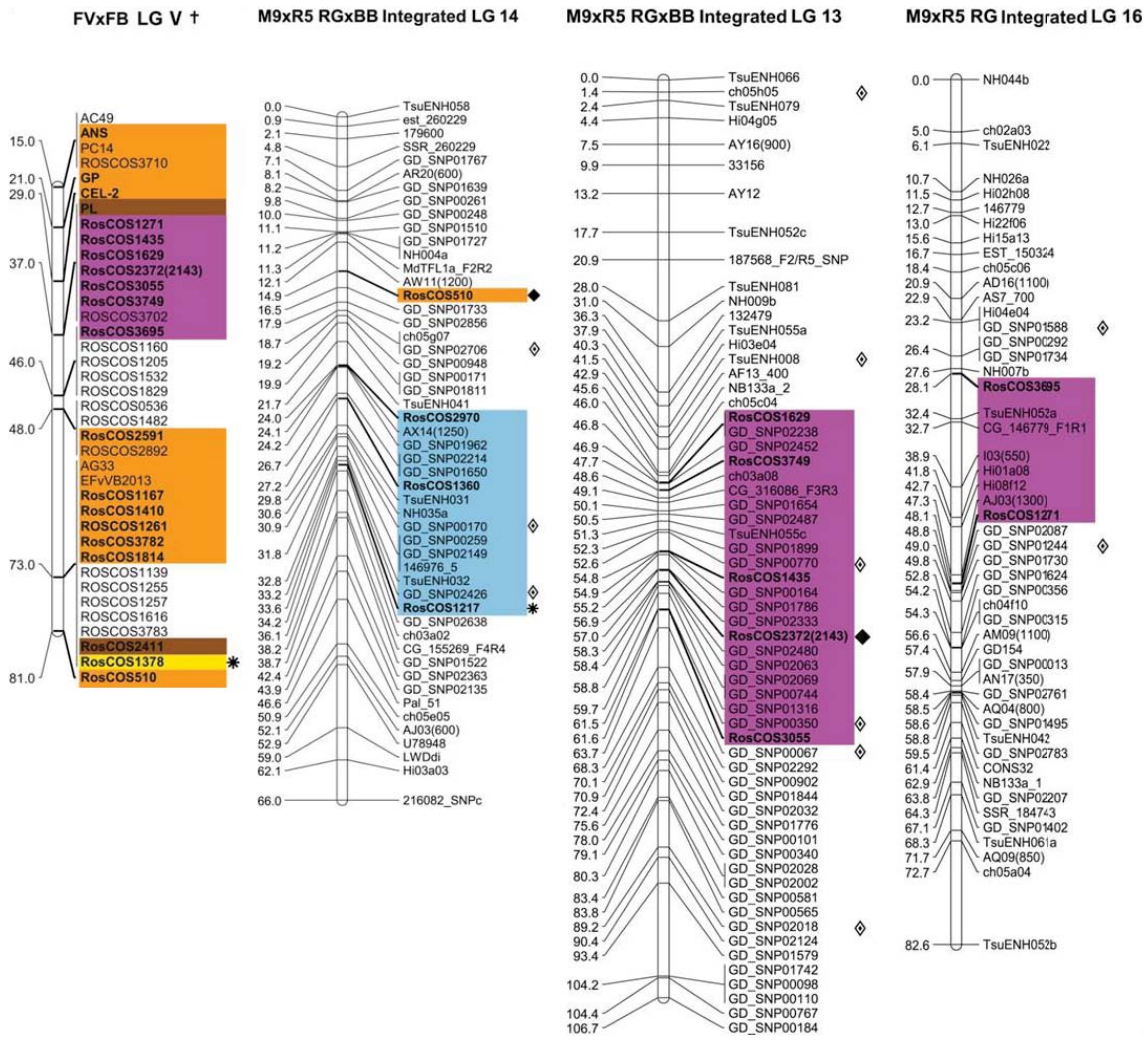

FVxFB LG II †
M9xR5 RGxBB Integrated LG 5 Inverted †□
M9xR5 StK Integrated LG 10 †■
M9xR5 Consensus LG 8 ×○
M9 A689 Integrated LG 15
M9 Fuji Integrated LG 7 †■×
EKO
APB
RosCOS530
RosCOS1222(596)
RosCOS2134
RosCOS1338
RosCOS52
RosCOS1272
ADH
RosCOS1341
RosCOS2054
RosCOS2596
RosCOS3781
RosCOS2118
RosCOS1125
ACL5
AMY
DFR
RosCOS1549(2673)
RosCOS1378
RosCOS2105
RosCOS1367
Supplementary
Figure S1b



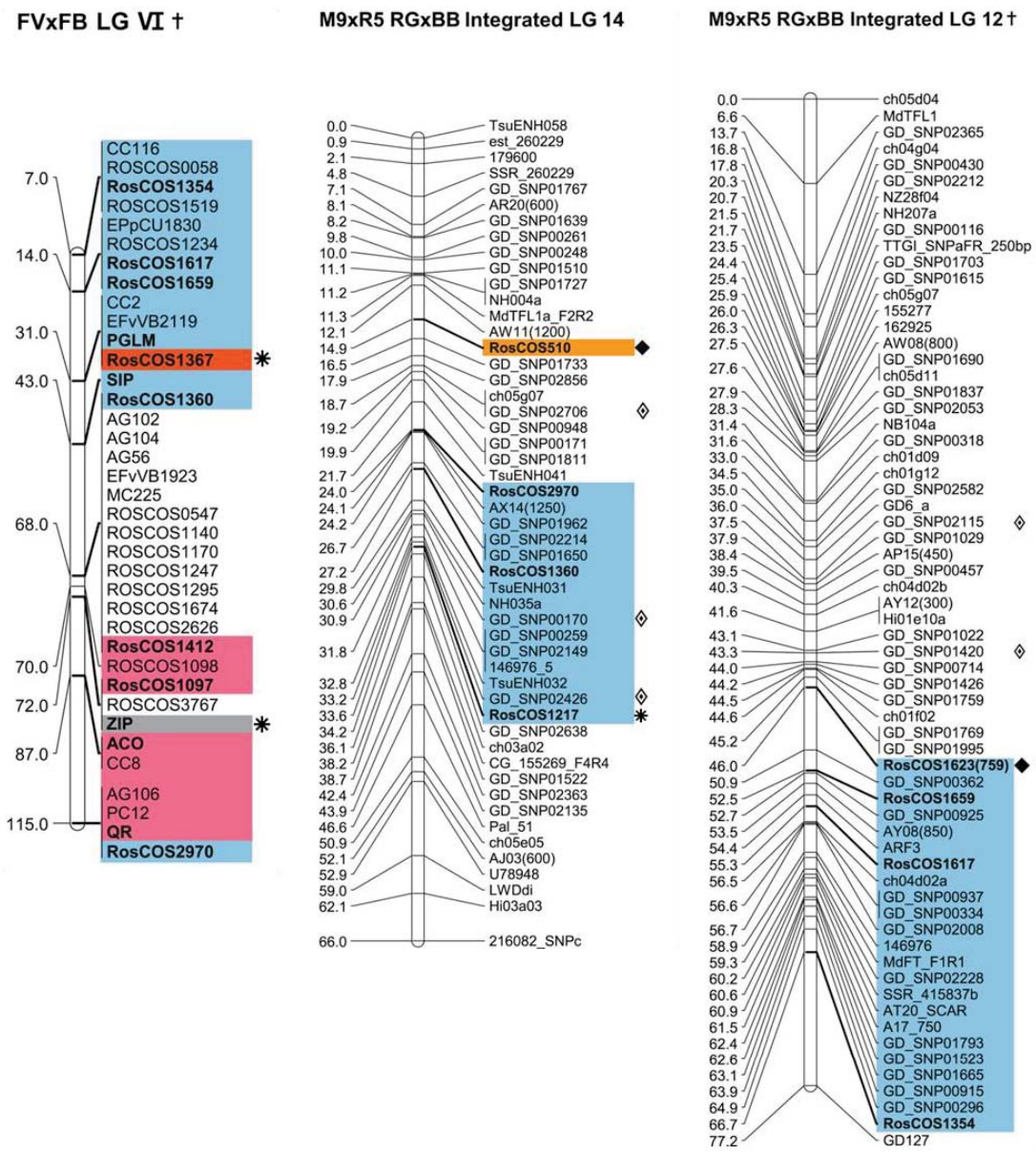
Supplementary
Figure S1c



Supplementary
Figure S1d



Supplementary
Figure S1e



Supplementary
Figure S1f

2.10 Addendum

I have included this addendum in response to comments made by the thesis examiners that the paper as published is too focused on RosCOS markers over SSR, either genomic or cDNA, as a useful source of polymorphism for genetic mapping in Rosaceae genera. In response to this comment, I note that in a search for SSR in a subset of RosCOS markers used for primer design in construction of the *Rubus* linkage map (Chapter 3), I identified only one marker (RiRosCOS510SSR) that contained an SSR and amplified a polymorphism in the progeny of 96395S1 × 'Latham'. In addition, during marker development for the *Rubus* linkage map, SSR markers derived from either genomic or cDNA had a low incidence of polymorphism [see Tables 3-1 and 3-4; see also Lewers et al. (2005) and Decrooq et al. (2003)]. Similar low polymorphism was also found when primer pairs designed from EST-SSR sequences for *Pinus* species were used for comparative analysis (Chagné et al. 2004). As concluded in Chapters 3 and 5, Rosaceae species are generally highly heterozygous, yet polymorphism between species and even genera is often difficult to detect. Therefore, I defend my time investment in developing and use of RosCOS markers as another source of polymorphism detection for comparative analysis within Rosaceae.

The use of RosCOS markers for the comparison and genome alignment of apple and strawberry support the goals of my thesis, which are to develop and assess new types of markers useful for genotyping members of Rosaceae and evaluate the effectiveness and efficiency of these markers.

CHAPTER THREE

This chapter published as:
***Bushakra JM**, Stephens MJ, Atmadjaja AN, Lewers KS, Symonds VV, Udall JA, Chagné D, Buck EJ, Gardiner SE (2012) Construction of black (*Rubus occidentalis*) and red (*R. idaeus*) raspberry linkage maps and their comparison to the genomes of strawberry, apple, and peach. Theoretical and Applied Genetics DOI: 10.1007/s00122-012-1835-5. See Appendix for reprint.*

3 Construction of black (*Rubus occidentalis*) and red (*R. idaeus*) raspberry linkage maps and their comparison to the genomes of strawberry, apple, and peach

3.1 Abstract

The genus *Rubus* belongs to the Rosaceae and is comprised of 600-800 species distributed world-wide. To date, genetic maps of the genus consist largely of non-transferable markers such as amplified fragment length polymorphisms (AFLP). An F_1 population developed from a cross between an advanced breeding selection of *Rubus occidentalis* (96395S1) and *R. idaeus* 'Latham' was used to construct a new genetic map consisting of DNA sequence-based markers. The genetic linkage maps presented here are constructed of 131 markers on at least one of the two parental maps. The majority of the markers are orthologous, including 14 Rosaceae conserved orthologous set (RosCOS) markers, and 60 new gene-based markers developed for raspberry. Thirty-four published raspberry simple sequence repeat (SSR) markers were used to align the new maps to published raspberry maps. The 96395S1 genetic map consists of six linkage groups (LG) and covers 309 cM with an average of 10 cM between markers; the 'Latham' genetic map consists of eight LG and covers 561 cM with an average of 5 cM between markers. We used BLAST analysis to align the orthologous sequences used to design primer pairs for *Rubus* genetic mapping with the genome sequences of *Fragaria vesca* 'Hawaii 4', *Malus* × *domestica* 'Golden Delicious', and *Prunus* 'Lovell'. The alignment of the orthologous markers designed here suggests that the genomes of *Rubus* and *Fragaria* have a high degree of synteny and that synteny decreases with phylogenetic distance. Our results give unprecedented insights into the genome evolution of raspberry from the putative ancestral genome of the single ancestor common to Rosaceae.

3.2 Introduction

The Rosaceae is a large and diverse plant family with many members of economic and nutritional importance. Members of the sub-family Rosoideae, including *Fragaria* (strawberry) and *Rubus* (raspberry and blackberry), have the same base chromosome number of $x = 7$, similar fruit form, and are closely related based on chloroplast and nuclear DNA phylogenies (Potter et al. 2007), as well as morphological data (Eriksson

et al. 2003). These lines of evidence suggest colinearity between the *Fragaria* and *Rubus* genomes; however, a lack of transferable markers has made genomic comparisons between these genera unachievable to date (Lewers et al. 2005).

The genus *Rubus*, with an estimated 600-800 species, includes red raspberry (*R. idaeus* L.), black raspberry (*R. occidentalis* L.) both of subgenus *Idaeobatus* (Focke) Focke, and blackberries such as *R. ursinus* Cham. & Schltdl., and *R. laciniatus* Willd., both of subgenus *Rubus* (Thompson 1995b). Sub-family Rosoideae is sister to Spiraeoideae, to which *Malus* (apple) and *Prunus* (peach and other stonefruits) belong (Potter et al. 2007) (Fig. 3.1). Comparative genetic mapping has been carried out between *Malus* and *Prunus* using restriction fragment length polymorphism (RFLP) markers and isozymes (Dirlewanger et al. 2004b); and between *Prunus* and *Fragaria* using RFLP and single nucleotide polymorphism (SNP) markers (Vilanova et al. 2008). A recent comparison of the *Malus* genome to the reference genetic maps of *Prunus* ['Texas' × 'Earlygold', TxE (Joobeur et al. 1998)] and *Fragaria* [*F. vesca* × *F. bucharica*, FVxFB (Sargent et al. 2006)] by Illa et al. (2011), and the comparison of genetic linkage maps between *Malus* ['Malling 9' × 'Robusta 5', M.9xR5 (Celton et al. 2009b)] and FVxFB (Sargent et al. 2006) by Bushakra et al. (2012b) utilized markers designed from transferable Rosaceae conserved orthologous set (RosCOS) sequences (Cabrera et al. 2009). The recent release of draft genome sequences for three Rosaceae genera [woodland strawberry 'Hawaii 4' (*F. vesca* ssp. *vesca*) (Shulaev et al. 2011), 'Golden Delicious' apple (*Malus* × *domestica* Borkh.) (Velasco et al. 2010), and peach doubled haploid 'Lovell' (*Prunus persica*) (Sosinski et al. 2010)], along with the availability of reference maps for *Prunus* and *Fragaria*, as well as RosCOS-derived markers, have provided tools that have allowed the discovery of regions of conserved genomic synteny. These discoveries have led to the proposal of a common ancestor for the family with a genome structure of nine chromosomes (Vilanova et al. 2008; Velasco et al. 2010).

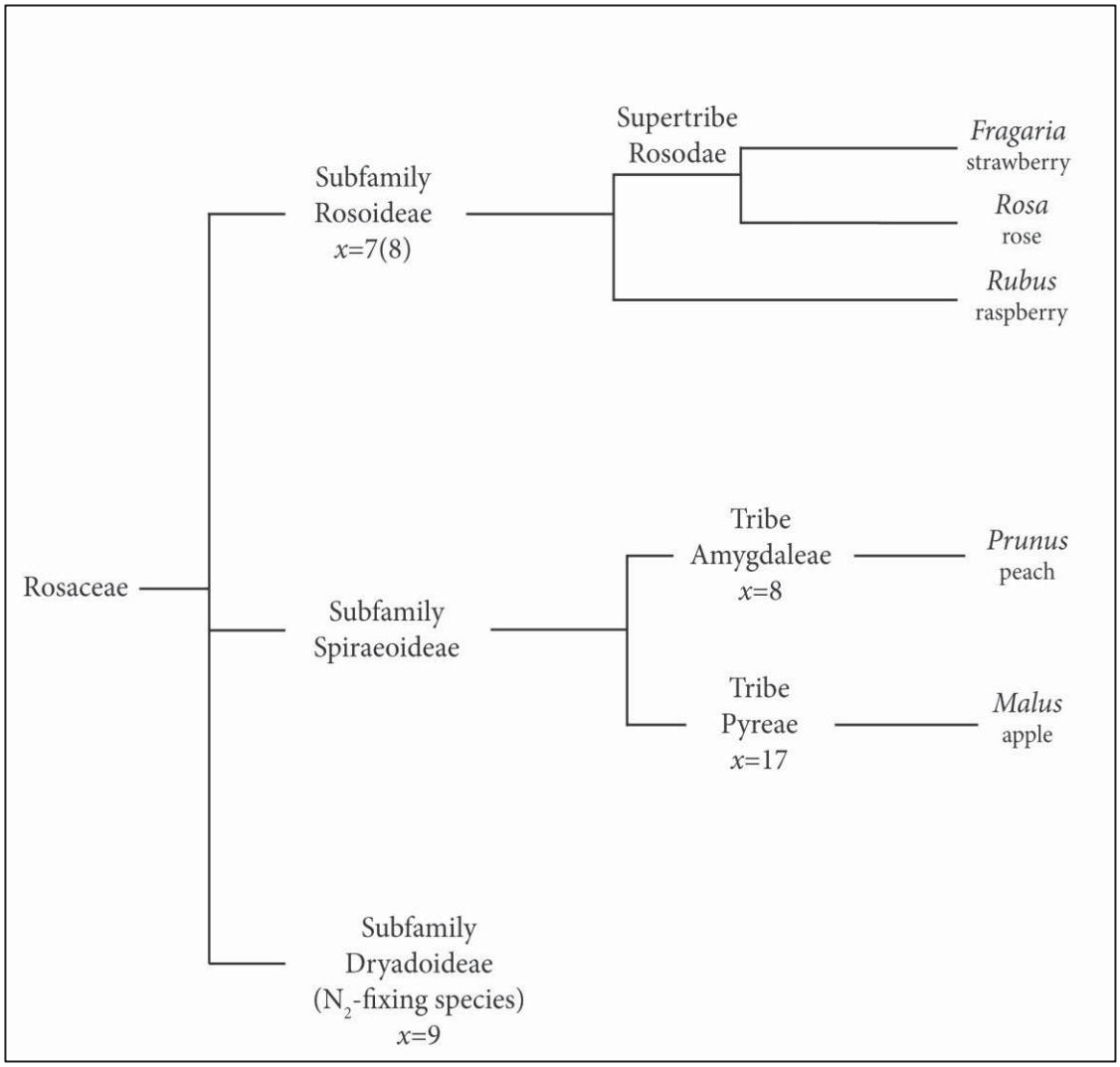


Fig. 3.1: Simplified and abbreviated Rosaceae phylogeny. The diagram represents the genera included in this study. Branches approximate shared ancestry and are not to scale nor all-inclusive. Base chromosome numbers (*x*) are indicated. The phylogeny illustrates the relative distances between the sub-families Rosoideae and Spiraeoideae, and between the genera within each sub-family.

Several genetic maps have been constructed for red raspberry using a variety of marker types, including amplified fragment length polymorphic (AFLP), genomic simple sequence repeat (SSR), expressed sequence tag-SSR (EST-SSR), and gene-based markers (Graham et al. 2004; Graham et al. 2006; Woodhead et al. 2008; Woodhead et al. 2010; Sargent et al. 2007b). The first published genetic linkage map of *Rubus* was developed using AFLP and SSR markers in progeny from an intraspecific cross between elite cultivars of red raspberry 'Latham' and 'Glen Moy'. This map of nine linkage groups (LG) consists of 273 markers, including 30 SSR and four EST-SSR

(Graham et al. 2004). Quantitative trait loci for two morphological traits – variation in extent of spines, and root sucker density and spread from mother plant – were genetically mapped. However, since two of the LG are composed of non-transferable AFLP markers only, the map is not useful for comparative analyses between genera.

Subsequently, the 'Latham' × 'Glen Moy' genetic linkage map was updated, and the population used to genetically map gene *H* controlling cane pubescence and associated disease resistance to some common *Rubus* fungal pathogens (Graham et al. 2006). Although the addition of 20 SSR markers reduced the number of LG from nine to eight, AFLP is still the predominant marker type. EST libraries from root tissue derived from 'Latham' and meristematic bud tissues derived from 'Glen Ample' have recently been generated (Woodhead et al. 2008). These libraries were mined to identify EST-SSR (Woodhead et al. 2008) and to develop functional gene-based markers (Woodhead et al. 2010). The newly developed markers were placed on the 'Latham' × 'Glen Moy' linkage map, bringing the total number of gene-based markers to 97 on seven LG, including 37 orthologous gene-based markers from *Prunus* Group 6 (G6).

A second genetic linkage map that consists of seven LG constructed of 95 AFLP, 22 SSR, and two gene-based markers, was developed from another intra-specific cross between elite cultivars of red raspberry 'Malling Jewel' and 'Malling Orion' (Sargent et al. 2007b). While this map is considered saturated, the high number of AFLP markers limits marker transferability.

The mapping of gene-based markers from *Prunus* G6 on to *Rubus* LG2 demonstrated by Woodhead et al. (2010), and the more efficient inter-generic transferability of EST-SSR relative to genomic SSR markers (Lewers et al. 2005; Zorrilla-Fontanesi et al. 2011b) suggests that orthologous markers are key to successful comparative genomic analyses in Rosaceae.

Two recent studies on transferability of *Fragaria*-derived markers to *Rubus* and *Rosa* illustrate the importance of marker source for successful transfer between Rosaceae genera. Lewers et al. (2005) assessed the transferability of GenBank-derived *Fragaria* EST-SSR to be 32% to blackberry, and 20% to raspberry. In the same study, the transferability of primers from GenBank-derived *Fragaria* genomic SSR was lower, at 26% in blackberry, and 18% in raspberry. Zorrilla-Fontanesi (2011b) reported that primer pairs developed from *Fragaria* EST-SSR successfully amplified a product 20%

of the time in *Rubus*, and 29% of the time in *Rosa*. In the same study, primer pairs developed from *Fragaria* genomic SSR successfully amplified a product 16% of the time in *Rubus*, and 19% of the time in *Rosa*. In general, these results indicate that the transferability of EST-SSR markers is higher than that of genomic SSR markers, though neither marker type is especially efficient; therefore additional PCR-based markers are needed to allow for the comparison of the *Rubus* genome with other members of Rosaceae.

The aim of this work was to construct genetic linkage maps of black and red raspberry based on an inter-specific progeny and enriched in orthologous molecular markers. This genetic map will be the first comprised entirely of molecular markers designed from sequenced DNA, as opposed to markers derived from anonymous DNA. The number of orthologous, gene-based markers will augment the existing number of transferable markers to facilitate the assessment of syntenic relationships of raspberry with strawberry, apple and peach. Comparative genomic studies among these economically important genera will provide new insights into the degree of genetic conservation at the family level. This will in turn provide new tools for Rosaceae crop geneticists and breeders, especially for breeders of less well-studied Rosaceae crops and ornamentals.

3.3 Methods

3.3.1 Plant material and DNA extraction

An inter-specific cross between black raspberry *R. occidentalis* 96395S1 (S1) and red raspberry *R. idaeus* 'Latham' made in 2005 resulted in 500 F_1 progeny (labeled S1xLatham) that were planted during 2005 at Plant & Food Research, Motueka, New Zealand. S1 lacks spines and has dark purple fruit, while 'Latham' has spines and red fruit. Pedigrees for the parents are illustrated in Fig. 3.2. The genetic mapping population is comprised of 155 individuals selected from the first 200 plants in the planting block (EA501-EA701) on the basis of production of sufficient fruit for analysis of polyphenolic content. DNA was extracted from young leaves using a modified CTAB method (Kobayashi et al. 1998).

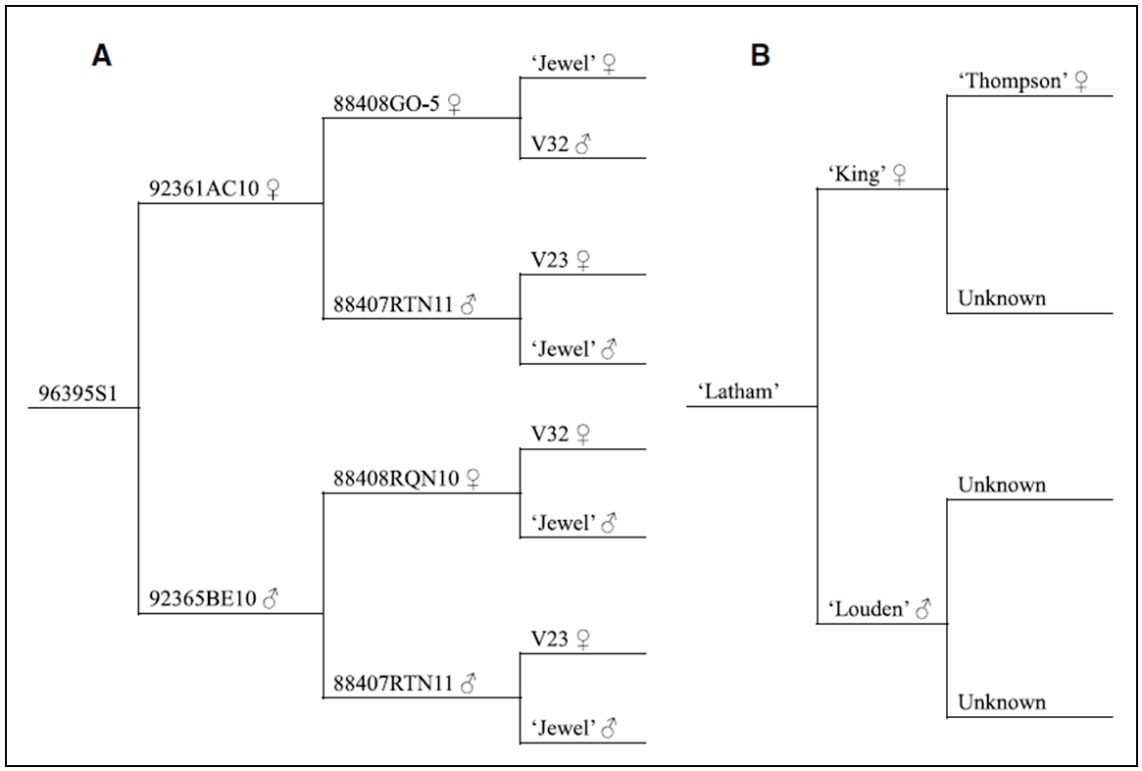


Fig. 3.2: Parental pedigrees. (A) Partial pedigree of black raspberry parent 96395S1 (*Rubus occidentalis*). Great grandparents V32 and V23 are full siblings; grandparents 88408GO-5 and 88408RQN10 are full siblings; R indicates the individual was produced by reciprocal cross. Individuals used in the crosses 88408 and 88407 are half-siblings as they are both derived from 'Jewel'. Parents 92361AC10 and 92365BE10 share the same pollen donor (88407RTN11). (B) Known pedigree of red raspberry parent 'Latham' (*Rubus idaeus*).

3.3.2 Genetic markers

Prior to screening the 155 progeny used for linkage analysis, an initial set of 269 primer pairs covering 250 loci were pre-screened for amplification of a polymorphic product over the mapping parents and a subset of 14 F_1 individuals using the high-resolution melting (HRM) technique (Wittwer et al. 2003). The 269 primer pairs were comprised of 194 published *Rubus* primer pairs (Amsellem et al. 2001; Sargent et al. 2009c; Woodhead et al. 2008; Woodhead et al. 2010; Zorrilla-Fontanesi et al. 2011b; Castillo et al. 2010; Graham et al. 2004; Graham et al. 2006; Lewers et al. 2008; Lopes et al. 2006), (28 of which did not include amplicon or GenBank accession information), 45 RosCOS primer pairs (Cabrera et al. 2009), and 30 previously unpublished primer pairs from *Prunus persica* (peach) and *Rubus* sp. (Set A) (Table 3-1). A subset of 45 primer pairs, including 11 RosCOS, that did not amplify a product under the conditions used

for HRM, were amplified under polymerase chain reaction (PCR) conditions used for direct PCR product sequencing and analyzed for amplicon size with agarose gel electrophoresis. Primer pairs that amplified a single PCR product in both parents were sequenced in both directions using BigDye Terminator v. 3.1 (BigDye™ Terminator Version 3.1 Ready Reaction Cycle Sequencing Kit, Applied Biosystems, Inc., Carlsbad, CA). The sequenced fragments were scanned for single nucleotide polymorphisms (SNP) using Sequencher™ 4.5 (demo version, Ann Arbor, MI) and new locus-specific primer pairs compatible with HRM were designed. As the pre-released *Rubus* genome sequence became available (J. Udall, personal communication), this was used to redesign the primer sequences for use with HRM of 195 published markers covering 72 loci (Set B), and to design 754 new HRM-based *Rubus*-specific primer pairs covering 198 loci (Set C) (Table 3-1). Sequences from *Arabidopsis thaliana*, *Cucumis melo*, *Fragaria* sp., *Malus* × *domestica*, *Prunus persica*, *Rubus* sp., *Solanum lycopersicum*, and Rosaceae conserved orthologous set, were searched using BLAST in the pre-released *Rubus* genome (Table 3-1; Supplemental Table S3-1; Supplemental Table S3-2). PCR primers were designed using Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) (Bushakra et al. 2012b).

Table 3-1: Results and sources of primer sequences used for marker development and molecular mapping in *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'.

	Type of sequence	Source of sequence	Primer pairs tested	Loci covered	Loci mapped	Markers polymorphic, not mapped	Markers monomorphic	Markers no amplification	Markers complex pattern	Primers not used, alternative primer sequence	% amplification success†	% mapping success‡	Reference
Set A: Initial primer pairs screened	DNA	Fsp	15	6	2	1	1	5	5	1	64	22	Sargent et al. 2007
	DNA	Fsp	9	9			1	6	2		33	0	Zorrilla-Fontanesi et al. 2011
	DNA	Mxd	1	1			1				100	0	Sargent et al. 2007
	DNA	Ra	2	2				2			0	0	Amsellem et al. 2001
	DNA	Rh	3	3			2		1		100	0	Castillo et al. 2010
	DNA	Rh	10	10			4	4	2		60	0	Lopes et al. 2006
	DNA	Ri	1	1			1				100	0	Castillo et al. 2010
	DNA	Ri	5	2	2		1	2			60	67	Previously unpublished
	DNA	RL	50	50	20	4	10	9	7		86	47	Graham et al. 2004
	DNA	RL	35	35	1	4	8	15	7		57	5	Graham et al. 2006
	DNA	RL	1	1				1			0	0	Zorrilla-Fontanesi et al. 2011
	Sub-total		132	120	25	9	29	42	26	1			
	EST	Fsp	2	1	1			1			50	100	Zorrilla-Fontanesi et al. 2011
	EST	Pp	22	22	1	1		13	7		41	11	Previously unpublished
	EST	RCOS	45	45	1	1	8	24	11		47	5	Cabrera et al. 2009
	EST	Rh	20	20	4	1	2	12	1		40	50	Lewers et al. 2008
	EST	Ri	3	1	1		1	1			67	50	Previously unpublished
	EST	RL	2	2			2				100	0	GDR
	EST	RL	7	7	2	1	2	2			71	40	Graham et al. 2004
	EST	RL	25	25	8	1	7	6	3		76	42	Woodhead et al. 2008
	mRNA	Fsp	11	7	1		1	8	1		27	33	Sargent et al. 2007
	Sub-total		137	130	19	5	23	67	23	0			
	Total		*269*	*250*	*44*	*14*	*52*	*109*	*49*	*1*			
Set B: Primers redesigned	DNA	Ra	2	2	1		1				100	50	Amsellem et al. 2001
	EST	RL	8	8	3		4		1		0	0	GDR
	DNA	Ri	2	1			2				100	0	Graham et al. 2004
	EST	RCOS	100	22	14	1	44	19	11	11	79	20	Cabrera et al. 2009
	EST	Ri	7	7	3		1	3			57	75	Woodhead et al. 2008
	EST	Ri	76	32	16	4	22	4	25	5	94	24	Woodhead et al. 2010
	Total		*195*	*72*	*37*	*5*	*74*	*26*	*37*	*16*			
Set C: Primers newly designed	DNA	Mxd, Pp, Ri, Fsp	41	13	6	1	15	9	4	6	74	23	
	EST, cDNA, mRNA, protein	At, Mxd, Ri, Ro, Rsp, Fsp	166	75	27	2	94	19	12	12	88	20	
	EST, cDNA, mRNA	At, Sl, Cm, Rsp	496	89	22	1	250	98	67	58	78	6	
	DNA	At, Sl, Cm	51	21	0	3	20	18	4	6	60	0	
	Total		*754*	*198*	*55*	*7*	*379*	*144*	*87*	*82*			
	Grand Totals		**1218**	**520**	**136**	**26**	**505**	**279**	**173**	**99**			
939 markers amplified			**77.1%**		**14.5%**		**53.8%**		**18.4%**				

†Calculated as the number of primer pairs that amplified a product, divided by the number of primer pairs tested, minus duplicates.
‡Calculated as the number of mapped loci, divided by the number of primer pairs that amplified a product, minus duplicates.
Abbreviations: At: *Arabidopsis thaliana*; Cm: *Cucumis melo*; RCOS: Rosaceae conserved orthologous set; Fsp: *Fragaria* sp.; Fxa: *Fragaria* × *ananassa*; Mxd: *Malus* × *domestica*; Pp: *Prunus persica*; Rsp: *Rubus* sp.; Ra: *Rubus alceifolius*; Rh: *Rubus hybrida*; Ri: *Rubus idaeus*; RL: 4*n Rubus* L.; Ro: *Rubus occidentalis*; Sl: *Solanum lycopersicum*; GDR: Genome Database for Rosaceae (www.rosaceae.org)

3.3.3 PCR and HRM conditions

HRM is a high-throughput, closed-tube melting curve analysis that uses a high-affinity double-stranded DNA-binding fluorescent dye to discriminate between genotypes relative to the presence or absence of a SNP (Reed and Wittwer 2004; Wittwer et al. 2003) or small insertion/deletion (indel) (Montgomery et al. 2007). The sensitivity of resolution, through-put capacity, and ease of use of HRM make this a useful technique for genotyping. PCR prior to HRM analyses were performed in a Roche LightCycler® 480 (Foster City, CA) in volumes of either 7 µl (384-well plates) or 10 µl (96-well plates) as in Bushakra et al. (2012b). The PCR cycles were followed by HRM analysis as in Chagné (2008). PCR for direct sequencing and amplicon visualizations were carried out as in Bushakra et al. (2012b).

3.3.4 Data collection and linkage analysis

Primer pairs that amplified a polymorphic product in the mapping population were scored based on the differences in HRM melting curves (Wittwer et al. 2003). JoinMap® v3.0 (van Ooijen and Voorrips 2001) software was used to construct the linkage map using the double pseudo-testcross mapping strategy (Grattapaglia and Sederoff 1994). The LOD threshold for grouping was 5 and the Kosambi mapping function was used to convert recombination units into genetic distances.

3.3.5 Genetic map comparison to published Rosaceae genomes

Sequences orthologous to the gene-based *Rubus* markers were searched in the genomes of *Fragaria vesca* (Shulaev et al. 2011), *Malus* (Velasco et al. 2010), and *Prunus* (Sosinski et al. 2010) using BLAST with a default expected value of E^{-10} (Altschul et al. 1990). The physical locations of the orthologous regions in the queried genomes were used to construct physical maps for *F. vesca*, *Malus*, and *Prunus* using MapChart© (Voorrips 2002). The physical map of each genus was aligned independently to the S1 and 'Latham' linkage maps using the orthologous markers in common. Percentage of

map similarity between the 'Latham' parental map and the physical maps for *F. vesca*, *Malus*, and *Prunus* was determined by calculating the distance between the top- and bottom-most markers in common. The BLAST-generated physical maps were compared with the published reference genetic maps for FV×FB (FLGI-FLGVII) and T×E (G1-G8) (www.rosaceae.org) to check the accuracy of the BLAST-derived map.

3.3.6 Ancestral chromosome contribution determination

To provide an estimate of the ancestral chromosome contribution to and proportion of each *Rubus* LG, the orthologous markers genetically mapped in S1×Latham were compared with the orthologous sequence positions, as determined by BLAST analysis, in the genomes of *F. vesca*, *Malus*, and *Prunus*. Markers in common were used to align the 'Latham' genetic linkage map to the BLAST-generated physical maps of *F. vesca*, *Malus*, and *Prunus*, and with the published ancestral representations (Illa et al. 2011; Vilanova et al. 2008).

3.4 Results

3.4.1 *Rubus* genetic linkage map construction

In total, 1218 primer pairs were evaluated in the S1×Latham population (Table 3-1). Of the 269 initial primer pairs including the 45 whose products were directly sequenced (Set A), 44 (16%) amplified a product that was genetically mapped; 14 (5%) amplified a polymorphic product that was not mapped; 52 (19%) produced a monomorphic amplicon, 109 (40.5%) did not amplify a PCR product; 49 (18%) exhibited complex or unclear melting curves in the progeny screening set, and one pair was not used because it was an alternative primer sequence for a mapped marker.

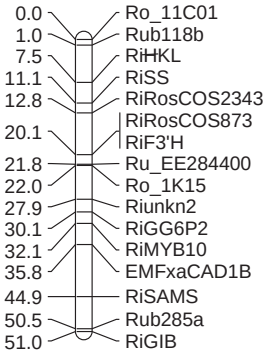
Of the 195 primer pairs that were redesigned using the *Rubus* genome sequence (Set B), 37 (19%) amplified a product that was genetically mapped; 5 (2.5%) amplified a polymorphic product that was not mapped; 74 (38%) produced a monomorphic amplicon, 26 (13%) did not amplify a PCR product, and 37 (19%) exhibited complex or unclear melting curves in the progeny screening set, and 16 (8%) were not used because they were alternative primer sequences for mapped markers.

Of the 754 newly designed primer pairs (Set C), 123 (16.3%) produced amplicons that were determined to be polymorphic and screened over the set of 155 progeny; 379 primer pairs (50.3%) produced a monomorphic amplicon, 144 (19%) did not amplify a

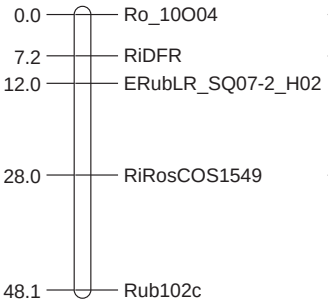
PCR product, and 87 primer pairs (11.5%) produced amplicons that exhibited complex or unclear results in the progeny screening set. Of the 123 polymorphic amplicons, 55 (45%) were genetically mapped. In summary, of the total 754 primer pairs designed, 55 (7%) were genetically mapped.

The linkage groups illustrated in Fig. 3.3 are labeled with the original linkage group (OLG) number as assigned by Graham et al. (2004) shown first in parentheses, followed by the proposed *Rubus* linkage group (RLG) number to place the RLG in the same order as the published FVxFB LG. Of the 136 mapped markers, 131 were placed on at least one of the two parental maps. The remaining five markers were removed because of skewed segregation ratios. The parental genetic map for 96395S1 consists of 29 markers, 17 of which are unique to S1, averaging one marker every 10 cM, and covering 306 cM over six LG. S1 had only one heterozygous marker (Ro_1K15) in RLG1. The parental genetic map for 'Latham' consists of 114 markers, 101 of which are unique to 'Latham', averaging one marker every 5 cM, and covering 561 cM over seven LG with RLG7 in two parts. The two parents share 13 markers (Fig. 3.3; Supplemental Table S3-1). The 'Latham' genetic linkage map was used for comparative figures with *Fragaria*, *Malus*, and *Prunus* since it is the more complete map of the two parents. Consensus LG for S1 and 'Latham' could not be constructed due to a lack of markers in common between the parents.

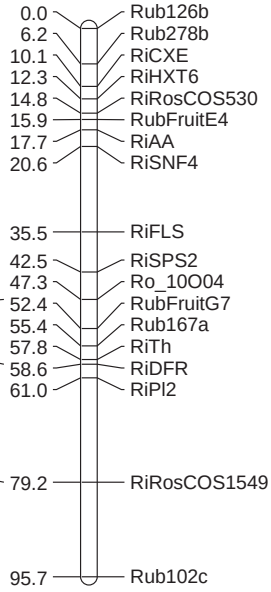
Latham (OLG6) RLG1



96395S1 (OLG4) RLG2

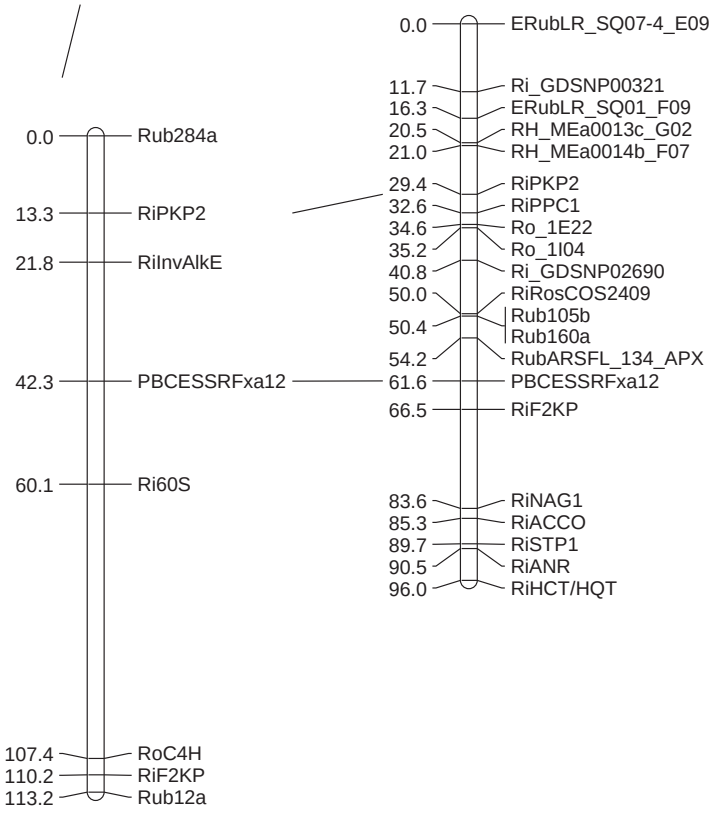


Latham (OLG4) RLG2

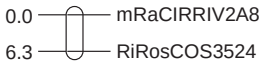


96395S1 (OLG2) RLG3

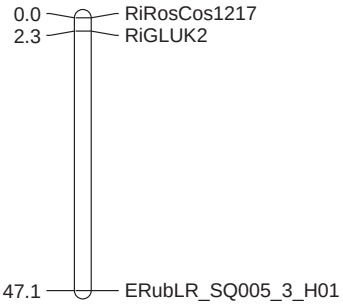
Latham (OLG2) RLG3



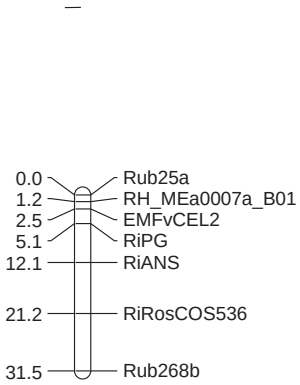
96395S1 (OLG7) RLG4



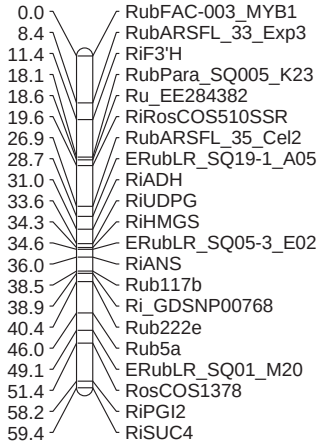
Latham (OLG7) RLG4



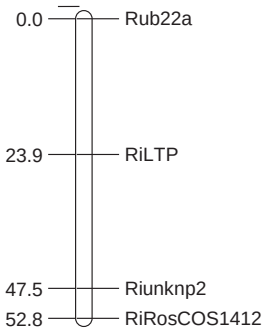
96395S1 (OLG5) RLG5



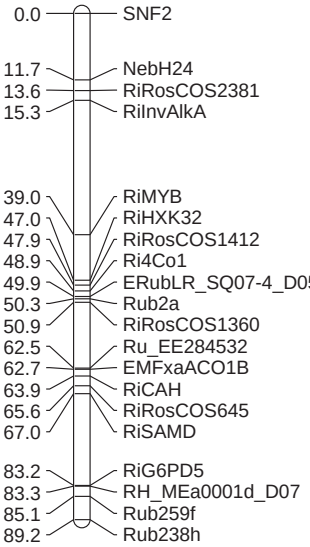
Latham (OLG5) RLG5



96395S1 (OLG3) RLG6



Latham (OLG3) RLG6



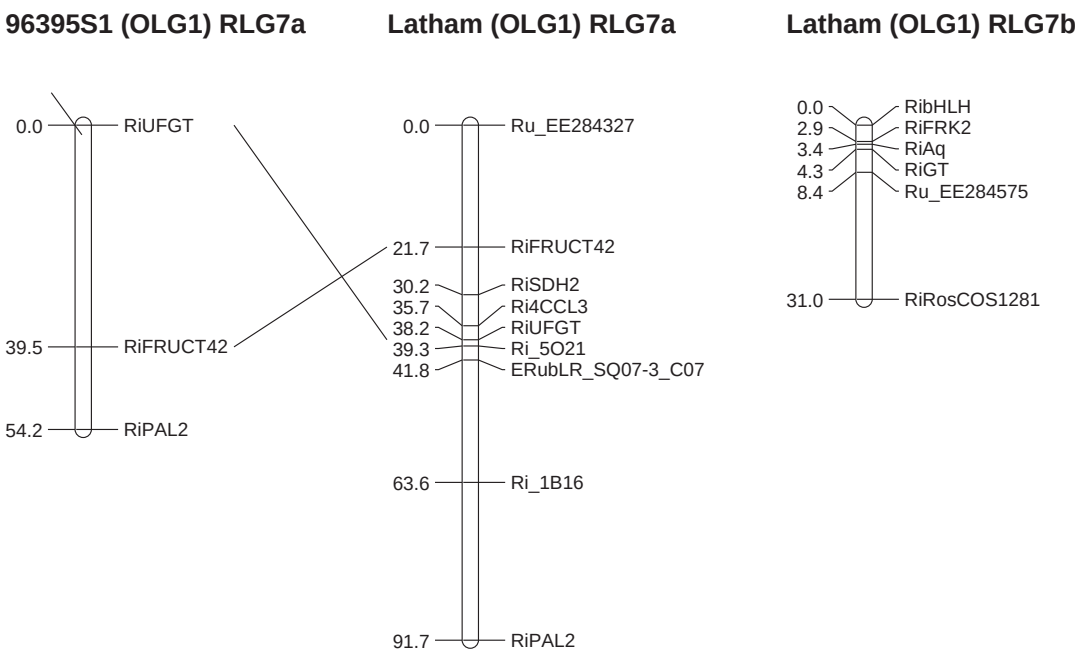


Fig. 3.3: Genetic linkage maps of *Rubus occidentalis* 96395S1 and *R. idaeus* 'Latham'. Original linkage group (OLG) nomenclature after Graham et al. (2004) is given in parentheses, followed by proposed *Rubus* linkage group (RLG) nomenclature corresponding with *Fragaria* LG nomenclature. Map distances are in centimorgans (cM). RLG2 through RLG7a are presented for 96395S1; RLG1 through RLG7a & 7b are presented for 'Latham'. Connecting lines indicate markers in common.

Table 3-2: Linkage group descriptions for the parental maps of *Rubus occidentalis* 96395S1 and *R. idaeus* 'Latham'.

	Markers mapped in 96395S1	Linkage group size (cM)	Markers mapped in 'Latham'	Linkage group size (cM)	Markers shared between parental maps†
RLG1			16	51	1
RLG2	5	48	18	96	4
RLG3	8	113	21	96	3
RLG4	2	6	3	47	0
RLG5	7	31	21	59	1
RLG6	4	53	20	89	1
RLG7a	3	54	9	92	3
RLG7b			6	31	0
Totals	**29**	**305**	**114**	**561**	**13**

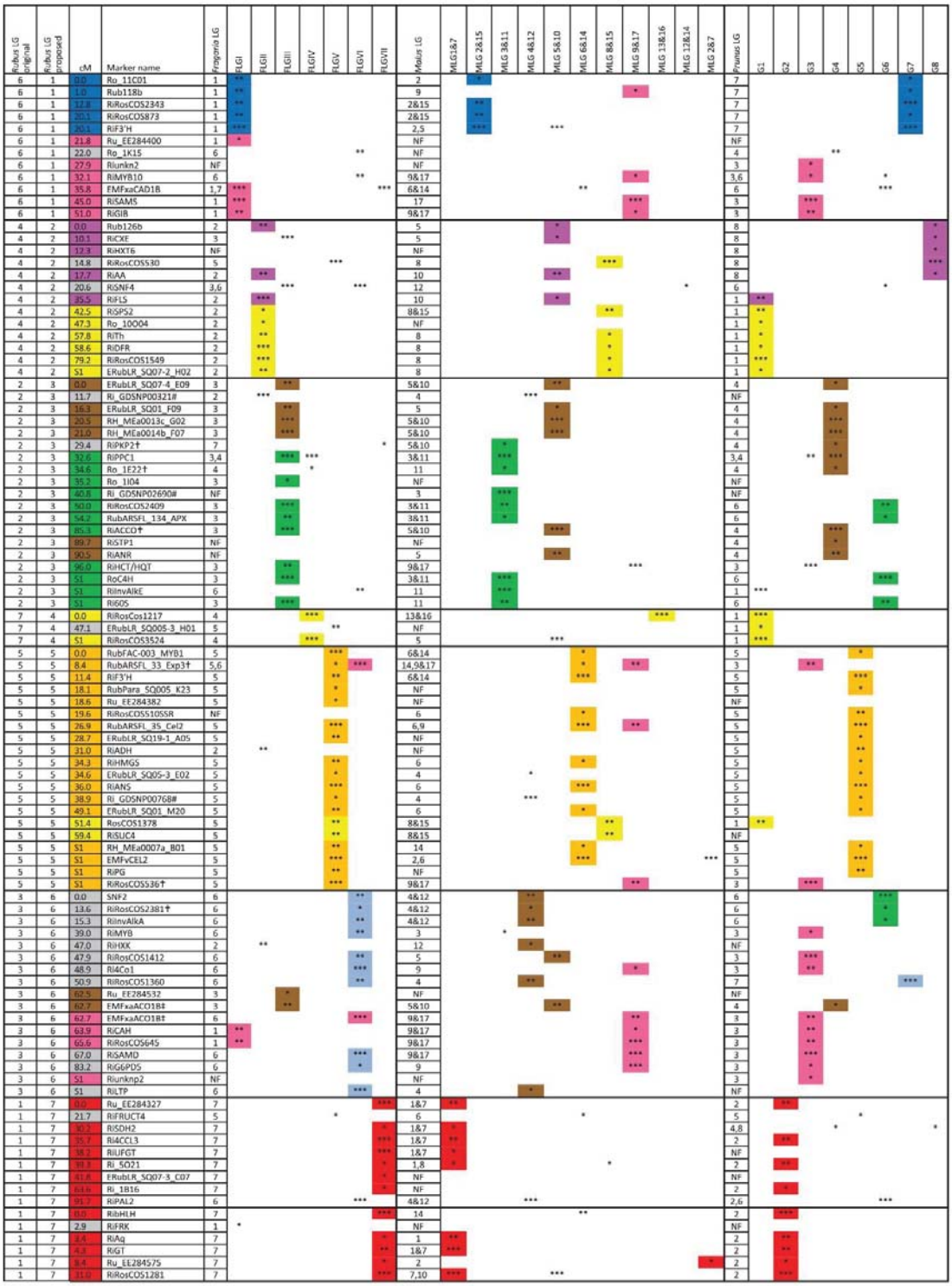
†Insufficient markers available for map integration.
Linkage group size is in centimorgans (cM).

3.4.2 Comparative genome mapping between *Rubus* and *Fragaria*

The comparison of the 131 markers that comprise the linkage maps of 96395S1 and 'Latham' with the draft genome assembly of *F. vesca* using BLAST analysis of marker sequences identified 90 markers in common between the two genera and demonstrates a nearly 1:1 LG relationship (Fig. 3.4: Comparison of *Rubus idaeus* 'Latham' genetic linkage map with *Fragaria*, *Malus*, and *Prunus*. ; Fig. 3.5: Ancestral chromosome (A1-A9) contribution to the genomes of *Rubus* linkage groups ; Supplemental Fig. S3.1). BLAST expected values (E-value) ranged from $2.00E^{-06}$ with 92% identity, to $0.00E^{+00}$ with 99% identity, with only five markers having E-values greater than E^{-15} (Table 3-3; Supplemental Fig. S3.1; Supplemental Table S3-2). Based on the positions of 77 markers common to both genera, each of the seven *Rubus* LG could be aligned to one of the seven *Fragaria* chromosomes, ranging from two markers on RLG4, to 12 markers on RLG3, and covering 96% (538 cM of 561 cM) of the 'Latham' genetic linkage map (Fig. 3.4; Supplemental Fig. S3.1). Colinearity was largely conserved between *Rubus* and *Fragaria* although a few differences were observed in gene order. Of the 90 markers in common, five (EMFxaCAD1B, RiSNF4, RiPPC1, RubARSFL_33_Exp3 and EMFxaACO1B) had significant BLAST scores on two different LG; and 15 (16.6%) did not map to the homologous LG (Fig. 3.3). These 15 markers were designed

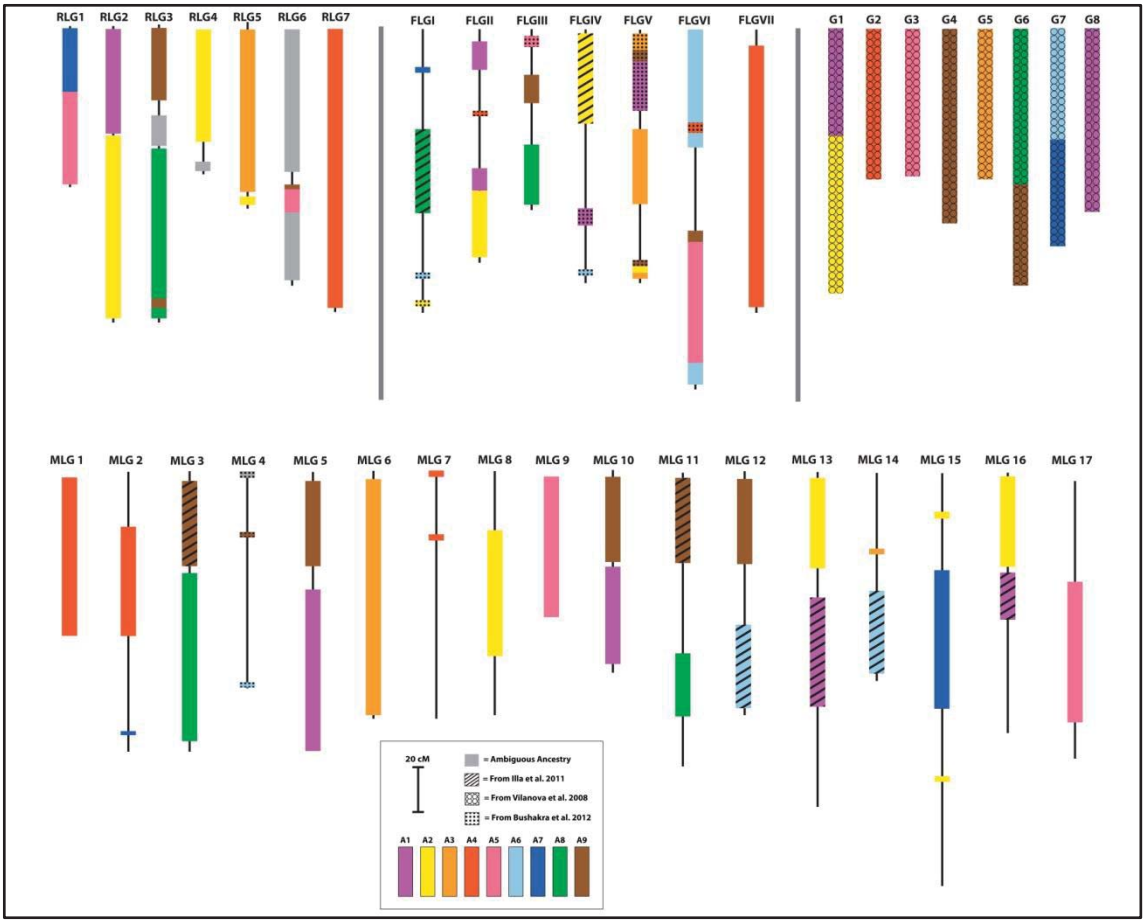
from candidate gene sequence (*RiMYB10*, *RiCXE*, *RiSNF4*, *RiPKP2*, *RiInvAlkE*, *RiADH*, *RiHXK*, *RiFRUCT4*, *RiPAL2* & *RiFRK*), *Rubus* SSR (Ro_1K15 & Ro_1E22), and *Rubus* EST (Ru_EE284365 & ERubLR_SQ005-3_H01). Marker Ri_GDSNP00321 was designed from a 'Golden Delicious' single nucleotide polymorphism (GDSNP) sequence.

Fig. 3.4: Comparison of *Rubus idaeus* 'Latham' genetic linkage map with *Fragaria*, *Malus*, and *Prunus*. Markers for which sequence was available for BLAST analysis are arranged based on the 'Latham' linkage map order; map distance is in centimorgans (cM). *Rubus* original linkage group (OLG) is based on numbering scheme by Graham et al. (2004). Colors indicate ancestral chromosome contribution (see Fig. 3.5 for color key based on Vilanova et al. 2008).



Apple LG separated by "&" indicate homeologous chromosomes; those separated by a comma are non-homeologous (Velasco et al. 2010).
Linkage groups separated by a comma in *Fragaria* and *Prunus* indicate markers located on non-homologous chromosomes.
Grey regions indicate ambiguous ancestral contribution.
Sequences for which similarity could not be found in the target genus are indicated as NF.
Markers unique to S1 are included at the end of each LG with "S1" in place of cM values.
***=E-values of $0.00E^{+00}$ to E^{-100}; **=E-values of E^{-99} to E^{-50}; *=E-values greater than E^{-50}.
‡ The BLAST analysis for marker EMFxaACO1B shows two equally likely loci in *Fragaria*, *Malus*, and *Prunus*. Only one locus is mapped in *Rubus*.
† These markers have two equally likely ancestral origins. When possible, the ancestral origin that agrees with *Fragaria* was chosen for the *Rubus* map.
# These markers were designed from *Malus* 'Golden Delicious' single nucleotide polymorphism (SNP) sequences.

Fig. 3.5: Ancestral chromosome (A1-A9) contribution to the genomes of *Rubus* linkage groups (RLG1-RLG7), *Fragaria* linkage groups (FLGI-FLGVII), *Prunus* groups (G1-G8), and *Malus* linkage groups (MLG1-MLG17).



The *Rubus* LG are in order to correspond to *Fragaria*, are to scale and represent the findings of this study. The ancestral contributions for all solid color regions in *Fragaria* and *Malus*, and for *Prunus* [circle pattern, Vilanova et al. (2008)] are supported by the findings of this study.
The *Fragaria* and *Malus* LG are based on information from Bushakra et al. (2012b) (dotted regions) and Illa et al. (cross-hatched regions).
For all genera: solid black lines indicate gaps where no markers were mapped genetically; colors follow Vilanova et al. (2008) with modifications to reflect Bushakra et al. (2012b). Grey regions represent areas of ambiguous ancestral contribution.

Table 3-3: Summary of the number and occurrence of markers in common between *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' and *Fragaria, Malus,* and *Prunus* genomes as determined by BLAST analyses.

	Markers in common with *Rubus*†	E-value range	Markers occurring once	Markers with significant second hits‡	Homeologous markers‡	Aligning markers ‡	Non-homologous markers	% of non-homologous markers
Fragaria	90	$2E^{-06}$ to 0	86	5		77	15	16.6
Malus	80	$5E^{-05}$ to 0	39	6	41	70	19	23.8
Prunus	86	$6E^{-13}$ to 0	81	5		80	15	17.4

†Does not include the second locus for EMFxaACO1B.
‡Includes EMFxaACO1B which has significant BLAST scores on two different LG in each of the three genera, and in apple occurs on two different sets of homeologous chromosomes, yet occurs in *Rubus* once only.

3.4.3 Comparative genome mapping between *Rubus* and *Malus*

The comparison of the 131 markers that comprise the linkage maps of 96395S1 and 'Latham' with the draft genome assembly of *Malus* using BLAST analysis of marker sequences identified 80 markers in common between the two genera, with E-values ranging from $5.00E^{-05}$ with 83% identity, to $0.00E^{+00}$ with 99% identity, with only five markers having E-values greater than E^{-15} (Table 3-3; Supplemental Fig. S3.2; Supplemental Table S3-2). Based on the positions of 70 markers (including the three GDSNP-derived markers) common to both genera, each of the seven *Rubus* LG could be aligned with one, two or three segments of the 11 homeologous chromosomes in *Malus* (Velasco et al. 2010), ranging from one marker on RLG4, to 15 markers on RLG3, and covering 77% (430 cM of 561 cM) of the 'Latham' genetic linkage map (Fig. 3.4: Comparison of *Rubus idaeus* 'Latham' genetic linkage map with *Fragaria*, *Malus*, and *Prunus*. ; Supplemental Fig. S3.2). Of the 80 markers in common, 41 (51%), including EMFxaACO1B, have significant BLAST scores on homeologous chromosomes; six markers (7.5%) (RiF3'H, RubARSFL_33_Exp3, RubARSFL_35_Cel2, EMFxaACO1B, Ri_5O21 & RiRosCOS1281) have significant BLAST scores on non-homeologous chromosomes; and 39 markers (47%) occur once only. Of the 80 markers in common, 19 (24%) did not map to the homologous LG (Fig. 3.3). These 19 markers are from *Rubus* EST (ERubLR_SQ05-3_E02), various genes (*RiSNF4*, *RiHCT*/*HQT*, *RiSNF2*, *RiInvAlkA*, *RiMYB*, *RiHXK*, *Ri4Co1*, *RiSAMD*, *RiLTP*, *RiG6PD*, *RiFRUCT4*, *RiPAL2* & *RibHLH*), *Fragaria* gene (*EMFxaCAD1B*), and RosCOS (RiRosCOS3524, RiRosCOS2381, RiRosCOS1412, RiRosCOS1360).

3.4.4 Comparative genome mapping between *Rubus* and *Prunus*

The comparison of the 131 markers that comprise the linkage maps of 96395S1 and 'Latham' with the draft genome assembly of *Prunus* using BLAST analysis of marker sequences identified 86 markers in common between the two genera, with E-values ranging from $2.00E^{-06}$ with 85% identity, to $0.00E^{+00}$ with 100% identity, with only five markers having E-values of greater than E^{-15} (Table 3-3; Supplemental Fig. S3.3; Supplemental Table S3-2). Based on the positions of 86 markers common to both genera, each of the seven *Rubus* LG could be aligned with one, two or three segments of the *Prunus* chromosomes, ranging from three markers on RLG4, to 14 markers on RLG3, and covering 90% (505 cM of 561 cM) of the 'Latham' genetic linkage map (Fig. 3.4: Comparison of *Rubus idaeus* 'Latham' genetic linkage map with *Fragaria*, *Malus*, and *Prunus*. ; Supplemental Fig. S3.3). Of the 86 markers in common, five (6%) (RiMYB10, RiPPC1, EMFxaACO1B, RiSDH2 & RiPAL2) had significant BLAST scores on two different LG; and 15 (17%) did not map to the homologous LG (Fig. 3.3). These 15 markers are from *Rubus* SSR (Ro_1K15), *Fragaria* gene (*EMFxaCAD1B*), various genes (*RiSNF4*, *RiInvAlkE*, *RiHCT*/HQT, *SNF2*, *RiInvAlkA*, *RiMYB*, *Ri4Co1*, *RiSAMD*, *RiG6PD*, *RiSDH2* & *RiFRUCT4*), and RosCOS (RiRosCOS2381 & RiRosCOS1412).

3.4.5 Comparative genome mapping between *Rubus* and the Rosaceae ancestral genome

The ancestral genome structure of nine chromosomes (A1-A9) and their proposed distribution among and contribution to the genomes of *Rubus* (RLG1-RLG7), *Fragaria vesca* (FLGI-FLGVII), *Malus* (MLG1-MLG17), and *Prunus* (G1-G8) are illustrated in Fig. 3.5. The nine ancestral chromosomes are distributed in 14 segments across the *Rubus* genome, with two *Rubus* linkage groups (including both segments of RLG7) composed of one ancestral chromosome and the remaining five *Rubus* linkage groups composed of two or more ancestral chromosomes.

RLG1 and FLGI both have A7 as a contributor (Fig. 3.5). A7 also contributes to G7 (Vilanova et al. 2008), and to segments of apple homeologous chromosomes MLG2 & MLG15 (Illa et al. 2011). To obtain a consensus for ancestral contribution, a change in orientation of RLG1 is needed relative to the representation from Graham et al.'s (2004) original LG6 (Fig. 3.3), and also to G7. Five markers in common between RLG1 and FLGI support the contribution of A5; however, Illa et al. (2011) shows A8 to be a

contributor to FLGI. The set of markers used in this study, and the results of Bushakra et al. (2012b), do not demonstrate a contribution of A5 to FLGI. While there are two markers found with homology to G6, which is derived from A8, the comparison of the four genera supports A5 as the second major contributor to RLG1 and FLGI.

RLG2 and FLGII both show strong support for derivation from A1 and A2 (Fig. 3.5). This is in agreement with the results of Bushakra et al. (2012b) and contrary to the results of Illa et al. (2011) which suggest that FLGII is derived entirely from A1. A1 also contributes to apple homeologous chromosomes MLG5 & MLG10 (Illa et al. 2011) and *Prunus* G1 and G8 (Vilanova et al. 2008). The contribution of A2 to apple homeologous chromosomes MLG8 & MLG15 (Illa et al. 2011) and *Prunus* G1 (Vilanova et al. 2008) is also supported by these data.

RLG3 and FLGIII show strong support for derivation from A8 and A9 (Fig. 3.5). A change in orientation of RLG3 has been made relative to the representation from Graham et al.'s (2004) original LG2 (Fig. 3.3). This change in orientation is also supported by the results of Bushakra et al. (2012b) for apple homeologous chromosomes MLG3 & MLG11. The contributions of A8 and A9 to G4 and G6 (Vilanova et al. 2008) is also supported with these data, and a change in LG orientation brings all genera into alignment.

RLG4 and FLGIV have only three markers in common; however, the contribution of A2 is suggested for all genera (Fig. 3.5).

RLG5, FLGV, G5 and apple homeologous chromosomes MLG6 & MLG14 show a high degree of conservation (Fig. 3.5). All are derived primarily from A3. There appears to be a small segment of A2 contributing to each LG and this is supported by the results of Bushakra et al. (2012b). To bring RLG5 and FLGV into alignment, a change in orientation of RLG5 has been made relative to the representation from Graham et al.'s (2004) original LG5 (Fig. 3.3). Marker RubARSFL_33_RiExp3 appears to be derived from both A3 and A5, as this gene codes for an enzyme involved in expansin biosynthesis, and could be a member of a gene family.

RLG6 and FLGVI are interesting in that they appear to contain elements of three ancestral chromosomes A5, A6, and A9 (Fig. 3.5). Alignment of *Rubus* and *Fragaria* has required a change in orientation of RLG6 relative to the representation from

Graham et al.'s (2004) original LG3 (Fig. 3.3). The contribution of A6 to FLGVI, the contribution of A5 to FLGVI and apple homeologous chromosomes MLG9 & MLG17, as well as the contribution of A9 to MLG4 & MLG12 and homeologous chromosomes MLG5 & MLG10 is supported by the results of Illa et al. (2011). The contribution of A5 to G3 is supported by the results of Vilanova et al. (2008). A change in orientation of FLGVI is suggested by Bushakra et al. (2012b).

RLG7, FLGVII, G2 and apple homeologous chromosomes MLG1 & MLG7 show a high degree of conservation (Fig. 3.5). All are derived primarily from A4, and this single origin is supported by the results of Illa et al. (2011), Vilanova et al. (2008), and Bushakra et al. (2012b).

3.5 Discussion

3.5.1 Inter-specific *Rubus* genetic map construction

We present here the first genetic linkage map of *Rubus* developed with *R. occidentalis* as a parent of the mapping progeny. Our map is composed entirely of markers derived from non-anonymous DNA sequences, the majority of which are EST-based or gene-derived. As this map was developed from an inter-specific progeny with 'Latham' as a parent, it can be aligned with existing *R. idaeus* 'Latham'-based maps (Graham et al. 2009; Graham et al. 2004; Graham et al. 2006; Kassim et al. 2009; McCallum et al. 2010; Woodhead et al. 2008; Sargent et al. 2007b) to obtain a wider view of genetic diversity in *Rubus*. Overall, our map aligns with the existing maps. This new map, derived from parents with different fruit color, and enriched for orthologous markers, provides a resource for future studies on variation in fruit color and polyphenolic compounds.

Our HRM-based marker design provides a technique for high-throughput genotyping not previously utilized in *Rubus* genetic mapping studies. The direct sequencing of 45 PCR products demonstrated the ability of HRM to distinguish accurately between homoduplex and heteroduplex PCR products, providing confidence in determining the parental source of alleles. Our use of existing EST together with the pre-released genome sequence resulted in development of new orthologous markers that will be useful as anchor loci for the on-going assembly of the *Rubus* genome.

The ‘Latham’ genetic linkage map only was used for the comparative analyses because of the much higher number of heterozygous markers between ‘Latham’ (114 mapped markers) compared to 96395S1 (29 mapped markers). The difference in number of heterozygous loci in the S1 parental map compared to the ‘Latham’ map is likely to be the result of the degree of inbreeding in the S1 parent. Using methods described by Falconer and Mackay (1996), we calculated the inbreeding coefficient for S1 to be 0.35, indicating that 35% of alleles at a locus are identical by descent, resulting in a sparsely populated S1 genetic map. The degree of inbreeding in S1 is higher than the average of 21% reported for raspberry cultivars (Dale et al. 1993), and the 0-25% reported in both blackberry (*Rubus* subgenus *Rubus*) (Stafne and Clark 2004) and almond (*Prunus dulcis*) cultivars (Lansari et al. 1994). The parent 96395S1 has a pedigree dominated by black raspberry (*R. occidentalis*); however it was derived from crosses of black raspberry with red raspberry ‘Burnethholm’ with the intention to introgress spinelessness from this cultivar. The spineless type was achieved via selections V23 and V32 (Fig. 3.2a). From available pedigree information we calculated selection 96395S1 to have approximately 9.4% genetic contribution from red raspberry with the remaining 90.6% from black raspberry.

From known pedigree information (Fig. 3.2b) we used the methods described by (Falconer and Mackay 1996) to calculate the inbreeding coefficient of *R. idaeus* ‘Latham’ as 0, indicating that the ‘Latham’ parent is not inbred and is likely to be significantly more heterozygous than S1. The parent ‘Latham’ was included in this cross because it provides a link to the Graham et al. (2004) genetic map. The degree of genetic homozygosity of the S1 parent was not considered when it was selected for crossing with ‘Latham’; rather, the emphasis was placed on developing a progeny with a range of fruit colors for color- and polyphenolic-based quantitative trait locus (QTL) analyses.

3.5.2 Development of orthologous markers for use in *Rubus*

During the screening of published markers (Set A), we found 42 markers to be monomorphic or with both parents homozygous (aaxbb) in our population. Of these, 29 (71%) were previously reported as heterozygous in ‘Latham’ (Graham et al. 2004). As a trial screen of a subset of the published SSR markers using traditional SSR screening methods (ABI 3100 polyacrylamide gel, data not shown) indicated similar degrees of monomorphism to that exhibited following HRM analysis, we believe that this

discrepancy could be caused by the use of different 'Latham' sources in the respective studies, rather than by the differences in marker analysis techniques employed. Unfortunately, synonyms for genotypes and errors in labeling and misidentifications in germplasm collections are not uncommon and spontaneous mutations and sports can arise in clonally propagated crops (Thompson 1995a, b; Bassil et al. 2009; Gygax et al. 2004; Gardiner et al. in prep). To resolve this issue, samples from different 'Latham' sources need to be genotyped. However, the variation in heterozygosity for some markers found in this study does not appear to influence greatly the ability to compare the linkage maps, as the overall alignment of anchor loci agree among the mapping studies.

The recent studies on marker transferability by Lewers et al. (2005), Zorrilla-Fontanesi et al. (2011b), and our own results, demonstrate the difficulty of identifying polymorphic loci in *Rubus*, and also illustrate that the amplification of a PCR product does not necessarily make a marker useful for map construction. In general, amplification of polymorphic loci in *Rubus* using PCR primer pairs designed from *Fragaria* sequence was more successful if the sequence was from EST (8%-13%, and a total of 18 markers) than if the sequence was from genomic SSR (5%-6%, and a total of four markers) (Table 3-4). Between the two studies, only 22 of a total of 222 markers (10%) amplified a polymorphic product in their respective *Rubus* species. In our study, 939 of 1218 primer pairs (77%) amplified a product. Of these, 136 (14.5%) mapped, 505 (54%) were monomorphic, and 177 (18%) had complex HRM curves. The percentage of monomorphic markers in our population is similar to what was reported by Lewers et al. (2005) (40-83%) and Zorrilla-Fontanesi (2011b) (34-60%) and suggests that the cultivars and selections most often used for mapping studies of *Rubus* may have high gene homozygosity. Our marker design technique focused on the discovery of *de novo* SNP using HRM to detect sequence polymorphisms, rather than amplification of SSR repeats. As the genome sequences of more *Rubus* cultivars become available and the positions of putative SNP loci can be identified, the HRM technique will be an efficient genotyping method.

Table 3-4: Summary of transferability of markers derived from *Fragaria* DNA sequence to raspberry and blackberry.

Source of sequence for primer design	Amplification results	Raspberry	Blackberry	Markers with potential for mapping
GenBank *Fragaria* EST-SSR primer pairs tested (Lewers et al. 2005)		28	23	
	amplified	5 (18%)	6 (26%)	
	polymorphic	3 (11%)	3 (13%)	6
GenBank *Fragaria* genomic SSR primer pairs tested (Lewers et al. 2005)		20	19	
	amplified	4 (20%)	6 (26%)	
	polymorphic	1 (5%)	1 (5%)	2
Fragaria EST-SSR primer pairs tested (Zorrilla-Fontanesi et al. 2011)		143	NA	
	amplified	29 (20%)		
	polymorphic	12 (8%)		12
Fragaria genomic SSR primer pairs tested (Zorrilla-Fontanesi et al. 2011)		31	NA	
	amplified	5 (16%)		
	polymorphic	2 (6%)		2

NA: not applicable

3.5.3 Genetic map comparison between *Rubus* and other Rosaceae genera

The degree of genomic synteny and conservation demonstrated in this study is further evidence of the relatively close phylogenetic relationship between *Rubus* and *Fragaria* (Eriksson et al. 2003; Potter et al. 2007). These relationships are especially clear when the *Rubus* and *Fragaria* LG are oriented with each other, as they are presented in Figs. 3.4 and 3.5. This relationship leads us to propose the renumbering and reorientation of the *Rubus* LG to allow easier comparisons with *Fragaria* and to be more consistent with previous work in the family, especially necessary now as the genome information for both genera increases rapidly.

Our study identified several previously undetected differences in ancestral contribution to the genomes of *Rubus* and *Fragaria* that will require further analysis (Figs. 3.4 & 3.5). The data suggest that RLG 2, 3, 4, 5 and 7 and the corresponding *Fragaria* LG have retained a high degree of synteny. It appears that RLG1 varies from *Fragaria* in its ancestral makeup and that RLG6 and FLGVI both differ in their ancestral makeup relative to *Malus* and *Prunus*.

Comparison of the four genera is facilitated by the high DNA sequence transferability, as the sequences of 73 (56%) of the mapped markers were identified in all four genera,

and provide new insights into family-wide orthology. For example, the high degree of colinearity identified between RLG7 compared with FLG7, G2 and MLG2 suggests that these LG have remained relatively unchanged from the ancestral state (A4). Ancestral conservation is also suggested by the level of colinearity seen between RLG5 compared with both G5 and MLG6 with respect to A3. Additional insights come from markers that mapped to non-homologous LG, or with significant BLAST scores on two different LG in *F. vesca* and *Prunus*, and those in *Malus* that are located on non-homeologous chromosomes. These markers could represent paralogous loci or be simply the best BLAST hit based on the state of assembly of the *Rubus* genome at the time of the analysis (Jan.-Apr. 2011). Genes used to design the *Rubus*-specific markers are known to be from several metabolic pathways in other plant species; therefore, some are likely to be members of gene families and not necessarily single copies. Additionally, as BLAST analysis identifies regions of similarity, but not the functionality of the region, some of these markers may be duplicates but non-functional. The variation in marker order in the genetic maps of 96395S1 and 'Latham' compared with the BLAST-derived map of *Fragaria* could indicate the amplification of other members of the gene families or the presence of paralogous genes. The two occurrences of RiRosCOS1281 on non-homeologous chromosomes (MLG7 and MLG10) and the position of RiRosCOS530 on FLGV rather than FLGII could be examples of errors in genotyping or perhaps in the draft *Malus* and *Fragaria* genome assemblies, respectively. These ambiguities will become clearer as more markers are mapped and the genome assemblies are further developed. Two of the three markers designed from *Malus* GDSNP sequences, Ri_GDSNP02690 and Ri_GDSNP00768 map to homologous LG in *Rubus*; however, these could easily be coincidental similarities rather than true homologies.

The orthologous markers developed in this study can be used to explore the genomes of other less-well characterized Rosaceae species, such as ornamentals and members of the subfamily Dryadoideae. Genomic analysis of members of Dryadoideae will lead to a more complete picture of the family and its evolution.

3.6 Conclusions

The genetic maps of *Rubus* constructed in this study map provide a useful set of transferable and orthologous markers for comparative mapping, and a solid foundation for further studies such as quantitative trait analyses. Additionally, the maps illustrate the nearly one-to-one alignment of the linkage groups of *Rubus* and *Fragaria* and

support the current phylogeny that places *Rubus* and *Fragaria* more closely to each other than either is to *Malus* or *Prunus*. The results of this study suggest that RLG1 and RLG6 have undergone rearrangements relative to the homologous *Fragaria* LG. The analysis of ancestral chromosome contribution suggests that of the four genera, *Prunus* shows the fewest rearrangements from the ancestral state. The insight into family-wide genome conservation gained by comparative mapping among these economically important genera provides useful information for Rosaceae crop geneticists and breeders. For example, gene homology location can be extrapolated from mapped orthologous markers and can be used to identify genes of interest in less well studied Rosaceae crops prior to the availability of genome sequences. The chromosome comparisons developed here will facilitate the assembly and annotation of future Rosaceae genome sequence projects.

3.7 Acknowledgements

This work was part of a Ph.D. thesis and supported in part by the PFR Excellence Program (JMB), the "New Berries" FRST Programme (CO6XO807) (EJB, MJS), PFR funding (MJS), and USDA-ARS Project 1275-21220-185-00D (KSL). The authors would like to thank Jared C. Price of Brigham Young University, Computer Science Department for his assistance with the *Rubus* genome; the International Peach Genome Initiative for access to the *Prunus* genome; Dr Ross Crowhurst and Roy Storey of PFR for bioinformatic support; Tony Corbett of PFR for graphic design of all figures; the Editor and two anonymous reviewers for helpful suggestions on the original manuscript. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by any of the agencies referenced.

3.8 Authors' contributions

JMB and MJS designed all primer sequences; JMB, MJS & ANA performed all experiments; JMB constructed the genetic map, interpreted the results and wrote the manuscript; MJS calculated the inbreeding coefficients; KSL provided *Rubus* SSR DNA sequence; JAU provided access to the pre-release *Rubus* genome sequence; VVS, DC, EJB, and SEG provided advice on experimental design and manuscript editing.

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215 (3):403-410

Amsellem L, Dutech C, Billotte N (2001) Isolation and characterization of polymorphic microsatellite loci in *Rubus alceifolius* Poir. (Rosaceae), an invasive weed in La Réunion island. Molecular Ecology Notes 1 (1-2):33-35. doi:10.1046/j.1471-8278.2000.00013.x

Bassil N, Hummer K, Postman J, Fazio G, Baldo A, Armas I, Williams R (2009) Nomenclature and genetic relationships of apples and pears from Terceira Island. Genetic Resources and Crop Evolution 56 (3):339-352

Bushakra JM, Sargent DJ, Cabrera A, Crowhurst RN, Lopez Girona E, Velasco R, Symonds VV, van der Knaap E, Troggio M, Gardiner SE, Chagné D (2012b) Rosaceae conserved orthologous set (RosCOS) markers as a tool to assess genome synteny between *Malus* and *Fragaria*. Tree Genetics & Genomes:1-16. doi:10.1007/s11295-011-0450-y

Cabrera A, Kozik A, Howad W, Arús P, Iezzoni A, van der Knaap E (2009) Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. BMC Genomics 10 (1):562

Castillo NRF, Reed BM, Graham J, Fernandez-Fernandez F, Bassil NV (2010) Microsatellite markers for raspberry and blackberry. Journal of the American Society for Horticultural Science 135 (3):271-278

Celton J-M, Tustin DS, Chagné D, Gardiner SE (2009b) Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. Tree Genetics & Genomes 5:93-107

Chagné D (2008) Amazing LightCycler 480 applications: Application of the High-Resolution Melting technique for mapping genes in apple. Roche Applied Science

Dale A, Moore PP, McNicol RJ, Sjulin TM, Burmistrov LA (1993) Genetic diversity of red raspberry varieties throughout the world. Journal of the American Society for Horticultural Science 118 (1):119-129

Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldera F, Cosson P, Howad W, Arús P (2004b) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proceedings of the National Academy of Sciences of the United States of America 101 (26):9891-9896. doi:10.1073/pnas.0307937101

Eriksson T, Hibbs MS, Yoder AD, Delwiche CF, Donoghue MJ (2003) The phylogeny of Rosoideae (Rosaceae) based on sequences of the internal transcribed spacers (ITS) of nuclear ribosomal DNA and the *TrnL/F* region of chloroplast DNA. International Journal of Plant Science 164 (2):197-211. doi:1058-5893/2003/16402-0001

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th Ed. Pearson Education Limited, Longmans Green, Harlow, UK

Gardiner SE, Norelli JL, de Silva N, Fazio G, Peil A, Malnoy M, Horner M, Bowatte D, Carlisle C, Wiedow C, Wan Y, Bassett CL, Baldo AM, Celton J-M, Richter K, Aldwinckle HS, Bus VGM (in prep) Candidate gene markers associated with quantitative trait loci for fire blight resistance in *Malus* 'Robusta 5' accessions. BMC Plant Biology

Graham J, Hackett C, Smith K, Woodhead M, Hein I, McCallum S (2009) Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit.

Theoretical and Applied Genetics 118 (6):1143-1155. doi:10.1007/s00122-009-0969-6

Graham J, Smith K, MacKenzie K, Jorgenson L, Hackett C, Powell W (2004) The construction of a genetic linkage map of red raspberry (*Rubus idaeus* subsp. *idaeus*) based on AFLPs, genomic-SSR and EST-SSR markers. Theoretical and Applied Genetics 109 (4):740-749

Graham J, Smith K, Tierney I, MacKenzie K, Hackett C (2006) Mapping gene H controlling cane pubescence in raspberry and its association with resistance to cane botrytis and spur blight, rust and cane spot. Theoretical and Applied Genetics 112 (5):818-831

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

Gygax M, Gianfranceschi L, Liebhard R, Kellerhals M, Gessler C, Patocchi A (2004) Molecular markers linked to the apple scab resistance gene *Vbj* derived from *Malus baccata jackii*. Theoretical and Applied Genetics 109 (8):1702-1709. doi:10.1007/s00122-004-1803-9

Illa E, Sargent D, Lopez Girona E, Bushakra J, Cestaro A, Crowhurst R, Pindo M, Cabrera A, van der Knaap E, Iezzoni A, Gardiner S, Velasco R, Arús P, Chagne D, Troggio M (2011) Comparative analysis of rosaceous genomes and the reconstruction of a putative ancestral genome for the family. BMC Evolutionary Biology 11 (1):9

Joobeur T, Viruel MA, de Vicente MC, Jáuregui B, Ballester J, Dettori MT, Verde I, Truco MJ, Messeguer R, Batlle I, Quarta R, Dirlewanger E, Arús P (1998) Construction of a saturated linkage map for *Prunus* using an almond × peach F_2 progeny. Theoretical and Applied Genetics 97 (7):1034-1041

Kassim A, Poette J, Paterson A, Zait D, McCallum S, Woodhead M, Smith K, Hackett C, Graham J (2009) Environmental and seasonal influences on red raspberry anthocyanin antioxidant contents and identification of quantitative traits loci (QTL). Molecular Nutrition & Food Research 53 (5):625-634

Kobayashi N, Horikoshi T, Katsuyama H, Handa T, Takayanagi K (1998) A simple and efficient DNA extraction method for plants, especially woody plants. Plant Tissue Culture and Biotechnology 4 (2):76-80

Lansari A, Kester DE, Iezzoni AF (1994) Inbreeding, coancestry, and founding clones of almonds of California, Mediterranean shores, and Russia. Journal of the American Society for Horticultural Science 119 (6):1279-1285

Lewers K, Saski C, Cuthbertson B, Henry D, Staton M, Main D, Dhanaraj A, Rowland L, Tomkins J (2008) A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. BMC Plant Biology 8 (1):69-76

Lewers KS, Styan SMN, Hokanson SC, Bassil NV (2005) Strawberry GenBank-derived and genomic simple sequence repeat (SSR) markers and their utility with strawberry, blackberry, and red and black raspberry. Journal of the American Society for Horticultural Science 130 (1):102-115

Lopes MS, Maciel GB, Mendonça D, Gil FS, Da Câmara Machado A (2006) Isolation and characterization of simple sequence repeat loci in *Rubus hochstetterorum* and their use in other species from the Rosaceae family. Molecular Ecology Notes 6 (3):750-752. doi:10.1111/j.1471-8286.2006.01329.x

McCallum S, Woodhead M, Hackett C, Kassim A, Paterson A, Graham J (2010) Genetic and environmental effects influencing fruit colour and QTL analysis in

raspberry. Theoretical and Applied Genetics 121 (4):611-627. doi:10.1007/s00122-010-1334-5

Montgomery J, Wittwer CT, Palais R, Zhou L (2007) Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. Nature Protocols 2 (1):59-66. doi:10.1038/nprot.2007.10

Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, Campbell CS (2007) Phylogeny and classification of Rosaceae. Plant Systematics and Evolution 266 (1):5-43

Reed GH, Wittwer CT (2004) Sensitivity and specificity of single-nucleotide polymorphism scanning by High-Resolution Melting analysis. Clinical Chemistry 50 (10):1748-1754. doi:10.1373/clinchem.2003.029751

Sargent DJ, Clarke J, Simpson DW, Tobutt KR, Arús P, Monfort A, Vilanova S, Denoyes-Rothan B, Rousseau M, Folta KM, Bassil NV, Battey NH (2006) An enhanced microsatellite map of diploid *Fragaria*. Theoretical and Applied Genetics 112 (7):1349-1359

Sargent DJ, Davis TM, Simpson DW (2009c) Strawberry (*Fragaria* spp.) Structural Genomics. In: Folta KM, Gardiner SE (eds) Genetics and Genomics of Rosaceae, vol 6. Plant Genetics and Genomics: Crops and Models. Springer New York, pp 437-456. doi:10.1007/978-0-387-77491-6_21

Sargent DJ, Fernández-Fernández F, Rys A, Knight VH, Simpson DW, Tobutt KR (2007b) Mapping of A_1 conferring resistance to the aphid *Amphorophora idaei* and *dw* (dwarfing habit) in red raspberry (*Rubus idaeus* L.) using AFLP and microsatellite markers. BMC Plant Biology 7 (15)

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton J-M, Rees DJG, Williams KP, Holt SH, Rojas JJR, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman T-L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Girona EL, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arús P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). Nature Genetics 43 (2):109-116. doi:http://www.nature.com/ng/journal/v43/n2/abs/ng.740.html#supplementary-information

Sosinski B, Verde I, Rokhsar DS (2010) International Peach Genome Initiative Peach Genome v1.0. http://www.rosaceae.org/peach/genome.

Stafne ET, Clark JR (2004) Genetic relatedness among eastern North American blackberry cultivars based on pedigree analysis. Euphytica 139 (2):95-104. doi:10.1007/s10681-004-2436-4

Thompson MM (1995a) Chromosome numbers of *Rubus* cultivars at the National Clonal Germplasm Repository. HortScience 30 (7):1453-1456

Thompson MM (1995b) Chromosome numbers of *Rubus* species at the National Clonal Germplasm Repository. HortScience 30 (7):1447-1452

van Ooijen JW, Voorrips RE (2001) JoinMap(R) 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nature Genetics 42 (10):833-839. doi:http://www.nature.com/ng/journal/v42/n10/abs/ng.654.html#supplementary-information

Vilanova S, Sargent D, Arús P, Monfort A (2008) Synteny conservation between two distantly-related Rosaceae genomes: *Prunus* (the stone fruits) and *Fragaria* (the strawberry). BMC Plant Biology 8 (1):67

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93 (1):77-78

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-Resolution Genotyping by Amplicon Melting Analysis using LCGreen. Clinical Chemistry 49 (6):853-860. doi:10.1373/49.6.853

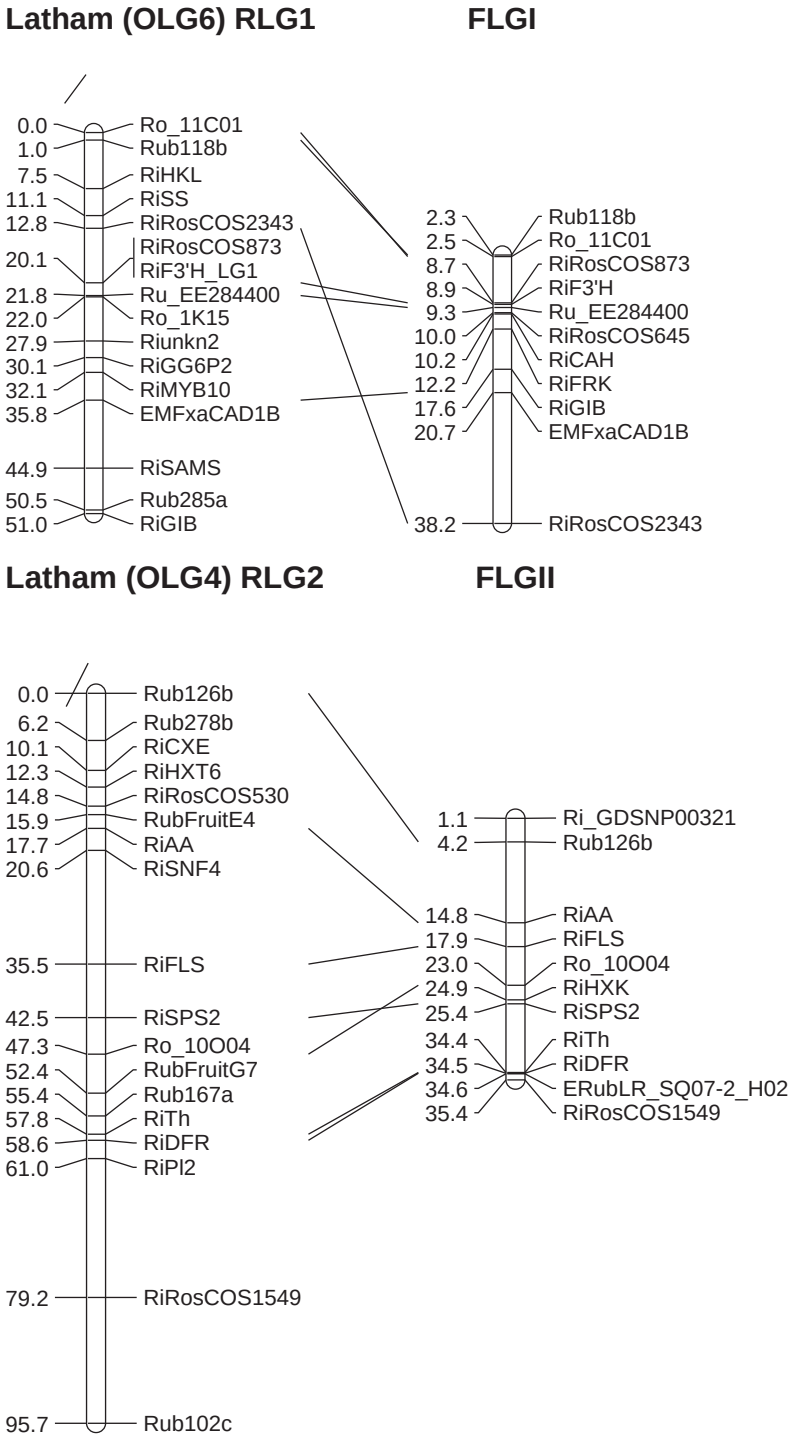
Woodhead M, McCallum S, Smith K, Cardle L, Mazzitelli L, Graham J (2008) Identification, characterisation and mapping of simple sequence repeat (SSR) markers from raspberry root and bud ESTs. Molecular Breeding 22 (4):555-563

Woodhead M, Weir A, Smith K, McCallum S, MacKenzie K, Graham J (2010) Functional markers for red raspberry. Journal of the American Society for Horticultural Science 135 (5):418-427

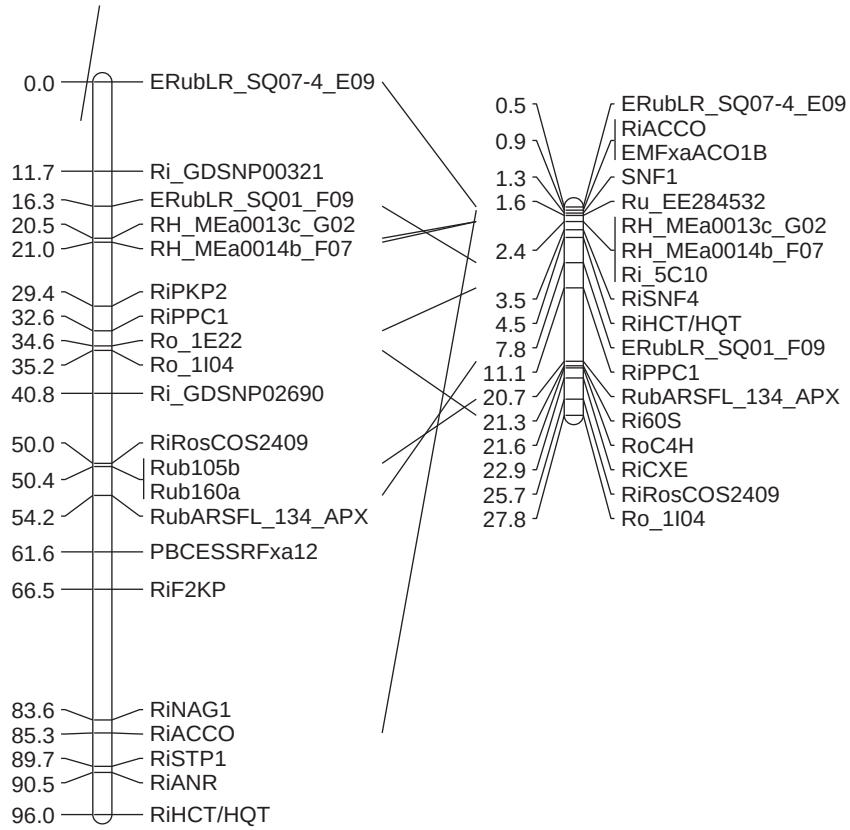
Zorrilla-Fontanesi Y, Cabeza A, Torres A, Botella M, Valpuesta V, Monfort A, Sánchez-Sevilla J, Amaya I (2011b) Development and bin mapping of strawberry genic-SSRs in diploid *Fragaria* and their transferability across the Rosoideae subfamily. Molecular Breeding 27 (2):137-156. doi:10.1007/s11032-010-9417-1

3.9 Supplemental files

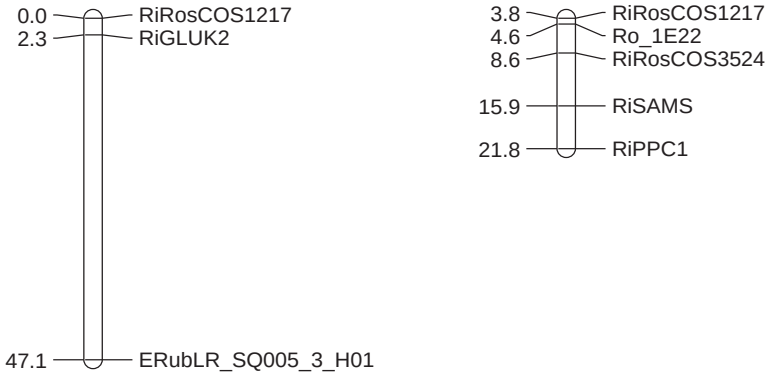
Supplemental Fig. S3.1: Comparison between *Rubus idaeus* 'Latham' genetic map and *Fragaria* physical map constructed using BLAST analysis. 'Latham' map distance is measured in centimorgans (cM); *Fragaria* linkage group (FLG) map distance is measured in megabase pairs (Mbp). 'Latham' linkage groups arranged in proposed order (RLG) with original numbering (OLG) in parentheses. Connecting lines indicate markers in common.



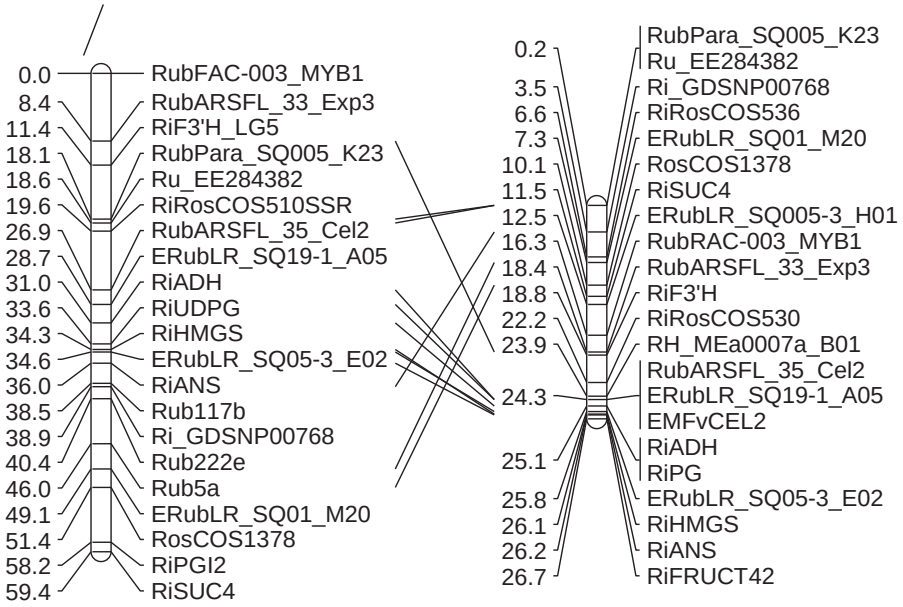
Latham (OLG2) RLG3 **FLGIII**



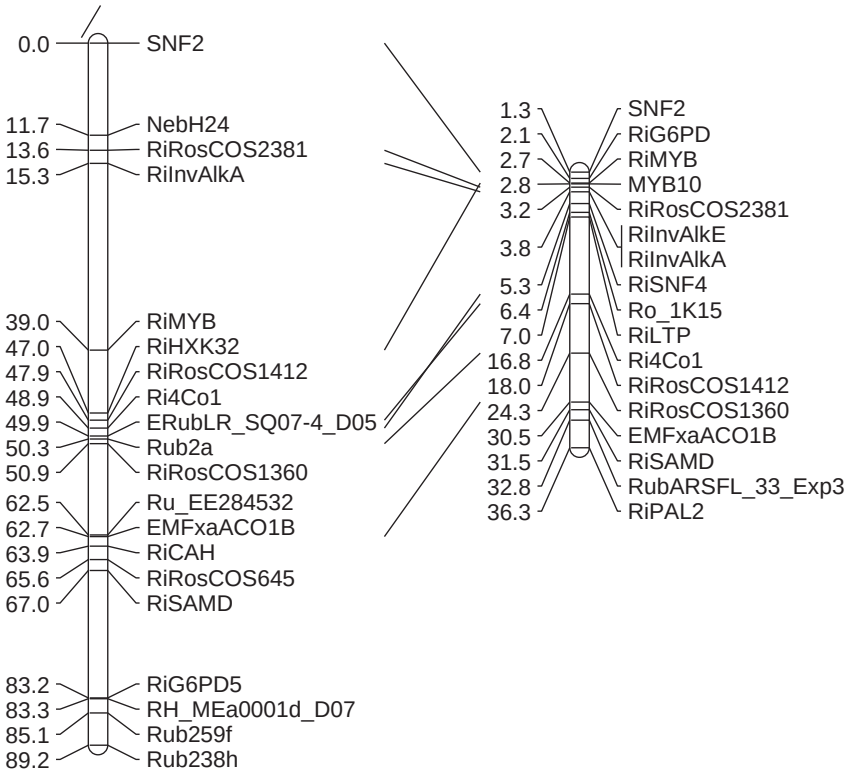
Latham (OLG7) RLG4 **FLGIV**

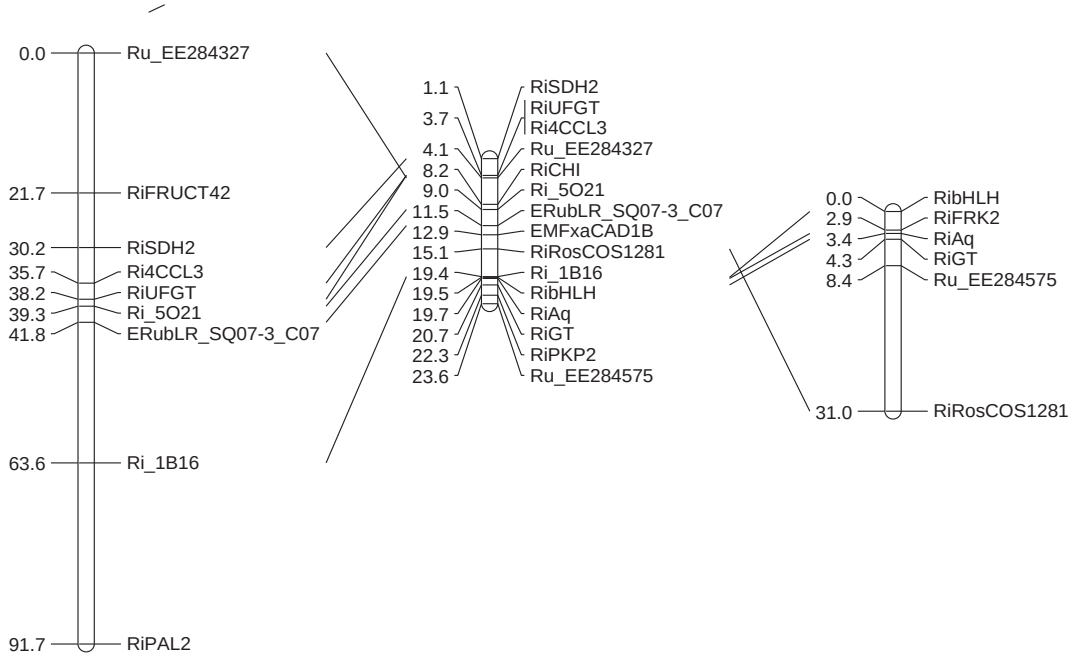


Latham (OLG5) RLG5 **FLGV**

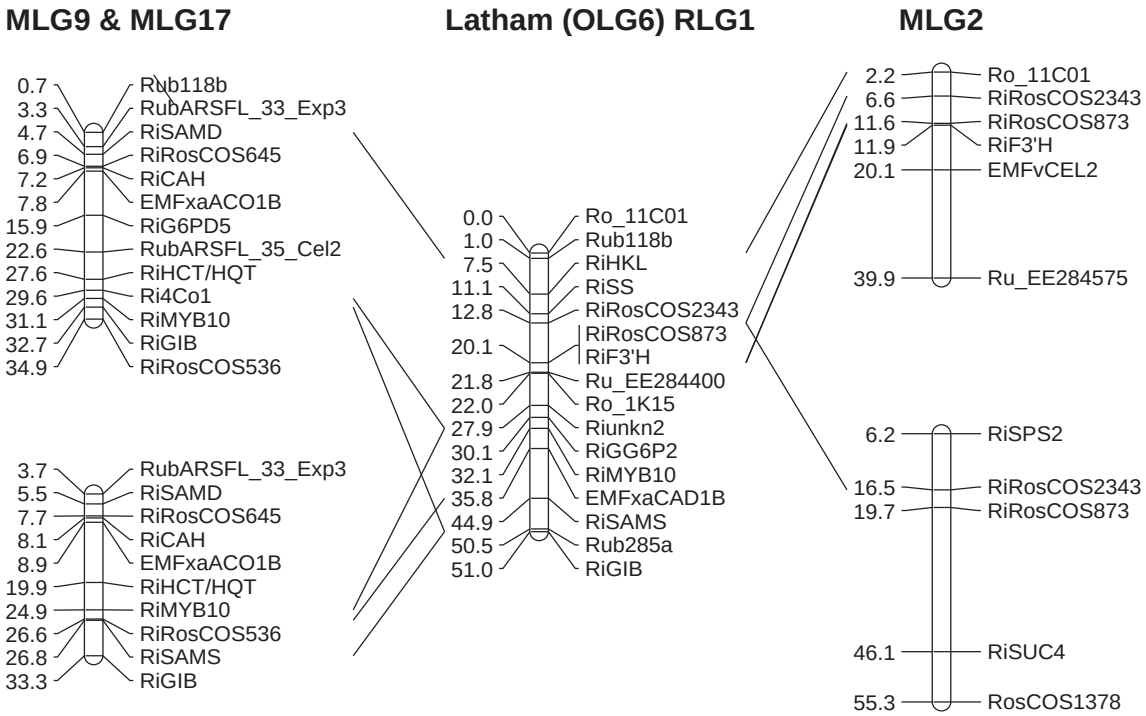


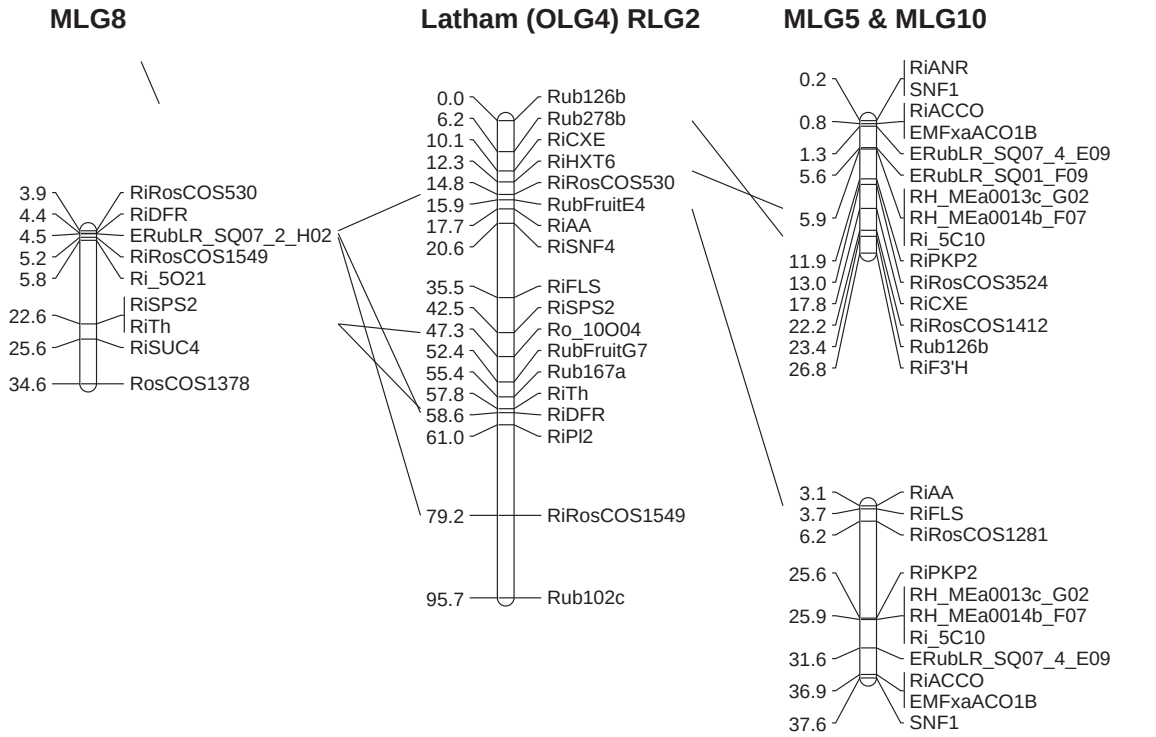
Latham (OLG3) RLG6 **FLGVI**

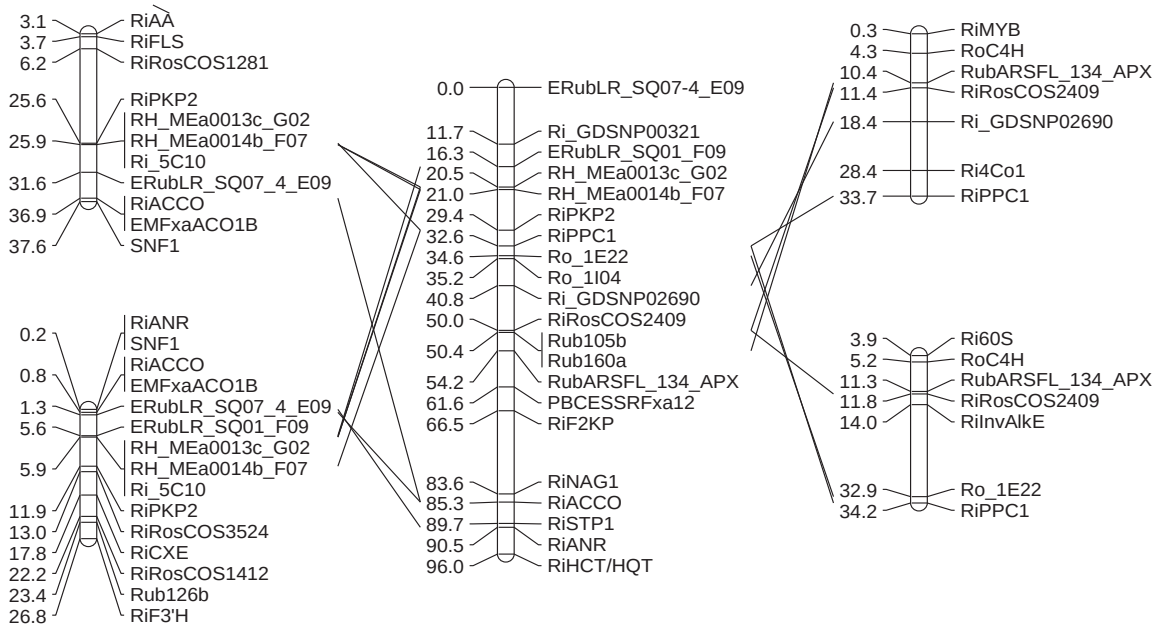



Latham (OLG1) RLG7a
FLGVII
Latham (OLG1) RLG7b
0.0 Ru_EE284327
21.7 RiFRUCT42
30.2 RiSDH2
35.7 Ri4CCL3
38.2 RiUFGT
39.3 Ri_5O21
41.8 ERubLR_SQ07-3_C07
63.6 Ri_1B16
91.7 RiPAL2
1.1 RiSDH2
3.7 RiUFGT
Ri4CCL3
4.1 Ru_EE284327
8.2 RiCHI
9.0 Ri_5O21
11.5 ERubLR_SQ07-3_C07
12.9 EMFxaCAD1B
15.1 RiRosCOS1281
19.4 Ri_1B16
19.5 RibHLH
19.7 RiAq
20.7 RiGT
22.3 RiPKP2
23.6 Ru_EE284575
0.0 RibHLH
2.9 RiFRK2
3.4 RiAq
4.3 RiGT
8.4 Ru_EE284575
31.0 RiRosCOS1281

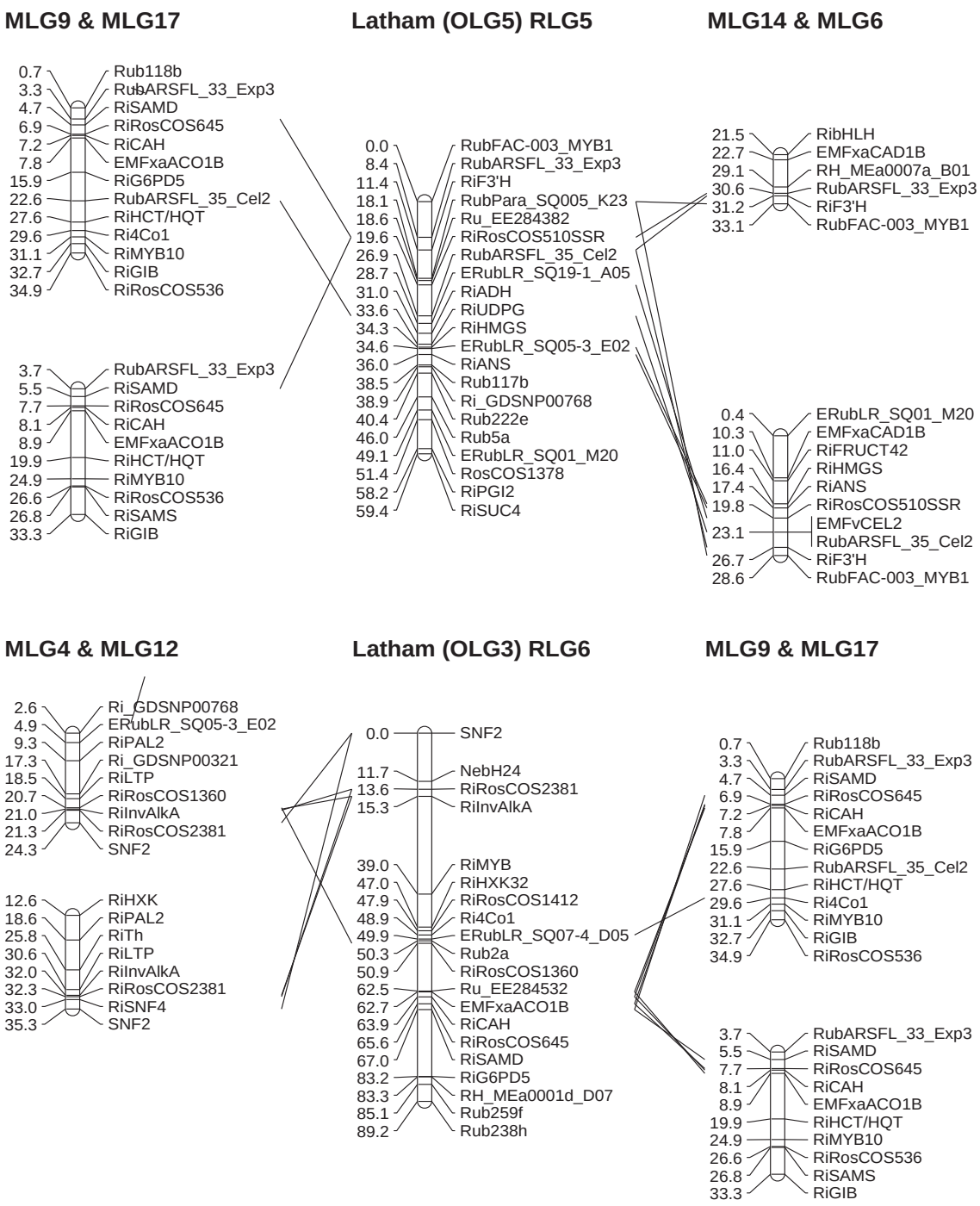
Supplemental Fig. S3.2: Comparison between *Rubus idaeus* 'Latham' genetic map and *Malus* physical map constructed using BLAST analysis. 'Latham' map distance is measured in centimorgans (cM); *Malus* linkage group (MLG) map distance is measured in megabase pairs (Mbp). 'Latham' linkage groups arranged in proposed order (RLG) with original numbering (OLG) in parentheses. Connecting lines indicate markers in common.



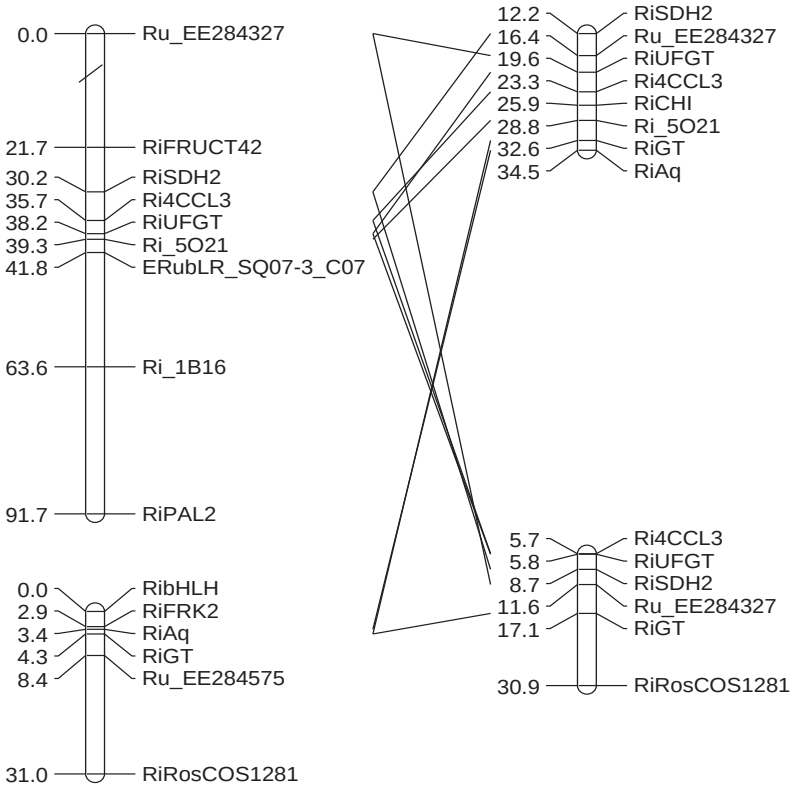




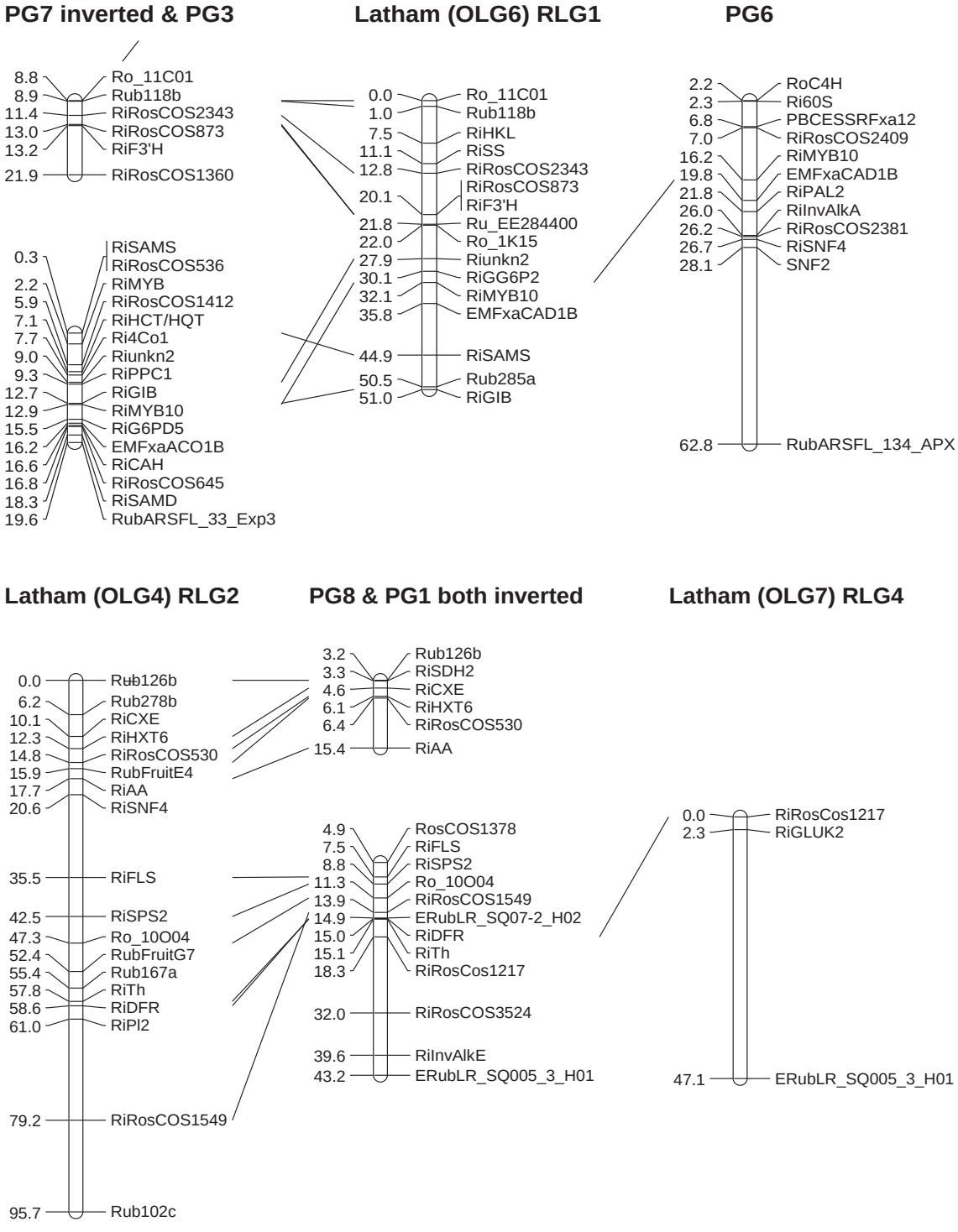
Note: RLG4 comparison not illustrated; only a single marker is in common with *Malus*

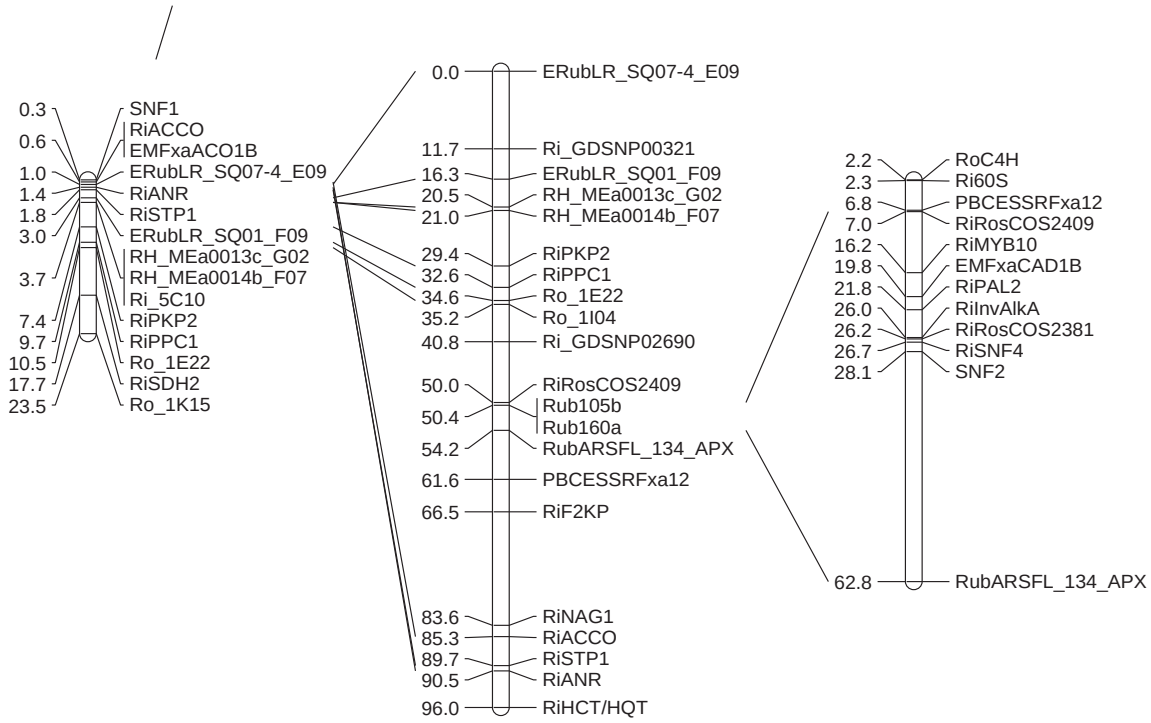

MLG9 & MLG17
Latham (OLG5) RLG5
MLG14 & MLG6
0.7 Rub118b
3.3 RubARSFL_33_Exp3
4.7 RiSAMD
6.9 RiRosCOS645
7.2 RiCAH
7.8 EMFxaACO1B
15.9 RiG6PD5
22.6 RubARSFL_35_Cel2
27.6 RiHCT/HQT
29.6 Ri4Co1
31.1 RiMYB10
32.7 RiGIB
34.9 RiRosCOS536
3.7 RubARSFL_33_Exp3
5.5 RiSAMD
7.7 RiRosCOS645
8.1 RiCAH
8.9 EMFxaACO1B
19.9 RiHCT/HQT
24.9 RiMYB10
26.6 RiRosCOS536
26.8 RiSAMS
33.3 RiGIB
0.0 RubFAC-003_MYB1
8.4 RubARSFL_33_Exp3
11.4 RiF3'H
18.1 RubPara_SQ005_K23
18.6 Ru_EE284382
19.6 RiRosCOS510SSR
26.9 RubARSFL_35_Cel2
28.7 ERubLR_SQ19-1_A05
31.0 RiADH
33.6 RiUDPG
34.3 RiHMGS
34.6 ERubLR_SQ05-3_E02
36.0 RiANS
38.5 Rub117b
38.9 Ri_GDSNP00768
40.4 Rub222e
46.0 Rub5a
49.1 ERubLR_SQ01_M20
51.4 RosCOS1378
58.2 RiPGI2
59.4 RiSUC4
21.5 RibHLH
22.7 EMFxaCAD1B
29.1 RH_MEa0007a_B01
30.6 RubARSFL_33_Exp3
31.2 RiF3'H
33.1 RubFAC-003_MYB1
0.4 ERubLR_SQ01_M20
10.3 EMFxaCAD1B
11.0 RiFRUCT42
16.4 RiHMGS
17.4 RiANS
19.8 RiRosCOS510SSR
23.1 EMFvCEL2
RubARSFL_35_Cel2
26.7 RiF3'H
28.6 RubFAC-003_MYB1
MLG4 & MLG12
Latham (OLG3) RLG6
MLG9 & MLG17
2.6 Ri_GDSNP00768
4.9 ERubLR_SQ05-3_E02
9.3 RiPAL2
17.3 Ri_GDSNP00321
18.5 RiLTP
20.7 RiRosCOS1360
21.0 RiInvAlkA
21.3 RiRosCOS2381
24.3 SNF2
12.6 RiHXK
18.6 RiPAL2
25.8 RiTh
30.6 RiLTP
32.0 RiInvAlkA
32.3 RiRosCOS2381
33.0 RiSNF4
35.3 SNF2
0.0 SNF2
11.7 NebH24
13.6 RiRosCOS2381
15.3 RiInvAlkA
39.0 RiMYB
47.0 RiHXK32
47.9 RiRosCOS1412
48.9 Ri4Co1
49.9 ERubLR_SQ07-4_D05
50.3 Rub2a
50.9 RiRosCOS1360
62.5 Ru_EE284532
62.7 EMFxaACO1B
63.9 RiCAH
65.6 RiRosCOS645
67.0 RiSAMD
83.2 RiG6PD5
83.3 RH_MEa0001d_D07
85.1 Rub259f
89.2 Rub238h
0.7 Rub118b
3.3 RubARSFL_33_Exp3
4.7 RiSAMD
6.9 RiRosCOS645
7.2 RiCAH
7.8 EMFxaACO1B
15.9 RiG6PD5
22.6 RubARSFL_35_Cel2
27.6 RiHCT/HQT
29.6 Ri4Co1
31.1 RiMYB10
32.7 RiGIB
34.9 RiRosCOS536
3.7 RubARSFL_33_Exp3
5.5 RiSAMD
7.7 RiRosCOS645
8.1 RiCAH
8.9 EMFxaACO1B
19.9 RiHCT/HQT
24.9 RiMYB10
26.6 RiRosCOS536
26.8 RiSAMS
33.3 RiGIB

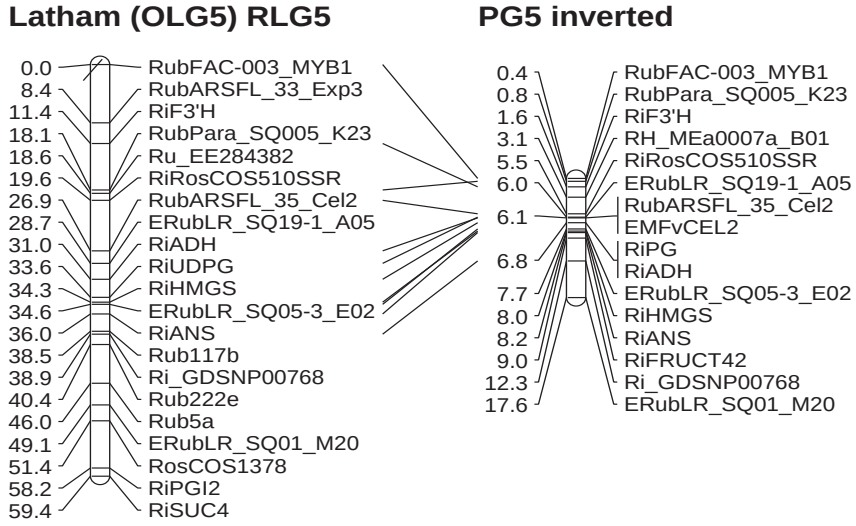
Latham (OLG1) RLG7a & b **MLG1 & MLG7 inverted**

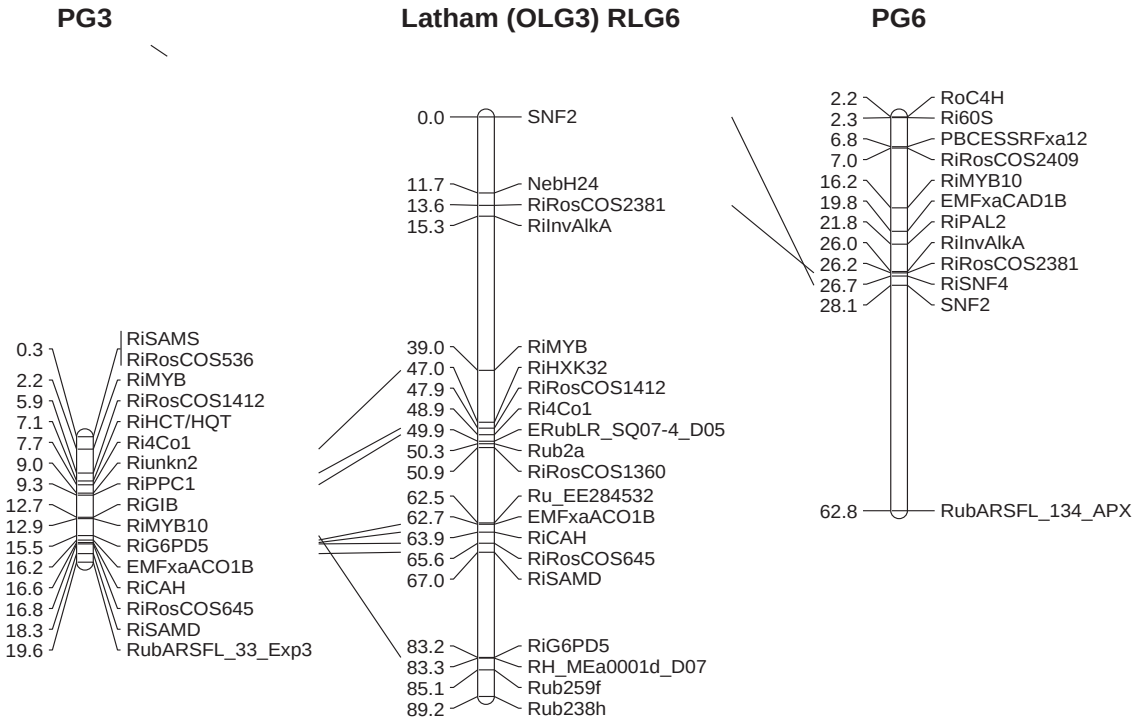


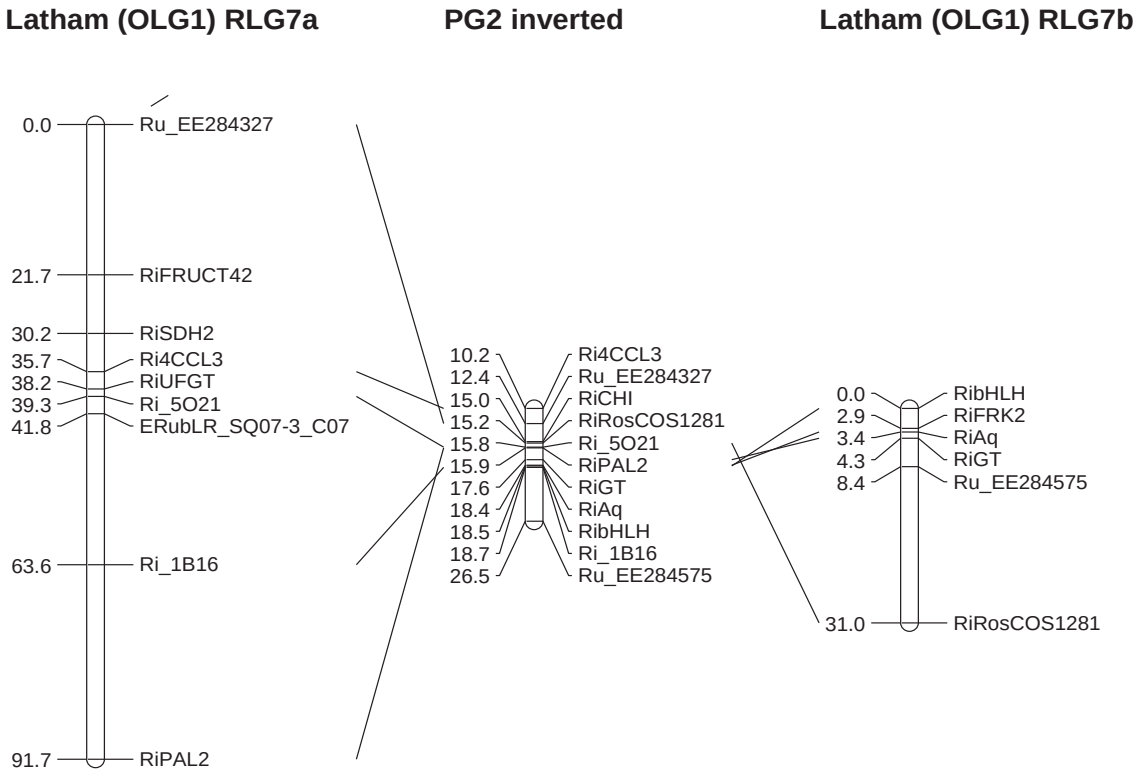
Supplemental Fig. S3.3: Comparison between *Rubus idaeus* 'Latham' genetic map and *Prunus* physical map constructed using BLAST analysis. 'Latham' map distance is measured in centimorgans (cM); *Prunus* linkage group (PG) map distance is measured in megabase pairs (Mbp). 'Latham' LG arranged in proposed order (RLG) with original numbering (OLG) in parentheses. Connecting lines indicate markers in common.




PG4
Latham (OLG2) RLG3
PG6
0.3 SNF1
0.6 RiACCO
EMFxaACO1B
1.0 ERubLR_SQ07-4_E09
1.4 RiANR
1.8 RiSTP1
3.0 ERubLR_SQ01_F09
RH_MEa0013c_G02
3.7 RH_MEa0014b_F07
Ri_5C10
7.4 RiPKP2
9.7 RiPPC1
10.5 Ro_1E22
17.7 RiSDH2
23.5 Ro_1K15
0.0 ERubLR_SQ07-4_E09
11.7 Ri_GDSNP00321
16.3 ERubLR_SQ01_F09
20.5 RH_MEa0013c_G02
21.0 RH_MEa0014b_F07
29.4 RiPKP2
32.6 RiPPC1
34.6 Ro_1E22
35.2 Ro_1I04
40.8 Ri_GDSNP02690
50.0 RiRosCOS2409
50.4 Rub105b
Rub160a
54.2 RubARSFL_134_APX
61.6 PBCESSRFxa12
66.5 RiF2KP
83.6 RiNAG1
85.3 RiACCO
89.7 RiSTP1
90.5 RiANR
96.0 RiHCT/HQT
2.2 RoC4H
2.3 Ri60S
6.8 PBCESSRFxa12
7.0 RiRosCOS2409
16.2 RiMYB10
19.8 EMFxaCAD1B
21.8 RiPAL2
26.0 RiInvAlkA
26.2 RiRosCOS2381
26.7 RiSNF4
28.1 SNF2
62.8 RubARSFL_134_APX


Latham (OLG5) RLG5
PG5 inverted
0.0 RubFAC-003_MYB1
8.4 RubARSFL_33_Exp3
11.4 RiF3'H
18.1 RubPara_SQ005_K23
18.6 Ru_EE284382
19.6 RiRosCOS510SSR
26.9 RubARSFL_35_Cel2
28.7 ERubLR_SQ19-1_A05
31.0 RiADH
33.6 RiUDPG
34.3 RiHMGS
34.6 ERubLR_SQ05-3_E02
36.0 RiANS
38.5 Rub117b
38.9 Ri_GDSNP00768
40.4 Rub222e
46.0 Rub5a
49.1 ERubLR_SQ01_M20
51.4 RosCOS1378
58.2 RiPGI2
59.4 RiSUC4
0.4 RubFAC-003_MYB1
0.8 RubPara_SQ005_K23
1.6 RiF3'H
3.1 RH_MEa0007a_B01
5.5 RiRosCOS510SSR
6.0 ERubLR_SQ19-1_A05
6.1 RubARSFL_35_Cel2
EMFvCEL2
RiPG
6.8 RiADH
7.7 ERubLR_SQ05-3_E02
8.0 RiHMGS
8.2 RiANS
9.0 RiFRUCT42
12.3 Ri_GDSNP00768
17.6 ERubLR_SQ01_M20


PG3
Latham (OLG3) RLG6
PG6
0.3 RiSAMS
RiRosCOS536
2.2 RiMYB
5.9 RiRosCOS1412
7.1 RiHCT/HQT
7.7 Ri4Co1
9.0 Riunkn2
9.3 RiPPC1
12.7 RiGIB
12.9 RiMYB10
15.5 RiG6PD5
16.2 EMFxaACO1B
16.6 RiCAH
16.8 RiRosCOS645
18.3 RiSAMD
19.6 RubARSFL_33_Exp3
0.0 SNF2
11.7 NebH24
13.6 RiRosCOS2381
15.3 RiInvAlkA
39.0 RiMYB
47.0 RiHXK32
47.9 RiRosCOS1412
48.9 Ri4Co1
49.9 ERubLR_SQ07-4_D05
50.3 Rub2a
50.9 RiRosCOS1360
62.5 Ru_EE284532
62.7 EMFxaACO1B
63.9 RiCAH
65.6 RiRosCOS645
67.0 RiSAMD
83.2 RiG6PD5
83.3 RH_MEa0001d_D07
85.1 Rub259f
89.2 Rub238h
2.2 RoC4H
2.3 Ri60S
6.8 PBCESSRFxa12
7.0 RiRosCOS2409
16.2 RiMYB10
19.8 EMFxaCAD1B
21.8 RiPAL2
26.0 RiInvAlkA
26.2 RiRosCOS2381
26.7 RiSNF4
28.1 SNF2
62.8 RubARSFL_134_APX


Latham (OLG1) RLG7a
PG2 inverted
Latham (OLG1) RLG7b
0.0 Ru_EE284327
21.7 RiFRUCT42
30.2 RiSDH2
35.7 Ri4CCL3
38.2 RiUFGT
39.3 Ri_5O21
41.8 ERubLR_SQ07-3_C07
63.6 Ri_1B16
91.7 RiPAL2
10.2 Ri4CCL3
12.4 Ru_EE284327
15.0 RiCHI
15.2 RiRosCOS1281
15.8 Ri_5O21
15.9 RiPAL2
17.6 RiGT
18.4 RiAq
18.5 RibHLH
18.7 Ri_1B16
26.5 Ru_EE284575
0.0 RibHLH
2.9 RiFRK2
3.4 RiAq
4.3 RiGT
8.4 Ru_EE284575
31.0 RiRosCOS1281

Supplemental Table S3-1: Redesigned and newly designed primer pair sequences of markers polymorphic in *Rubus occidentalis* 96395S1 x *R. idaeus* 'Latham'.

Primer name	Primer F 5'-3'	Primer R 5'-3'
ERubLR_SQ07_2_H02_HRM	TTTGAGATTGAGCCGTCGTT	CAAACTGACAAACGCTCTTCC
ERubLR_SQ07_3_C07_HRM	GGATCAAGGAGTGAGGATGG	CCGTGGTGGTTGTTATGTTG
ERubLR_SQ07_4_E09_HRM	GCTGCCATAGAAACAAACGA	AGAGCCTTCACCATTCTCTCC
mRaCIRRIV2A8	TAAAAAGGCGCAACAGTCG	AGACACAGAAACAGGCATCG
Ri_1B16	CTTGGGCAGCTTTAGCCTTT	AAGAAGAAGGGTGGGTTTCA
Ri_5C10	GAACCTTGCTGACCACCTTG	AAAACAAATTCCAGGCTCTCTC
Ri_5O21	ACGTTGAAGTTGCAGAGAAGG	AGAGGAAGAGAAGGGTTGCT
Ri_GDSNP00321	GCTGCAAAGAAGGAATTTGG	CAGCGTTCATTATCGAGAGG
Ri_GDSNP00768	CTAAGCCATGTCCCAGCTAA	GTGTTGTGCCCTTGTGGTAA
Ri_GDSNP02690	TCTGTCAAGTGGGCTCTCATACT	GTCATAACGGCAAGGACCA
Ri4CCL3	TGCTGCTACTCCAAATCCTC	ATAGCCGTCTCTGTGGCTCT
Ri4Co1snp1	CTTCTTGGGAGCCTCGTTCT	GCGAGTCAATGCACATCAA
Ri60S	CTCCGAATCAACAAGATGCT	TGAACTTAGCCCTCCTGAGA
RiAA	ACCACCATGTCTGAGGAGAA	GTCTCGCCGCTGTACTTAAC
RiABP_920	ACTATGGGAGGGGAGGTTTG	CCTAAGCAGAGGGAAGGATG
RiACCO	CCACTCCAACATATCCCAAA	GTTCCACTGCAAATTCCTTC
RiADH	GTGTGAGACATTGCCAAAAG	CGGTATTACTTTGATGGGAGAG
RiANR_00016	ACCAAGCCTGTTCCAGTGAG	GAAATCATGTGTCAAAGCCAAG
RiANS_00063	CGGGCAAAATTCAAGGCTAC	TGTAATCGGCAGGTGTTTGA
RiAPX	GCTGTAGCTCATCAGAAGCTC	TGCCTAAAACCCCAACACTA
RiAq	ATTGTCATCAGAGCCATTCC	TGCACATCAACACCTCAAAC
RibHLH_00475	CCTTGGAACATTCATCAGCA	CCTCTGCACCTCACAAATCA
RiCAH	CTTACAAGGGACGGAGGAAT	ACAAGCAACAATCGGAAGAG
RiCHI_1296	CCAATGTTGAAATGGTTGGA	CGCGTCAGGATTAGGATAGG
RiCXE	GCAGATTTAGACCGGGTTTT	TAGGGATGCACCAGAACAAT
RiDFR_0801	TTCCACAGCGTCAAGTGAGT	TCCGAATCCGTTTGTGTGA
RiF2KP	CCTTACATCCAGCCTCAAAA	GAAGAGTATGGCTTTTCTCCAG
RiF3'H_0002	TCGAGGGAGGGGATAAAGTC	CGCAACTGCTGAATCTGTGT
RiF3'H_4731	CTGTCCATAGCCACATTCCA	TGACCTCAAGTTTGGCCTTC
RiFLS_1828	ATTGACGAGCTTTGGGTGAG	TGAGCAATGGGAAGTACAAGG
RiFRK2	TTGCTAAGAAGGCTGGTTGT	ACAGTTCCCTCGATGTTGAA
RiFRUCT42	TAGTGGGAAAGGGTGAATGA	GGGTTTATCCAACAGAAGCA
RiG6PD4	GATGGTTATGGGATCATTCG	CTTGGCCTGGAAAGAAGTTT
RiG6PD5	GCAGGTTCTTTGTTTGGTTG	GGGGAACACATTTCAGATTG
RiGG6P2	CTCTGGAGGGTTTCCTCTCT	TAGCACACCATTCGTAACCA
RiGIB	CTTTTGCTGGTGCTTTGTCT	GCATTTGGAACAGCACTTCT
RiGLUK2	GCCCCAATACATTCAACTCA	ATAGACATTGGTTGCCTTGC
RiGT_03785	CATCCCCGTAGAGGTGAAGA	CCGATGCTGTTATGGAAACC
RiHCT/HQT_1896	ACTTCAAATGCGGAGGAGTG	GAGGGGTTTTCATCAGTGGA
RiHKL12	TTCAGTTTGGGTGGAACAGT	GCCTATGGTAGATGGCTTGA
RiHMGS	GAAGATCATGGAGCACAGGT	ACTTTCCCCATCCTTCTGAG
RiHXK32	AGAGAAGCCTTTCGTCCACT	CAGGCGCTGTTTAACTTCAT
RiHXT6	TAGAAGAACACCCAGCCAAC	CGTCAACAGAGTCGTGAGTG
RiINHI	TTATGGACTGATGCCTCCAC	TCACCATAGACCTTCCGAAC
RiInvAlkA2	AATCCACCTTCCTCTGGAAC	GTAAGGGCCATCAACAACAG
RiInvAlkE	CCCAGTTGACTCTGGTAGGA	CCTGGAACAAAAAGCACAAG
RiLTP	GTGTCAGCGTTCCCTACAAG	GGATGTCATAAACACCCACA
RiMYB_0209	GTTCGGACCACGTTTGAGTT	TCAAGATAAGCCCTCCTCCA
RiMYB10_snp	AAGCCAAACATCAAGAGAGGA	GCTTGTTCGCGCTACCTACT
RiNAG1	AAGTTTCACCCTTCCACCAT	GCCTTGTGAATCCTGAGAAA
RiPAL2_00457	CTGTGACCAACCATGTCCAA	GTTCTCCTCCAAATGCCTCA
RiPGI2	TGACCATCCTTAGGTTTGGA	AATACTTCTCCGGCTGCTTT
RiPG	GTGTGAGACATTGCCAAAAG	CGGTATTACTTTGATGGGAGAG
RiPI2	CTTAGGCCCTCTATTTGTGC	GGTAAGCCAACAATGCAAAG

RiPKP2	GCTGACTTGAATGTGCCTTT	GCCCCCTTCTCAATCAAATA
RiPPC1	GTGTGCCTATCGGATTCTTG	TCACTCTGCTCTTGTGGTGA
RiRosCOS1217	TTGGACGGGACAAAGTGGTT	TTGCTTGGCTCTGGAAAAAG
RiRosCOS1248	CGTTCCCAGAAACAAGTACCA	TGTGGGCAAAGCTCAAAATC
RiRosCOS1281	TTGGTATGTGAAGCGGACTG	TGCCTCTTGGATGGTCTTCT
RiRosCOS1360	TATTTCCCTCCTTCCCCTTC	CCCACATAAACCAAGGCTGT
RiRosCOS1412	AATTTTTGGGAACCCTTTGG	AATTTTTGGGAACCCTTTGG
RiRosCOS1549	TTTAAGCTGCGTTTGGTCAG	GGTCTGCTCCGATTCTGTAA
RiRosCOS2343	TTGCTTCTCATTTGCTGAGG	GCAACGCCAAACAATCAGT
RiRosCOS2381	TCGGGGTCAATTTCAGTTTC	TCTTTGTTTGGCCTGATTTG
RiRosCOS2409	CTGGGAGACAATGGCCTAAA	TCTGACGAATGGGAAGAGAGA
RiRosCOS3524	TGCATTGATGGGAACACTTT	AGCAGAGGAGGATCGAGTTG
RiRosCOS510SSR	GTGTGTAGGCGTGCAGATAA	ACCACCAGTAGGGAAACCAA
RiRosCOS530	TCAAGCAACATCAGGTCAGG	AGGCTCCGCACAAATTAAGA
RiRosCOS536	CTTTGGCCTATTACATCTTGTTCTC	AATTGTTTGCATGGATCTGG
RiRosCOS645	TGTGAATTGATTTGGGCTGT	TTCATAAAACCGTGGGGAAA
RiRosCOS873	CCCTACAGGTTTGGTCTATGCT	TCCCCATGTCAAGTGATTGT
RiSAMD	AGGTACTCTCGTGGGAGCTT	TGAGATTTGTCAGGGCTACC
RiSAMS	GATGAGATTGCTGCTGACCT	AGTGAGACCTGCATCACCAT
RiSDH2	CACATGGAGATGACCAACAA	ACATATCCCAACAGCCTTCA
RiSNF4	GTTACCCTGGAGACCAATCC	AGGGTCAATTTGTGGGAGTT
RiSPS2	AGCACGAATCAAAAGAGGTG	TGCAGAAAGGAGGACAAAAC
RiSS	GAGTGGTATGGCAAGAATGC	TCAGCCTTCTCCTCATTGTC
RiSTP1	GATCCTCTGGAGTTGGGATT	TTTGAGTTTGGGAGGTGCTA
RiSUC4	GTTTATGGCGGTTGGTAATG	CCTGAGCTGCTGATATGCTT
RiTh	AACCCTAAGCCAAGGACCAT	CACCACCCATGACAGTCAGA
RiUDPG	AGGCAATCTTCCAACAGCTA	GGGGAATTATTTTGCTGAGG
RiUFGT_00035	CACGTTGAGGAACCTGTTGA	CGGAAGGAATCATCTTTGGA
Riunkn2	TCCCGAGAGTTATTCTGGAT	TTTAGACCGAGCTTCAGTTG
Riunkn	GCTTGGAAAGAAACAGAGGA	AACACACGACCTTTGGAAGA
Ro_10O04	ACGACGAGACCACCGTAAAG	TCGGATCACAAACAAACGAC
Ro_11C01	TGAAGCAAGCAACGTCCTC	CGATCTTAACACAGCCTCACA
Ro_1E22	GCTGCTCTTGCTTCTGCTTT	ATCGGCATCGTCACTCTTTC
Ro_1I04	ATGTCTGGAAAGTGCGACAG	AAAACGGGGAAACGTATGAA
Ro_1K15	GCGTTTGGTTGGGATCTATG	GTATGCCCGATCATTTACTTCT
RoC4H_6627	GACAACTCCTCCTCCCAACA	AACCCTGCACACTGGAAGAA
Ru_EE284327	CAGTAATTGAGAACCATGCCTAC	CTGCACTGATGGGTTTGCT
Ru_EE284365	ATTGTTGAGGCGGAATGAAA	GGCAAAGTTGACCGAAGCA
Ru_EE284382	ACGGAGGATGACAGAGAACC	AGGTGAGGTGGGAGATGATG
Ru_EE284400	TACGCACGCTTATGAACCAC	ACACAGCACCACTCGATGAA
Ru_EE284532	TCATGGACGACTAGGGTCTG	CCAAGCCGGAGTAGTAGGTT
Ru_EE284575	AAAACACCCAAGCCCAGAGT	GGAATGCCAGAAGTGAAAGG
RubARSFL_134	CTTTCAGGACAAGCCCAATC	GCCAAACACATCCCTCAAGT
RubARSFL_33	AGGGTACGGGACAAACACAG	AATGTTGCTGAGGAGGGTTG
RubARSFL_35	CTGGTTGGGGGTTACTACGA	AACATTGGGTTCTGGGTGAG
RubARSFL_35	CTGGTTGGGGGTTACTACGA	AACATTGGGTTCTGGGTGAG
RubFAC-003	CCTTTCCCCTTTCCCAATCT	GCTAGGTACAGAAGAGCCAAAAA
SNF1	CCCTTACAAAGACCGGATGA	TTCCCACCATAGCTCCATGT
SNF2	TGGGAGGAAGTGCAGAGAGT	TGCTCATCTCTGGTGCAGAT

Supplemental Table S3-2: Marker design information. Information on origin and type of sequence, locations in strawberry, apple and peach genomes, expected values (E-values), and percent identity for each marker mapped in *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'.

(Three A3 pages following)

CHAPTER FOUR

This chapter is in preparation as:

***Jill M. Bushakra**, Célia Krieger, Dawei Deng, M. Joseph Stephens, Andrew C. Allan, Roy Storey, V. Vaughan Symonds, David Stevenson, Tony McGhie, David Chagné, Emily J. Buck, Susan E. Gardiner (in prep). QTL involved in the modification of cyanidin compounds in black and red raspberry fruit.*

4 QTL involved in the modification of cyanidin compounds in black and red raspberry fruit

4.1 Abstract

Fruit from *Rubus* species are highly valued for their flavor and nutritive qualities. Anthocyanin content contribute to these qualities, and although many studies have been conducted to identify and quantify the major anthocyanin compounds from various *Rubus* species, the genetic control of the accumulation of these complex traits in *Rubus* is not yet well understood. The identification of the regions of the genome involved in the production of anthocyanins is an important first step in identifying the genes underlying their expression. In this study, ultra and high performance liquid chromatography (UHPLC and HPLC) and two newly developed *Rubus* linkage maps were used to conduct QTL analyses to explore the presence of associations between concentrations of five anthocyanins in fruit and genotype. In total, 27 QTL were identified on the *Rubus* linkage maps, four of which are associated with molecular markers designed from transcription factors and three of which are associated with molecular markers designed from anthocyanin biosynthetic pathway candidate genes. The results of this study suggest that, while QTL for anthocyanin accumulation have been identified on six of seven *Rubus* linkage groups (RLG), the QTL on RLG2 and RLG7 may be very important for genetic control of cyanidin modification in *Rubus*.

4.2 Introduction

One of the largest and most diverse of plant genera, *Rubus* contains between 600 and 800 species, including blackberry and red and black raspberry, is distributed world-wide (Thompson 1995b) and has a long history of use for food and medicinal purposes (reviewed by Hummer 2010). Biologically active (bioactive) compounds found in red and black raspberry have been extracted for individual evaluation, in either animal models or *in vitro* assays, for their health-promoting characteristics including the reduction of cell proliferation in rat models (Chen et al. 2011), anti-inflammatory effect on ulcers (Montrose et al. 2011), and reduction of blood cholesterol levels (Ash et al. 2011). These reports, and others, have led to the marketing of bioactive compounds as ingredients in "functional foods", that is products reported to have a specific health benefit (Finley et al. 2011; Espín et al. 2007).

Anthocyanins (ACY) are a class of water-soluble flavonoids that are derived from phenylalanine and stored in the plant vacuole. They provide plant tissues with a range of colors from orange/red to violet/blue, are widely distributed in flowering plants (Tanaka et al. 2008; Holton and Cornish 1995), and are among the bioactive compounds found in *Rubus* plants. Although the ACY profiles of cultivated red raspberry (*Rubus idaeus*), black raspberry (*R. occidentalis*), tetraploid blackberry and hybrid berry (*Rubus* sp.) fruit are the best characterized of the genus (Connor et al. 2005; Dossett et al. 2010; Jennings and Carmichael 1980; McGhie et al. 2002; Scalzo et al. 2008; Torre and Barritt 1977), the ACY biosynthetic pathway has yet to be examined in *Rubus*. However, the biosynthetic pathway of ACY production appears to be well conserved in *Arabidopsis thaliana*, maize, snapdragon, and petunia (reviewed by Holton and Cornish 1995), as well as in apple (Takos et al. 2006), and many of the genes encoding the enzymes involved are conserved in both sequence and function (reviewed by Tanaka et al. 2008; Holton and Cornish 1995; Grotewold 2006).

Several studies have been conducted to determine the predominant ACY compounds in *Rubus* species (Ozgen et al. 2008; Dossett et al. 2010; Määttä-Riihinen et al. 2004; Connor et al. 2005; Cooney et al. 2004; Torre and Barritt 1977; Seeram et al. 2001; Jennings and Carmichael 1980; Dossett et al. 2011; Tulio et al. 2008). Black raspberry fruits are characterized by two xylose-containing pigments, cyanidin 3-*O*-sambubioside (C3Sb) and cyanidin 3-*O*-2^{G}-xylosylrutinoside (C3XR), and overall high ACY concentrations (200 mg/100 g fruit) (Tulio et al. 2008; Ozgen et al. 2008); whereas red raspberry fruits are characterized by the presence of cyanidin 3-*O*-sophoroside (C3S), cyanidin 3-*O*-2^{G}-glucosylrutinoside (C3GR), and overall lower ACY concentrations (60 mg/100 g fruit) (Torre and Barritt 1977; Jennings and Carmichael 1980). Other ACY compounds found in *Rubus* fruits are cyanidin 3-*O*-rutinoside (C3R), cyanidin 3-*O*-glucoside (C3G), and pelargonidin 3-*O*-rutinoside (P3R) (Dossett et al. 2011; Kassim et al. 2009; Torre and Barritt 1977). Each of the cyanidin-containing compounds is a variation on the molecule cyanidin 3-*O*-glucoside, resulting from the enzyme-catalyzed addition of different sugar moieties, either glucose, rhamnose or xylose (Grotewold 2006; Holton and Cornish 1995). Pelargonidin 3-*O*-rutinoside (P3R), and cyanidin 3-*O*-glucoside (C3G) have similar structures, with the basic pelargonidin and cyanidin three-ring molecule differing by the presence of a second hydroxyl on the cyanidin B-ring 3'-

carbon, the addition of which is catalyzed by flavanone 3'hydroxylase (F3'H) early in the biosynthetic pathway (Grotewold 2006; Holton and Cornish 1995).

Four studies have been conducted to detect QTL in a population of *Rubus idaeus* 'Latham' × 'Glen Moy', both red-fruited cultivars. An assessment of fruit development from bud break to ripeness identified QTL involved in stages of fruit ripening and cane height to be associated with regions on linkage groups (LG) 2, 3, 5 and 6 (Graham et al. 2009). Two QTL involved in the variation of the accumulation of specific ACY compounds were mapped to regions of LG 1 and 4 (Kassim et al. 2009). QTL involved in the variation in fruit color during ripening, and the accumulation of total ACY were mapped to LG 2, 3, 4 and 6 (McCallum et al. 2010). Each of these studies identified unique QTL, although the QTL on LG3 show some overlap, and the QTL associated with marker RubFruitE4 on LG4 are in common between the studies of McCallum and colleagues, and Kassim and colleagues. The most recent study assesses QTL associated with total ACY content along with other fruit quality traits (Dobson et al. 2012). This study identified two QTL associated with total ACY on LG1 and LG2. The linkage group nomenclature used in these four studies follows Graham et al. (2004). These four studies provide insight into QTL of several fruit quality traits in a single population of red raspberry progeny over several years, as well as under different growing locations and growing conditions.

New genetic linkage maps constructed from a progeny of black raspberry (*Rubus occidentalis*) 96395S1 (deep purple fruit), and red raspberry (*R. idaeus*) 'Latham' (red fruit), have been developed recently (Bushakra et al. 2012a). These maps, deliberately enriched for markers derived from orthologous sequences and candidate genes coding for enzymes involved in the polyphenolic biosynthetic pathway, allow the association of candidate genes with QTL for variations in amounts of ACY compounds in this study. Improving our understanding of the genetics controlling the production and interactions between levels of ACY compounds will assist in the development of marker assisted breeding (MAB) to improve the efficiency of breeding new cultivars with the most beneficial balance of compounds for human health. To this end, ACY phenotype data were collected over three fruiting seasons and analyzed for marker associations using the 'Latham' (red raspberry) and 96395S1 (black raspberry) genetic linkage maps (Bushakra et al. 2012a) to identify QTL.

4.3 Methods

4.3.1 Plant material

The diploid purple raspberry mapping population consists of 500 full-sib F_1 individuals derived from an inter-specific cross between parents of contrasting fruit color, *Rubus occidentalis* selection 96395S1 (deep purple fruit) and *R. idaeus* 'Latham' (red fruit) (Bushakra et al. 2012a). The fruit from a sub-set of 155 individuals (Bushakra et al. 2012a) was selected for ACY extraction and QTL analysis.

4.3.2 Marker nomenclature and linkage groups

Markers were designed previously, and

Table 4-1 and Table 4-2 provide information on the sources of the orthologous sequence used for primer design (Bushakra et al. 2012a). *Rubus* linkage groups (RLG) are named to correspond to the diploid *Fragaria vesca* LG nomenclature (Bushakra et al. 2012a).

Table 4-1: Derivation of markers mapped on *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' that were designed from transcription factors and genes for enzymes involved in the polyphenolic biosynthetic pathway (Bushakra et al. 2012a).

Marker name	*Sequence source species*	*Accession*	*Putative gene function*
EMFxaCAD1B*	strawberry	AF320110	cinnamyl alcohol dehydrogenase
Ri4CCL3	raspberry	AF239685.1	4-coumarate:coA ligase 3
Ri4Co1	raspberry	AF239687	4-coumarate:coA ligase 1
RiANR	strawberry	DQ664193	anthocyanidin reductase
RiANS	raspberry	AY695818	anthocyanidin synthase
RiCHI	apple	CN946541	chalcone isomerase
RiDFR	strawberry	AY695813.1	dihydroflavonol 4-reductase
RiF3H	raspberry	EU078685	flavanone-3-hydroxylase
RiF3'H	apple	FJ919631	flavonoid 3' hydroxylase
RiFLS	strawberry	DQ087252.1	flavonol synthase
RiGT	apple	EB124403	UDP-glucosyltransferase
RiHCT/HQT	apple	NP_199704	hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase
RiPAL2	raspberry	AF237955.1	phenylalanine ammonia-lyase 2
RiUFGT	strawberry	AY575056	UDP glucose:flavonoid-3-O-glucosyltransferase
RoC4H	raspberry	FJ554629.1	cinnamate-4-hydroxylase
RibHLH	apple	ABB84474.1	b-helix loop helix transcription factor
RiMYB_209	raspberry	FJ527833	Myb-related transcription factor
RiMYB10	raspberry	unpublished	MYB10 transcription factor
RubFAC-003_MYB1	strawberry	AF401220	MYB1 transcription factor
RubFruitE4†	raspberry	unpublished	bZIP transcription factor

*(Sargent et al. 2007a)
†(Graham et al. 2004)

Table 4-2: Derivation of markers mapped on *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' that were designed from genes from pathways other than the polyphenolic pathway, and other sequences (Bushakra et al. 2012a).

Marker name	*Sequence source species*	*Accession*	*Putative gene function*
EMFxaACO1B*	strawberry	AJ851828	1-aminocyclopropane-1-carboxylate oxidase
RiABP	strawberry	X91839	auxin-binding protein
RiAq	raspberry	GT119101	aquaporin
RiCAH	raspberry	GT119097	cytoplasmic aconitate hydratase (Aco2)
RiF2KP	*A. thaliana*	AT1G07110	fructose-6-phosphate 2-kinase
RiFRK	*A. thaliana*	AT1G06020	fructokinase-3
RiFRUCT4	*A. thaliana*	NM_101096.2	b-fructofuranosidase
RiG6PD5	*A. thaliana*	NM_121314	glucose-6-phosphate dehydrogenase 2
RiHKL	*A. thaliana*	AT1G50460	hexokinase
RiPPC1	tomato	AJ243416.1	phosphoenolpyruvate carboxylase 1
RiSAMS	raspberry	GT119107	s-adenosylmethionine synthetase
RiSDH2	*A. thaliana*	AY133848.1	sorbitol dehydrogenase
RiSNF4	raspberry	GU056038	protein kinase
RiSUC4	*A. thaliana*	AT1G09960	sucrose transporter 4
RubARSFL_134_APX	strawberry	AF158652	cytosolic ascorbate peroxidase (ApxSC)
RubARSFL_35_Cel2	strawberry	AF054615	cellulase
SNF2	peach	BU044922	protein kinase
RiRosCOS1549	Rosaceae	ROSC_FMLY_CSA1_1549†	similar to Tsi1-interacting protein
RiRosCOS2343	Rosaceae	ROSC_FMLY_CSA1_2343†	uncharacterized protein
RiRosCOS2381	Rosaceae	ROSC_FMLY_CSA1_2381†	uncharacterized protein
RiRosCOS2409	Rosaceae	ROSC_FMLY_CSA1_2409†	putative 1-hydroxy-2-methyl-2-(E)-butenyl 4- diphosphate synthase, GcpE family protein
RosCOS1378‡	Rosaceae	ROSC_FMLY_CSA1_1378	putative UDP-glucose:glycoprotein glucosyltransferase
PBCESSRFxa12§	strawberry	DQ180317	SSR
Ri_GDSNP00321	apple	GD_SNP00321	SNP
Ri_GDSNP02690	apple	GD_SNP02690	SNP

*(Sargent et al. 2007a)
†
http://cgpdb.ucdavis.edu/rosaceae_assembly/rosaceae_sequences_412832_Dec_2007.Clean.COS.CDS.assembly
‡(Cabrera et al. 2009)
§(Keniry et al. 2006)

4.3.3 Experimental design, sample collection, extraction, preparation and analysis

All fruit samples were collected from a single, non-replicated population of the F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'. The population is located

at the PFR research site near Motueka (41°058'S, 172°584'E). For 2009, 2010 and 2011, field-grown fruit was collected three times, once per week over three weeks to represent the typical harvest season for raspberry in New Zealand (mid-December through early-January). Individuals were arranged in rows and after planting were assigned a unique identification number (EA501 – EA701). Plants were grown following standard New Zealand commercial raspberry practice.

Fruit samples were collected on ice and then stored at -20 °C until processing. The three samples collected per individual were pooled for processing either by whole fruit solvent extraction (2009) or by juice extraction (2010, 2011) methods. In 2009, the first 200 plants that were producing fruit were selected for analysis. Of those 200 individuals, 155 had sufficient fruit production accumulated over the three harvests for analysis. In 2010 and 2011, the fruit from all 500 F_1 progeny were netted to protect the fruit from bird predation, and samples were collected from all individuals, with the sub-set of 155 F_1 progeny being used for the QTL analysis.

4.3.3.1 Sample extraction, 2009: Solvent extraction method

Frozen whole fruit samples were shipped overnight on dry ice from Plant & Food Research (PFR), Motueka, New Zealand to PFR, Palmerston North for processing. The three replicates of fruit from each individual were pooled. Twenty grams, including seeds, of the pooled fruit samples, were homogenized with 100 mL ethanol/water/formic acid (80:20:1, v/v/v) using an Omni GLH homogenizer (Omni International, Marietta, GA), and stored at 1˚C for 48 h before centrifugation at 1000 g for 10 min with a Jouan® Hema-C centrifuge (Jouan, Saint Herblain, France). An aliquot (180 µL) of the supernatant was transferred to a 96-well plate for analysis by reversed-phase ultra-high performance liquid chromatography (UHPLC).

4.3.3.2 UHPLC analysis of fruit polyphenolics

Analyses of ACY composition and concentration in individuals in the sub-set of 155 F_1 progeny of 96395S1 × 'Latham', were performed on a Dionex Ultimate® 3000 Rapid Separation LC system equipped with a SRD-3400 solvent rack with four degasser channels, HPR-3400RS binary pump, WPS-3000RS thermostated autosampler, TCC-3000RS thermostated column compartment, and DAD-3000RS Diode Array Detector (Sunnyvale, CA, USA). A volume of 1.0 µL supernatant was injected into a Zorbax™ SB-C18 2.1 x 150 mm, 1.8 µm analytical column (Agilent Technologies, Santa Clara,

CA, USA) and analyzed at a mobile flow rate of 0.5 mL/min with column oven temperature maintained at 45 °C. Solvents were (A) 0.5% phosphoric acid in water, and (B) acetonitrile/water (80:20 v/v); a binary gradient elution was performed as follows: isocratic at 90% A:10% B (0–0.5 min); linear gradient to 85% A:15% B (0.5–3.5 min); linear gradient to 50% A:50% B (3.5–9.0 min); linear gradient to 100% B (9.0–9.5 min), isocratic at 100% B (9.5–10.5 min); linear gradient to 90% A:10% B (10.5–11.0 min). Total sample run time was 13.5 min and spectral data were collected for the entire run.

4.3.3.3 Compound quantification, 2009

Quantification of ACY extracted from the sub-set of 155 F_1 progeny was carried out at 520 nm as cyanidin 3-*O*-glucoside equivalents (Mullen et al. 2003). Chromatographic data were collected and manipulated using the Chromeleon® Chromatography Management System version 6.8 (Dionex Corporation, Sunnyvale, CA, USA). External calibration curves were constructed for cyanidin 3-*O*-glucoside using standards from Extrasynthese (Genay, France). Components for which standards were not available were quantified using the standard curve of a related compound. The results are presented as μg/mL.

4.3.3.4 Sample extraction, 2010 and 2011: Juice extraction method

The three replicates of fruit from each individual from the F_1 progeny and the parents (fruit from both parents collected 2010, fruit from 'Latham' only collected 2011) were pooled to give an average fruit weight per individual of 348 g. Juice was extracted using a manual juicer with fine-screen mesh to exclude pulp and seeds. The juicer was washed well and dried between samples. For each extraction, a 1.0 mL sample was collected from the total juice, transferred to a 1.5 mL Eppendorf® tube, and stored at -80°C until analysis. Juice samples were shipped overnight on dry ice from PFR, Motueka to PFR, Ruakura for analysis.

4.3.3.5 HPLC analysis of fruit anthocyanins: 2010 and 2011

Analyses of ACY composition and concentration for fruit from the sub-set of 155 F_1 progeny of 96395S1 × 'Latham', were performed on a Shimadzu 20-series analytical HPLC with a column oven, auto-sampler, 4-channel vacuum solvent degas module and diode-array detector (Portland, OR, USA). A volume of 10 μL supernatant was injected into a 150 × 2 mm, Synergi Polar-RP, 4 μ particle size, 80Å pore size analytical column,

fitted with a Security-Guard 3 × 2 mm Polar-RP guard cartridge (Phenomenex, Auckland, New Zealand). Flow rate was 0.6 mL/min and column oven temperature 50ºC. Solvents were (A) methanol (neat) and (B) 2% aqueous formic acid, and the initial mobile phase was 5% (A):95% (B). The time program for pump B concentration was set up as 92% at 2.5 min, 82% at 5 min, 72% at 7 min, 58% at 9.5 min, 45% at 11 min, 30% at 12 min, 20% from 12.5–13.3 min, returning to 95% at 13.8 min and holding at that concentration until the end of the run at 15 min.

Quantification of ACY extracted from the 500 F_1 progeny was carried out at 520 nm, in comparison with standard solutions of cyanidin 3-*O*-glucoside (Mullen et al. 2003). UV/visible spectra were recorded from 245-600 nm in 1.2 nm steps. Results were expressed as mg cyanidin 3-*O*-glucoside equivalent/100 mL of juice.

4.3.3.6 Confirmation of compound identity, 2010 and 2011

To assist identification of compounds observed during HPLC runs, some samples were rerun on the same HPLC instrument, with the addition of a Shimadzu LCMS 2020, single quadrupole mass spectrometer, fitted with an electrospray ionization (ESI) source [Shimadzu Scientific Instruments (Oceania) Pty Limited, Auckland, NZ], following manufacturer's default parameter settings. MS scans were carried out in both positive and negative mode in the same run, using a mass range of 140-2000 Da. Confirmation of identity was achieved both through comparison of expected and observed molecular ion masses and "neutral losses" of sugars during ionization to leave aglycones or partially glycosylated ions. For example, cyanidin 3-*O*-glucoside exhibited a positive molecular ion at 449 Da and an ion at 287 Da, corresponding to loss of the glucose residue (162 Da), to leave the cyanidin ion.

4.3.4 Statistical analyses

The GenStat 14th edition software package (VSN International 2010) was used to analyze the phenotypic data. The mean concentration values for the 155 F_1 progeny for all three years were calculated for each compound. Within-year, pair-wise compound correlations were conducted using the concentration data. A probability distribution plot was applied to verify the normality of the trait distributions, and the means, ranges, standard deviations (SD), skewness and kurtosis were calculated for these distributions. Transgressive segregants (those progeny whose values exceed or are less than, the

parental values by at least two SD) were identified for the compounds analyzed in 2010, as the parental data was available only for that year.

As the methods of analysis resulted in the presentation of data with different units for 2009 and 2010/2011, each compound was evaluated as a percentage of the total ACY for each year to assess the proportion of each compound per year. Restricted maximum likelihood (REML) was conducted on the mean percent data to access the trait variance contributed by year-to-year variation and by differences in genotype. The response variate was the trait's percent of total ACY, the fixed model was constant, the random model was Year+Genotype+Year.Genotype, and there were 462 units (3 years x 155 individuals – 3 individuals with no data for 2011), with 3 units excluded for missing data.

4.3.5 QTL map construction

Markers for the QTL mapping were selected from the parental linkage maps (Bushakra et al. 2012a) with the goal of developing linkage groups with one marker every 2–5 centimorgans (cM) and including as many gene-based markers as possible (Tables 4-1 and 4-2). Marker selection was based on map location, least missing data, best segregation ratio, and when possible, markers in common with other *Rubus* linkage maps. The linkage maps were generated using MapChart® 2.1 (Voorrips 2002).

4.3.6 QTL mapping

QTL analyses were conducted for the concentration of each ACY compound and for total ACY for each year using MapQTL® 5 (van Ooijen 2004). To limit the genotypic classes to two, QTL detection was performed on the parental maps (Bushakra et al. 2012a). Logarithm of odds (LOD) thresholds for each map were estimated with a 1,000-permutation test for each trait and year. An interval mapping (IM) analysis was then performed to locate preliminary QTL. QTL with LOD scores greater than the genome-wide threshold of 95% were considered significant. Markers with significant association with a trait were selected for use as cofactors in the Multiple QTL Method (MQM) to uncover the presence of minor QTL that might be hidden by large-effect QTL. A step size of 1 cM was used for both IM and MQM analyses. The square of the partial correlation coefficient (R^2) was used to calculate the phenotypic variance explained by a single QTL. The 95% and 99% confidence intervals for QTL positions were obtained by taking the two positions left and right of the point estimate of the QTL for the LOD±1 interval, and

the next two positions out, left and right, for the LOD±2 interval (van Ooijen 1992), respectively for QTL detected with IM and with MQM.

4.3.7 Epistasis calculations

An evaluation of potential epistasis was conducted on the traits C3S, C3XR C3XR_C3R, and C3GR, as these traits had at least two QTL that were stable over 2009 and 2010. Simple fixed-effect models, whereby the particular ACY was modeled by the putative QTL (as fixed effects), and all possible interactions were fitted. Residual plots were checked for the presence of outliers and to ascertain the validity of the assumption of normality (Supplemental Fig. S4.1). All analyses and figures were undertaken in R 2.13.0 with the models using the asreml 3.0.1 (P. Alspach, pers. comm.).

4.3.8 Comparative gene analysis

The annotated genes from the syntenic region of *Fragaria vesca* LGII (1:4 Mb) were mined to identify candidates potentially involved in the expression of the large effect QTL located at the top of 'Latham' RLG2. Gene models identified in *Fragaria* were compared to the *Malus* × *domestica* 'Golden Delicious' (Velasco et al. 2010) and *Fragaria vesca* 'Hawaii 4' (Shulaev et al. 2011) draft genome sequences. Over 400 gene models were represented. The best basic local alignment search tool (BLAST) matches for the gene models were then compared to The *Arabidopsis* Information Research (TAIR) database (www.arabidopsis.org) to validate the predicted gene function.

4.4 Results

4.4.1 Description of traits

Seven ACY compounds were detected in the F_1 progeny in all three years and in both parents in 2010. Five of the compounds, co-eluting compounds cyanidin 3-*O*-glucoside and cyanidin 3-*O*-sambubioside (C3G_C3Sb), cyanidin 3-*O*-sophoroside (C3S), co-eluting compounds cyanidin 3-*O*-2^{G}-xylosylrutinoside and cyanidin 3-*O*-rutinoside (C3XR_C3R), cyanidin 3-*O*-2^{G}-glucosylrutinoside (C3GR), and pelargonidin 3-*O*-rutinoside (P3R) were detected in all three years, were identified based on UHPLC or HPLC retention times and confirmed with mass spectrometry data, and therefore selected for QTL analysis (Table 4-3). In the samples from 2009 and 2011, compounds C3G and C3Sb co-eluted as a single peak (C3G_C3Sb), while in the samples from 2010, C3G only was detected. In the samples from 2009, compounds C3XR and C3R were detected as

separate peaks, while in the samples from 2010 and 2011, the two compounds co-eluted as a single peak (C3XR_C3R). Therefore, the concentrations of C3G_C3Sb 2009 and 2011 and C3G 2010, as well as the sum of C3XR 2009 and C3R 2009 (C3XR+C3R) and co-eluting compounds C3XR_C3R 2010 and 2011 were used for within-year comparisons and the proportions of total ACY for each compound were used for between-year comparisons. As the co-eluting compounds are each considered as a single trait, five traits in total were considered for analysis. The two unknown ACY, present in trace amounts, were excluded from further analysis. Total ACY was calculated as the sum of all analyzed compounds for each year.

Table 4-3: Anthocyanin compounds extracted from purple raspberry fruit collected from 155 F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham', analyzed with ultra high performance liquid chromatography (UHPLC) in 2009 and HPLC in 2010 and 2011, and used for quantitative trait locus (QTL) analysis.

Chemical name	*Common name*	*Abbreviation*
cyanidin 3-*O*-glucoside	cyanidin 3-*O*-glucoside	C3G
cyanidin 3-*O*-[2-(xylosyl)glucoside]	cyanidin 3-*O*-sambubioside	C3Sb*
cyanidin 3-*O*-[2-(glucosyl)glucoside]	cyanidin 3-*O*-sophorside	C3S
cyanidin 3-*O*-[6(rhamnosyl)glucoside]	cyanidin 3-*O*-rutinoside	C3R†
cyanidin 3-*O*-[2-(xylosyl)-6-(rhamnosyl)glucoside]	cyanidin 3-*O*-2^{G}-xylosylrutinoside	C3XR†
cyanidin 3-*O*-[2-(glucosyl)-6-(rhamnosyl)glucoside]	cyanidin 3-*O*-2^{G}-glucosylrutinoside	C3GR
pelargonidin 3-*O*-[6-(rhamnosyl)glucoside]	pelargonidin 3-*O*-rutinoside	P3R

*Cyanidin 3-*O*-sambubioside co-eluted with C3G and was detected in 2009 and 2011 only.
†Cyanidin 3-*O*-rutinoside and cyanidin 3-*O*-2^{G}-xylosylrutinoside co-eluted as a single peak in 2010 and 2011

Relative proportions of total ACY content of each analyzed compound, calculated from the mean compound values of the sub-set of 155 F_1 progeny, were similar for all three years (Table 4-4). C3GR had the highest overall proportion of total analyzed ACY in the progeny (33%, 2009; 39%, 2010; 37%, 2011), followed by co-eluting compounds C3XR_C3R (sum of compounds 39% 2009; 33% 2010; 34% 2011). The mean compound concentrations in 2010 and 2011 were not significantly different with the exception of C3GR, which was higher in 2010 (2010 81.7 mg/100mL; 2011 71.1 mg/100 mL; $p<0.05$).

As the mean differences between 2010 and 2011 were largely not significant, all between-year comparisons are performed using data from 2009 and 2010 only.

Table 4-4: Mean proportions of anthocyanins detected in the F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' in 2009, 2010, and 2011.

Trait	*Mean*	*SD*	*SEM*
C3G_C3Sb 2009	10%	3%	0%
C3G 2010	7%	2%	0%
C3G_C3Sb 2011	8%	2%	0%
C3S 2009	17%	11%	1%
C3S 2010	19%	10%	1%
C3S 2011	21%	10%	1%
C3XR+C3R 2009	39%	18%	1%
C3XR_C3R 2010	33%	14%	1%
C3XR_C3R 2011	34%	13%	1%
C3GR 2009	33%	11%	1%
C3GR 2010	39%	8%	1%
C3GR 2011	37%	2%	0%
P3R 2009	1%	2%	0%
P3R 2010	2%	2%	0%
P3R 2011	0%	1%	0%

Abbreviations: SD: Standard deviation; SEM: Standard error of the mean; C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^{G}-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^{G}-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside

Weather data, compiled by MetWatch Online (www.metwatch.co.nz), for the typical New Zealand raspberry growing season (1 August through 8 January) for 2009, 2010 and 2011, are shown in Table 4-5. Total rain fall (millimeters, mm), mean daily temperature (degrees Celsius, °C), and growing degree days [GDD, number of hours with day-time temperatures above the base temperature (McMaster and Wilhelm 1997)] were all higher in 2010 than in 2009 and 2011.

Table 4-5: Summary of growing conditions at Plant & Food Research, Motueka Research Centre for the *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' parents and F_1 progeny.

Year	Rain Total (mm)	Solar Energy Total (MJ/m^2)	Temperature (Dry Bulb) Mean (°C)	Growing Degree Days (GDD)*
2009/10	549.4	2537.31	11.9	882.5
2010/11	612.0	2726.08	13.4	1125.3
2011/12	607.2	2744.54	11.6	869.3

*GDD for 1 Aug-8 Jan, representing the typical growing season for raspberry; base temperature 7°C.
Abbreviations: mm: millimeter; MJ/m^2: megajoules per square meter; °C: degrees Celsius

Table 4-6: Mean phenotypic values and basic statistical parameters for the five traits analyzed in 155 F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' in 2009 and 2011, and the progeny and parents in 2010.

Trait	F_1 Mean	F_1 Maximum	F_1 Minimum	Standard Deviation	Skewness	Kurtosis	Probability distribution	96395S1 *	'Latham'*	Transgressive seqreqants †
C3G_C3Sb_09	7.3	21.8	2.6	3.4	1.8	4.6	p≤0.01	n.m.	n.m.	
C3S_09	12.6	40.4	0.7	9.1	0.7	-0.3	p≥0.05	n.m.	n.m.	
C3XR_09	10.0	27.0	2.0	5.4	0.9	0.3	p≤0.01	n.m.	n.m.	
C3R_09	19.7	61.5	4.6	11.7	1.2	1.2	p≤0.01	n.m.	n.m.	
C3GR_09	24.8	54.7	6.6	11.0	0.5	-0.5	p≥0.05	n.m.	n.m.	
P3R_09	0.8	10.2	0.1	1.1	5.4	40.9	p≤0.01	n.m.	n.m.	
C3G_10	15.1	27.0	4.6	4.3	0.3	-0.3	p≥0.05	46.0	8.4	0
C3S_10	40.9	99.7	4.1	25.7	0.4	-0.9	p≤0.01	1.3	34.6	10
C3XR_C3R_10‡	69.5	160.3	19.0	28.0	0.8	0.5	p≤0.01	239.6	11.5	0
C3GR_10	81.8	178.8	32.1	29.6	0.4	-0.3	p≥0.05	3.0	28.8	68
P3R_10	3.6	18.3	0.2	2.7	2.5	10.5	p≤0.01	9.5	.83	2
C3G_C3Sb_11	14.9	32.0	1.3	4.5	0.5	1.3	p≥0.05	n.m.	5.6	
C3S_11	41.4	107.3	0.6	24.7	0.4	-0.8	p≤0.01	n.m.	21.8	
C3XR_C3R_11‡	65.2	139.2	9.5	29.7	0.5	-0.4	p≤0.01	n.m.	6.3	
C3GR_11	71.3	163.4	19.8	26.2	0.6	0.5	p≥0.05	n.m.	18.7	
P3R_11	0.8	8.1	0.0	1.4	2.7	8.1	p≤0.01	n.m.	1.1	

4.4.2 Descriptive statistics

The probability distribution plot *p*-values performed on the concentration data indicate that the phenotypic distributions of co-eluting compounds C3G_C3Sb (2009), C3S (all three years), C3R (2009), C3XR (2009), co-eluting compounds C3XR_C3R (2010 and

2011), and P3R (all three years) deviated significantly ($p \leq 0.01$) from normal (Table 4-6, Supplemental Fig. S 4.2). The phenotypic distributions of ACY compound C3G (2010), co-eluting compounds C3G_C3Sb (2011), and C3GR (all three years) were approximately normal ($p \geq 0.05$; Table 4-6).

In 2010, fruit collected from the black raspberry female parent, 96395S1, had higher concentrations of compounds P3R, C3G, and C3XR_C3R than the red raspberry male parent, 'Latham', which had higher concentrations of compounds C3S and C3GR (Table 4-6, columns 96395S1 and 'Latham'). In 2010 and 2011, 'Latham' contained moderate amounts of C3S and C3GR, both red raspberry-specific compounds, as well as relatively low quantities of black raspberry-specific compounds C3XR_C3R and C3G, while in 2010, 96395S1 had trace amounts of red raspberry-specific C3S and C3GR along with concentrations of C3XR_C3R and C3G typical for black raspberry. Concentrations of C3GR were greatest of all individual compounds in all three years. In 2010 both parents had concentrations of C3GR less than all of the progeny, 68 of which had concentrations greater than two SD of 'Latham', which had the higher concentration of the two parents (Table 4-6). Transgressive segregants (TS) were identified for traits P3R 2010, with two TS (96395S1 9.5 mg/100 mL, 'Latham' 0.8 mg/100 mL; progeny range 0.25 to 18.3 mg/100 mL); C3S 2010, with 10 TS (96395S1 1.3 mg/100 mL, 'Latham' 34.6 mg/100 mL; progeny range 4.1 to 99.9 mg/100 mL); and C3GR 2010 with 68 TS (96395S1 3.0 mg/100 mL, 'Latham' 28.8 mg/100 mL; progeny range 32.1 to 178.8 mg/100 mL).

REML variance component analysis conducted on the mean percentage of each compound for each year suggests that the amount of variation contributed by differences between years has much less influence on compound proportion of total ACY accumulation than does genotype (Table 4-7). This observation is supported by the standard error (s.e.) of each of the components. The s.e. for component "year" is nearly the same as the variance of the component itself, whereas the s.e. for component "genotype" is 6 – 7 times less than the variation contributed by genotype.

Table 4-7: Variance components for five analyzed anthocyanin compounds for three years (2009, 2010, 2011) detected in 155 F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'. Values are based on the mean percentage of the total anthocyanin accumulation for each year.

		component	*s.e.*
% C3GR All	Year	7.74	7.97
	Genotype	42.01	6.23
	Residual	35.63	2.89
% C3G_C3Sb All	Year	1.39	1.41
	Genotype	3.28	0.49
	Residual	2.82	0.23
% C3S All	Year	4.05	4.26
	Genotype	71.84	9.45
	Residual	31.8	2.57
% C3XR_C3R All	Year	9.78	10.31
	Genotype	150.59	20.46
	Residual	82.49	6.68
% P3R All	Year	0.48	0.48
	Genotype	0.81	0.14
	Residual	1.22	0.10

Abbreviations: s.e.: standard error; C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^G-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^G-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside

Pearson's correlation coefficients calculated for each pair of compounds are shown inTable 4-8. A number of compounds correlated significantly with each other, with the strongest positive correlation ($r > 0.70$) between C3S and C3GR for all three years; the strongest negative correlation ($r = -0.55$) was between C3XR_C3R and C3S in 2010. In general, the correlation coefficients agreed between years. The results of the correlation analysis suggest that C3G (and C3G_C3Sb) is positively correlated with all other compounds, except a slightly negative correlation with C3GR 2009. Compound C3S has a strong positive correlation with C3GR and is negatively correlated with all other compounds, except a slightly positive correlation with P3R 2011. The correlations of compound C3GR generally agree with those of C3S with the exception of a positive correlation with both C3G_C3Sb and C3XR_C3R in 2011. All remaining compounds show a mix of positive and negative correlations relative to the other compounds. As some of the data distributions were non-normal, a Spearman's ranking analysis was conducted and showed similar results as the Pearson's correlations (Supplemental Table S 4-1, Supplemental Fig. S 4.3)

Table 4-8: Pearson's correlation coefficient calculations and statistical significance (I-value) for each pair of the six anthocyanin compounds analyzed for 2009 and 2010 in 155 F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'.

Trait	*C3GR 09*	*C3G_C3Sb 09*	*C3S 09*	*C3XR+C3R 09*	*C3GR 10*	*C3G 10*	*C3S 10*	*C3XR_C3R 10*	*C3GR 11*	*C3G_C3Sb 11*	*C3S 11*	*C3XR_C3R 11*
C3G_C3Sb 09	-0.05											
C3S 09	**0.78**	0.08†										
C3XR+C3R 09	**-0.30†**	**0.63**	**-0.50†**									
P3R 09	**-0.35**	0.16	**-0.36**	**0.37**								
C3G 10					0.17							
C3S 10					**0.77**	0.24†						
C3XR_C3R 10					-0.19	**0.36**	**-0.55**					
P3R 10					-0.18	0.08	*-0.25*	*0.22*				
C3G_C3Sb 11									**0.57**			
C3S 11									**0.72**	**0.43†**		
C3XR_C3R 11									**0.32**	**0.62**	-0.14	
P3R 11									0.16	0.05	0.05	-0.10

N=155

*Missing data replaced by mean value (0.849)

†These correlations changed when the data were considered as Sets A and B relative to marker RubFruitE4 on RLG2

Bold font indicates significance at $p<0.001$. *Italic* font indicates significance at $p<0.05$

Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^G-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^G-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside; RLG: *Rubus* linkage group

Scatter plots for compound associations between C3S 2009 and C3G_C3Sb 2009, C3S 2009 and C3XR+C3R 2009, C3GR 2009 and C3XR+C3R 2009, C3S 2010 and C3G 2010, as well as C3S 2011 and C3G_C3Sb 2011 show two distinct groupings. The occurrence of subgroups was investigated and an association was detected between the genotype of the progeny at locus RubFruitE4 on 'Latham' RLG2, and the compound associations that segregate into two subgroups (Fig. 4.1, Supplemental Fig. S 4.4). The two subgroups consist of 61 heterozygous ("lm"; Set A) and 94 homozygous ("ll"; Set B) individuals, where "lm" and "ll" represent allele combinations of marker RubFruitE4. Individual EA650 groups with its opposite phenotype when C3S 2009 is correlated with C3G_C3Sb 2009. Individuals EA509 ("ll") consistently groups with the phenotypes of Set A, while EA611 and EA697 (both "lm") consistently group with the phenotypes of Set B.

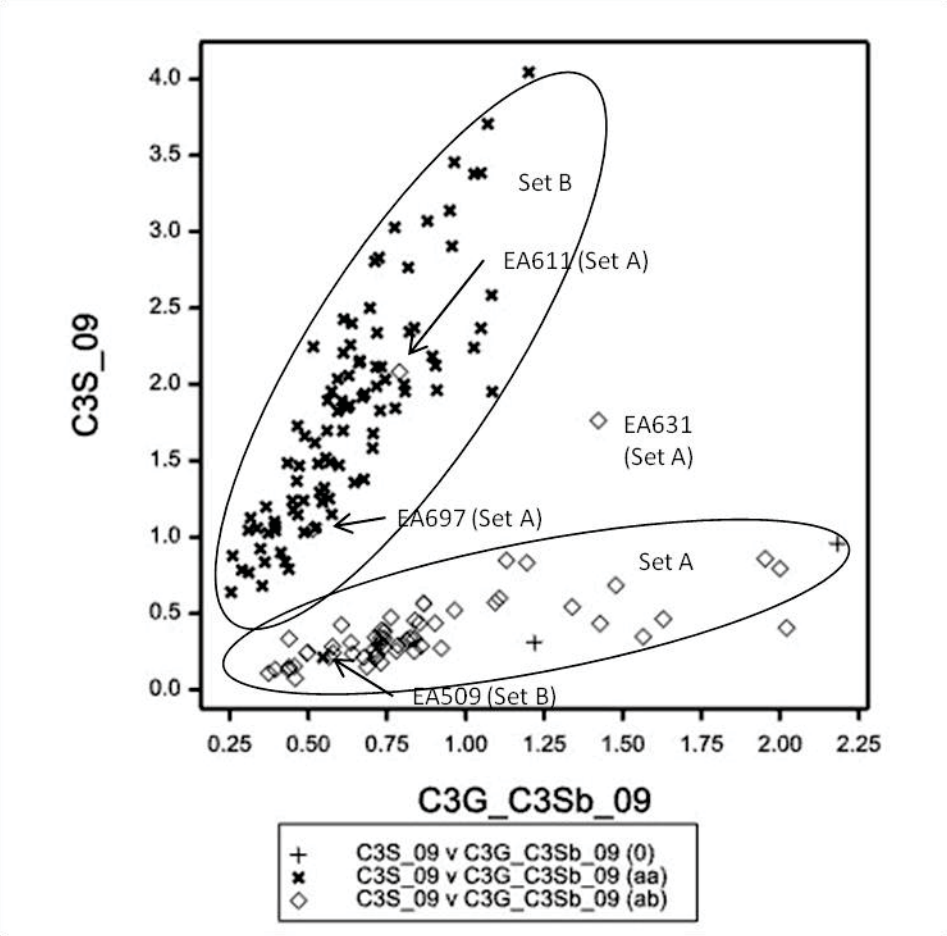


Fig. 4.1: Scatter plot for compounds cyanidin 3-O-sophoroside (C3S) and cyanidin 3O-glucoside and cyanidin 3O-sambubioside (C3G_C3Sb) in 2009. The subset of F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' segregates into two groups according to genotype at locus RubFruitE4 on 'Latham' *Rubus* linkage group (RLG) 2. Set A (open diamond symbols) consists of 61 heterozygous ("lm") individuals; Set B (solid cross symbols) consists of 94 homozygous ("ll") individuals, where "ll" and "lm" represent allele combinations. The correlation coefficient for the entire data set is r = 0.08; the correlation coefficient for Set A is r = 0.57, and the correlation coefficient for set B is r = 0.85. Arrows indicate progeny that group opposite of their genotype.

When the Pearson's correlation coefficient analysis was conducted on data Set A (Table 4-9) and Set B (Table 4-10), most of the correlations became positive, some strongly so, with r values as high as 0.85 for the correlation between C3G_C3Sb and C3S for 2009, in contrast to the finding of no significant correlation when the data were considered as a whole. The negative correlations between C3S 2009 and C3R 2009, and between C3S 2009 and C3XR 2009, when the data were considered as a whole, changed to non-significant correlations when the data were considered as separate sets.

A comparison of average, minimum and maximum values of the four compounds that formed two distinct groups based on genotype at RubFruitE4, showed that though the compound concentrations overlapped, the mean concentration of Set B ("ll") was higher for red raspberry compounds C3GR and C3S, and the mean concentration of Set A ("lm") was higher for black raspberry compounds C3XR_C3R and C3G_C3Sb. A Student *t*-test performed on the differences between the means of the two sets indicates that the differences are statistically significant ($p<0.001$) for all compounds except C3S 2010 (Table 4-11).

Table 4-9: Pearson's correlation coefficients calculated for each pair of compounds that segregated into two distinct groups in the F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'. The 61 individuals comprising Set A are heterozygous ("lm") at locus RubFruitE4 on 'Latham' *Rubus* linkage group (RLG) 2.

Trait	*C3S 2009*	*C3R 2009*	*C3XR 2009*	*C3S 2010*
C3R 2009	**.025**			
C3XR 2009	**-.073**			
C3GR 2009		**.11**	**.44**	
C3G 2010				**.81**

N=155
Bold font indicates significance at $p<0.001$.
Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^{G}-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3GR: cyanidin 3-*O*-2^{G}-glucosylrutinoside

Table 4-10: Pearson's correlation coefficients calculated for each pair of compounds that segregated into two distinct groups in the F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'. The 94 individuals comprising Set B are homozygous ("ll") at locus RubFruitE4 on 'Latham' *Rubus* linkage group (RLG) 2.

Trait	*C3S 2009*	*C3R 2009*	*C3XR 2009*	*C3S 2010*
C3R 2009	**.22**			
C3XR 2009	**.19**			
C3G_C3Sb 2009	**.85**			
C3GR 2009		**.64**	**.76**	
C3G 2010				**.64**

N=155
Bold font indicates significance at $p<0.001$.
Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^{G}-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^{G}-glucosylrutinoside

Table 4-11: Comparison of Set A and Set B for the four compounds that segregated into two distinct groups in the F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' during correlation analysis.

		Set A			*Set B*					
Found in:	*Trait*	*Ave*	*Min*	*Max*	*Ave*	*Min*	*Max*	*'Latham'*	*96395S1*	*p-value**
RR	C3GR 2009	15.3	6.6	28.3	31.0	14.7	54.7			p<.001
	C3GR 2010	56.7	32.1	119.6	98.0	46.1	178.8	28.8	3.0	p<.001
BR	C3XR 2009	14.5	5.7	27.0	7.0	2.0	17.4			p<.001
BR_RR&BR†	C3XR_C3R 2010	88.0	0.0	160.3	55.0	19.0	130.5	11.5	239.6	p<.001
RR	C3S 2009	4.1	0.7	17.6	18.2	6.4	40.4			p<.001
	C3S 2010	14.5	4.1	26.2	58.0	17.7	99.7	34.6	1.3	p<.001
RR&BR†_BR	C3G_C3Sb 2009	8.9	3.7	21.8	6.3	2.6	12.0			p<.001
	C3G 2010	15.6	6.9	27.0	14.9	4.6	26.4	8.4	46.0	0.386

*Probability associated with paired Student *t*-test for the significance of the difference between the mean values for Set A and Set B.
†C3R and C3G are found in both species but in higher concentrations in BR.
Abbreviations: Ave: average; Min: minimum; Max: maximum; RR: red raspberry only compound; BR: black raspberry only compound; C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^G-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^G-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside

4.4.3 Identification of QTL

4.4.3.1 Individual compounds

QTL were identified using the mean trait concentration values for each individual progeny from each year separately on the parental genetic maps, the genotype data set for each individual compound, and IM followed by MQM. Markers RubFruitE4 and RiF2KP were selected for use as co-factors for MQM analysis conducted on the 'Latham' parental map; marker Ri60S was selected for use as the co-factor used for MQM analysis conducted on the 96395S1 parental map. Co-location of traits with markers was detected, at LOD scores greater than the genome-wide 95% threshold, on all *Rubus* Linkage Group (RLG) except RLG4 (Supplemental Fig. S 4.5). The 2009 phenotype data taken alone showed that all five traits had significant LOD scores associated with markers on the 'Latham' genetic map, while four traits (C3S, C3XR, C3G_C3Sb, and C3GR) had significant LOD scores associated with markers on the 96395S1 genetic map. Similarly, the 2010 and 2011 phenotype data, when each was taken alone, showed that all five traits had significant LOD scores associated with markers on the 'Latham' genetic map. On the

96395S1 map, traits C3S, C3G, C3GR, and P3R for 2010 had significant LOD scores associated with markers, while in 2011 traits C3XR_C3R, C3G_C3Sb, and C3GR had significant LOD scores associated with markers.

In total, 49 QTL were detected (Table 4-12, Supplemental Fig. S 4.5), using IM followed by MQM, on all 'Latham' linkage groups except RLG4, and on 96395S1 RLG 3, 5 and 6. The 49 QTL can be summarized as 28 significant associations (21 on the 'Latham' map, and seven on the 96395S1 map), as QTL that mapped to the same interval, or with overlapping confidence intervals, in all three years are considered the same. Of the 21 QTL mapped to 'Latham', four (19%) were stable over all three years, while of the seven QTL mapped to 96395S1, two (29%) were stable in all three years; these six stable QTL are indicated in bold font on Table 4-12.

QTL were identified on seven of the nine RLG, varying from one each on 'Latham' RLG1, RLG7a and 96395S1 RLG6, to 12 on 'Latham' RLG2 (Supplemental Fig. S 4.5). For the compounds analyzed in 2010, red raspberry-specific compounds C3S and C3GR, 50% (5 of 10) of the loci displayed gene effects in the same direction as predicted by the phenotype of the parental lines, while 50% of the loci showed opposite effects (Table 4-12). For traits C3XR_C3R and C3G_C3Sb, of which C3XR and C3Sb are black raspberry-specific compounds, predicted trait direction agreed with the parental phenotypes in 5 of 7 instances (71%). Predicted trait direction for P3R was opposite of the parental phenotypes in both instances. One to six QTL have been identified for each trait, with the total phenotypic variation (R^2) explained by the QTL for each trait ranging from 0.07 (one QTL for P3R 2010 on 96395S1 RLG5) to 0.88 (four QTL for C3S 2010 on 'Latham' RLG 2, 3, 6 and 7). The largest cluster of QTL, representing all five traits over all three years, is located on 'Latham' RLG2 in association with marker RubFruitE4 and its LOD ±2 confidence interval spans approximately 32 cM.

Table 4-12: Mapping and significance information for five anthocyanin compounds mapped in *Rubus occidentalis* 96395S1 and *R. idaeus* 'Latham', % variation explained per locus, and the nearest mapped marker.

Trait	*Year*	*Parental Map*	*Original LG**	*RLG†*	*IM LOD max*	*cM IM LOD max*	*MQM LOD max‡*	*cM MQM LOD max*	*GW LOD 95% significance*	*% explanation*	*Effect§*	*Nearest marker*
C3S	**2009**	**'Latham'**	4	**2**	**25.8**	**14.1**			**2.4**	**53.5**	**13.6**	**RubFruitE4**[23]
C3S	**2010**	**'Latham'**	4	**2**	**33.4**	**15.1**			**2.5**	**64.3**	**41.3**	**RubFruitE4**[2]
C3S	**2011**	**'Latham'**	4	**2**	**8.0**	**16.1**			**2.4**	**22.6**	**23.1**	**RubFruitE4**[2]
C3S	2009	'Latham'	2	3	3.8	70.4			2.4	10.6	-5.9	RiF2KP[3]
C3S	2010	'Latham'	2	3	3.3	70.4			2.5	9.4	-15.7	RiF2KP[3]
C3S	2009	'Latham'	3	6			3.5	30.9	2.4	6.5	-3.5	Ri4Co1[2]
C3S	2010	'Latham'	3	6			4.0	16.9	2.5	4.7	-6.5	RiRosCOS2381
C3S	2009	'Latham'	1	7	4.0	0.0			2.4	11.2	5.7	RibHLH[23]
C3S	2010	'Latham'	1	7	3.3	4.6			2.5	9.2	15.5	RiGT[3]
C3S	2009	96395S1	2	3	4.1	53.3			1.9	13.6	-5.7	Ri60S
C3S	2010	96395S1	2	3	3.2	63.9			1.9	9.5	-13.8	Ri60S
C3S	2011	96395S1			NS				1.9			
C3XR+C3R[1]	**2009**	**'Latham'**	4	**2**	**26.5**	**15.1**			**2.5**	**55.3**	**-24.2**	**RubFruitE4**[2]
C3XR_C3R	**2010**	**'Latham'**	4	**2**	**18.8**	**13.0**			**2.3**	**44.8**	**-36.8**	**RubFruitE4**[2]
C3XR_C3R	**2011**	**'Latham'**	4	**2**	**5.6**	**16.1**			**2.3**	**17.5**	**-21.7**	**RubFruitE4**[2]
C3XR+C3R	2009	'Latham'	2	3			3.8	70.4	2.5	4.9	-3.3	RiF2KP
C3XR+C3R	2009	'Latham'	3	6			4.4	58.1	2.5	6.2	6.5	EMFxaACO1B
C3XR_C3R	2010	'Latham'	3	6			4.7	62.1	2.3	7.4	14.5	EMFxaACO1B
C3XR_C3R	2010	'Latham'	1	7	4.5	3.2			2.3	12.5	-20.0	RiFRK2[3]
C3XR+C3R	2009	96395S1	3	6	2.0	24.1			1.9	5.8	7.5	Riunknp2[3]
C3XR_C3R	2010	96395S1			NS				2.1			
C3XR_C3R	2011	96395S1	5	5	2.0	25.8			2.0	5.9	13.6	RiRosCOS536[3]
C3G_C3Sb	2009	'Latham'	6	1			2.8	0.0	2.2	7.2	-1.7	Rub285a[3]
C3G_C3Sb	2009	'Latham'	4	2	5.4	15.1			2.2	15.4	12.9	RubFruitE4[2]
C3G_C3Sb	**2009**	**'Latham'**	2	**3**	**3.9**	**70.4**			**2.2**	**11.0**	**68.2**	**RiF2KP**[3]
C3G	**2010**	**'Latham'**	2	**3**	**3.2**	**70.4**			**2.4**	**9.0**	**68.0**	**RiF2KP**[3]
C3G_C3Sb	**2011**	**'Latham'**	2	**3**	**2.7**	**70.4**			**2.3**	**7.9**	**68.1**	**RiF2KP**[3]
C3G_C3Sb	2009	'Latham'	3	6	3.8	10.0			2.2	12.5	7.8	RiRosCOS2381
C3G	2010	'Latham'	3	6	4.9	13.9			2.4	13.5	11.5	RiRosCOS2381[3]
C3G_C3Sb	2009	'Latham'	1	7			2.5	20.7	2.2	6.1	1.6	ERubLR_SQ07-3_C07[3]
C3G	2010	'Latham'	1	7			2.4	0.0	2.4	6.8	2.2	RibHLH[23]
C3G_C3Sb	**2009**	**96395S1**	2	**3**	**4.2**	**66.4**			**1.9**	**12.0**	**-2.3**	**Rub12a**[3]
C3G	**2010**	**96395S1**	2	**3**	**3.5**	**66.4**			**2.0**	**10.0**	**-2.5**	**Rub12a**[3]

C3G_C3Sb	**2011**	**96395S1**	2	**3**	**2.4**	**54.3**			**1.9**	**8.3**	**-2.3**	**Ri60S**
C3GR	**2009**	**'Latham'**	4	**2**	**21.1**	**14.1**			**2.4**	**46.6**	**11.7**	**RubFruitE4[2,3]**
C3GR	**2010**	**'Latham'**	4	**2**	**18.1**	**14.1**			**2.3**	**41.6**	**11.8**	**RubFruitE4[2,3]**
C3GR	**2011**	**'Latham'**	4	**2**	**3.0**	**13.0**			**2.3**	**9.1**	**10.7**	**RubFruitE4[2]**
C3GR	2009	'Latham'	2	3	5.2	69.1			2.4	15.8	66.7	RiF2KP
C3GR	2010	'Latham'	2	3	3.3	70.4			2.3	9.3	68.1	RiF2KP[3]
C3GR	2010	'Latham'	5	5			5.1	49.6	2.3	8.2	14.0	RiF3'H[3]
C3GR	2010	'Latham'	3	6	2.7	75.3			2.3	9.7	16.2	RiG6PD5
C3GR	2011	'Latham'	1	7			2.4	3.7	2.3	6.3	-11.3	RiAq[3]
C3GR	**2009**	**96395S1**	2	**3**	**6.5**	**50.3**			**1.9**	**23.6**	**-7.9**	**Ri60S**
C3GR	**2010**	**96395S1**	2	**3**	**3.8**	**66.4**			**1.9**	**10.8**	**-19.2**	**Rub12a[3]**
C3GR	**2011**	**96395S1**	2	**3**	**2.4**	**47.3**			**1.9**	**8.5**	**-10.7**	**Ri60S**
P3R	2009	'Latham'	4	2	5.7	14.1			1.8	15.6	-0.8	RubFruitE4[2]
P3R	2009	'Latham'	4	2			35.8	19.1	1.8	68.1	0.4	RiSNF4
P3R	2009	'Latham'	5	5	2.0	46.9			1.8	6.1	-0.5	RiF3'H
P3R	2010	'Latham'	5	5	3.6	53.6			2.3	10.7	-1.7	RiF3'H
P3R	2011	'Latham'			NS				2.3			
P3R	2009	96395S1			NS				1.5			
P3R	2010	96395S1	5	5	2.3	6.2			1.9	6.6	-1.4	RH_MEa0007a_B01[3]
P3R	2011	96395S1			NS				1.8			

*Linkage group (LG) numbering from Graham et al. (2004)
†Rubus linkage group (RLG) numbering from Bushakra et al. (2012a)
‡Multiple QTL Model (MQM) co-factors for traits mapped on the 'Latham' parental map were either RubFruitE4 (RLG2) or RiF2KP (RLG3). MQM co-factor for traits mapped on the 96395S1 parental map was Ri60S (RLG3).
§Estimated additive effects calculated as the μ ll – μ lm, where μ ll and μ lm are the estimated mean of the distribution of the quantitative trait associated with the homozygous or heterozygous genotypes, respectively.
Bold font indicates that the QTL was stable over all three years.
[1]Sum of C3XR and C3R compounds for 2009
[2]Marker in common with Kassim et al. (2009)
[3]Marker at peak of QTL

Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^{G}-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^{G}-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside; RLG: *Rubus* Linkage Group; IM: interval mapping; GW: Genome-wide; LOD: logarithm of odds; NS: no significant LOD; LG: linkage group; cM: centimorgan

While the composition and overall proportions of the compounds were similar among years, the concentrations of the five ACY compounds analyzed varied widely among individuals in the progeny of interspecific cross 96395S1 × 'Latham' (Table 4-6). Table 4-13 summarizes the proportions of individual ACY in the sub-set of 155 analyzed progeny of 96395S1 × 'Latham' compared with published results for red, black, and

purple raspberry. Cyanidin 3-*O*-rutinoside (C3R) and cyanidin 3-*O*-2^{G}-xylosylrutinoside (C3XR) were the major ACY compounds identified in the black raspberry cultivars analyzed by others (Dossett et al. 2010; Tulio et al. 2008) comprising 39–90% of the total ACY. The proportions of individual ACY in the sub-set of 96395S1 × 'Latham' analyzed progeny were more evenly distributed among the five identified compounds when compared with the other black raspberry cultivars (Dossett et al. 2010), but with a greater proportion of the co-eluting compounds cyanidin 3-*O*-glucoside and cyanidin 3-*O*-sambubioside (C3G_C3Sb) and smaller proportions of cyanidin 3-*O*-rutinoside (C3R), and co-eluting compounds cyanidin 3-*O*-2^{G}-xylosylrutinoside and cyanidin 3-*O*-rutinoside (C3XR_C3R). The sub-set of 96395S1 × 'Latham' analyzed progeny has smaller proportions of cyanidin 3-*O*-sophoroside (C3S) and cyanidin 3-*O*-glucoside (C3G), with greater proportions of cyanidin 3-*O*-rutinoside (C3R) and cyanidin 3-*O*-2^{G}-glucosylrutinoside (C3GR) when compared with 'Latham' × 'Glen Moy' (Kassim et al. 2009). The proportion of pelargonidin 3-*O*-rutinoside (P3R) was consistent among studies.

Table 4-13: Proportions of the total anthocyanin for each compound detected in the analyzed set of F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' in 2009 and 2010. Proportions of anthocyanins from populations of black raspberry cultivars, red raspberry 'Latham' × 'Glen Moy', and purple raspberry are provided for comparison.

Trait	*Compound proportion of total anthocyanins 2009*[*]	*Trait*	*Compound proportion of total anthocyanins 2010*[*†]	*[a]black raspberry cultivars*	*[b]'Latham' × 'Glen Moy'*	*[c]purple raspberry populations*
C3S	17%	**C3S**	20%	na	64%	14%
C3XR+C3R[1]	39%	**C3XR_C3R**	33%	32-64%	1.6%	14-33%
C3G_C3Sb	10%	**C3G**	7%	2-6% (C3Sb)	23% (C3G)	12% (C3G) 5% (C3Sb)
C3GR	33%	**C3GR**	39%	na	23%	26%
P3R	1%	**P3R**	2%	1-5%	1.3%	na

4.4.3.2 Ranking individuals

Fig. 4.2 (and Supplemental Fig. S 4.6) shows the relative accumulation of percent C3S per genotype with individuals ranked from lowest to highest percent for 2009. In general,

rankings were consistent between 2009 and 2010 (those individuals with low accumulation in 2009 also had low accumulation in 2010). Percent accumulation in 2011 was the most variable and at the extremes appears to be opposite of the percent accumulation in 2009.

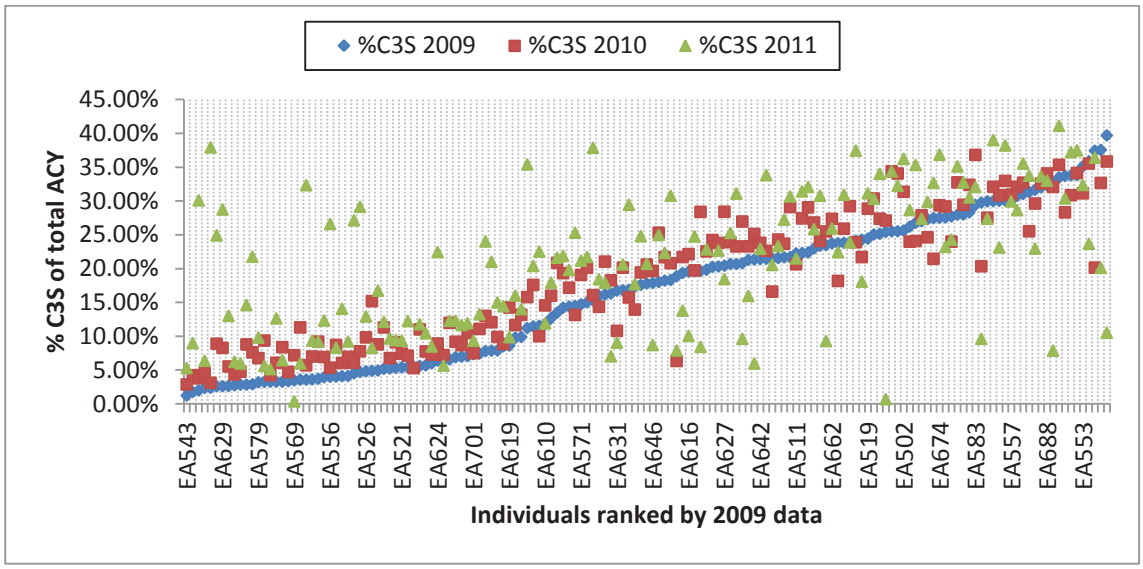


Fig. 4.2: Ranking of individuals based on the percentage of cyanidin 3-O-sophoroside (C3S) of the total anthocyanin (ACY) accumulated in 2009. This graph suggests that, generally, percent C3S accumulation in 2009 (diamonds) agreed with percent C3S accumulation in 2010 (squares). Percent C3S accumulation in 2011 was the most variable (triangles), with the low and high extremes trending to the opposite of 2009.

4.4.3.3 Total anthocyanin

QTL for total ACY were identified using the sum of the mean trait values for each individual progeny for each year separately on the 'Latham' parental genetic map, the genotype data set for total ACY, and IM. Two QTL were identified, one on 'Latham' RLG3 associated with marker RiF2KP in all three years (2009: LOD=8.3, R^2=0.22; 2010: LOD=5.1, R^2=0.14; 2011: LOD=2.6, R^2=0.08), and one on 'Latham' RLG2 associated with marker RubFruitE4 (2010 only, LOD=5.8, R^2=0.16) (Supplemental Fig. S 4.7).

4.4.4 Candidate genes co-locating with QTL

On the 'Latham' parental map, two polyphenolic biosynthetic pathway gene-derived markers (RiF3'H and RiPAL) and four transcription factor (TF)-derived markers

(RubFruitE4, RibHLH, RiMYB, RubFAC-003_MYB1) are associated with QTL on three different RLG (Table 4-12). No QTL on the 96395S1 parental map are associated with polyphenolic biosynthetic pathway gene- or TF-based markers.

4.4.4.1 Association of QTL with markers derived from polyphenolic biosynthetic pathway genes and transcription factors

Marker RiF3'H on RLG5 was designed from *flavanone 3'-hydroxylase* (*F3'H*) sequence encoding the enzyme that catalyzes an early step in flavonoid metabolism, provides precursors for many classes of flavonoid compounds, and regulates the flavonoid pathway (Pelletier and Shirley 1996). On 'Latham' RLG5, QTL for P3R (2009, 2010) and C3GR 2010 co-located with the candidate gene-based marker RiF3'H. On 'Latham' RLG6, marker Ri4Co1, designed from the gene encoding 4-coumarate:coA ligase 1 which catalyzes the formation of 4-coumaroyl-CoA during the formation of flavonoids (Kuhn et al. 1984), is associated with, or is within the LOD±1 confidence interval of, the QTL for C3S 2009. On 'Latham' RLG7a, marker RiPAL2, designed from the candidate gene *phenylalanine lyase* (*PAL*) that encodes the enzyme for biosynthesis of a wide range of phenylpropanoid products from phenylalanine (Liang et al. 1989), is within 2 cM of marker Ri_5O21, as well as within the LOD±1 confidence interval of the QTL for co-eluting C3G_C3Sb 2009.

QTL for all traits analyzed co-located with marker RubFruitE4, a putative *bZIP* TF (Kassim et al. 2009), on 'Latham' RLG2. On 'Latham' RLG7b, the marker RibHLH, designed from a TF similar to apple *MdbHLH33*, co-located with QTL for C3S (2009, 2010), co-eluting compounds C3XR_C3R 2010, C3G 2010, and C3GR 2011. Marker RiMYB, derived from a putative *MYB* TF, on 'Latham' RLG6 co-locates with, or is within the LOD±1 confidence interval of, the QTL for C3XR+C3R 2009.

4.4.4.2 Association of QTL with markers derived from additional gene sequences

On 'Latham' RLG3, QTL for traits C3GR (2009, 2010), C3G (2010) and C3G_C3Sb (2009, 2011), C3S (2009, 2010), and C3XR+C3R (2009) co-located with marker RiF2KP, designed from a regulatory enzyme-encoding gene (Draborg et al. 1999). On 96395S1 RLG3, marker Ri60S, which is similar to the 60S ribosomal protein L10 (Woodhead et al. 2010), co-located with three of the same four QTL (C3S, C3G,

C3G_C3Sb, and C3GR) mapped on 'Latham' RLG3. Finally, on 'Latham' RLG6, marker RiRosCOS2381 co-located with QTL for C3G_C3Sb 2009, C3G 2010, and C3S 2010.

4.4.5 Comparative gene analysis of the top of RLG2

The genomic region surrounding the QTL associated with marker RubFruitE4 on RLG2 was examined for positional candidate genes based on synteny with *Fragaria vesca* (alpine strawberry) (Bushakra et al. 2012a). The gene model prediction identified 409 gene models within the 4 Mb region of *Fragaria* LGII surrounding the homolog of *Rubus* marker RubFruitE4. Of these 409 gene models, 59 were considered as potential candidates for, or regulators of, genes involved in the phenylpropanoid biosynthetic pathway (Supplemental Fig. S7). Within this region, genes T51560 (probable *flavonol 3-O-glucosyltransferase* (EC 2.4.1.91) F2K13_200 [similarity] - *Arabidopsis thaliana*) and AAD17392.1 (putative *glucosyltransferase* [*Arabidopsis thaliana*]) are the most promising candidates as their function is to catalyze the addition of sugar moieties on to the cyanidin molecule.

4.5 Discussion

4.5.1 Individual anthocyanin concentrations in the F_1 progeny of 96395S1 × 'Latham'

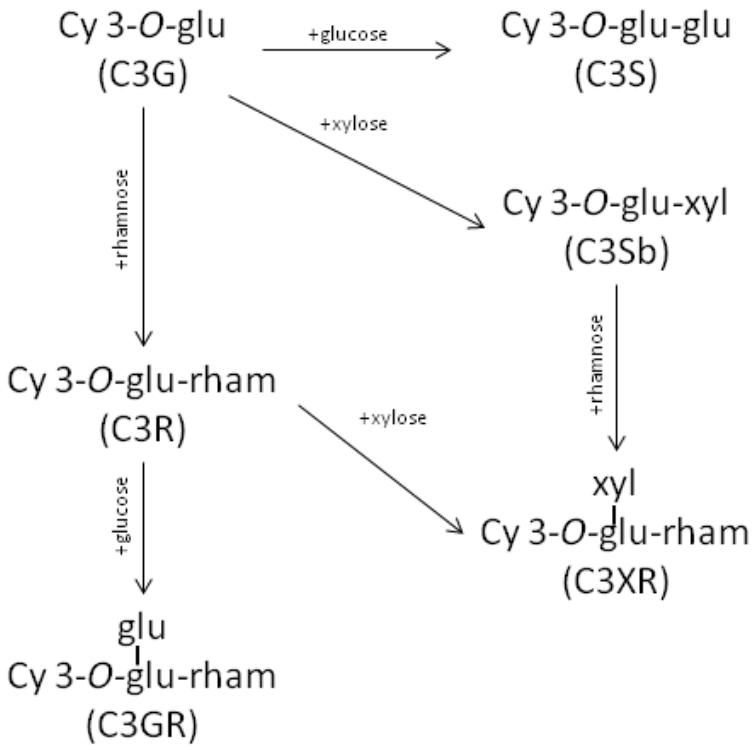
The interspecific progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' provides another genotype, other than the 'Latham' × 'Glen Moy', for analysis of the genetic and environmental controls of the ACY biosynthetic pathway and subsequent compound accumulation in the fruit of purple raspberry. The comparison of the 96395S1 × 'Latham' F_1 progeny sub-set with other black raspberry, red raspberry and purple raspberry populations suggests that the concentrations of individual ACY differ widely between individuals, although the proportions and composition of the compounds appear to be consistent within a population. This indicates that the proportions and compositions of ACY are under strong genetic control, while ACY concentration is more subject to environmental influences.

4.5.2 Pearson's correlations coefficient analyses

The positive correlation between C3G and the other compounds suggests that the enzyme-mediated addition of sugar moieties to this base molecule to produce other compounds does not deplete C3G (Fig. 4.3). The strong positive correlation between C3S and C3GR is interesting as they are both end products resulting from the enzyme-

mediated addition of glucose. This positive correlation suggests that glucose is prevalent in the fruit (Kafkas et al. 2008).

Fig. 4.3: Six cyanidin-based anthocyanin compounds found in red and black raspberry fruit. The addition of sugar moieties (above the arrows) modifies cyanidin 3-O-glucoside (C3G).



Abbreviations: C3S: cyanidin 3-O-sophoroside; C3XR: cyanidin 3-O-2G-xylosylrutinoside; C3R: cyanidin 3-O-rutinoside; C3G: cyanidin 3-O-glucoside; C3Sb: cyanidin 3-O-sambubioside; C3GR: cyanidin 3-O-2G-glucosylrutinoside

4.5.3 'Latham'*Rubus* linkage groups 2 and 7 and cyanidin modification

The comparisons conducted are based on markers in common between this and the other four *Rubus* QTL studies as presented in the introduction, however we use the linkage group nomenclature which is based on the alignment of *Rubus* with *Fragaria vesca* (Bushakra et al. 2012a). Dobson et al. (2012), conducted an analysis for the accumulation of total ACY in the 'Latham' × 'Glen Moy' progeny and identified two QTL on LG2 and LG6 (our RLG3 and RLG1, respectively). The analysis of individual compounds

conducted in the same population by Kassim et al. (2009), identified two QTL on LG4 and LG1 (our RLG2 and RLG7, respectively). Our study validates these QTL and suggests that the QTL associated with markers RubFruitE4 and RubbHLH are involved in the sugar modification of cyanidin, rather than the production of total anthocyanins.

No TF or biosynthetic gene-based marker co-located with the QTL on RLG2. Marker RubbHLH was designed from *Rubus* sequence similar to *MdbHLH33*, which encodes a component of the MYB/bHLH/WD protein complex involved in the regulation of the expression of ACY biosynthetic pathway genes and the subsequent accumulation of ACY in plant tissues (Allan et al. 2008; Baudry et al. 2004). Comparative gene analysis of RLG2 with the homologous region of the *Fragaria* whole genome sequence (Shulaev et al. 2011) identified two genes that may be responsible for the addition of sugars to the cyanidin molecule: a putative flavonol 3-*O*-glucosyltransferase (F3OGT) involved in the glycosylation of kaempferol (Tohge et al. 2005), and a putative flavonol 7-*O*-glucosyltransferase (F7OGT), which catalyzes the transfer of a glucosyl group from UDP-glucose to the 7-hydroxy group of a quercetin molecule (Lim et al. 2004). These two genes, along with *Rubus bHLH*, would be candidates for gene expression analysis in tobacco (Espley et al. 2007), and *in planta* gene expression in *Rubus* progeny exhibiting the extremes of specific ACY compound composition.

4.5.4 Comparison of QTL identified in the progeny of 96395S1 × 'Latham' with those identified previously in the progeny of 'Latham' × 'Glen Moy'

The QTL for total ACY accumulation in all three years that is located on 'Latham' RLG3 and associated with marker RiF2KP may be the same as the QTL identified by Dobson et al. (2012) (LG2 in their nomenclature). The QTL for total ACY accumulation in 2010 located on 'Latham' RLG2 and associated with marker RubFruitE4 was not identified by Dobson and colleagues and in the current study has a much lower LOD score and explains a much lower percentage of variation than the QTL for individual compounds. This finding could indicate that the large-effect QTL associated with marker RubFruitE4 on 'Latham' RLG2 may be associated with the modification of the base cyanidin molecule, rather than production of the cyanidin molecule, supporting F3OGT and F7OGT as candidates for enzyme-mediated sugar addition.

QTL for concentrations of eight ACY compounds [cyanidin 3-*O*-sophoroside (C3S), cyanidin 3-*O*-2^{G}-glucosylrutinoside (C3GR), cyanidin 3-*O*-glucoside (C3G), cyanidin 3-

O-rutinoside (C3R), pelargonidin 3-*O*-rutinoside (P3R), pelargonidin 3-*O*-sophoroside (P3S), pelargonidin 3-*O*-glucoside (P3G) and pelargonidin 3-*O*-glucosylrutinoside (P3GR)] were previously mapped in association with marker RubFruit E4 and marker bHLH on the linkage map of 'Latham' × 'Glen Moy' (Kassim et al. 2009). QTL for five (C3S, C3GR, C3G, C3R, and P3R) of these eight compounds have been identified in this study and are also associated with the locus RubFruitE4. Neither C3Sb nor C3XR were reported as being detected in the progeny of 'Latham' × 'Glen Moy'. In the current study, the greatest amount of variation in expression of all six traits (14.9–64.3%) is explained by locus RubFruitE4 on 'Latham' RLG2 (Table 10). This locus on the 'Latham' × 'Glen Moy' linkage map explains less (between 10% and 40%) of the variation in concentration of the eight compounds (Kassim et al. 2009). The difference between these two studies in percent variation explained by this locus could result from the presence of black raspberry-specific xylose-containing compounds in the progeny of 96395S1 × 'Latham', as the 'Latham' × 'Glen Moy' is lacking in these compounds. The enzyme-mediated addition of another sugar moiety to the base cyanidin molecule is hence supported by this finding.

In the current study, markers RibHLH and RiFRK2 map within 3.2 cM of each other on 'Latham' RLG7b and are associated with 6.8–11.2% of the variation in expression of C3S 2009 and C3G 2010 in the set of 96395S1 × 'Latham' progeny analyzed. Locus bHLH on the 'Latham' × 'Glen Moy' linkage map explains 35–60% of the variation in the expression of the eight compounds detected in the progeny of 'Latham' × 'Glen Moy' by Kassim et al. (2009). Compound C3XR was not reported as being detected in the progeny of 'Latham' × 'Glen Moy' by Kassim et al. (2009). The difference between these two studies in percent variation explained by this locus could be the result of the more consistent proportions of compounds in the progeny of 96395S1 × 'Latham' compared to 'Latham' × 'Glen Moy'. For example, excluding P3R, the range of percent compounds in the progeny of 96395S1 × 'Latham' is 10 – 39%, whereas for the progeny of 'Latham' × 'Glen Moy' the range is 1.6 – 64% (Table 4-13). This difference could be a reflection of the population sizes used for these two studies (155 vs. 320).

On 'Latham' RLG6, the QTL for C3XR+C3R 2009, C3XR_C3R 2010, and C3GR 2010 are all associated with, or within the LOD±2 confidence interval of, marker RiMYB, which was designed from *Rubus* sequence homologous to TF *MdMYB10* (Lin-Wang et al.

2010) which, like *bHLH*, encodes a component of the MYB/bHLH/WD protein complex (Baudry et al. 2004). *MdMYB10* was isolated from red-fleshed apple and is responsible for controlling ACY biosynthesis in apple flesh (Espley et al. 2007). Additionally, *PavMYB10* from sweet cherry (*Prunus avium*) is located within the confidence interval for a QTL associated with fruit color on the 'New York 54' × 'Emperor Francis' consensus linkage map (Sooriyapathirana et al. 2010). RiMYB maps within 7 cM of candidate gene marker Ri4Co1 and within 5 cM of marker Rub2a1, which is shared between 'Latham' RLG6 and 'Latham' × 'Glen Moy' LG3. Marker Rub2a1 falls within the 95% confidence interval of the QTL for total ACY identified by McCallum et al. (2010) and within 8 cM of Ri4CL1, which is derived from the same sequence as marker Ri4Co1 on 'Latham' RLG6.

QTL have been mapped to neither 'Latham' nor 96395S1 RLG4, nor its equivalent in the other *Rubus* studies. In summary, each linkage group, except RLG4, appears to carry one to several loci involved in the control of various aspects of fruit quality in *Rubus* species. The markers that are associated with stable QTL (Table 4-12), for example RubFruitE4, RibHLH, RiRosCOS2381 and RiF3'H, would be good candidates for marker assisted breeding (MAB) applications.

4.6 Conclusions

The proportion and composition of ACY compounds in the fruit of 96395S1 × 'Latham' progeny suggests a strong genetic component for control. The accumulation of total ACY and the accumulation of individual compounds appear to be controlled by different regions of the genome, with QTL on RLG2 and RLG7 important in the accumulation of individual compounds and QTL important for the accumulation of total ACY located on RLG2 and RLG3. The variation in ACY concentrations exhibited in the 96395S1 × 'Latham' progeny suggests an environmental influence as the individual ACY concentrations within genotypes varies from year to year while the ACY proportions stay consistent. The differences in the production of cyanidin 3-*O*-2^{G}-glucosylrutinoside (C3GR), as demonstrated by the division of the analyzed progeny into two sets of individuals, indicates the need for closer examination of the genetic control of that trait. Additional comparative gene analysis of the chromosome regions of interest, as well as analysis of the forthcoming *Rubus idaeus* 'Heritage' genome sequence and gene

annotation, will help to identify candidate genes that may encode the enzymes necessary to catalyze the reactions that produce individual ACY compounds.

4.7 Authors' contributions

JMB constructed the QTL map, performed the MapQTL analyses, interpreted the data and wrote the manuscript. CK and TM extracted the compounds and performed the UHPLC analyses for 2009. DD and DS extracted the compounds and performed the UHPLC analyses for 2010. MJS collected the fruit samples and cared for the population. VVS, DC, EJB, SEG provided guidance in experimental design and manuscript editing.

4.8 Acknowledgements

This work was completed as part of a Ph.D. thesis and supported in part by the PFR Excellence Program (JMB) and the "New Berries" FRST Programme (CO6XO807) (EJB). The authors would like to thank Andrew McLachlan for help with the statistics and Peter Alspach for help with the epistasis analysis.

REFERENCES

Allan AC, Hellens RP, Laing WA (2008) MYB transcription factors that colour our fruit. Trends in Plant Science 13 (3):99-102

Ash MM, Wolford KA, Carden TJ, Hwang KT, Carr TP (2011) Unrefined and refined black raspberry seed oils significantly lower triglycerides and moderately affect cholesterol metabolism in male Syrian hamsters. J Med Food 14 (9):1032-1038. doi:10.1089/jmf.2010.0181

Baudry A, Heim M, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L (2004) TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant Journal 39 (3):366 - 380

Bushakra J, Stephens M, Atmadjaja A, Lewers K, Symonds V, Udall J, Chagné D, Buck E, Gardiner S (2012a) Construction of black (*Rubus occidentalis*) and red (*R. idaeus*) raspberry linkage maps and their comparison to the genomes of strawberry, apple, and peach. Theoretical and Applied Genetics:1-17. doi:10.1007/s00122-012-1835-5

Cabrera A, Kozik A, Howad W, Arús P, Iezzoni A, van der Knaap E (2009) Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. BMC Genomics 10 (1):562

Chen HS, Liu M, Shi LJ, Zhao JL, Zhang CP, Lin LQ, Liu Y, Zhang SJ, Jin JC, Wang L, Shen BZ, Liu JR (2011) Effects of raspberry phytochemical extract on cell proliferation, apoptosis, and serum proteomics in a rat model. Journal of Food Science 76 (8):T192-T198. doi:10.1111/j.1750-3841.2011.02373.x

Connor AM, Finn CE, McGhie TK, Alspach PA (2005) Genetic and environmental variation in anthocyanins and their relationship to antioxidant activity in blackberry and hybridberry cultivars. Journal of the American Society for Horticultural Science 130 (5):680-687

Cooney JM, Jensen DJ, McGhie TK (2004) LC-MS identification of anthocyanins in boysenberry extract and anthocyanin metabolites in human urine following dosing. Journal of the Science of Food and Agriculture 84 (3):237-245. doi:10.1002/jsfa.1645

Dobson P, Graham J, Stewart D, Brennan R, Hackett CA, McDougall GJ (2012) Over-seasons analysis of quantitative trait loci affecting phenolic content and antioxidant capacity in raspberry. Journal of Agricultural and Food Chemistry 60 (21):5360-5366. doi:10.1021/jf3005178

Dossett M, Lee J, Finn CE (2010) Variation in anthocyanins and total phenolics of black raspberry populations. Journal of Functional Foods 2 (4):292-297. doi:10.1016/j.jff.2010.10.004

Dossett M, Lee J, Finn CE (2011) Characterization of a novel anthocyanin profile in wild black raspberry mutants: An opportunity for studying the genetic control of pigment and color. Journal of Functional Foods 3 (3):207-214. doi:10.1016/j.jff.2011.04.003

Draborg H, Villadsen D, Hamborg Nielsen T (1999) Cloning, characterization and expression of a bifunctional fructose-6-phosphate, 2-kinase/fructose-2,6-bisphosphatase from potato. Plant Molecular Biology 39 (4):709-720. doi:10.1023/a:1006102412693

Espín JC, García-Conesa MT, Tomás-Barberán FA (2007) Nutraceuticals: Facts and fiction. Phytochemistry 68 (22-24):2986-3008. doi:10.1016/j.phytochem.2007.09.014

Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC (2007) Red colouration in apple fruit is due to the activity of the *MYB* transcription factor, *MdMYB10*. Plant Journal 49 (3):414-427

Finley JW, Kong A-N, Hintze KJ, Jeffery EH, Ji LL, Lei XG (2011) Antioxidants in foods: State of the science important to the food industry. Journal of Agricultural and Food Chemistry:6837-6846. doi:10.1021/jf2013875

Graham J, Hackett C, Smith K, Woodhead M, Hein I, McCallum S (2009) Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit. Theoretical and Applied Genetics 118 (6):1143-1155. doi:10.1007/s00122-009-0969-6

Graham J, Smith K, MacKenzie K, Jorgenson L, Hackett C, Powell W (2004) The construction of a genetic linkage map of red raspberry (*Rubus idaeus* subsp. *idaeus*) based on AFLPs, genomic-SSR and EST-SSR markers. Theoretical and Applied Genetics 109 (4):740-749

Grotewold E (2006) The genetics and biochemistry of floral pigments. Annual Review of Plant Biology 57:761-780. doi:10.1146/annurev.arplant.57.032905.105248

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7 (7):1071-1083. doi:10.1105/tpc.7.7.1071

Hummer KE (2010) Rubus pharmacology: Antiquity to the present. HortScience 45 (11):1587-1591

Jennings DL, Carmichael E (1980) Anthocyanin variation in the genus *Rubus*. New Phytologist 84 (3):505-513

Kafkas E, Özgen M, Özoğui Y, Türemiş N (2008) Phytochemical and fatty acid profile of selected red raspberry cultivars: a comparative study. Journal of Food Quality 31 (1):67-78. doi:10.1111/j.1745-4557.2007.00184.x

Kassim A, Poette J, Paterson A, Zait D, McCallum S, Woodhead M, Smith K, Hackett C, Graham J (2009) Environmental and seasonal influences on red raspberry anthocyanin antioxidant contents and identification of quantitative traits loci (QTL). Molecular Nutrition & Food Research 53 (5):625-634

Keniry A, Hopkins CJ, Jewell E, Morrison B, Spangenberg GC, Edwards D, Batley J (2006) Identification and characterization of simple sequence repeat (SSR) markers from Fragaria×ananassa expressed sequences. Molecular Ecology Notes 6 (2):319-322. doi:10.1111/j.1471-8286.2005.01215.x

Kuhn DN, Chappell J, Boudet A, Hahlbrock K (1984) Induction of phenylalanine ammonia-lyase and 4-coumarate:CoA ligase mRNAs in cultured plant cells by UV light or fungal elicitor. Proceedings of the National Academy of Sciences 81 (4):1102-1106

Liang XW, Dron M, Cramer CL, Dixon RA, Lamb CJ (1989) Differential regulation of phenylalanine ammonia-lyase genes during plant development and by environmental cues. Journal of Biological Chemistry 264 (24):14486-14492

Lim E-K, Ashford DA, Hou B, Jackson RG, Bowles DJ (2004) Arabidopsis glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. Biotechnology and Bioengineering 87 (5):623-631. doi:10.1002/bit.20154

Lin-Wang K, Bolitho K, Grafton K, Kortstee A, Karunairetnam S, McGhie T, Espley R, Hellens R, Allan A (2010) An R2R3 MYB transcription factor associated with

regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC Plant Biology 10 (1):50

Määttä-Riihinen KR, Kamal-Eldin A, Törrönen AR (2004) Identification and quantification of phenolic compounds in berries of *Fragaria* and *Rubus* species (Family Rosaceae). Journal of Agricultural and Food Chemistry 52 (20):6178-6187. doi:10.1021/jf049450r

McCallum S, Woodhead M, Hackett C, Kassim A, Paterson A, Graham J (2010) Genetic and environmental effects influencing fruit colour and QTL analysis in raspberry. Theoretical and Applied Genetics 121 (4):611-627. doi:10.1007/s00122-010-1334-5

McGhie TK, Hall HK, Ainge GD, Mowat AD (2002) Breeding *Rubus* cultivars for high anthocyanin content and high antioxidant capacity. Acta Horticulturae (ISHS) 585:495-500

Montrose DC, Horelik NA, Madigan JP, Stoner GD, Wang L-S, Bruno RS, Park HJ, Giardina C, Rosenberg DW (2011) Anti-inflammatory effects of freeze-dried black raspberry powder in ulcerative colitis. Carcinogenesis 32 (3):343-350. doi:10.1093/carcin/bgq248

Mullen W, Yokota T, Lean MEJ, Crozier A (2003) Analysis of ellagitannins and conjugates of ellagic acid and quercetin in raspberry fruit by LC-MS^n. Phytochemistry 64:617-624

Ozgen M, Wyzgoski FJ, Tulio AZ, Gazula A, Miller AR, Scheerens JC, Reese RN, Wright SR (2008) Antioxidant capacity and phenolic antioxidants of midwestern black raspberries grown for direct markets are influenced by production site. HortScience 43 (7):2039-2047

Pelletier MK, Shirley BW (1996) Analysis of flavanone 3-hydroxylase in *Arabidopsis* seedlings (coordinate regulation with chalcone synthase and chalcone isomerase). Plant Physiology 111 (1):339-345. doi:10.1104/pp.111.1.339

Sargent D, Rys A, Nier S, Simpson D, Tobutt K (2007a) The development and mapping of functional markers in *Fragaria* and their transferability and potential for mapping in other genera. Theoretical and Applied Genetics 114 (2):373-384

Scalzo J, Currie A, Stephens J, McGhie T, Alspach P (2008) The anthocyanin composition of different *Vaccinium*, *Ribes*, and *Rubus* genotypes. BioFactors 34:13-21

Seeram NP, Momin RA, Nair MG, Bourquin LD (2001) Cyclooxygenase inhibitory and antioxidant cyanidin glycosides in cherries and berries. Phytomedicine 8 (5):362-369. doi:10.1078/0944-7113-00053

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton J-M, Rees DJG, Williams KP, Holt SH, Rojas JJR, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman T-L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Girona EL, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arús P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). Nature Genetics 43 (2):109-116. doi:http://www.nature.com/ng/journal/v43/n2/abs/ng.740.html#supplementary-information

Sooriyapathirana S, Khan A, Sebolt A, Wang D, Bushakra J, Lin-Wang K, Allan A, Gardiner S, Chagné D, Iezzoni A (2010) QTL analysis and candidate gene mapping for skin and flesh color in sweet cherry fruit (Prunus avium L.). Tree Genetics & Genomes 6 (6):821-832

Takos AM, Ubi BE, Robinson SP, Walker AR (2006) Condensed tannin biosynthesis genes are regulated separately from other flavonoid biosynthesis genes in apple fruit skin. Plant Science 170:487 - 499

Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant Journal 54 (4):733 - 749

Thompson MM (1995b) Chromosome numbers of *Rubus* species at the National Clonal Germplasm Repository. HortScience 30 (7):1447-1452

Tohge T, Nishiyama Y, Hirai M, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe D, Kitayama M, Noji M, Yamazaki M, Saito K (2005) Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants over-expressing an MYB transcription factor. Plant Journal 42 (2):218 - 235

Torre LC, Barritt BH (1977) Quantitative evaluation of *Rubus* fruit anthocyanin pigments. Journal of Food Science 42 (2):488-490. doi:10.1111/j.1365-2621.1977.tb01528.x

Tulio AZ, Reese RN, Wyzgoski FJ, Rinaldi PL, Fu R, Scheerens JC, Miller AR (2008) Cyanidin 3-rutinoside and cyanidin 3-xylosylrutinoside as primary phenolic antioxidants in black raspberry. Journal of Agricultural and Food Chemistry 56 (6):1880-1888. doi:10.1021/jf072313k

van Ooijen JW (1992) Accuracy of mapping quantitative trait loci in autogamous species. Theoretical and Applied Genetics 84 (7):803-811. doi:10.1007/bf00227388

van Ooijen JW (2004) MapQTL® 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma, B.V., Wageningen, Netherlands

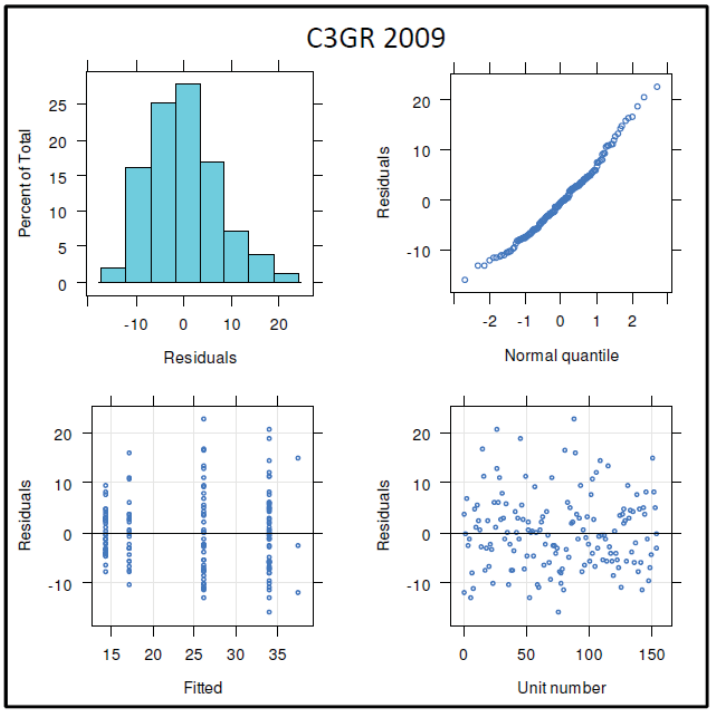
Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nature Genetics 42 (10):833-839. doi:http://www.nature.com/ng/journal/v42/n10/abs/ng.654.html#supplementary-information

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93 (1):77-78

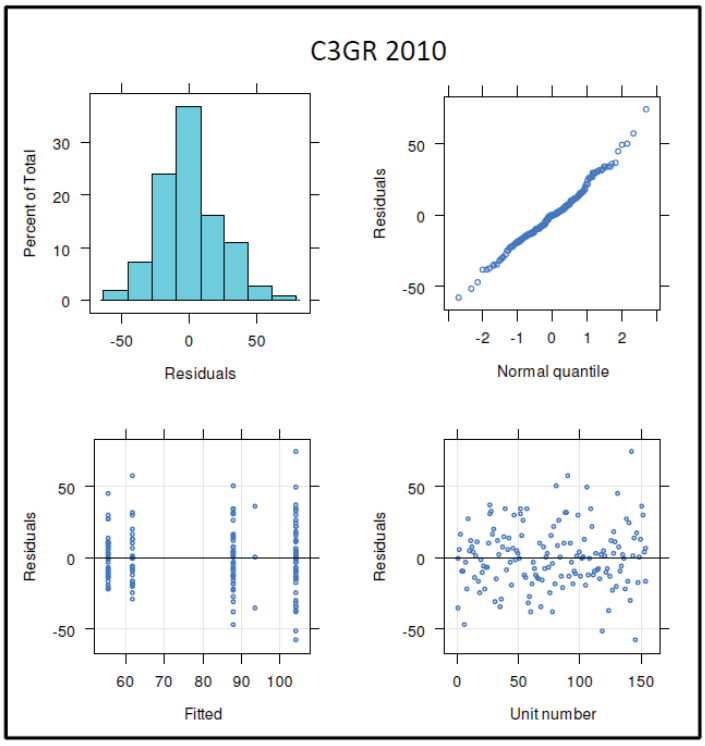
VSN International L (2010) GenStat 12.2.0.3717 edn.,

Woodhead M, Weir A, Smith K, McCallum S, MacKenzie K, Graham J (2010) Functional markers for red raspberry. Journal of the American Society for Horticultural Science 135 (5):418-427

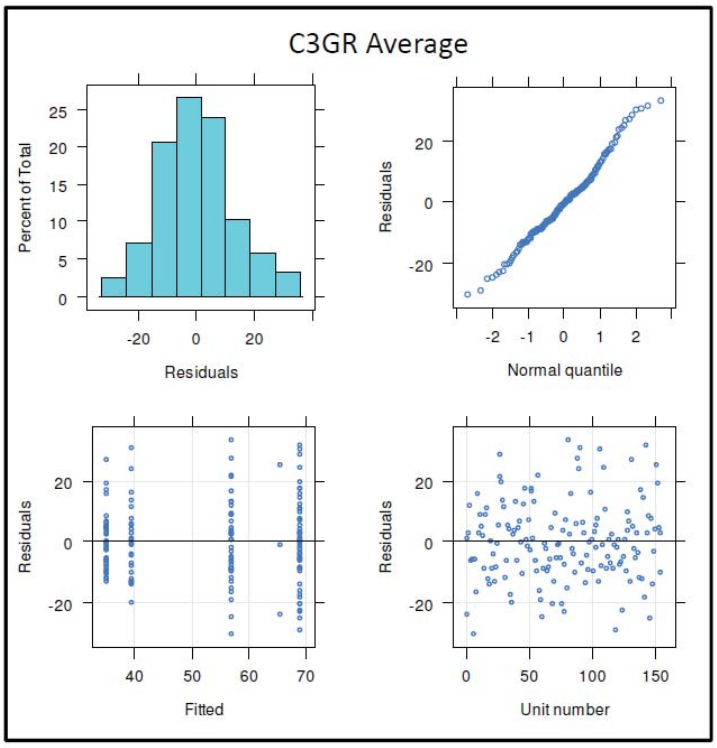
4.9 Supplemental files



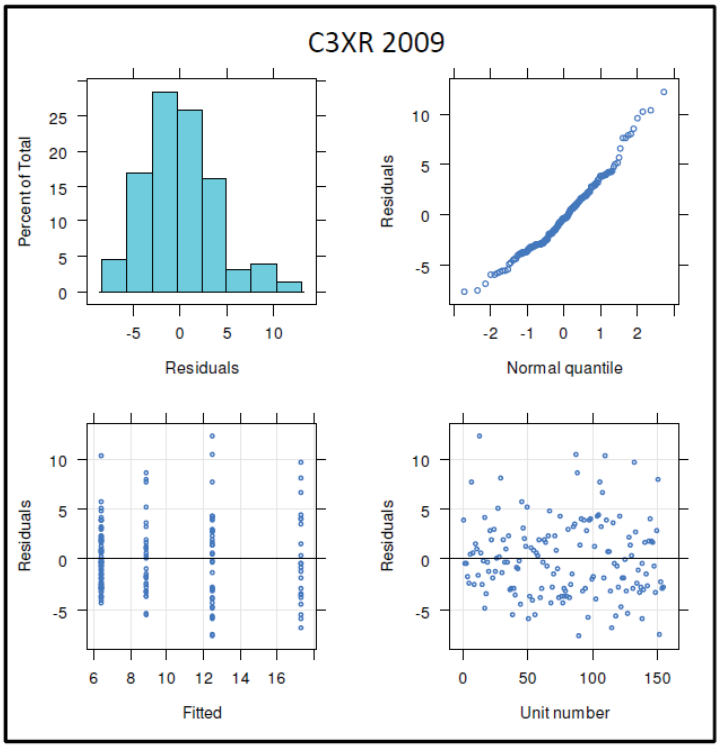
Supplemental Fig. S1a



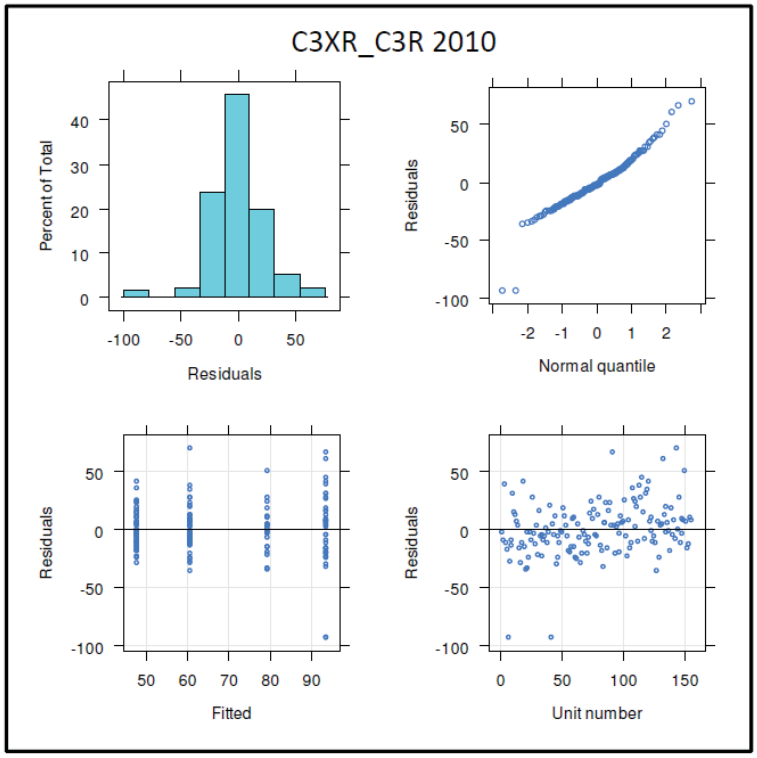
Supplemental Fig. S1b



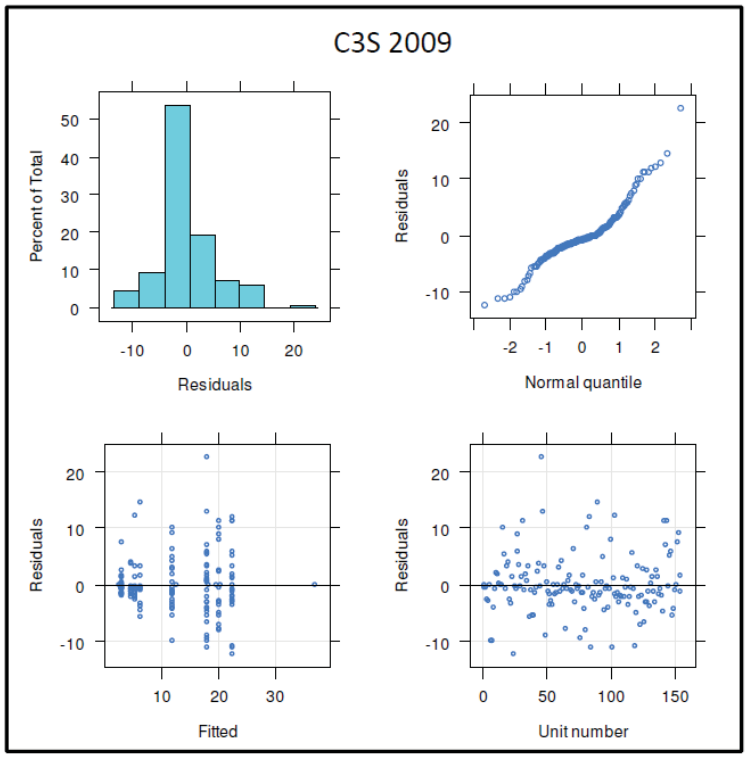
Supplemental Fig. S1c



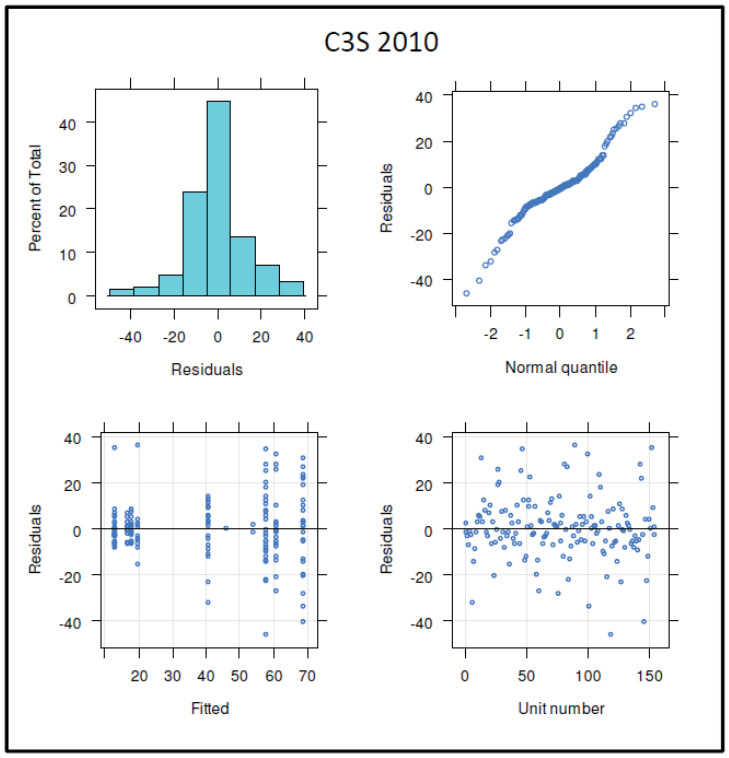
Supplemental Fig. S1d



Supplemental Fig. S1e



Supplemental Fig. S1f



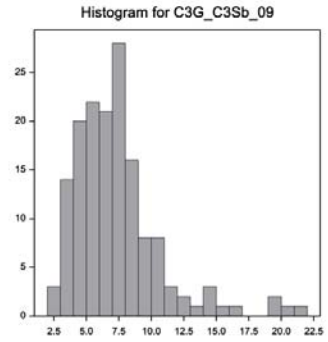
Supplemental Fig. S1g

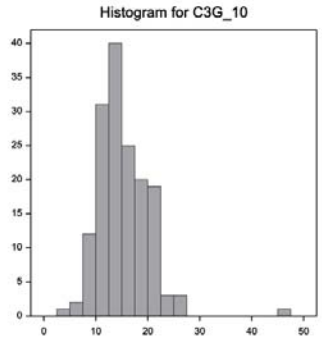
Supplemental Fig. S4.1: Residual plots from simple fixed effects analysis for traits for which epistasis analyses were conducted.

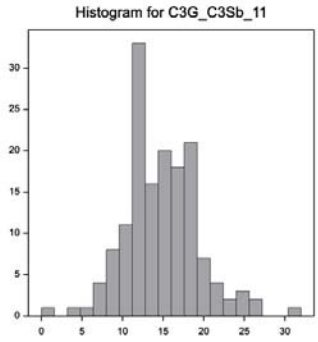
a) C

Supplemental files 3-*O*-2^{G}-glucosylrutinoside 2009
b) C3GR 2010 = cyanidin 3-*O*-2^{G}-glucosylrutinoside 2010
c) C3GR Average = cyanidin 3-*O*-2^{G}-glucosylrutinoside averaged over both years
d) C3XR 2009 = cyanidin 3-*O*-2^{G}-xylosylrutinoside 2009
e) C3XR_C3R 2010 = cyanidin 3-*O*-2^{G}-xylosylrutinoside & cyanidin 3-*O*-rutinoside compounds co-eluted in 2010
f) C3S 2009 = cyanidin 3-*O*-sophoroside 2009
g) C3S 2010 = cyanidin 3-*O*-sophoroside 2010

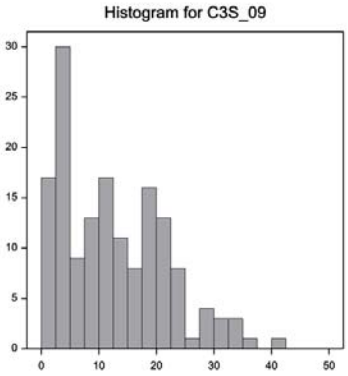
Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^{G}-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3GR: cyanidin 3-*O*-2^{G}-glucosylrutinoside.

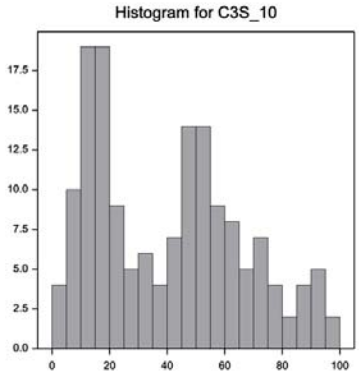


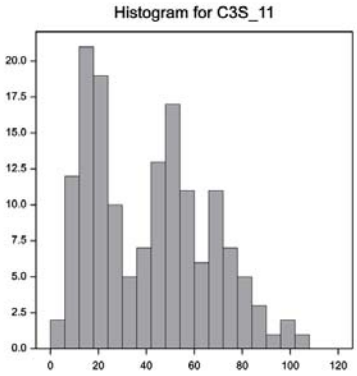




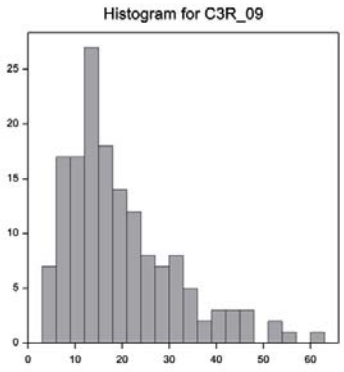
	C3G_C3Sb_09 µg/mL	C3G_10 mg/100 mL	C3G_C3Sb_11 mg/100 mL
Mean (no parents)	7.3	15.1	14.9
Max (no parents)	21.8	27.0	32.0
Min (no parents)	2.6	4.6	1.3
sd (no parents)	3.38	4.31	4.45
skewness	1.83	0.34	0.51
kurtosis	4.60	-0.26	1.35

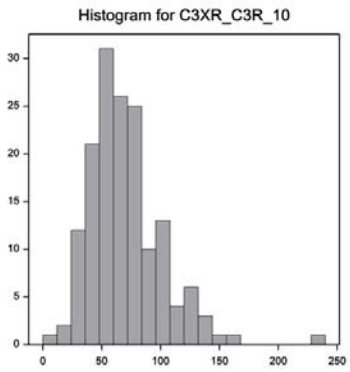


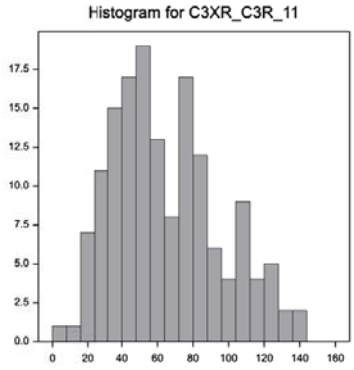


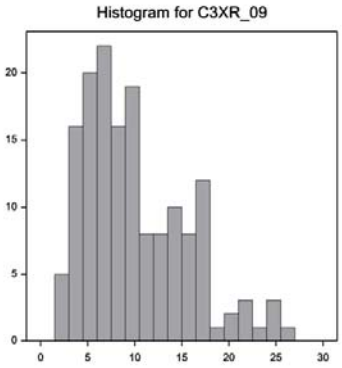


	C3S 2009 µg/mL	C3S 2010 mg/100 mL	C3S2011 mg/100 mL
Mean (no parents)	12.6	40.9	41.4
Max (no parents)	40.4	99.7	107.3
Min (no parents)	0.7	4.1	0.6
sd (no parents)	9.14	25.66	24.65
skewness	0.67	0.36	0.37
kurtosis	-0.26	-0.94	-0.77

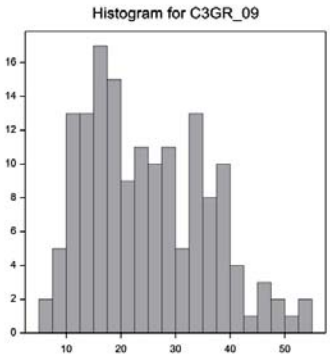


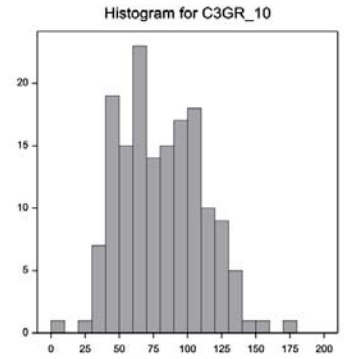


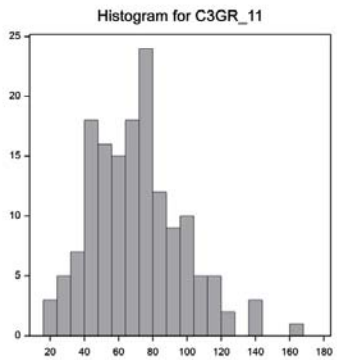




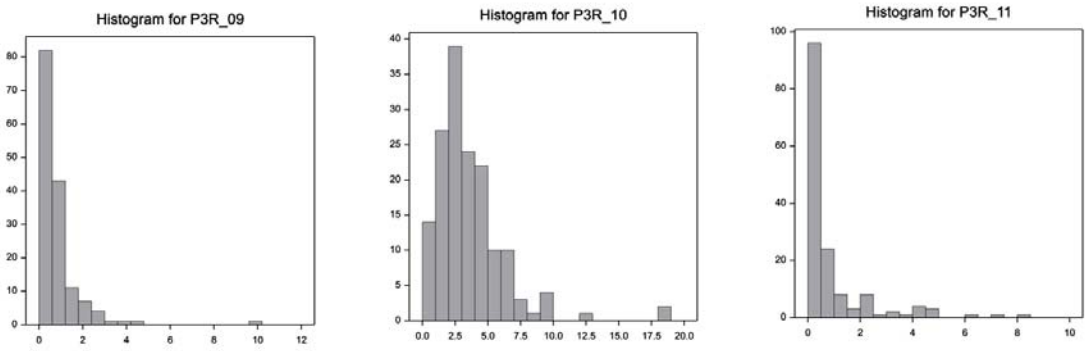
	C3R_09 µg/mL	C3XR_09 µg/mL	C3XR_C3R_10 mg/100 mL	C3XR_C3R_11 mg/100 mL
Mean (no parents)	19.7	10.0	69.5	65.2
Max (no parents)	61.5	27.0	160.3	139.2
Min (no parents)	4.6	2.0	19.0	9.5
sd (no parents)	11.65	5.43	28.01	29.73
skewness	1.20	0.88	0.79	0.51
kurtosis	1.20	0.32	0.47	-0.45







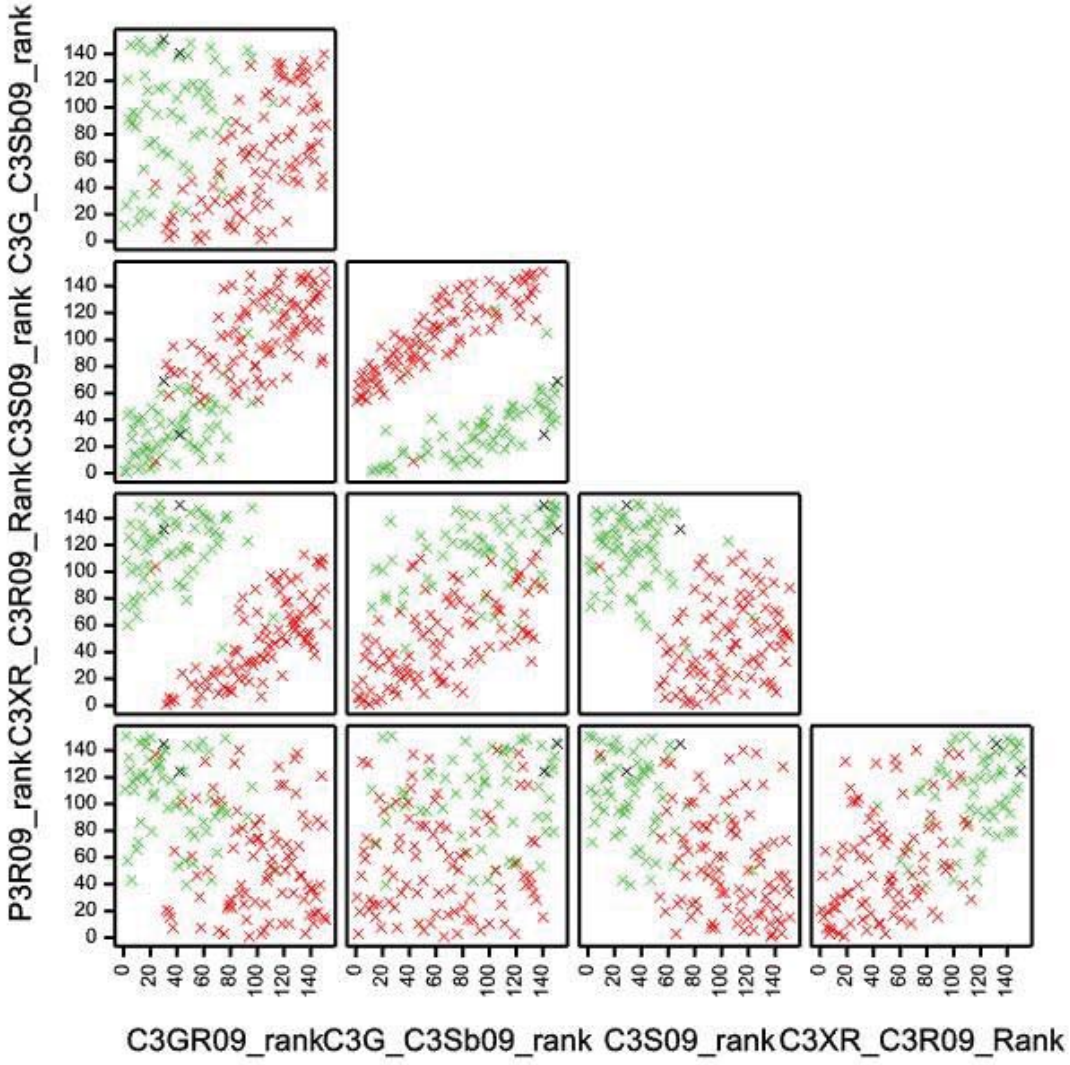
	C3GR_09 µg/mL	C3GR_10 mg/100 mL	C3GR_11 mg/100 mL
Mean (no parents)	24.8	81.8	71.3
Max (no parents)	54.7	178.8	163.4
Min (no parents)	6.6	32.1	19.8
sd (no parents)	11.00	29.62	26.21
skewness	0.50	0.42	0.63
kurtosis	-0.51	-0.32	0.50

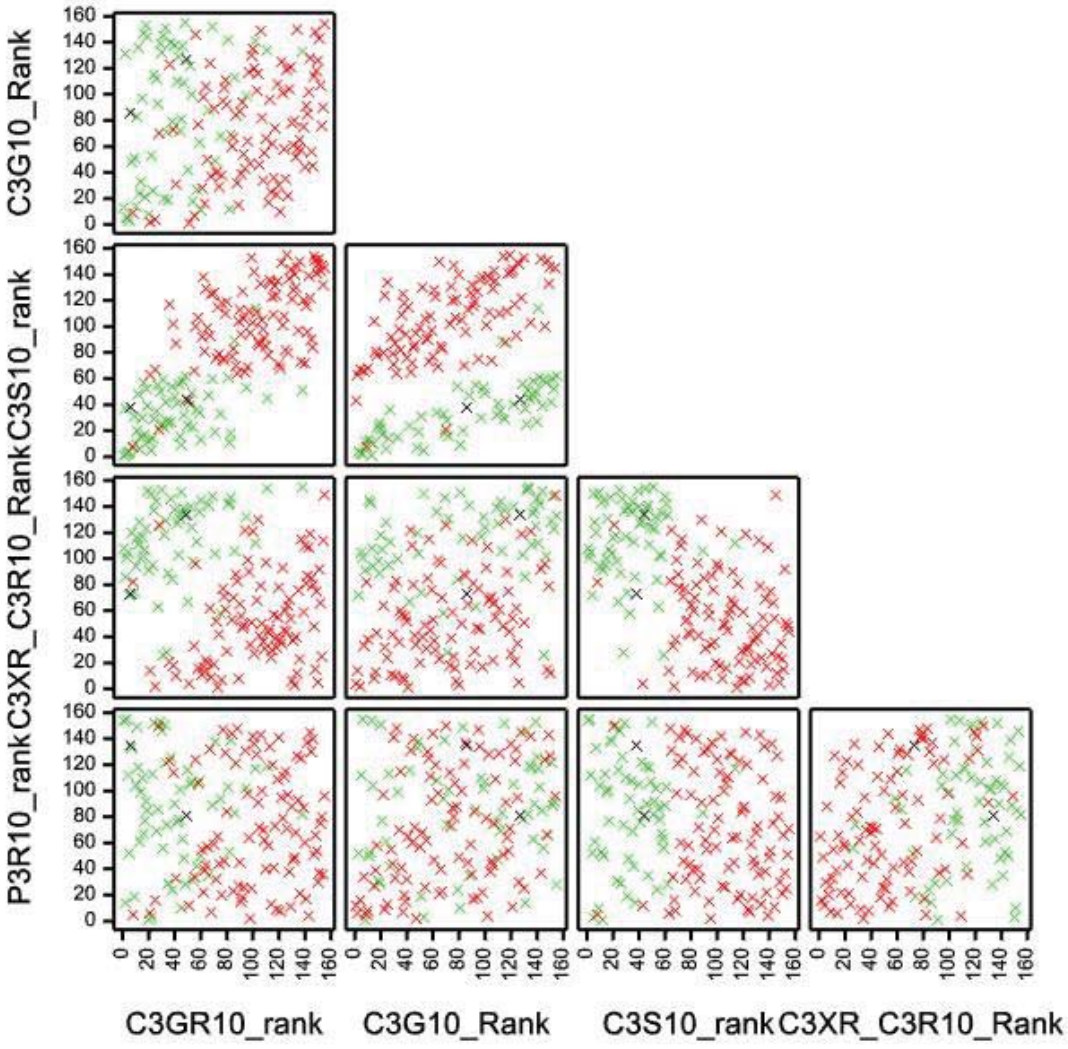


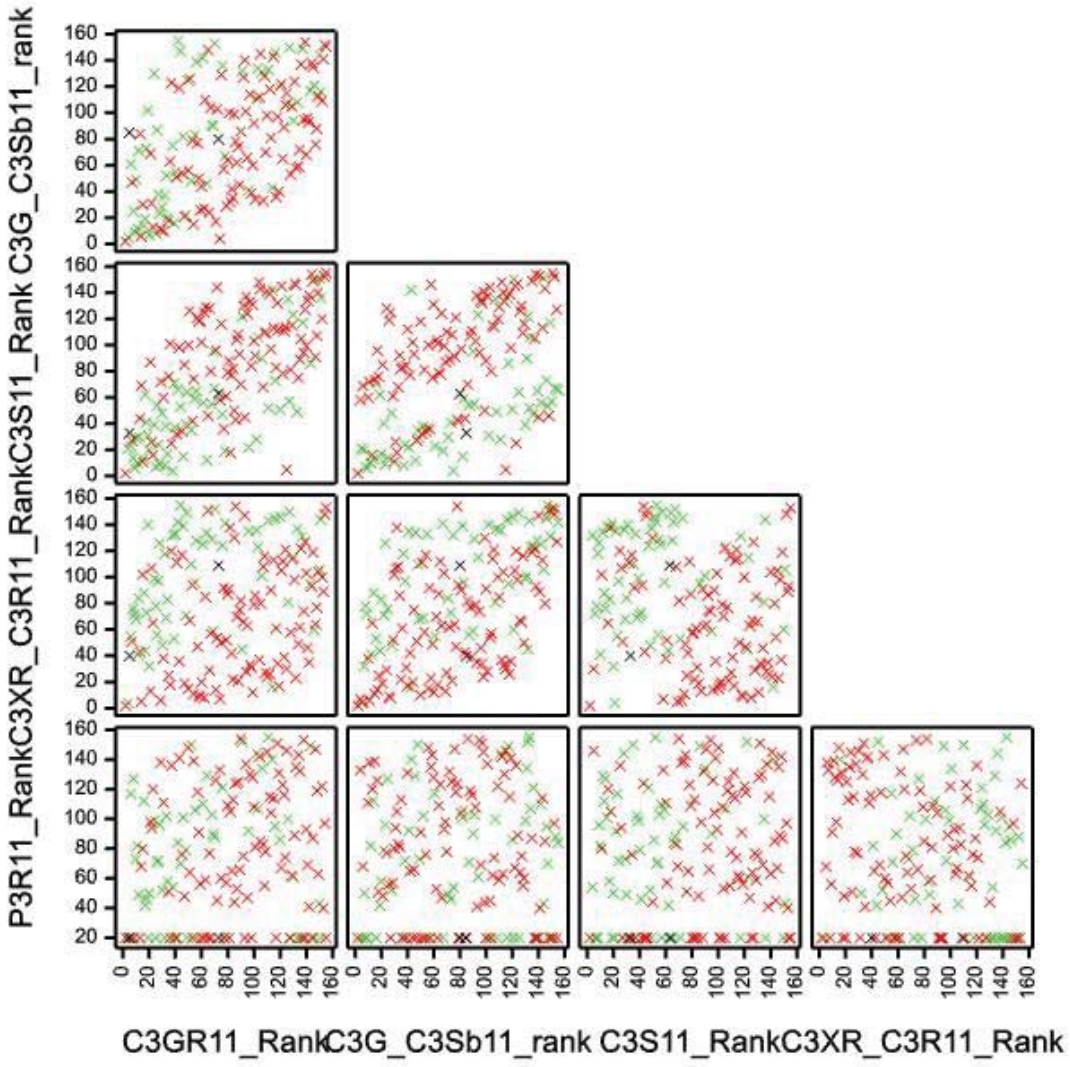
	P3R_09 µg/mL	P3R_10 mg/100 mL	P3R_11 mg/100 mL
Mean (no parents)	0.8	3.6	0.8
Max (no parents)	10.2	18.3	8.1
Min (no parents)	0.1	0.2	0.0
sd (no parents)	1.06	2.68	1.44
skewness	5.37	2.52	2.74
kurtosis	40.90	10.53	8.09

Supplemental Fig. S 4.2: Graphical summaries of the descriptive statistics of each trait detected in the F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' in 2009, 2010 and 2011.

Abbreviations: Max: maximum; Min: minimum; sd: standard deviation; C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^G-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^G-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside.

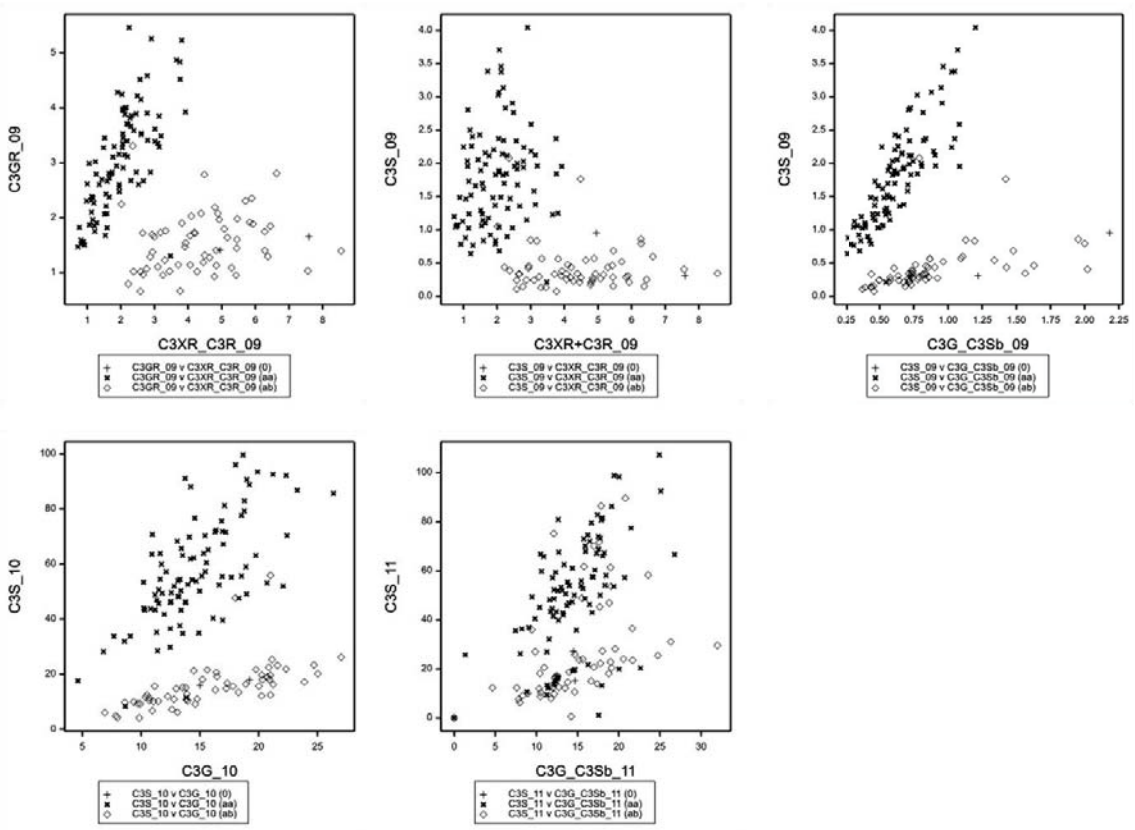

P3R09_rank
C3XR_C3R09_Rank
C3S09_rank
C3G_C3Sb09_rank
C3GR09_rank
C3G_C3Sb09_rank
C3S09_rank
C3XR_C3R09_Rank
0
20
40
60
80
100
120
140


C3G10_Rank
C3S10_rank
C3XR_C3R10_Rank
P3R10_rank
C3GR10_rank
C3G10_Rank
C3S10_rank
C3XR_C3R10_Rank



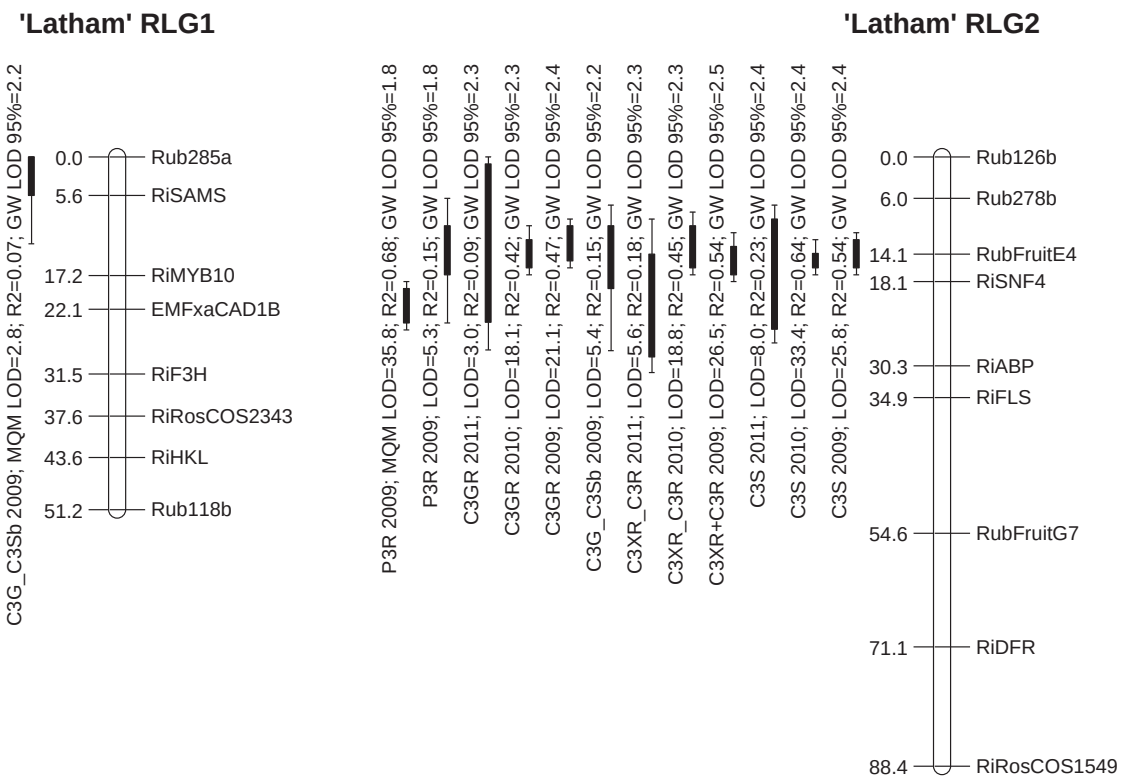
Supplemental Fig. S 4.3: Spearman's ranking scatter plots for anthocyanin compounds detected in the sub-set of 155 F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' for 2009, 2010, and 2011.

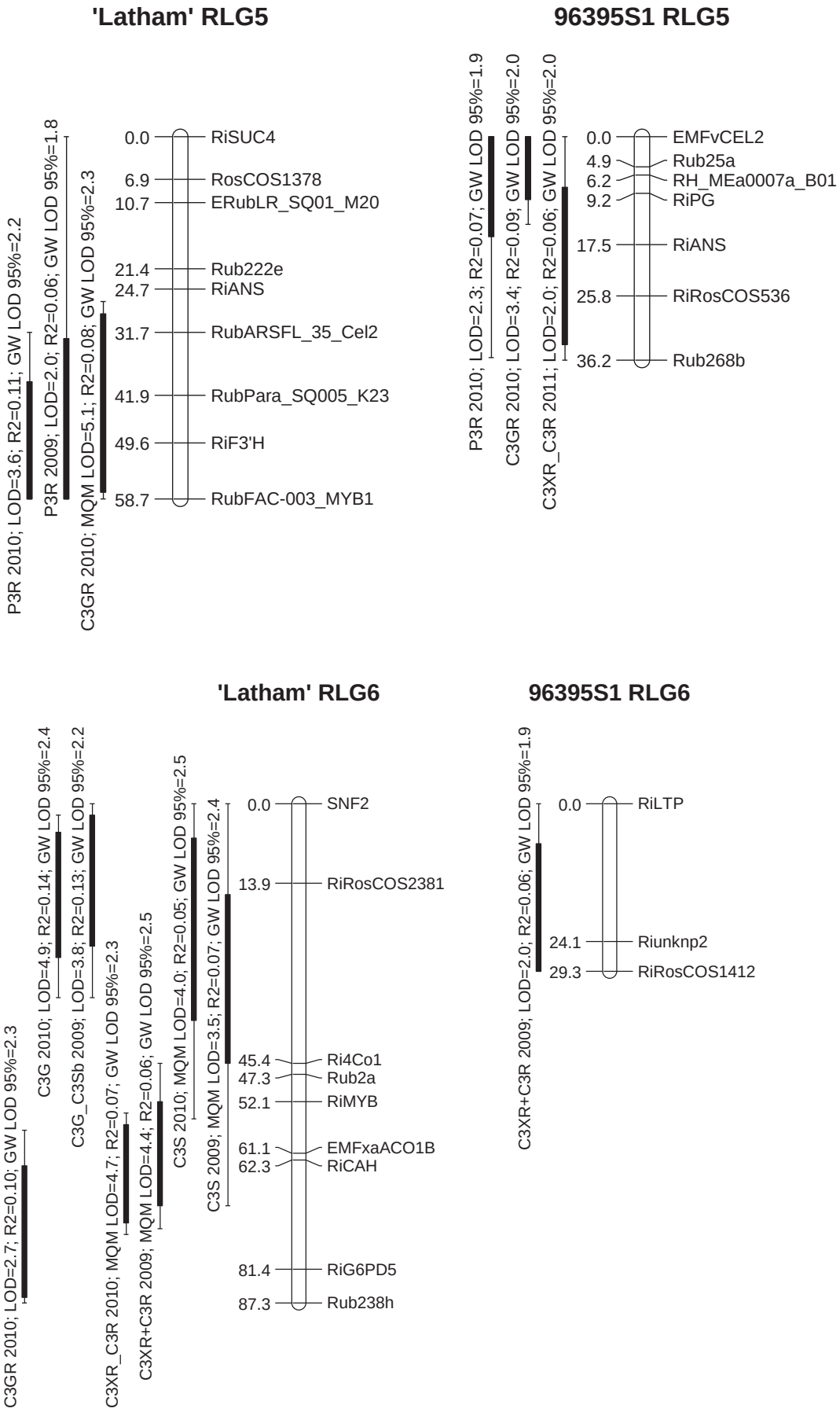
Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^G-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^G-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside

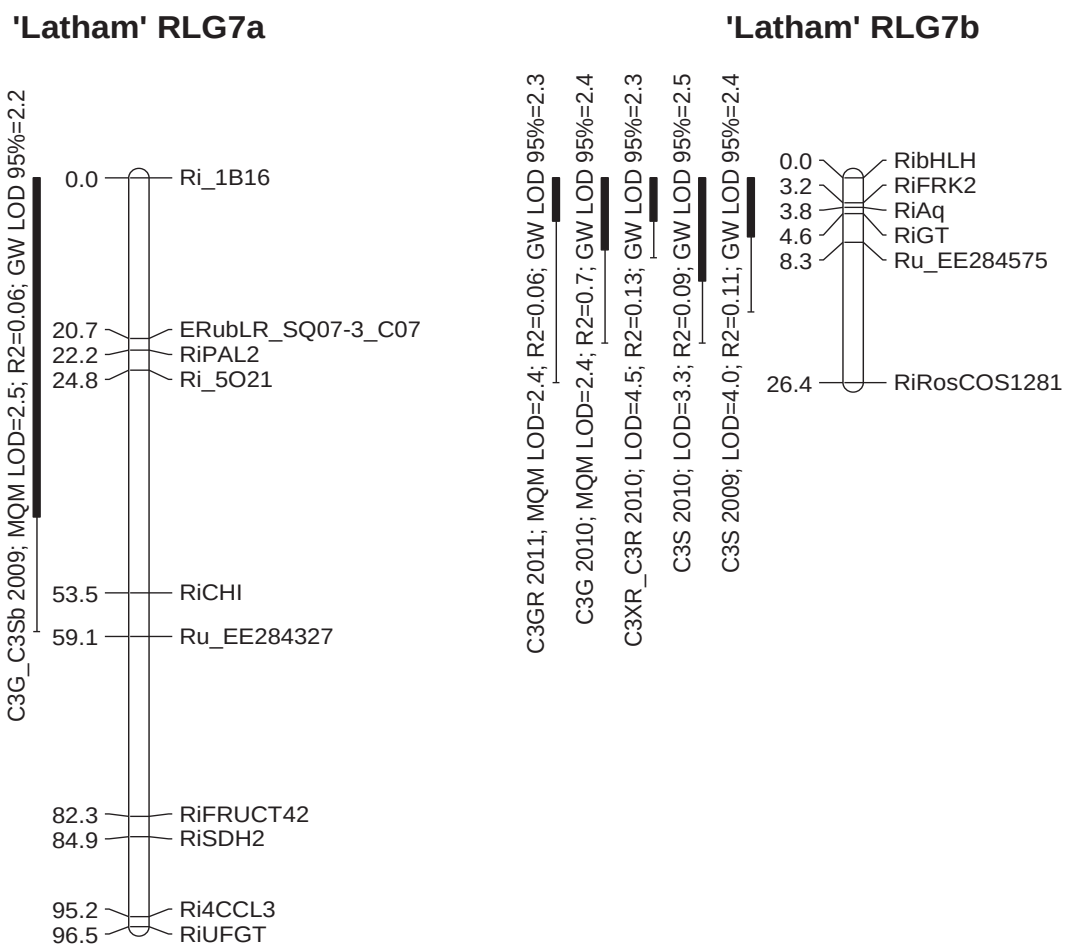


Supplemental Fig. S 4.4: Scatter plots for anthocyanin compounds detected in the sub-set of 155 F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham'.

Set A: 61 heterozygous individuals (open diamond symbols); Set B: 94 homozygous individuals (black cross symbols).
Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^G-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^G-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside


'Latham' RLG1
C3G_C3Sb 2009; MQM LOD=2.8; R2=0.07; GW LOD 95%=2.2
0.0 Rub285a
5.6 RiSAMS
17.2 RiMYB10
22.1 EMFxaCAD1B
31.5 RiF3H
37.6 RiRosCOS2343
43.6 RiHKL
51.2 Rub118b
'Latham' RLG2
P3R 2009; MQM LOD=35.8; R2=0.68; GW LOD 95%=1.8
P3R 2009; LOD=5.3; R2=0.15; GW LOD 95%=1.8
C3GR 2011; LOD=3.0; R2=0.09; GW LOD 95%=2.3
C3GR 2010; LOD=18.1; R2=0.42; GW LOD 95%=2.3
C3GR 2009; LOD=21.1; R2=0.47; GW LOD 95%=2.4
C3G_C3Sb 2009; LOD=5.4; R2=0.15; GW LOD 95%=2.2
C3XR_C3R 2011; LOD=5.6; R2=0.18; GW LOD 95%=2.3
C3XR_C3R 2010; LOD=18.8; R2=0.45; GW LOD 95%=2.3
C3XR+C3R 2009; LOD=26.5; R2=0.54; GW LOD 95%=2.5
C3S 2011; LOD=8.0; R2=0.23; GW LOD 95%=2.4
C3S 2010; LOD=33.4; R2=0.64; GW LOD 95%=2.4
C3S 2009; LOD=25.8; R2=0.54; GW LOD 95%=2.4
0.0 Rub126b
6.0 Rub278b
14.1 RubFruitE4
18.1 RiSNF4
30.3 RiABP
34.9 RiFLS
54.6 RubFruitG7
71.1 RiDFR
88.4 RiRosCOS1549
'Latham' RLG3
C3GR 2010; LOD=3.3; R2=0.09; GW LOD 95%=2.3
C3GR 2009; LOD=5.2; R2=0.16; GW LOD 95%=2.4
C3G_C3Sb 2011; LOD=3.0; R2=0.08; GW LOD 95%=2.3
C3G 2010; LOD=3.2; R2=0.09; GW LOD 95%=2.4
C3G_C3Sb 2009; LOD=3.9; R2=0.11; GW LOD 95%=2.4
C3XR+C3R 2009; MQM LOD=3.8; R2=0.05; GW LOD 95%=2.3
C3S 2010; LOD=3.3; R2=0.09; GW LOD 95%=2.5
C3S 2009; LOD=3.8; R2=0.11; GW LOD 95%=2.4
0.0 ERubLR_SQ07-4_E09
15.2 ERubLR_SQ01_F09
20.2 RH_MEa0013c_G02
26.8 Ri_GDSNP00321
35.5 RiPPC1
38.3 Ro_1E22
44.8 Ri_GDSNP02690
54.1 RiRosCOS2409
57.6 RubARSFL_134_APX
64.1 PBCESSRFxa12
70.4 RiF2KP
96395S1 RLG3
C3GR 2011; LOD=2.4; R2=0.09; GW LOD 95%=2.0
C3GR 2010; LOD=3.8; R2=0.11; GW LOD 95%=1.9
C3GR 2009; LOD=6.5; R2=0.24; GW LOD 95%=2.0
C3G_C3Sb 2011; LOD=2.4; R2=0.08; GW LOD 95%=1.9
C3G 2010; LOD=3.5; R2=0.10; GW LOD 95%=2.0
C3G_C3Sb 2009; LOD=4.2; R2=0.12; GW LOD 95%=1.9
C3S 2010; LOD=3.2; R2=0.10; GW LOD 95%=1.9
C3S 2009; LOD=4.1; R2=0.14; GW LOD 95%=1.9
0.0 Rub284a
13.2 RiPKP2
21.6 RiInvAlkE
41.3 PBCESSRFxa12
58.0 Ri60S
66.4 Rub12a

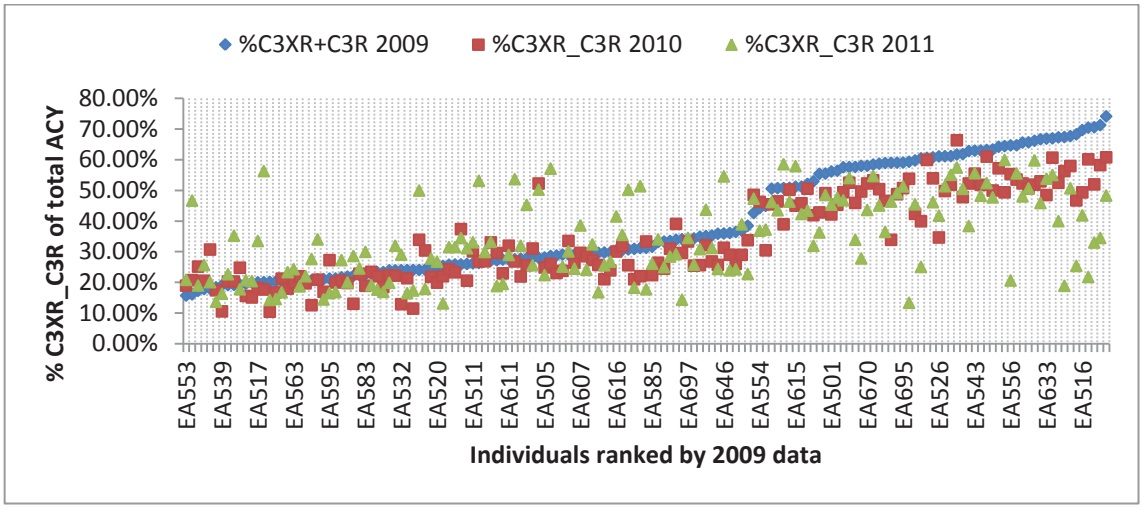

'Latham' RLG5
96395S1 RLG5
P3R 2010; LOD=3.6; R2=0.11; GW LOD 95%=2.2
P3R 2009; LOD=2.0; R2=0.06; GW LOD 95%=1.8
C3GR 2010; MQM LOD=5.1; R2=0.08; GW LOD 95%=2.3
0.0 RiSUC4
6.9 RosCOS1378
10.7 ERubLR_SQ01_M20
21.4 Rub222e
24.7 RiANS
31.7 RubARSFL_35_Cel2
41.9 RubPara_SQ005_K23
49.6 RiF3'H
58.7 RubFAC-003_MYB1
P3R 2010; LOD=2.3; R2=0.07; GW LOD 95%=1.9
C3GR 2010; LOD=3.4; R2=0.09; GW LOD 95%=2.0
C3XR_C3R 2011; LOD=2.0; R2=0.06; GW LOD 95%=2.0
0.0 EMFvCEL2
4.9 Rub25a
6.2 RH_MEa0007a_B01
9.2 RiPG
17.5 RiANS
25.8 RiRosCOS536
36.2 Rub268b
'Latham' RLG6
96395S1 RLG6
C3GR 2010; LOD=2.7; R2=0.10; GW LOD 95%=2.3
C3G 2010; LOD=4.9; R2=0.14; GW LOD 95%=2.4
C3G_C3Sb 2009; LOD=3.8; R2=0.13; GW LOD 95%=2.2
C3XR_C3R 2010; MQM LOD=4.7; R2=0.07; GW LOD 95%=2.3
C3XR+C3R 2009; MQM LOD=4.4; R2=0.06; GW LOD 95%=2.5
C3S 2010; MQM LOD=4.0; R2=0.05; GW LOD 95%=2.5
C3S 2009; MQM LOD=3.5; R2=0.07; GW LOD 95%=2.4
0.0 SNF2
13.9 RiRosCOS2381
45.4 Ri4Co1
47.3 Rub2a
52.1 RiMYB
61.1 EMFxaACO1B
62.3 RiCAH
81.4 RiG6PD5
87.3 Rub238h
C3XR+C3R 2009; LOD=2.0; R2=0.06; GW LOD 95%=1.9
0.0 RiLTP
24.1 Riunknp2
29.3 RiRosCOS1412

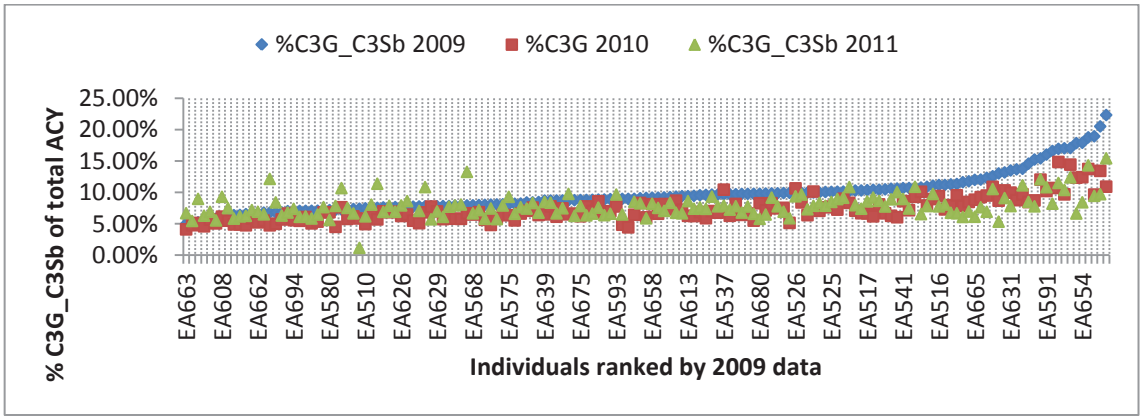


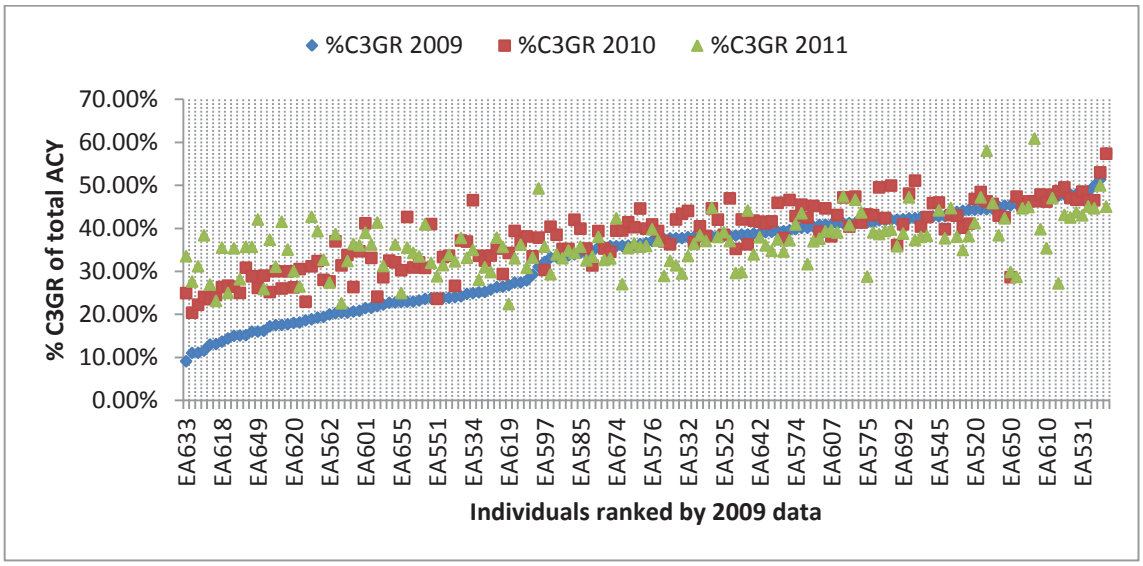
Supplemental Fig. S 4.5: Linkage maps of *Rubus idaeus* 'Latham' and *R. occidentalis* 96395S1 displaying quantitative trait loci (QTL) for variation in anthocyanin trait expression. QTL were mapped to all 'Latham' *Rubus* linkage groups (RLG) except RLG4. QTL were mapped to RLG 3, 5, and 6 for 96395S1.

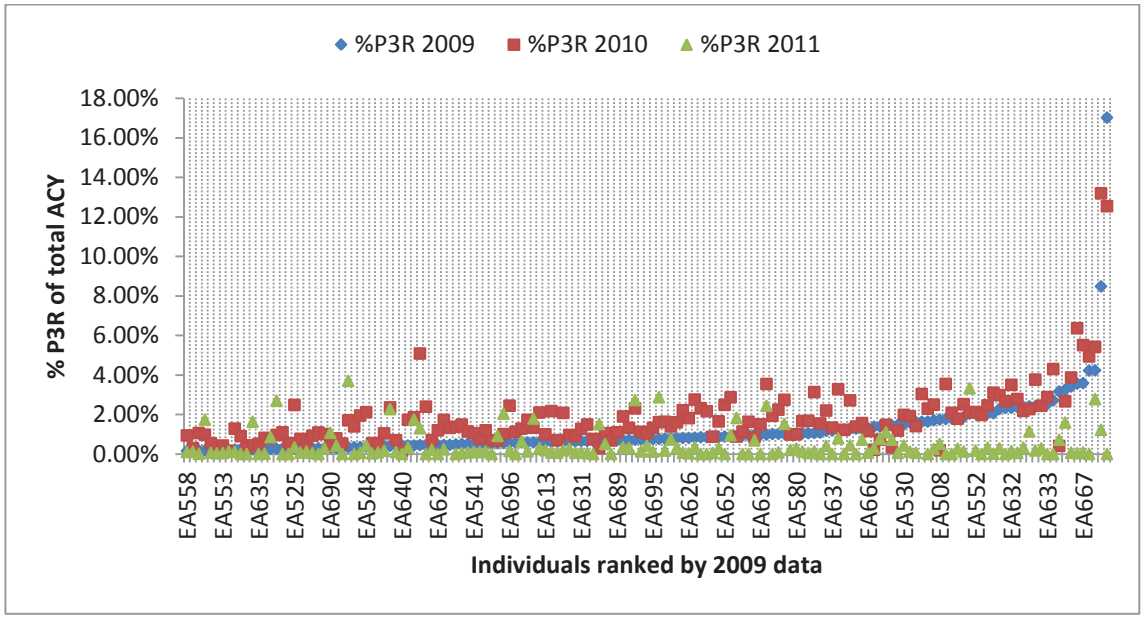
To the left of the RLG, trait abbreviation is followed by the year analyzed, the maximum logarithm of odds (LOD) score, the R^2 value representing the amount of variation explained by the locus, and the genome-wide (GW) LOD for 95% (LOD±1; dark bars) determined by 1000 permutations, the GW LOD for 99% (LOD±2) is indicated by the thin lines. *Rubus* linkage group order follows Bushakra et al. (2012a).

Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^G-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^G-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside.

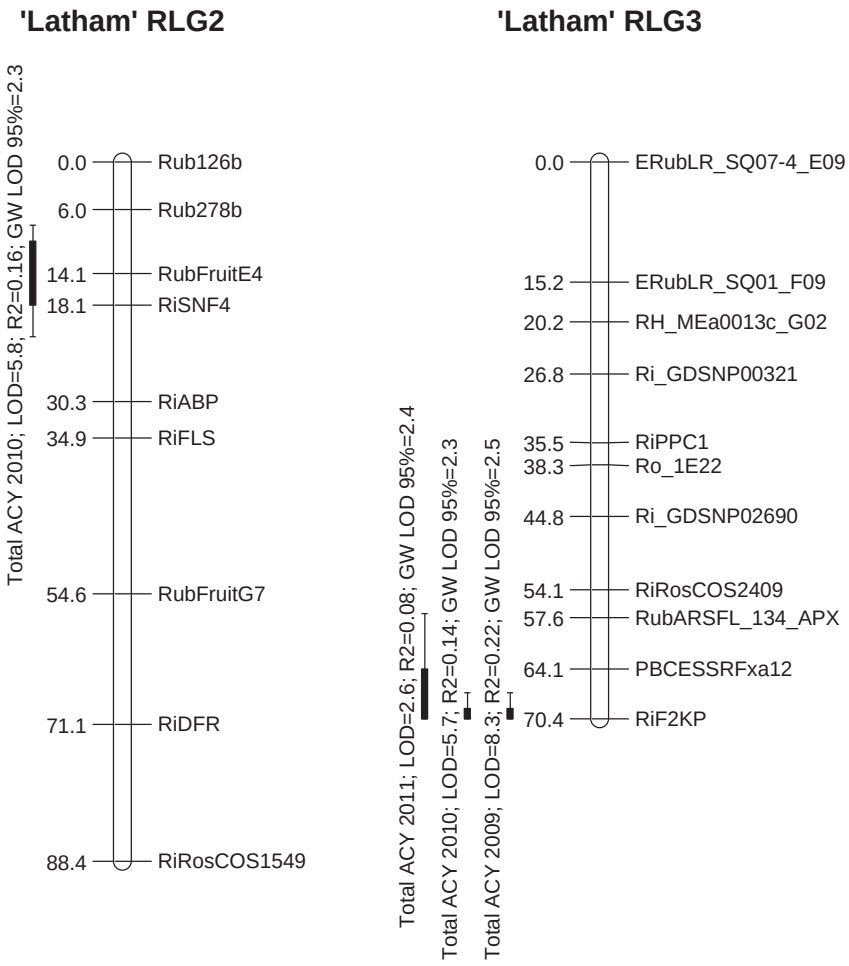

%C3XR+C3R 2009
%C3XR_C3R 2010
%C3XR_C3R 2011
% C3XR_C3R of total ACY
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
EA553 EA539 EA517 EA563 EA595 EA583 EA532 EA520 EA511 EA611 EA505 EA607 EA616 EA585 EA697 EA646 EA554 EA615 EA501 EA670 EA695 EA526 EA543 EA556 EA633 EA516
Individuals ranked by 2009 data


%C3G_C3Sb 2009
%C3G 2010
%C3G_C3Sb 2011
% C3G_C3Sb of total ACY
25.00%
20.00%
15.00%
10.00%
5.00%
0.00%
EA663 EA608 EA662 EA694 EA580 EA510 EA626 EA629 EA568 EA575 EA639 EA675 EA593 EA658 EA613 EA537 EA680 EA526 EA525 EA517 EA541 EA516 EA665 EA631 EA591 EA654
Individuals ranked by 2009 data


%C3GR 2009
%C3GR 2010
%C3GR 2011
% C3GR of total ACY
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
EA633 EA618 EA649 EA620 EA562 EA601 EA655 EA551 EA534 EA619 EA597 EA585 EA674 EA576 EA532 EA525 EA642 EA574 EA607 EA575 EA692 EA545 EA520 EA650 EA610 EA531
Individuals ranked by 2009 data



Supplemental Fig. S 4.6: Ranking of 155 individual F_1 progeny of *Rubus occidentalis* 96395S1 × *R. idaeus* 'Latham' by their percent accumulation of each of the compounds analyzed in 2009 (diamonds), compared with their accumulation in 2010 (squares) and 2011 (triangles).



Supplemental Fig. S 4.7: Linkage map of *Rubus idaeus* 'Latham' displaying quantitative trait loci (QTL) for variation in total anthocyanin trait expression. QTL were mapped to just two 'Latham' *Rubus* linkage groups (RLG).

Abbreviations: ACY: anthocyanin; C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^{G}-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^{G}-glucosylrutinoside; GW: Genome wide; LOD: log of odds; P3R: pelargonidin 3-*O*-rutinoside.

Supplemental Table S 4-1: Spearman's correlation coefficient calculations and statistical significance (p-value) for each pair of the six anthocyanin compounds analyzed for two consecutive years in 155 F1 progeny of Rubus occidentalis 96395S1 × R. idaeus 'Latham'.

	C3GR 09	*C3G_ C3Sb 09*	*C3S 09*	*C3XR+ C3R 09*	*C3GR 10*	*C3G 10*	*C3S 10*	*C3XR_ C3R 10*	*C3GR 11*	*C3G_ C3Sb 11*	*C3S 11*	*C3XR_ C3R 11*
C3G_C3Sb 09	0.05											
C3S 09	**0.82**	*0.16*										
C3XR+C3R 09	**-0.30**	**0.60**	**-0.55**									
P3R 09	**-0.47**	*0.23*	**-0.58**	**0.61**								
C3G 10					0.15							
C3S 10					**0.77**	**0.31**						
C3XR_C3R 10					**-0.28**	**0.30**	**-0.60**					
P3R 10					-0.07	*0.25*	*-0.18*	**0.32**				
C3G_C3Sb 11									**0.55**			
C3S 11									**0.70**	**0.46**		
C3XR_C3R 11									*0.24*	**0.54**	*-0.2*	
P3R 11									*0.19*	0.04	0.1	-0.13

N=155

*Missing data replaced by mean value (0.849)

†These correlations changed when the data were considered as Sets A and B

Bold font indicates significance at $p<0.001$. *Italic* font indicates significance at $p<0.05$

Abbreviations: C3S: cyanidin 3-*O*-sophoroside; C3XR: cyanidin 3-*O*-2^G-xylosylrutinoside; C3R: cyanidin 3-*O*-rutinoside; C3G: cyanidin 3-*O*-glucoside; C3Sb: cyanidin 3-*O*-sambubioside; C3GR: cyanidin 3-*O*-2^G-glucosylrutinoside; P3R: pelargonidin 3-*O*-rutinoside

CHAPTER FIVE

5 Conclusions

In the Introduction, I posed several questions that could be addressed using comparative genome mapping:

1. Can we use the information gained from the study of one organism to expand our knowledge of another organism?
2. Would the comparison of closely related organisms allow the discovery of lineage-specific differences?
3. Is it possible to reconstruct an ancestral genome by comparing extant genomes?

The studies presented in Chapters 2-4 describe how I endeavored to address these questions using the tools available to me.

5.1 Using the known to delve into the unknown

In Chapters 2 & 3, I explored the development and genetic mapping of molecular markers designed from known sources, such as orthologous sequences from genes and RosCOS, and made use of whole genome sequence, to assess the homology and degree of synteny among the genomes of *Fragaria*, *Malus*, *Prunus* and *Rubus*. In Chapter 2, I used bioinformatic tools to align orthologous sequences from *Fragaria*, *Malus*, and *Prunus* with the genomes of *M.* × *domestica* 'Golden Delicious' and *F. vesca* 'Hawaii 4' in order to design species-specific primer pairs for use in developing HRM-based genetic markers. HRM was used to detect polymorphisms among PCR products and those products that were polymorphic were mapped to existing linkage maps for *Malus*, which were then compared with a bin map of *F. vesca*. This allowed me to identify 21 regions of genomic synteny between 15 of the 17 apple linkage groups and the seven linkage groups of *Fragaria*, as well as to assess the transferability of the primer pairs designed from orthologous sequence for genetic mapping in the wider Rosaceae. Fortuitously, I was also able to identify some potential errors in the apple genome assembly where the genetic locations of some markers on the apple linkage maps were different from their genome physical positions.

In Chapter 3, I used a similar approach to construct new gene-based linkage maps of *Rubus occidentalis* and *R. idaeus*. For this, I used the genome sequences of *Fragaria*, *Malus*, *Prunus persica* 'Lovell', pre-released *Rubus idaeus* 'Heritage', and orthologous sequences to design *Rubus*-specific primer pairs. This study was not only the first to

compare the genome of *Rubus* with other Rosaceae genera, but also the first to make use of the pre-released *Rubus* genome for marker development. These *Rubus* genetic linkage maps are also being used to assist in orienting and assembling the 'Heritage' genome (J. Udall, pers. comm.), and therefore expanding on the knowledge of the *Rubus* genome using information from other genera.

5.2 Ancestral genome contribution

In Chapters 2 & 3, I tested the hypothesis that members of Rosaceae are derived from a common ancestor. In Chapter 2, the alignment of linkage maps and genome sequences of *Malus* and *F. vesca* was used to assess the ancestral contribution to homologous chromosomes. This analysis identified several previously unknown regions of homology between *Malus* and *Fragaria*. For example, it suggests that portions of homeologous apple linkage groups (LG) 8 & 15 and *Fragaria* LGII are derived from Ancestral chromosome 2 (A2), and that A5 has contributed to both apple LG17 and *Fragaria* LGIII (see Chapter 2, Fig. 2.3). While corroborating previous findings, my results also suggest that a more detailed examination of the genomes of these genera may uncover regions that have undergone additional rearrangements.

In Chapter 3, the use of orthologous markers, genetically mapped in *Rubus* and compared with physical maps of *Fragaria*, *Malus*, and *Prunus*, provided insight into the genome homology of *Rubus*, *Fragaria*, *Malus*, and *Prunus*. This work was the first to assess the ancestral genome contribution to *Rubus*. I was able to show that some regions of the genomes of all four genera are highly conserved, and also identified regions that appear to have undergone substantial rearrangement, which will be further discussed in section "5.4 *Impressions and interpretations*". In addition, the high level of genome homology detected between *Rubus* and *Fragaria* led me to suggest that the *Rubus* linkage groups be renamed to agree with *Fragaria*, to make future comparisons easier.

5.3 Lineage-specific differences

In Chapter 3, the genome comparison of the four genera suggests that lineage-specific differences might be associated with RLG6, FLGVI, apple homeologous LG 4 & 12 and 9 & 17, and *Prunus* G3, as these have different ancestral chromosome contributions. These contributions will be further discussed in section "5.4 *Impressions and interpretations*". In Chapter 4, I used the orthologous marker-based *Rubus* linkage maps to identify QTL involved in the variation in expression of six anthocyanins. I then

aligned the *Rubus* linkage map, the *Fragaria vesca* physical map, and a published QTL *F.* × *ananassa* map, to assess the homology of QTL between genera. I expanded this comparison to include published anthocyanin and fruit quality QTL from *Prunus* and *Malus* and identified potential regions of homology. This work suggests that though specific QTL may not be homologous, certain regions of the genome appear to be important across the different genera for "big picture" traits, such as fruit quality.

5.4 Impressions and interpretations

The overall knowledge base of Rosaceae genomics and genetics has been extended by the alignment of the maps of *Rubus*, *Fragaria*, *Malus*, and *Prunus*, which provide a first look at the degree of genome synteny and ancestral genome contribution among these four economically important genera (Bushakra et al. 2012a). However, even with these new developments and the existing resources, comparative genome mapping in Rosaceae still lags behind similar work performed in Brassicaceae, Solanaceae and Fabaceae.

One possible reason for the lag in progress in Rosaceae, is the generally low transferability of primer pairs among genera, as demonstrated by the *in silico* assessment of SSR transferability conducted between published apple SSR sequences aligned with the *F. vesca* draft genome sequence (Chapter 2), as well as the difficulties encountered in the development of markers for the *Rubus* linkage map (Chapter 3). Even though primer pairs derived from orthologous sequence are more likely to amplify a PCR product, the detection of polymorphism can be low. This was demonstrated by the work performed in the present study, as well as by Gasic et al. (2009b), Lewers et al. (2005), and Decroocq et al. (2003). These findings seem counter-intuitive, since if Rosaceae is monophyletic and has undergone a recent and rapid radiation as is strongly suggested (Wikström et al. 2001; Bortiri et al. 2001; Potter et al. 2007; Eriksson et al. 2003; Shaw and Small 2004; Campbell et al. 2007; Crepet et al. 2004), then one would expect a high level of family-wide sequence similarity. High sequence similarity, and therefore low sequence divergence, has been found in both *Prunus* (Bortiri et al. 2001) and Pyrinae (Campbell et al. 2007), yet at the same time, a comparison of nuclear internal transcribed spacer (nrITS) and chloroplast *trn*L-*trn*F sequences of 25 members of Pyrinae and 25 members of Rosoideae showed that mean sequence divergence among the Rosoideae genera was significantly greater than the mean sequence divergence found in the genera of Pyrinae (Campbell et al. 2007). These findings

suggest that "the Devil is in the details" when genetic mapping is performed in Rosaceae, since while genome synteny exists to the extent that polymorphism can be difficult to detect, there are enough differences in the genome sequence to reduce primer pair transfer efficiency.

By considering the overall picture presented by the comparisons of apple and strawberry, and the comparison of the four genera, I have identified some genome regions that appear to be highly conserved, as evidenced by marker order. While marker order is not a conclusive means of determining synteny, since order can vary depending on genotyping accuracy, as well as potential genome assembly errors, when four genera show conservation, it deserves attention. By reading across Table 5-1, where each box represents a region of the genome over which the marker order is maintained between two or more genera, and comparing it to the linkage maps of *Rubus* aligned with *Fragaria*, *Malus*, and *Prunus* (see Chapter 3, Supplemental Figs. S3.1, S3.2 & S3.3), one can identify three regions that are conserved in marker order in all four genera relative to the marker order in *Rubus*. On RLG2, the conserved region is flanked by markers Rub126b and RiFLS; on RLG5, the conserved region is flanked by markers RubFAC-003_MYB1 and RiF3'H; and on RLG7b, the conserved region is flanked by markers RiAq and RoRosCOS1281. On RLG3, the conserved region flanked by markers ERubLR_SQ_1_F09 and RH_MEa0014b_F07 is found in *Rubus*, *Malus*, and *Prunus*, however in *Fragaria* this region appears to be inverted. There are an additional 12 areas that are conserved between *Rubus* (RLG1-RLG7) and at least one other genus.

Along with these specific regions where marker order is maintained, larger segments exist where markers are all (or mostly) present but the order is variable. In addition, when the ancestry of these regions is considered, it appears that some ancestral chromosome regions are more likely to be inherited as single blocks (see Chapter 3, Figs. 3.4 & 3.5) resulting in some ancestral chromosomes appearing to be less fragmented than others (Table 5-2). A new bioinformatic analysis to detect orthology among the genomes of apple, peach, and strawberry with the outgroup grape supports this observation (Jung et al. in press). The observations resulting from these two investigations suggest that Ancestral chromosomes A1, A3, A4, A7, and A8 have undergone the least amount of fragmentation, while A2, A5, A6, and A9 appear to have undergone more rearrangements. These observations also suggest that *Prunus* has maintained the most ancestor-like genome. The pattern of ancestral chromosome

inheritance suggests that the maintenance of the ancestral state is important in some instances, and therefore may be the genome regions that define what it means to be "rosaceous" (i.e. Potter and colleagues' (2007) Rosaceae ancestral state: shrubs with alternate simple leaves, stipules present, stamens numerous, etc.) while the rearrangements and variations found in each genus may determine their uniqueness.

Table 5-1: Genome regions conserved in terms of marker order among *Rubus* (RLG1-RLG7), *Fragaria* (FLGI-FLGVII), *Prunus* (PG1-PG8) and *Malus* (MLG1-MLG17). Markers on RLG linkage groups are genetically mapped and distances spanned by the markers are in centimorgans (cM); markers on the other three genera are physically mapped and distances spanned by the markers are in megabases (Mb). Shaded rows indicate three regions conserved in all four genera.

RLG1	**cM**	**FLGI**	**Mb**	**PG7**	**Mb**	**PG6**	**Mb**	**MLG17**	**Mb**	**MLG9**	**Mb**	**MLG2**	**Mb**	**MLG15**	**Mb**
Ro_11C01		Ro_11C01		Ro_11C01								Ro_11C01		RosCOS2343	
Ru_EE284400	21.8	Ru_EE284400	7	RiF3'H	4.4							RiF3'H	9.7	RosCOS873	3.2
RiMYB10						RiMYB10		RiMYB10		RiMYB10					
EMFxa						EMFxa									
CAD1B	3.7					CAD1B	3.6	RiGIB	8.4	RiGIB	1.6				
RLG2	**cM**	**FLGII**	**Mb**	**PG1**	**Mb**	**PG8**	**Mb**	**MLG5**	**Mb**	**MLG10**	**Mb**				
Rub126b		Rub126b		RiFLS		Rub126b		RiCXE		RiAA					
RiFLS	35.5	RiFLS	13.7	Ro_10O04	3.8	RiAA	2.2	Rub126b	5.6	RiFLS	0.6				
RiTh		RiTh													
RiDFR	0.8	RiDFR	0.1												
RLG3	**cM**	**FLGIII**	**Mb**	**PG4**	**Mb**			**MLG5**	**Mb**	**MLG3**	**Mb**	**MLG11**	**Mb**		
ERubLR_SQ1_F09		possible inversion		ERubLR_SQ1_F09				ERubLR_SQ07_4_E09		RubARSFL_134_APX		Ro_1E22			
RH_MEa0014b_F07	4.7		5.4	RH_MEa0014b_F07	0.7			RH_MEa0014b_F07	4.6	Ri_GDSNP02690	8	RiPPC1	1.3		
RiPKP2				RiPKP2											
Ro_1E22	5.2			Ro_1E22	3.1										
RLG4	**cM**			**PG1**	**Mb**										
RiRosCOS1217				RiRosCOS1217											
ERubLR_SQ005_3_H01	47.1			ERubLR_SQ005_3_H01	24.9										
RLG5	**cM**	**FLGV**	**Mb**	**PG5**	**Mb**			**MLG6**	**Mb**						
RubFAC-003_MYB1		RubFAC-003_MYB1		RubFAC-003_MYB1				RubFAC-003_MYB1							
RiF3'H	11.4	RiF3'H	2.5	RiF3'H	1.2			RiF3'H	1.9						
Ri_GDSNP00768				Ri_GDSNP00768											
ERubLR_SQ01_M20	10.2			ERubLR_SQ01_M20	5.3										
RLG6	**cM**	**FLGI**	**Mb**	**PG6**	**Mb**			**MLG4**	**Mb**	**MLG9**	**Mb**	**MLG17**	**Mb**		
RiRosCOS2381				RiRosCOS2381											
RiInvAlkA	1.7			RiInvAlkA	0.2										
SNF2								SNF2							
RiInvAlkA	15.3							RiInvAlkA	3.3						
EMFxa										EMFxa		EMFxa			
ACO1B										ACO1B		ACO1B			
RiSAMD	4.3									RiSAMD	3.1	RiSAMD	3.4		
RiCAH		RiCAH													
RiRosCOS645	1.7	RiRosCOS645	0.2												
RLG7a	**cM**	**FLGVII**	**Mb**					**MLG7**	**Mb**						
RiSDH2		RiSDH2													
RiUFGT	8	RiUFGT	2.6												
Ri4CCL3								Ri4CCL3							
RiUFGT	2.5							RiUFGT	0.1						
RLG7b	**cM**	**FLGVII**	**Mb**	**PG2**	**Mb**			**MLG7**	**Mb**						
RiAq		RiAq		RiAq											
RiGT	0.9	RiGT	1	RiGT	0.1			RiGT							
RiRosCOS1281	26.7							RiRosCOS1281	13.8						

Table 5-2: Distribution of ancestral chromosomes (A1-A9) among the linkage groups (LG) of *Rubus*, *Fragaria*, *Malus*, and *Prunus*.

	Rubus LG	*Fragaria* LG	*Malus* LG†	*Prunus* LG
A1	2#	2#*	5/10#* 13/16*	1# 8#
A2	2# 4# 5#	1* 2#* 4# 5#*	8/15#* 13/16#?	1#
A3	5#	5#*	6/14#*	5#
A4	7#	6*? 7#*	1/7# 2#*	2#
A5	1# 6#	1# 3#	9/17#*	3#
A6	6#	4*? 6#*	4*? 12/14*	7#?
A7	1#	1#*	2/15#*	7#
A8	3#	3#*	3/11#*	6#
A9	3# 6#	3# 6#	5/10# 12/14#	3#? 4#

†*Malus* homeologs indicated by slash separating LG
*data from apple/strawberry comparison (Chapter 2)
#data from *Rubus* map construction (Chapter 3)
?insufficient data

5.5 Future work

To determine if the observed inheritance of the ancestral regions is maintained, and the extent of ancestral chromosome fragmentation, a greater number of genetic maps densely populated with orthologous markers and covering a wider range of genera should be constructed. This applies especially to members of subfamily Dryadoideae, which includes those genera that have a base chromosome number of nine ($x = 9$) and are nitrogen-fixing, and to *Gillenia*, the proposed ancestor of the Pyrodae (Potter et al. 2007). To date, no genetic maps have been constructed for any member of subfamily Dryadoideae, and only one genetic linkage map has been constructed for a genus outside of the economically important Rosaceae fruit crop genera, i.e. the ornamental *Physocarpus* (Spiraeoideae) (Sutherland et al. 2008).

The genome regions that are conserved among the four genera are good candidates for gene mining to identify the genes underlying these regions. Once identified, transient expression assays and gene modification (knock outs/over expression) could be

performed on strawberry (Folta et al. 2006) to assess the importance of the genes for determining "rosaceous" traits. These techniques could also be applied to the genome regions associated with QTL, especially those regions that seem to encompass QTL for "big picture" traits, such as fruit quality. It would also be interesting to compare these genome regions with genetic maps of non-domesticated Rosaceae species to determine if generations of selective breeding have influenced the distribution of the QTL through meiotic recombination.

Further characterization of the genomes of Rosaceae members will be facilitated by the use of many of the tools used during the research conducted here. Such tools include the set of orthologous markers developed for the genetic mapping of *Rubus*, the release and refinement of the *Malus*, *Prunus* and *F. vesca* genomes, the development of the *Rubus* genome, and the increase and improvement of genetic linkage maps. The relatively small genome size of Rosaceae genera and the morphological diversity that is represented by the family presents opportunities to discover regions important for plant development as Rosaceae encompasses species with many different fruit types and growth habits that span the range from annuals to perennials, and from herbaceous to woody, with *Rubus* providing an opportunity to study an "in-between" cane growth habit. As more populations representing a wider range of Roseaceae genetic diversity are established and used for QTL mapping, gene mining of regions of interest will be used to develop gene-based markers to conduct association mapping in more distantly related species, including members of sister clades. Genome studies in Rosaceae are on the threshold of expansion and with the recent investment in SNP chip technology for apple and peach (Micheletti et al. 2012; Chagné et al. 2012; Khan et al. 2011), as well as the accumulation of sequence data in the form of EST libraries and RosCOS, combined with existing genetic maps for the economically important members of the family, the Rosaceae research community has an extensive tool box for further research incorporating more genera, once appropriate segregating populations are established.

REFERENCES

Bortiri E, Oh S-H, Jiang J, Baggett S, Granger A, Weeks C, Buckingham M, Potter D, Parfitt DE (2001) Phylogeny and systematics of *Prunus* (Rosaceae) as determined by sequence analysis of ITS and the chloroplast *trnL-trnF* spacer DNA. Systematic Botany 26 (4):797-807

Bushakra J, Stephens M, Atmadjaja A, Lewers K, Symonds V, Udall J, Chagné D, Buck E, Gardiner S (2012a) Construction of black (*Rubus occidentalis*) and red (*R. idaeus*) raspberry linkage maps and their comparison to the genomes of strawberry, apple, and peach. Theoretical and Applied Genetics:1-17. doi:10.1007/s00122-012-1835-5

Campbell CS, Evans RC, Morgan DR, Dickinson TA, Arsenault MP (2007) Phylogeny of subtribe Pyrinae (formerly the Maloideae, Rosaceae): Limited resolution of a complex evolutionary history. Plant Systematics and Evolution 266 (1):119-145

Chagné D, Crowhurst RN, Troggio M, Davey MW, Gilmore B, Lawley C, Vanderzande S, Hellens RP, Kumar S, Cestaro A, Velasco R, Main D, Rees J, Iezzoni A, Mockler TC, Wilhelm LJ, Van de Weg E, Gardiner SE, Bassil NV, Peace C (2012) Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. Public Library of Science One

Crepet WL, Nixon KC, Gandolfo MA (2004) Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. American Journal of Botany 91 (10):1666-1682. doi:10.3732/ajb.91.10.1666

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theoretical and Applied Genetics 106 (5):912-922

Eriksson T, Hibbs MS, Yoder AD, Delwiche CF, Donoghue MJ (2003) The phylogeny of Rosoideae (Rosaceae) based on sequences of the internal transcribed spacers (ITS) of nuclear ribosomal DNA and the *TrnL/F* region of chloroplast DNA. International Journal of Plant Science 164 (2):197-211. doi:1058-5893/2003/16402-0001

Folta KM, Dhingra A, Howard L, Stewart PJ, Chandler CK (2006) Characterization of LF9, an octoploid strawberry genotype selected for rapid regeneration and transformation. Planta 224 (5):1058-1067. doi:10.1007/s00425-006-0278-0

Gasic K, Han Y, Kertbundit S, Shulaev V, Iezzoni A, Stover E, Bell R, Wisniewski M, Korban S (2009b) Characteristics and transferability of new apple EST-derived SSRs to other Rosaceae species. Molecular Breeding 23 (3):397-411

Jung S, Cestaro A, Troggio M, Main D, Zheng P, Cho I, Folta KM, Sosinski B, Abbott A, Celton J-M, Arús P, Shulaev V, Verde I, Morgante M, Rokhsar DS, Velasco R, Sargent DJ (in press) Whole genome comparisons of *Fragaria*, *Prunus*, and *Malus* reveal different modes of evolution between rosaceous subfamilies. BMC Genomics

Khan MA, Han Y, Zhao YF, Korban SS (2011) A high-throughput apple SNP genotyping platform using the GoldenGate™ assay. Gene. doi:10.1016/j.gene.2011.12.001

Lewers KS, Styan SMN, Hokanson SC, Bassil NV (2005) Strawberry GenBank-derived and genomic simple sequence repeat (SSR) markers and their utility with strawberry, blackberry, and red and black raspberry. Journal of the American Society for Horticultural Science 130 (1):102-115

Micheletti D, Aramini V, Arús P, Banchi E, Barreneche T, Bassi D, Dirlewanger E, Gao Z, Gazza L, Lambert P, Laurens F, Li X, Pascal T, Quilot-Turion B, Rossini L, Troggio M, Van De Weg E, Verde I, Aranzana MJ (2012) Application of high throughput genotyping techniques in peach germplasm and breeding lines within the FruitBreedomics project. Paper presented at the Plant and Animal Genome XX, San Diego, CA, 14-18 Jan, 2012

Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, Campbell CS (2007) Phylogeny and classification of Rosaceae. Plant Systematics and Evolution 266 (1):5-43

Shaw J, Small RL (2004) Addressing the "hardest puzzle in American pomology:" Phylogeny of *Prunus* sect. Prunocerasus (Rosaceae) based on seven noncoding chloroplast DNA regions. American Journal of Botany 91 (6):985-996. doi:10.3732/ajb.91.6.985

Sutherland BG, Tobutt KR, Marchese A, Paternoster G, Simpson DW, Sargent DJ (2008) A genetic linkage map of *Physocarpus*, a member of the Spiraeoideae (Rosaceae), based on RAPD, AFLP, RGA, SSR and gene specific markers. Plant Breeding 127 (5):527-532

Wikström N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proceedings of the Royal Society B 268 (1482):2211-2220. doi:10.1098/repb.201.1782

BIBLIOGRAPHY

Acarkan A, Roßberg M, Koch M, Schmidt R (2000) Comparative genome analysis reveals extensive conservation of genome organisation for *Arabidopsis thaliana* and *Capsella rubella*. Plant Journal 23 (1):55-62

Adams KL, Wendel JF (2005) Novel patterns of gene expression in polyploid plants. Trends in Genetics 21 (10):539-543

Aharoni A, O'Connell AP (2002) Gene expression analysis of strawberry achene and receptacle maturation using DNA microarrays. Journal of Experimental Botany 53 (377):2073-2087. doi:10.1093/jxb/erf026

Ahn S, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. Proceedings of the National Academy of Sciences of the United States of America 90 (17):7980-7984

Allan AC, Hellens RP, Laing WA (2008) MYB transcription factors that colour our fruit. Trends in Plant Science 13 (3):99-102

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215 (3):403-410

Amsellem L, Dutech C, Billotte N (2001) Isolation and characterization of polymorphic microsatellite loci in *Rubus alceifolius* Poir. (Rosaceae), an invasive weed in La Réunion island. Molecular Ecology Notes 1 (1-2):33-35. doi:10.1046/j.1471-8278.2000.00013.x

Andersen JR, Lübberstedt T (2003) Functional markers in plants. Trends in Plant Science 8 (11):554-560

Anderson JA, Churchill GA, Autrique JE, Tanksley SD, Sorrells ME (1993) Optimizing parental selection for genetic linkage maps. Genome 36 (1):181-186. doi:10.1139/g93-024

Andolfatto P, Davison D, Erezyilmaz D, Hu TT, Mast J, Sunayama-Morita T, Stern DL (2011) Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Research 21 (4):610-617. doi:10.1101/gr.115402.110

APG (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Botanical Journal of the Linnean Society 141 (4):399-436. doi:10.1046/j.1095-8339.2003.t01-1-00158.x

Aranzana M, Carbó J, Arús P (2003a) Microsatellite variability in peach [*Prunus persica* (L.) Batsch]: cultivar identification, marker mutation, pedigree inferences and population structure. Theoretical and Applied Genetics 106 (8):1341-1352

Aranzana MJ, Pineda A, Cosson P, Dirlewanger E, Ascasibar J, Cipriani G, Ryder C, Testolin R, Abbott A, King G, Iezzoni A, Arús P (2003b) A set of simple-sequence repeat (SSR) markers covering the *Prunus* genome. Theoretical and Applied Genetics 106 (5):819-825

Arumuganathan K, Earle E (1991) Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter 9 (3):208-218

Arús P, Gardiner S (2007) Genomics for improvement of Rosaceae temperate tree fruit. In: Varshney RK, Tuberosa R (eds) Genomics-Assisted Crop Improvement: Genomics Applications in Crops, vol 2. Springer, Dordrecht, The Netherlands,

Ash MM, Wolford KA, Carden TJ, Hwang KT, Carr TP (2011) Unrefined and refined black raspberry seed oils significantly lower triglycerides and moderately affect cholesterol metabolism in male Syrian hamsters. J Med Food 14 (9):1032-1038. doi:10.1089/jmf.2010.0181

Badenes ML, Parfitt DE (1995) Phylogenetic relationships of cultivated *Prunus* species from an analysis of chloroplast DNA variation. Theoretical and Applied Genetics 90 (7):1035-1041

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3 (10):e3376. doi:10.1371/journal.pone.0003376

Baldi P, Patocchi A, Zini E, Toller C, Velasco R, Komjanc M (2004) Cloning and linkage mapping of resistance gene homologues in apple. Theoretical and Applied Genetics 109 (1):231-239

Ballester J, Bošković R, Batlle I, Arús P, Vargas F, de Vicente MC (1998) Location of the self-incompatibility gene on the almond linkage map. Plant Breeding 117 (1):69-72. doi:10.1111/j.1439-0523.1998.tb01450.x

Ballester J, Socias i Company R, Arús P, De Vicente MC (2001) Genetic mapping of a major gene delaying blooming time in almond. Plant Breeding 120 (3):268-270. doi:10.1046/j.1439-0523.2001.00604.x

Bassil N, Hummer K, Postman J, Fazio G, Baldo A, Armas I, Williams R (2009) Nomenclature and genetic relationships of apples and pears from Terceira Island. Genetic Resources and Crop Evolution 56 (3):339-352

Bassil NV, Gunn M, Folta K, Lewers K (2006) Microsatellite markers for *Fragaria* from 'Strawberry Festival' expressed sequence tags. Molecular Ecology Notes 6 (2):473-476. doi:10.1111/j.1471-8286.2006.01278.x

Baudry A, Heim M, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L (2004) TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant Journal 39 (3):366 - 380

Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L (2006) Connected populations for detecting quantitative trait loci and testing for epistasis: An application in maize Theoretical and Applied Genetics 113 (2):206-224

Bolot S, Abrouk M, Masood-Quraishi U, Stein N, Messing J, Feuillet C, Salse J (2009) The 'inner circle' of the cereal genomes. Curr Opin Plant Biol 12 (2):119-125

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120 (4):1095-1103

Booi S, van Dyk MM, du Preez MG, Rees DJG, Labuschagné I (2005) Molecular typing of red and green phenotypes of 'Bon Rouge' pear trees, with the use of microsatellites. Acta Horticulturae (ISHS) 671:293-297

Bortiri E, Oh S-H, Jiang J, Baggett S, Granger A, Weeks C, Buckingham M, Potter D, Parfitt DE (2001) Phylogeny and systematics of *Prunus* (Rosaceae) as determined by sequence analysis of ITS and the chloroplast *trnL-trnF* spacer DNA. Systematic Botany 26 (4):797-807

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in Man using restriction fragment length polymorphisms. American Journal of Human Genetics 32 (3):314-331

Bus V, Bassett H, Bowatte D, Chagné D, Ranatunga C, Ulluwishewa D, Wiedow C, Gardiner S (2010) Genome mapping of an apple scab, a powdery mildew and a woolly apple aphid resistance gene from open-pollinated Mildew Immune Selection. Tree Genetics & Genomes 6 (3):477-487

Bus V, White A, Gardiner S, Weskett R, Ranatunga C, Samy A, Cook M, Rikkerink E (2002) An update on apple scab resistance breeding in New Zealand. Acta Horticulturae (ISHS) 595:43-47

Bushakra J, Stephens M, Atmadjaja A, Lewers K, Symonds V, Udall J, Chagné D, Buck E, Gardiner S (2012a) Construction of black (*Rubus occidentalis*) and red (*R. idaeus*) raspberry linkage maps and their comparison to the genomes of strawberry, apple, and peach. Theoretical and Applied Genetics:1-17. doi:10.1007/s00122-012-1835-5

Bushakra JM, Sargent DJ, Cabrera A, Crowhurst RN, Lopez Girona E, Velasco R, Symonds VV, van der Knaap E, Troggio M, Gardiner SE, Chagné D (2012b) Rosaceae conserved orthologous set (RosCOS) markers as a tool to assess genome synteny between *Malus* and *Fragaria*. Tree Genetics & Genomes:1-16. doi:10.1007/s11295-011-0450-y

Byrne D, Jelenkovic G (1976) Cytological diploidization in the cultivated octoploid strawberry *Fragaria* x *ananassa*. Canadian Journal of Genetics and Cytology 18 (4):653-659

Cabrera A, Kozik A, Howad W, Arús P, Iezzoni A, van der Knaap E (2009) Development and bin mapping of a Rosaceae Conserved Ortholog Set (COS) of markers. BMC Genomics 10 (1):562

Cabrera A, Rosyara U, De Franceschi P, Sebolt A, Sooriyapathirana S, Dirlewanger E, Quero-Garcia J, Schuster M, Iezzoni A, van der Knaap E (2011) Rosaceae conserved orthologous sequences marker polymorphism in sweet cherry germplasm and construction of a SNP-based map. Tree Genetics & Genomes:1-11. doi:10.1007/s11295-011-0436-9

Campbell CS, Evans RC, Morgan DR, Dickinson TA, Arsenault MP (2007) Phylogeny of subtribe Pyrinae (formerly the Maloideae, Rosaceae): Limited resolution of a complex evolutionary history. Plant Systematics and Evolution 266 (1):119-145

Cantini C, Iezzoni AF, Lamboy WF, Boritzki M, Struss D (2001) DNA fingerprinting of tetraploid cherry germplasm using simple sequence repeats. Journal of the American Society for Horticultural Science 126 (2):205-209

Cao K, Wang LR, Zhu GR, Fang WC, Chen CW, Zhao P (2011) Construction of a linkage map and identification of resistance gene analog markers for root-knot nematodes in wild peach, *Prunus kansuensis*. Journal of the American Society for Horticultural Science 136 (3):190-197

Castillo NRF, Reed BM, Graham J, Fernandez-Fernandez F, Bassil NV (2010) Microsatellite markers for raspberry and blackberry. Journal of the American Society for Horticultural Science 135 (3):271-278

Castle JC (2011) SNPs occur in regions with less genomic sequence conservation. Public Library of Science One 6 (6):e20660. doi:10.1371/journal.pone.0020660

Cato SA, Gardner RC, Kent J, Richardson TE (2001) A rapid PCR-based method for genetically mapping ESTs Theoretical and Applied Genetics 102 (2-3):296-306

Celton J-M, Chagné D, Tustin SD, Terakami S, Nishitani C, Yamamoto T, Gardiner SE (2009a) Update on comparative genome mapping between *Malus* and *Pyrus*. BMC Research Notes 2 (182)

Celton J-M, Tustin DS, Chagné D, Gardiner SE (2009b) Construction of a dense genetic linkage map for apple rootstocks using SSRs developed from *Malus* ESTs and *Pyrus* genomic sequences. Tree Genetics & Genomes 5:93-107

Chagné D (2008) Amazing LightCycler 480 applications: Application of the High-Resolution Melting technique for mapping genes in apple. Roche Applied Science

Chagné D, Brown G, Lalanne C, Madur D, Pot D, Neale D, Plomion C (2003) Comparative genome and QTL mapping between maritime and loblolly pines. Molecular Breeding 12 (3):185-195

Chagné D, Chaumeil P, Ramboer A, Collada C, Guevara A, Cervera MT, Vendramin GG, Garcia V, Frigerio J-M, Echt C, Richardson T, Plomion C (2004) Cross-species transferability and mapping of genomic and cDNA SSRs in pines. Theoretical and Applied Genetics 109: 1204–1214.

Chagné D, Crowhurst RN, Troggio M, Davey MW, Gilmore B, Lawley C, Vanderzande S, Hellens RP, Kumar S, Cestaro A, Velasco R, Main D, Rees J, Iezzoni A, Mockler TC, Wilhelm LJ, Van de Weg E, Gardiner SE, Bassil NV, Peace C (2012) Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. Public Library of Science One

Chagné D, Gasic K, Crowhurst RN, Han Y, Bassett HC, Bowatte DR, Lawrence TJ, Rikkerink EHA, Gardiner SE, Korban SS (2008) Development of a set of SNP markers present in expressed genes of the apple. Genomics 92 (5):353-358

Chagné D, Krieger C, Rassam M, Sullivan M, Fraser J, Andre C, Pindo M, Troggio M, Gardiner S, Henry R, Allan A, McGhie T, Laing W (2012) QTL and candidate gene mapping for polyphenolic composition in apple fruit. BMC Plant Biology 12 (1):12

Chakraborty R, Weiss KM (1988) Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. Proceedings of the National Academy of Sciences 85 (23):9119-9123

Chaparro JX, Werner DJ, O'Malley D, Sederoff RR (1994) Targeted mapping and linkage analysis of morphological, isozyme, and RAPD markers in peach. Theoretical and Applied Genetics 87 (7):805-815

Chen HS, Liu M, Shi LJ, Zhao JL, Zhang CP, Lin LQ, Liu Y, Zhang SJ, Jin JC, Wang L, Shen BZ, Liu JR (2011) Effects of raspberry phytochemical extract on cell proliferation, apoptosis, and serum proteomics in a rat model. Journal of Food Science 76 (8):T192-T198. doi:10.1111/j.1750-3841.2011.02373.x

Chen L, Zhang S, Illa E, Song L, Wu S, Howad W, Arús P, van de Weg E, Chen K, Gao Z (2008) Genomic characterization of putative allergen genes in peach/almond and their synteny with apple. BMC Genomics 9 (1):543

Chou L-S, Lyon E, Wittwer CT (2005) A comparison of High-Resolution Melting analysis with Denaturing High-Performance Liquid Chromatography for mutation scanning: Cystic fibrosis transmembrane conductance regulator gene as a model. American Journal of Clinical Pathology 124 (3):330-338

Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138 (3):963-971

Cipriani G, Lot G, Huang WG, Marrazzo MT, Peterlunger E, Testolin R (1999) AC/GT and AG/CT microsatellite repeats in peach [*Prunus persica* (L.) Batsch]: isolation, characterisation and cross-species amplification in *Prunus*. Theoretical and Applied Genetics 99 (1):65-72

Comai L (2005) The advantages and disadvantages of being polyploid. Nature Reviews Genetics 6:836-846

Condit R, Hubbell SP (1991) Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome 34 (1):66-71. doi:10.1139/g91-011

Connor AM, Finn CE, McGhie TK, Alspach PA (2005) Genetic and environmental variation in anthocyanins and their relationship to antioxidant activity in blackberry and hybridberry cultivars. Journal of the American Society for Horticultural Science 130 (5):680-687

Cooney JM, Jensen DJ, McGhie TK (2004) LC-MS identification of anthocyanins in boysenberry extract and anthocyanin metabolites in human urine following dosing. Journal of the Science of Food and Agriculture 84 (3):237-245. doi:10.1002/jsfa.1645

Costa F, Costa G, Sansavini S, Soglio V, Gianfranceschi L, Schouten HJ, Alba R, Giovannoni J (2009) Heterologous comparative genomics to identify candidate genes impacting fruit quality in apple (*Malus* x *domestica* Borkh.). Acta Horticulturae (ISHS) 814:517-522

Costantini L, Madini A, Battilana J, Grando MS (2005) Comparative mapping of quantitative traits in grape. In: Williams LE (ed) Proceedings of the Seventh International Symposium on Grapevine Physiology and Biotechnology. Acta Horticulturae, vol 689. pp 409-415

Crane MB, Lawrence WJC (1929) Genetical and cytological aspects of incompatibility and sterility in cultivated fruits. Journal of Pomology and Horticultural Science 7:276-301

Cregan PB, Bhagwat AA, Akkaya MS, Rongwen J (1994) Microsatellite fingerprinting and mapping of soybean. Methods in Molecular and Cellular Biology 5 (1):49-61

Crepet WL, Nixon KC, Gandolfo MA (2004) Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. American Journal of Botany 91 (10):1666-1682. doi:10.3732/ajb.91.10.1666

Crepieux S, Lebreton C, Flament P, Charmet G (2005) Application of a new IBD-based QTL mapping method to common wheat breeding population: Analysis of kernel hardness and dough strength Theoretical and Applied Genetics 111 (7):1409-1419

Crespel LC, Chirollet MC, Durel CD, Zhang DZ, Meynet JM, Gudin SG (2002) Mapping of qualitative and quantitative phenotypic traits in *Rosa* using AFLP markers. Theoretical and Applied Genetics 105 (8):1207-1214. doi:10.1007/s00122-002-1102-2

Crow JF (1990) Mapping Functions. Genetics 125 (4):669-671

Dale A, Moore PP, McNicol RJ, Sjulin TM, Burmistrov LA (1993) Genetic diversity of red raspberry varieties throughout the world. Journal of the American Society for Horticultural Science 118 (1):119-129

Davis TM, Yu H (1997) A linkage map of the diploid strawberry, *Fragaria vesca*. Journal of Heredity 88 (3):215-221

Debener T, Mattiesch L (1999) Construction of a genetic linkage map for roses using RAPD and AFLP markers. Theoretical and Applied Genetics 99 (5):891-899. doi:10.1007/s001220051310

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theoretical and Applied Genetics 106 (5):912-922

Deng C, Davis TM (2001) Molecular identification of the yellow fruit color (*c*) locus in diploid strawberry: A candidate gene approach. Theoretical and Applied Genetics 103 (2):316-322

Devos KM, Gale MD (2000) Genome relationships: The grass model in current research. Plant Cell 12 (5):637-646. doi:10.1105/tpc.12.5.637

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

Dirlewanger E, Cosson P, Howad W, Capdeville G, Bosselut N, Claverie M, Voisin R, Poizat C, Lafargue B, Baron O, Laigret F, Kleinhentz M, Arús P, Esmenjaud D (2004a) Microsatellite genetic linkage maps of myrobalan plum and an almond-peach hybrid—location of root-knot nematode resistance genes. Theoretical and Applied Genetics 109 (4):827-838

Dirlewanger E, Cosson P, Tavaud M, Aranzana M, Poizat C, Zanetto A, Arús P, Laigret F (2002) Development of microsatellite markers in peach [*Prunus persica* (L.) Batsch] and their use in genetic diversity analysis in peach and sweet cherry (*Prunus avium* L.). Theoretical and Applied Genetics 105 (1):127-138

Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldera F, Cosson P, Howad W, Arús P (2004b) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. Proceedings of the National Academy of Sciences of the United States of America 101 (26):9891-9896. doi:10.1073/pnas.0307937101

Dobson P, Graham J, Stewart D, Brennan R, Hackett CA, McDougall GJ (2012) Over-seasons analysis of quantitative trait loci affecting phenolic content and antioxidant capacity in raspberry. Journal of Agricultural and Food Chemistry 60 (21):5360-5366. doi:10.1021/jf3005178

Dossett M, Lee J, Finn CE (2010) Variation in anthocyanins and total phenolics of black raspberry populations. Journal of Functional Foods 2 (4):292-297. doi:10.1016/j.jff.2010.10.004

Dossett M, Lee J, Finn CE (2011) Characterization of a novel anthocyanin profile in wild black raspberry mutants: An opportunity for studying the genetic control of pigment and color. Journal of Functional Foods 3 (3):207-214. doi:10.1016/j.jff.2011.04.003

Downey SL, Iezzoni AF (2000) Polymorphic DNA markers in black cherry (*Prunus serotina*) are identified using sequences from sweet cherry, peach, and sour cherry. Journal of the American Society for Horticultural Science 125 (1):76-80

Doyle JJ, Davis JI (1998) Homology in molecular phylogenetics: A parsimony perspective. In: Soltis DE, Soltis PS, Doyle JJ (eds) Molecular Systematics of Plants II: DNA Sequencing. Klewer Academic Publishers, Norwell, MA,

Doyle JJ, Gaut BS (2000) Evolution of genes and taxa: a primer. Plant Molecular Biology 42 (1):1-23. doi:10.1023/a:1006349518932

Draborg H, Villadsen D, Hamborg Nielsen T (1999) Cloning, characterization and expression of a bifunctional fructose-6-phosphate, 2-kinase/fructose-2,6-bisphosphatase from potato. Plant Molecular Biology 39 (4):709-720. doi:10.1023/a:1006102412693

Durel CE, Denancé C, Brisset MN (2009) Two distinct major QTL for resistance to fire blight co-localize on linkage group 12 in apple genotypes 'Evereste' and *Malus floribunda* clone 821. Genome 52 (2):139-147. doi:10.1139/g08-111

Edwards KJ, Mogg R (2001) Plant genotyping by analysis of Single Nucleotide Polymorphisms. In: Henry RJ (ed) Plant Genotyping: The DNA Fingerprinting of Plants. CABI, Oxon and New York,

Ellwood S, Phan H, Jordan M, Hane J, Torres A, Avila C, Cruz-Izquierdo S, Oliver R (2008) Construction of a comparative genetic map in faba bean (*Vicia faba* L.); conservation of genome structure with *Lens culinaris*. BMC Genomics 9 (1):380

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6 (5):e19379. doi:10.1371/journal.pone.0019379

Eriksson T, Hibbs MS, Yoder AD, Delwiche CF, Donoghue MJ (2003) The phylogeny of Rosoideae (Rosaceae) based on sequences of the internal transcribed spacers (ITS) of nuclear ribosomal DNA and the *TrnL/F* region of chloroplast DNA. International Journal of Plant Science 164 (2):197-211. doi:1058-5893/2003/16402-0001

Espín JC, García-Conesa MT, Tomás-Barberán FA (2007) Nutraceuticals: Facts and fiction. Phytochemistry 68 (22-24):2986-3008. doi:10.1016/j.phytochem.2007.09.014

Espley RV, Hellens RP, Putterill J, Stevenson DE, Kutty-Amma S, Allan AC (2007) Red colouration in apple fruit is due to the activity of the *MYB* transcription factor, *MdMYB10*. Plant Journal 49 (3):414-427

Etienne C, Rothan C, Moing A, Plomion C, Bodénès C, Svanella-Dumas L, Cosson P, Pronier V, Monet R, Dirlewanger E (2002) Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]. Theoretical and Applied Genetics 105 (1):145-159. doi:10.1007/s00122-001-0841-9

Eujayl I, Sorrells ME, Baum M, Wolters P, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theoretical and Applied Genetics 104 (2):399-407

Evans RC, Campbell CS (2002) The origin of the apple subfamily (Maloideae; Rosaceae) is clarified by DNA sequence data from duplicated *GBSSI* genes. American Journal of Botany 89 (9):1478-1484. doi:10.3732/ajb.89.9.1478

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics, 4th Ed. Pearson Education Limited, Longmans Green, Harlow, UK

Fan S, Bielenberg DG, Zhebentyayeva TN, Reighard GL, Okie WR, Holland D, Abbott AG (2010) Mapping quantitative trait loci associated with chilling requirement, heat requirement and bloom date in peach (*Prunus persica*). New Phytologist 185 (4):917-930. doi:10.1111/j.1469-8137.2009.03119.x

Fernández-Fernández F, Evans K, Clarke J, Govan C, James C, Marič S, Tobutt K (2008) Development of an STS map of an interspecific progeny of *Malus*. Tree Genetics & Genomes 4 (3):469-479

Fernández i Martí À, Howad W, Tao R, Alonso Segura J, Arús P, Socias i Company R (2011) Identification of quantitative trait loci associated with self-compatibility in a *Prunus* species. Tree Genetics & Genomes 7 (3):629-639. doi:10.1007/s11295-010-0362-2

Feuillet C, Keller B (2002) Comparative genomics in the grass family: Molecular characterization of grass genome structure and evolution. Annals of Botany 89 (1):3-10. doi:10.1093/aob/mcf008

Finley JW, Kong A-N, Hintze KJ, Jeffery EH, Ji LL, Lei XG (2011) Antioxidants in foods: State of the science important to the food industry. Journal of Agricultural and Food Chemistry:6837-6846. doi:10.1021/jf2013875

Fitch WM (1970) Distinguishing homologous from analogous proteins. Systematic Zoology 19 (2):99-113

Folta KM, Dhingra A, Howard L, Stewart PJ, Chandler CK (2006) Characterization of LF9, an octoploid strawberry genotype selected for rapid regeneration and transformation. Planta 224 (5):1058-1067. doi:10.1007/s00425-006-0278-0

Foolad MR, Arulsekar S, Becerra V, Bliss FA (1995) A genetic map of *Prunus* based on an interspecific cross between peach and almond. Theoretical and Applied Genetics 91 (2):262-269

Fulton TM, Van der Hoeven R, Eannetta NT, Tanksley SD (2002) Identification, analysis, and utilization of Conserved Ortholog Set markers for comparative genomics in higher plants. Plant Cell 14 (7):1457-1467. doi:10.1105/tpc.010479

Gardiner S, Norelli J, de Silva N, Fazio G, Peil A, Malnoy M, Horner M, Bowatte D, Carlisle C, Wiedow C, Wan Y, Bassett C, Baldo A, Celton J-M, Richter K, Aldwinckle H, Bus V (2012) Putative resistance gene markers associated with quantitative trait loci for fire blight resistance in Malus 'Robusta 5' accessions. BMC Genetics 13 (1):25

Gardiner SE, Norelli JL, de Silva N, Fazio G, Peil A, Malnoy M, Horner M, Bowatte D, Carlisle C, Wiedow C, Wan Y, Bassett CL, Baldo AM, Celton J-M, Richter K, Aldwinckle HS, Bus VGM (in prep) Candidate gene markers associated with quantitative trait loci for fire blight resistance in *Malus* 'Robusta 5' accessions. BMC Plant Biology

Gasic K, Gonzalez DO, Thimmapuram J, Liu L, Malnoy M, Gong G, Han Y, Vodkin LO, Aldwinckle HS, Carroll NJ, Orvis KS, Goldsbrough P, Clifton S, Pape D, Fulton L, Martin J, Theising B, Wisniewski ME, Fazio G, Feltus FA, Korban SS (2009a) Comparative analysis and functional annotation of a large expressed sequence tag collection of apple. Plant Genetics 2 (1):23-38. doi:10.3835/plantgenome2008.11.0014

Gasic K, Han Y, Kertbundit S, Shulaev V, Iezzoni A, Stover E, Bell R, Wisniewski M, Korban S (2009b) Characteristics and transferability of new apple EST-derived SSRs to other Rosaceae species. Molecular Breeding 23 (3):397-411

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998) Simple sequence repeats for the genetic analysis of apple. Theoretical and Applied Genetics 96 (8):1069-1076

Gil-Ariza DJ, Amaya I, Botella MA, Blanco JM, Caballero JL, López-Aranda JM, Valpuesta V, Sánchez-Sevilla JF (2006) EST-derived polymorphic microsatellites from cultivated strawberry (*Fragaria* × *ananassa*) are useful for diversity studies and varietal identification among *Fragaria* species. Molecular Ecology Notes 6 (4):1195-1197. doi:10.1111/j.1471-8286.2006.01489.x

Gisbert A, Martínez-Calvo J, Llácer G, Badenes M, Romero C (2009) Development of two loquat [*Eriobotrya japonica* (Thunb.) Lindl.] linkage maps based on AFLPs and SSR markers from different Rosaceae species. Molecular Breeding 23 (3):523-538

Graham J, Hackett C, Smith K, Woodhead M, Hein I, McCallum S (2009) Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit. Theoretical and Applied Genetics 118 (6):1143-1155. doi:10.1007/s00122-009-0969-6

Graham J, Hackett C, Smith K, Woodhead M, MacKenzie K, Tierney I, Cooke D, Bayer M, Jennings N (2011) Towards an understanding of the nature of resistance to *Phytophthora* root rot in red raspberry. Theoretical and Applied Genetics 123 (4):585-601. doi:10.1007/s00122-011-1609-5

Graham J, Smith K, MacKenzie K, Jorgenson L, Hackett C, Powell W (2004) The construction of a genetic linkage map of red raspberry (*Rubus idaeus* subsp. *idaeus*) based on AFLPs, genomic-SSR and EST-SSR markers. Theoretical and Applied Genetics 109 (4):740-749

Graham J, Smith K, Tierney I, MacKenzie K, Hackett C (2006) Mapping gene H controlling cane pubescence in raspberry and its association with resistance to cane botrytis and spur blight, rust and cane spot. Theoretical and Applied Genetics 112 (5):818-831

Graham J, Smith K, Woodhead M, Russell J (2002) Development and use of simple sequence repeat SSR markers in *Rubus* species. Molecular Ecology Notes 2 (3):250-252. doi:10.1046/j.1471-8286.2002.00203.x

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

Grimplet J, Romieu C, Audergon J-M, Marty I, Albagnac G, Lambert P, Bouchet J-P, Terrier N (2005) Transcriptomic study of apricot fruit (*Prunus armeniaca*) ripening among 13 006 expressed sequence tags. Physiologia Plantarum 125 (3):281-292. doi:10.1111/j.1399-3054.2005.00563.x

Grotewold E (2006) The genetics and biochemistry of floral pigments. Annual Review of Plant Biology 57:761-780. doi:10.1146/annurev.arplant.57.032905.105248

Grube RC, Radwanski ER, Jahn M (2000) Comparative genetics of disease resistance within the Solanaceae. Genetics 155 (2):873-887

Guilford P, Prakash S, Zhu JM, Rikkerink E, Gardiner S, Bassett H, Forster R (1997) Microsatellites in *Malus* x *domestica* (apple): Abundance, polymorphism and cultivar identification. Theoretical and Applied Genetics 94 (2):249-254

Guitton B, Kelner J-J, Velasco R, Gardiner SE, Chagné D, Costes E (2012) Genetic control of biennial bearing in apple. Journal of Experimental Botany 63 (1):131-149. doi:10.1093/jxb/err261

Gupta P, Balyan H, Edwards K, Isaac P, Korzun V, Röder M, Gautier MF, Joudrier P, Schlatter A, Dubcovsky J, De la Pena R, Khairallah M, Penner G, Hayden M, Sharp P, Keller B, Wang R, Hardouin J, Jack P, Leroy P (2002a) Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. Theoretical and Applied Genetics 105 (2):413-422

Gupta P, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. Functional & Integrative Genomics 4 (3):139-162

Gupta PK, Roy JK, Prasad M (2001) Single nucleotide polymorphisms: A new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Current Science 80 (4):524-535

Gupta PK, Varshney RK, Prasad M (2002b) Molecular Markers: Principles and Methodology. In: Jain SM, Brar DS, Ahloowalia BS (eds) Molecular Techniques in Crop Improvement. Kluwer Academic Publishers, Dordrecht and Boston, pp 9-54

Gygax M, Gianfranceschi L, Liebhard R, Kellerhals M, Gessler C, Patocchi A (2004) Molecular markers linked to the apple scab resistance gene *Vbj* derived from *Malus baccata jackii*. Theoretical and Applied Genetics 109 (8):1702-1709. doi:10.1007/s00122-004-1803-9

Hackauf B, Rudd S, van der Voort J, Miedaner T, Wehling P (2009) Comparative mapping of DNA sequences in rye (*Secale cereale* L.) in relation to the rice genome. Theoretical and Applied Genetics 118 (2):371-384

Hagen LS, Chaib J, Fady B, Decroocq V, Bouchet JP, Lambert P, Audergon JM (2004) Genomic and cDNA microsatellites from apricot (*Prunus armeniaca* L.). Molecular Ecology Notes 4 (4):742-745

Hardison RC (2003) Comparative Genomics. Public Library of Science, Biology 1 (2):e58

Hartl DL, Jones EW (2009) Genetics. 7th edn. Jones and Bartlett, Sudbury

Hemmat M, Weedon NF, Manganaris AG, Lawson DM (1994) Molecular marker linkage map for apple. Journal of Heredity 85 (1):4-11

Hibrand-Saint Oyant L, Crespel L, Rajapakse S, Zhang L, Foucher F (2008) Genetic linkage maps of rose constructed with new microsatellite markers and locating QTL controlling flowering traits. Tree Genetics & Genomes 4:11-23

High-quality Disease Resistant Apples for a Sustainable Agriculture, http://users.unimi.it/hidras/. http://users.unimi.it/hidras/.

Hokanson SC, Szewc-McFadden AK, Lamboy WF, McFerson JR (1998) Microsatellite (SSR) markers reveal genetic identities, genetic diversity and relationships in a *Malus* × *domestica* Borkh. core subset collection. Theoretical and Applied Genetics 97 (5):671-683

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7 (7):1071-1083. doi:10.1105/tpc.7.7.1071

Howad W, Yamamoto T, Dirlewanger E, Testolin R, Cosson P, Cipriani G, Monforte AJ, Georgi L, Abbott AG, Arús P (2005) Mapping with a few plants: Using selective mapping for microsatellite saturation of the *Prunus* reference map. Genetics 171 (3):1305-1309. doi:10.1534/genetics.105.043661

Huehn M (2011) On the bias of recombination fractions, Kosambi's and Haldane's distances based on frequencies of gametes. Genome 54 (3):196-201. doi:10.1139/g10-109

Hummer KE (2010) Rubus pharmacology: Antiquity to the present. HortScience 45 (11):1587-1591

Hummer KE, Janick J (2009) Rosaceae: Taxonomy, economic importance, genomics. In: Folta K, Gardiner S (eds) Genetics and genomics of Rosaceae, vol 6. Springer, New York, pp 1-17

Igarashi M, Abe Y, Hatsuyama Y, Ueda T, Fukasawa-Akada T, Kon T, Kudo T, Sato T, Suzuki M (2008) Linkage maps of the apple (*Malus* × *domestica* Borkh.) cvs 'Ralls Janet' and 'Delicious' include newly developed EST markers. Molecular Breeding 22 (1):95-118

Illa E, Sargent D, Lopez Girona E, Bushakra J, Cestaro A, Crowhurst R, Pindo M, Cabrera A, van der Knaap E, Iezzoni A, Gardiner S, Velasco R, Arús P, Chagne D, Troggio M (2011) Comparative analysis of rosaceous genomes and the reconstruction of a putative ancestral genome for the family. BMC Evolutionary Biology 11 (1):9

Illumina I (2006) Illumina (R) SNP genotyping.

Iovene M, Wielgus SM, Simon PW, Buell CR, Jiang J (2008) Chromatin structure and physical mapping of chromosome 6 of potato and comparative analyses with tomato. Genetics 180 (3):1307-1317. doi:10.1534/genetics.108.093179

James CM, Wilson F, Hadonou AM, Tobutt KR (2003) Isolation and characterization of polymorphic microsatellites in diploid strawberry (*Fragaria vesca* L.) for mapping, diversity studies and clone identification. Molecular Ecology Notes 3 (2):171-173. doi:10.1046/j.1471-8286.2003.00365.x

Jansen RC, Stam P (1994) High resolution of quantitative traits into multiple loci via interval mapping. Genetics 136 (4):1447-1455

Jáuregui B, de Vicente MC, Messeguer R, Felipe A, Bonnet A, Salesses G, Arús P (2001) A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. Theoretical and Applied Genetics 102 (8):1169-1176

Jennings DL, Carmichael E (1980) Anthocyanin variation in the genus *Rubus*. New Phytologist 84 (3):505-513

Jonkers H (1971) An international experiment on juvenility in apple. Euphytica 20 (1):57-59

Joobeur T, Periam N, Vicente MCd, King GJ, Arús P (2000) Development of a second generation linkage map for almond using RAPD and SSR markers. Genome 43 (4):649-655. doi:10.1139/g00-040

Joobeur T, Viruel MA, de Vicente MC, Jáuregui B, Ballester J, Dettori MT, Verde I, Truco MJ, Messeguer R, Batlle I, Quarta R, Dirlewanger E, Arús P (1998) Construction of a saturated linkage map for *Prunus* using an almond × peach F_2 progeny. Theoretical and Applied Genetics 97 (7):1034-1041

Jung S, Cestaro A, Troggio M, Main D, Zheng P, Cho I, Folta KM, Sosinski B, Abbott A, Celton J-M, Arús P, Shulaev V, Verde I, Morgante M, Rokhsar DS, Velasco R, Sargent DJ (in press) Whole genome comparisons of *Fragaria*, *Prunus*, and *Malus* reveal different modes of evolution between rosaceous subfamilies. BMC Genomics

Kafkas E, Özgen M, Özoğui Y, Türemiş N (2008) Phytochemical and fatty acid profile of selected red raspberry cultivars: a comparative study. Journal of Food Quality 31 (1):67-78. doi:10.1111/j.1745-4557.2007.00184.x

Kaló P, Seres A, Taylor S, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis T, Kiss G (2004) Comparative mapping between *Medicago sativa* and *Pisum sativum*. Molecular Genetics and Genomics 272 (3):235-246

Kassim A, Poette J, Paterson A, Zait D, McCallum S, Woodhead M, Smith K, Hackett C, Graham J (2009) Environmental and seasonal influences on red raspberry anthocyanin antioxidant contents and identification of quantitative traits loci (QTL). Molecular Nutrition & Food Research 53 (5):625-634

Kellerhals M (2009) Introduction to Apple (*Malus* x *domestica*). In: Folta KM, Gardiner SE (eds) Plant Genetics and Genomics: Crops and Models, vol 6. Plant Genetics and Genomics: Crops and Models. Springer, New York, pp 73-84

Keniry A, Hopkins CJ, Jewell E, Morrison B, Spangenberg GC, Edwards D, Batley J (2006) Identification and characterization of simple sequence repeat (SSR) markers from Fragaria×ananassa expressed sequences. Molecular Ecology Notes 6 (2):319-322. doi:10.1111/j.1471-8286.2005.01215.x

Khan MA, Durel C-E, Duffy B, Drouet D, Kellerhals M, Gessler C, Patocchi A (2007) Development of molecular markers linked to the 'Fiesta' linkage group 7 major QTL for fire blight resistance and their application for marker-assisted selection. Genome 50 (6):568-577. doi:10.1139/g07-033

Khan MA, Han Y, Zhao YF, Korban SS (2011) A high-throughput apple SNP genotyping platform using the GoldenGate™ assay. Gene. doi:10.1016/j.gene.2011.12.001

Khan SA, Chibon P-Y, de Vos RCH, Schipper BA, Walraven E, Beekwilder J, van Dijk T, Finkers R, Visser RGF, van de Weg EW, Bovy A, Cestaro A, Velasco R, Jacobsen E, Schouten HJ (2012) Genetic analysis of metabolites in apple fruits indicates an mQTL hotspot for phenolic compounds on linkage group 16. Journal of Experimental Botany 63 (8):2895-2908. doi:10.1093/jxb/err464

Kimura T, Shi YZ, Shoda M, Kotobuki K, Matsuta N, Hayashi T, Ban Y, Yamamoto T (2002) Identification of Asian pear varieties by SSR analysis. Breeding Science 52 (2):115-121

Kobayashi N, Horikoshi T, Katsuyama H, Handa T, Takayanagi K (1998) A simple and efficient DNA extraction method for plants, especially woody plants. Plant Tissue Culture and Biotechnology 4 (2):76-80

Koonin E (2001) An apology for orthologs - or brave new memes. Genome Biology 2 (4):comment1005.1001 - comment1005.1002

Koonin EV, Aravind L, Kondrashov AS (2000) The impact of comparative genomics on our understanding of evolution. Cell 101 (6):573-576

Kouassi AB, Durel CE, Mathis F, Gianfranceschi L, Komjanc M, Mott D, Cova V, Patocchi A, Gobbin D, Rezzonico F, Evans K, Fernandez-Fernandez F, Dunemann F, Boudichevskaja A, Stankiewicz-Kosyl M, Antofie A, Weg Evd, Bink M, Laurens F, de Weg Ev, van de Weg E (2008) QTL mapping for fruit firmness in apple using a pedigree-based Bayesian approach. Modern variety breeding for present and future needs. Proceedings of the 18th EUCARPIA general congress, Valencia, Spain, 9-12 September, 2008.

Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB (2004) Comparative mapping in the Pinaceae. Genetics 168 (1):447-461. doi:10.1534/genetics.104.028381

Kuhn DN, Chappell J, Boudet A, Hahlbrock K (1984) Induction of phenylalanine ammonia-lyase and 4-coumarate:CoA ligase mRNAs in cultured plant cells by UV light or fungal elicitor. Proceedings of the National Academy of Sciences 81 (4):1102-1106

Kumar S, Bink MCAM, Volz RK, Bus VGM, Chagné D (2012) Towards genomic selection in apple (*Malus* x *domestica* Borkh.) breeding programmes: Prospects, challenges and strategies. Tree Genetics & Genomes 8 (1):1-14. doi:10.1007/s11295-011-0425-z

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150 (3):1217-1228

Lagercrantz U, Putterill J, Coupland G, Lydiate D (1996) Comparative mapping in *Arabidopsis* and *Brassica*, fine scale genome collinearity and congruence of genes controlling flowering time. Plant Journal 9 (1):13-20

Lambert P, Hagen LS, Audergon JM, Arús A (2006) Evaluation of a mapping strategy based on synteny: Using the *Prunus* reference map for apricot. Acta Horticulturae (ISHS) 701:295-301

Lander ES, Botstein D (1989) Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121 (1):185-199

Lander ES, Kruglyak L (1995) Genetic dissection of complex traits:guidelines for interpreting and reporting linkage results. Nature Genetics 11:241-247

Lansari A, Kester DE, Iezzoni AF (1994) Inbreeding, coancestry, and founding clones of almonds of California, Mediterranean shores, and Russia. Journal of the American Society for Horticultural Science 119 (6):1279-1285

Lerceteau-Kohler E, Moing A, Guerin G, Renaud C, Courlit S, Camy D, Praud K, Parisy V, Bellec F, Rolin D, Roudeillac P, Denoyes-Rothan B (2004) QTL analysis for fruit quality traits in octoploid strawberry (*Fragaria* x *ananassa*). Acta Horticulturae (ISHS) 663:331-335

Lewers K, Saski C, Cuthbertson B, Henry D, Staton M, Main D, Dhanaraj A, Rowland L, Tomkins J (2008) A blackberry (*Rubus* L.) expressed sequence tag library for the development of simple sequence repeat markers. BMC Plant Biology 8 (1):69-76

Lewers KS, Styan SMN, Hokanson SC, Bassil NV (2005) Strawberry GenBank-derived and genomic simple sequence repeat (SSR) markers and their utility with strawberry, blackberry, and red and black raspberry. Journal of the American Society for Horticultural Science 130 (1):102-115

Liang XW, Dron M, Cramer CL, Dixon RA, Lamb CJ (1989) Differential regulation of phenylalanine ammonia-lyase genes during plant development and by environmental cues. Journal of Biological Chemistry 264 (24):14486-14492

Liebhard R, Gianfranceschi L, Koller B, Ryder CD, Tarchini R, Van De Weg E, Gessler C (2002) Development and characterisation of 140 new microsatellites in apple (*Malus* x *domestica* Borkh.). Molecular Breeding 10 (4):217-241

Liebhard R, Koller B, Gianfranceschi L, Gessler C (2003a) Creating a saturated reference map for the apple (*Malus* x *domestica* Borkh.) genome. Theoretical and Applied Genetics 106:1497 - 1508

Liebhard R, Koller B, Patocchi A, Kellerhals M, Pfammatter W, Jermini M, Gessler C (2003b) Mapping quantitative field resistance against apple scab in a 'Fiesta' x 'Discovery' progeny. Phytopathology 93:493 - 501

Liewlaksaneeyanawin C, Zhuang J, Tang M, Farzaneh N, Lueng G, Cullis C, Findlay S, Ritland C, Bohlmann J, Ritland K (2009) Identification of COS markers in the Pinaceae. Tree Genetics & Genomes 5 (1):247-255

Lim E-K, Ashford DA, Hou B, Jackson RG, Bowles DJ (2004) Arabidopsis glycosyltransferases as biocatalysts in fermentation for regioselective synthesis of diverse quercetin glucosides. Biotechnology and Bioengineering 87 (5):623-631. doi:10.1002/bit.20154

Lin-Wang K, Bolitho K, Grafton K, Kortstee A, Karunairetnam S, McGhie T, Espley R, Hellens R, Allan A (2010) An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC Plant Biology 10 (1):50

Lin H, Walker MA (1998) Identifying grape rootstocks with simple sequence repeat (SSR) DNA markers. American Journal of Enology and Viticulture 49 (4):403-407

Lin J-J, Kuo J, Ma J, Saunders J, Beard H, MacDonald M, Kenworthy W, Ude G, Matthews B (1996) Identification of molecular markers in soybean comparing RFLP, RAPD and AFLP DNA mapping techniques. Plant Molecular Biology Reporter 14 (2):156-169. doi:10.1007/bf02684905

Livingstone III D, Royaert S, Mockaitis K, May GD, Farmer A, Saski C, Schnell RJ, Motamayor JC, Kuhn D (2012) Identification, analysis and mapping of SNP markers in *Theobroma cacao*. Paper presented at the Plant & Animal Genome XX, San Diego, CA, 14-18 January 2012

Livingstone KD, Lackney VK, Blauth JR, van Wijk R, Jahn MK (1999) Genome mapping in *Capsicum* and the evolution of genome structure in the Solanaceae. Genetics 152 (3):1183-1202

Lopes MS, Maciel GB, Mendonça D, Gil FS, Da Câmara Machado A (2006) Isolation and characterization of simple sequence repeat loci in *Rubus hochstetterorum* and their use in other species from the Rosaceae family. Molecular Ecology Notes 6 (3):750-752. doi:10.1111/j.1471-8286.2006.01329.x

Lu Y, Shah T, Hao Z, Taba S, Zhang S, Gao S, Liu J, Cao M, Wang J, Prakash AB, Rong T, Xu Y (2011) Comparative SNP and haplotype analysis reveals a higher genetic diversity and rapider LD decay in tropical than temperate germplasm in maize. Public Library of Science One 6 (9):e24861. doi:10.1371/journal.pone.0024861

Lu Z-X, Sosinski B, Reighard GL, Baird WV, Abbott AG (1998) Construction of a genetic linkage map and identification of AFLP markers for resistance to root-knot nematodes in peach rootstocks. Genome 41 (2):199-207. doi:10.1139/g98-008

Määttä-Riihinen KR, Kamal-Eldin A, Törrönen AR (2004) Identification and quantification of phenolic compounds in berries of *Fragaria* and *Rubus* species (Family Rosaceae). Journal of Agricultural and Food Chemistry 52 (20):6178-6187. doi:10.1021/jf049450r

Mackay TFC (2001) The genetic architecture of quantitative traits. Annual Review of Genetics 35:303 - 339

Maliepaard C, Alston FH, Van Arkel G, Brown LM, Chevreau E, Dunemann F, Evans KM, Gardiner S, Guilford P, Van Heusden AW, Janse J, Laurens F, Lynn JR, Manganaris AG, Den Nijs APM, Periam N, Rikkerink E, Roche P, Ryder C, Sansavini S, Schmidt H, Tartarini S, Verhaegh JJ, Vrielink-van Ginkel M, King GJ (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theoretical and Applied Genetics 97:60 - 73

Martinez Garcia PJ, Parfitt D, Ogundiwin E, Fass J, Chan HM, Gradziel T, Crisosto CH (2012) High-density SNP mapping on peach (*Prunus persica* L. Batsch) populations. Paper presented at the Plant and Animal Genome XX, San Diego, CA, 14-18 Jan 2012

Mauricio R (2001) Mapping quantitative trait loci in plants: Uses and caveats for evolutionary biology. Nature Genetics 2:370-381

McCallum S, Woodhead M, Hackett C, Kassim A, Paterson A, Graham J (2010) Genetic and environmental effects influencing fruit colour and QTL analysis in raspberry. Theoretical and Applied Genetics 121 (4):611-627. doi:10.1007/s00122-010-1334-5

McCouch SR, Doerge RW (1995) QTL mapping in rice. Trends in Genetics 11 (12):482-487. doi:10.1016/s0168-9525(00)89157-x

McGhie TK, Hall HK, Ainge GD, Mowat AD (2002) Breeding *Rubus* cultivars for high anthocyanin content and high antioxidant capacity. Acta Horticulturae (ISHS) 585:495-500

McGregor CE, Lambert CA, Greyling MM, Louw JH, Warnich L (2000) A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. Euphytica 113 (2):135-144

Medina-Escobar N, Cárdenas J, Moyano E, Caballero JL, Muñoz-Blanco J (1997) Cloning, molecular characterization and expression pattern of a strawberry ripening-specific cDNA with sequence homology to pectate lyase from higher plants. Plant Molecular Biology 34 (6):867-877. doi:10.1023/a:1005847326319

Micheletti D, Aramini V, Arús P, Banchi E, Barreneche T, Bassi D, Dirlewanger E, Gao Z, Gazza L, Lambert P, Laurens F, Li X, Pascal T, Quilot-Turion B, Rossini L, Troggio M, Van De Weg E, Verde I, Aranzana MJ (2012) Application of high throughput genotyping techniques in peach germplasm and breeding lines within the FruitBreedomics project. Paper presented at the Plant and Animal Genome XX, San Diego, CA, 14-18 Jan, 2012

Milbourne D, Meyer R, Bradshaw JE, Baird E, Bonar N, Provan J, Powell W, Waugh R (1997) Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato. Molecular Breeding 3 (2):127-136

Mittelsten Scheid O, Jakovleva L, Afsar K, Maluszynska J, Paszkowski J (1996) A change of ploidy can modify epigenetic silencing. Proceedings of the National Academy of Sciences 93 (14):7114-7119

Montgomery J, Wittwer CT, Palais R, Zhou L (2007) Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. Nature Protocols 2 (1):59-66. doi:10.1038/nprot.2007.10

Montrose DC, Horelik NA, Madigan JP, Stoner GD, Wang L-S, Bruno RS, Park HJ, Giardina C, Rosenberg DW (2011) Anti-inflammatory effects of freeze-dried black raspberry powder in ulcerative colitis. Carcinogenesis 32 (3):343-350. doi:10.1093/carcin/bgq248

Moore G, Devos KM, Wang Z, Gale MD (1995) Grasses, line up and form a circle. Current Biology 5:737-739

Mullen W, Yokota T, Lean MEJ, Crozier A (2003) Analysis of ellagitannins and conjugates of ellagic acid and quercetin in raspberry fruit by LC-MS^n. Phytochemistry 64:617-624

Newcomb RD, Crowhurst RN, Gleave AP, Rikkerink EHA, Allan AC, Beuning LL, Bowen JH, Gera E, Jamieson KR, Janssen BJ, Laing WA, McArtney S, Nain B, Ross GS, Snowden KC, Souleyre EJF, Wallon EF, Yauk YK (2006) Analyses of expressed sequence tags from apple. Plant Physiology 141:147 - 166

Ogundiwin E, Peace C, Gradziel T, Parfitt D, Bliss F, Crisosto C (2009) A fruit quality gene map of *Prunus*. BMC Genomics 10 (1):587

Olmstead J, Sebolt A, Cabrera A, Sooriyapathirana S, Hammar S, Iriarte G, Wang D, Chen C, van der Knaap E, Iezzoni A (2008) Construction of an intra-specific sweet cherry (*Prunus avium* L.) genetic linkage map and synteny analysis with the *Prunus* reference map. Tree Genetics & Genomes 4 (4):897-910

Oraguzie NC, Rikkerink EHA, Gardiner SE, de Silva HN (2007) Association mapping in plants. Springer-Verlag GmbH, Heidelberg

Ozgen M, Wyzgoski FJ, Tulio AZ, Gazula A, Miller AR, Scheerens JC, Reese RN, Wright SR (2008) Antioxidant capacity and phenolic antioxidants of midwestern black raspberries grown for direct markets are influenced by production site. HortScience 43 (7):2039-2047

Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theoretical and Applied Genetics 85 (8):985-993. doi:10.1007/bf00215038

Paterson AH (1996) Genome mapping in Plants. Academic Press,

Paterson AH, Bennetzen JL (2001) Comparative mapping of plant chromosomes. In: Phillips RL, Vasil IK (eds) DNA-Based Markers in Plants, vol 6. Advances in Cellular and Molecular Biology of Plants, 2nd edn. Kluwer Academic Publishers, Dordrecht, pp 101-114

Paterson AH, Bowers JE, Burow MD, Draye X, Elsik CG, Jiang C-X, Katsar CS, Lan T-H, Lin Y-R, Ming R, Wright RJ (2000) Comparative genomics of plant chromosomes. Plant Cell 12 (9):1523-1540. doi:10.1105/tpc.12.9.1523

Paterson AH, Bowers JE, Chapman BA (2004) Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proceedings of the National Academy of Sciences of the United States of America 101 (26):9903-9908. doi:10.1073/pnas.0307901101

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M (1998) Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theoretical and Applied Genetics 97 (8):1248-1255

Pelletier MK, Shirley BW (1996) Analysis of flavanone 3-hydroxylase in *Arabidopsis* seedlings (coordinate regulation with chalcone synthase and chalcone isomerase). Plant Physiology 111 (1):339-345. doi:10.1104/pp.111.1.339

Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce-Jacino M (1999) Mining SNPs from EST databases. Genome Research 9:167-174. doi:PMC310719

Pierantoni L, Cho KH, Shin IS, Chiodini R, Tartarini S, Dondini L, Kang SJ, Sansavini S (2004) Characterisation and transferability of apple SSRs to two European pear F_1 populations. Theoretical and Applied Genetics 109 (7):1519-1524

Pierantoni L, Dondini L, Cho KH, Shin IS, Gennari F, Chiodini R, Tartarini S, Kang SJ, Sansavini S (2007) Pear scab resistance QTLs via a European pear (*Pyrus communis*) linkage map. Tree Genetics & Genomes 3 (4):311-317

Potter D, Eriksson T, Evans RC, Oh S, Smedmark JEE, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA, Campbell CS (2007) Phylogeny and classification of Rosaceae. Plant Systematics and Evolution 266 (1):5-43

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends in Plant Science 1 (7):215-222

Rajapakse S, Belthoff LE, He G, Estager AE, Scorza R, Verde I, Ballard RE, Baird WV, Callahan A, Monet R, Abbott AG (1995) Genetic linkage mapping in peach using morphological, RFLP and RAPD markers. Theoretical and Applied Genetics 90:503-510

Reed GH, Wittwer CT (2004) Sensitivity and specificity of single-nucleotide polymorphism scanning by High-Resolution Melting analysis. Clinical Chemistry 50 (10):1748-1754. doi:10.1373/clinchem.2003.029751

Riechmann J, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe O, Samaha R (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290:2105 - 2110

Romero C, Pedryc A, Munoz V, Llácer G, Badenes M (2003) Genetic diversity of different apricot geographical groups determined by SSR markers. Genome 46 (2):244-252

Rusholme Pilcher RL, Celton JM, Gardiner SE, Tustin DS (2008) Genetic markers linked to the dwarfing trait of apple rootstock 'Malling 9'. Journal of the American Society for Horticultural Science 133 (1):100-106

Sargent D, Fernandéz-Fernandéz F, Ruiz-Roja J, Sutherland B, Passey A, Whitehouse A, Simpson D (2009a) A genetic linkage map of the cultivated strawberry (*Fragaria* x *ananassa*) and its comparison to the diploid *Fragaria* reference map. Molecular Breeding 24 (3):293-303. doi:10.1007/s11032-009-9292-9

Sargent D, Marchese A, Simpson D, Howad W, Fernández-Fernández F, Monfort A, Arús P, Evans K, Tobutt K (2009b) Development of "universal" gene-specific markers from *Malus* spp. cDNA sequences, their mapping and use in synteny studies within Rosaceae. Tree Genetics & Genomes 5 (1):133-145

Sargent D, Rys A, Nier S, Simpson D, Tobutt K (2007a) The development and mapping of functional markers in *Fragaria* and their transferability and potential for mapping in other genera. Theoretical and Applied Genetics 114 (2):373-384

Sargent DJ, Cipriani G, Vilanova S, Gil-Ariza D, Arús P, Simpson DW, Tobutt KR, Monfort A (2008) The development of a bin mapping population and the selective mapping of 103 markers in the diploid *Fragaria* reference map. Genome 51 (2):120-127

Sargent DJ, Clarke J, Simpson DW, Tobutt KR, Arús P, Monfort A, Vilanova S, Denoyes-Rothan B, Rousseau M, Folta KM, Bassil NV, Battey NH (2006) An enhanced microsatellite map of diploid *Fragaria*. Theoretical and Applied Genetics 112 (7):1349-1359

Sargent DJ, Davis TM, Simpson DW (2009c) Strawberry (*Fragaria* spp.) Structural Genomics. In: Folta KM, Gardiner SE (eds) Genetics and Genomics of Rosaceae, vol 6. Plant Genetics and Genomics: Crops and Models. Springer New York, pp 437-456. doi:10.1007/978-0-387-77491-6_21

Sargent DJ, Davis TM, Tobutt KR, Wilkinson MJ, Battey NH, Simpson DW (2004) A genetic linkage map of microsatellite, gene-specific and morphological markers in diploid *Fragaria*. Theoretical and Applied Genetics 109 (7):1385-1391

Sargent DJ, Fernández-Fernández F, Rys A, Knight VH, Simpson DW, Tobutt KR (2007b) Mapping of A_1 conferring resistance to the aphid *Amphorophora idaei* and *dw* (dwarfing habit) in red raspberry (*Rubus idaeus* L.) using AFLP and microsatellite markers. BMC Plant Biology 7 (15)

Sargent DJ, Hadonou AM, Simpson DW (2003) Development and characterization of polymorphic microsatellite markers from *Fragaria viridis*, a wild diploid strawberry. Molecular Ecology Notes 3 (4):550-552. doi:10.1046/j.1471-8286.2003.00507.x

Sauge MH, Lambert P, Pascal T (2012) Co-localisation of host plant resistance QTLs affecting the performance and feeding behaviour of the aphid Myzus persicae in the peach tree. Heredity 108 (3):292-301

Sax K (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. Genetics 8 (6):552-560

Scalzo J, Currie A, Stephens J, McGhie T, Alspach P (2008) The anthocyanin composition of different *Vaccinium*, *Ribes*, and *Rubus* genotypes. BioFactors 34:13-21

Schlötterer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Research 20 (2):211-215

Scorza R, Mehlenbacher SA, Lightner GW (1985) Inbreeding and coancestry of freestone peach cultivars of the eastern United States and implications for peach germplasm improvement. Journal of the American Society for Horticultural Science 110:547-552

Scott DH (1951) Cytological studies on polyploids derived from tetraploid *Fragaria vesca* and cultivated strawberries. Genetics 36 (4):311-331

Seeram NP, Momin RA, Nair MG, Bourquin LD (2001) Cyclooxygenase inhibitory and antioxidant cyanidin glycosides in cherries and berries. Phytomedicine 8 (5):362-369. doi:10.1078/0944-7113-00053

Segura V, Denance C, Durel C-E, Costes E (2007) Wide range QTL analysis for complex architectural traits in a 1-year-old apple progeny. Genome 50 (2):159-171. doi:10.1139/G07-002

Seipp MT, Durtschi JD, Liew MA, Williams J, Damjanovich K, Pont-Kingdon G, Lyon E, Voelkerding KV, Wittwer C (2007) Unlabeled oligonuclotides as internal temperature controls for genotyping by amplicon melting. Journal of Molecular Diagnostics 9 (3):284-289

Shaw J, Small RL (2004) Addressing the "hardest puzzle in American pomology:" Phylogeny of *Prunus* sect. Prunocerasus (Rosaceae) based on seven noncoding chloroplast DNA regions. American Journal of Botany 91 (6):985-996. doi:10.3732/ajb.91.6.985

Shenghua L, Chengquan F (1994) Studies on chromosome of *Pyrus* in China. Acta Horticulturae (ISHS) 367:27-32

Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, Burns P, Davis TM, Slovin JP, Bassil N, Hellens RP, Evans C, Harkins T, Kodira C, Desany B, Crasta OR, Jensen RV, Allan AC, Michael TP, Setubal JC, Celton J-M, Rees DJG, Williams KP, Holt SH, Rojas JJR, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin SA, Troggio M, Viola R, Ashman T-L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest HD, Bryant DW, Fox SE, Givan SA, Wilhelm LJ, Naithani S, Christoffels A, Salama DY, Carter J, Girona EL, Zdepski A, Wang W, Kerstetter RA, Schwab W, Korban SS, Davik J, Monfort A, Denoyes-Rothan B, Arús P, Mittler R, Flinn B, Aharoni A, Bennetzen JL, Salzberg SL, Dickerman AW, Velasco R, Borodovsky M, Veilleux RE, Folta KM (2011) The genome of woodland strawberry (*Fragaria vesca*). Nature Genetics 43 (2):109-116. doi:http://www.nature.com/ng/journal/v43/n2/abs/ng.740.html#supplementary-information

Silfverberg-Dilworth E, Matasci CL, van de Weg WE, Van Kaauwen MPW, Walser M, Kodde LP, Soglio V, Gianfranceschi L, Durel CE, Costa F, Yamamoto T, Koller B, Gessler C, Patocchi A (2006) Microsatellite markers spanning the apple (*Malus* x *domestica* Borkh.) genome. Tree Genetics & Genomes 2:202 - 224

Sjulin TM, Dale A (1987) Genetic diversity of North American strawberry cultivars. Journal of the American Society for Horticultural Science 112 (2):375-385

Soller M, Brody T, Genizi A (1976) On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. Theoretical and Applied Genetics 47 (1):35-39. doi:10.1007/bf00277402

Sooriyapathirana S, Khan A, Sebolt A, Wang D, Bushakra J, Lin-Wang K, Allan A, Gardiner S, Chagné D, Iezzoni A (2010) QTL analysis and candidate gene mapping for skin and flesh color in sweet cherry fruit (Prunus avium L.). Tree Genetics & Genomes 6 (6):821-832

Soriano JM, Romero C, Vilanova S, Llácer G, Badenes M (2005) Genetic diversity of loquat germplasm (*Eriobotrya japonica* (Thunb.) Lindl) assessed by SSR markers. Genome 48 (1):108-114

Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD, Miftahudin, Mahmoud A, Ma X, Gustafson PJ, Qi LL, Echalier B, Gill BS, Matthews DE, Lazo GR, Chao S, Anderson OD, Edwards H, Linkiewicz AM, Dubcovsky J, Akhunov ED, Dvorak J, Zhang D, Nguyen HT, Peng J, Lapitan NLV, Gonzalez-Hernandez JL, Anderson JA, Hossain K, Kalavacharla V, Kianian SF, Choi D-W, Close TJ, Dilbirligi M, Gill KS, Steber C, Walker-Simmons MK, McGuire PE, Qualset CO (2003) Comparative DNA sequence analysis of wheat and rice genomes. Genome Research 13 (8):1818-1827. doi:10.1101/gr.1113003

Sosinski B, Gannavarapu M, Hager LD, Beck LE, King GJ, Ryder CD, Rajapakse S, Baird WV, Ballard RE, Abbott AG (2000) Characterization of microsatellite markers in peach [*Prunus persica* (L.) Batsch]. Theoretical and Applied Genetics 101 (3):421-428

Sosinski B, Verde I, Rokhsar DS (2010) International Peach Genome Initiative Peach Genome v1.0. http://www.rosaceae.org/peach/genome.

Stafne ET, Clark JR (2004) Genetic relatedness among eastern North American blackberry cultivars based on pedigree analysis. Euphytica 139 (2):95-104. doi:10.1007/s10681-004-2436-4

Stafne ET, Clark JR, Weber CA, Graham J, Lewers KS (2005) Simple sequence repeat (SSR) markers for genetic mapping of raspberry and blackberry. Journal of the American Society for Horticultural Science 130 (5):722-728

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: Join Map. Plant Journal 3 (5):739-744. doi:10.1111/j.1365-313X.1993.00739.x

Sturtevant AH (1913) The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of inheritance. Journal of Experimental Zoology 14:43-59

Sutherland BG, Tobutt KR, Marchese A, Paternoster G, Simpson DW, Sargent DJ (2008) A genetic linkage map of *Physocarpus*, a member of the Spiraeoideae (Rosaceae), based on RAPD, AFLP, RGA, SSR and gene specific markers. Plant Breeding 127 (5):527-532

Takos AM, Ubi BE, Robinson SP, Walker AR (2006) Condensed tannin biosynthesis genes are regulated separately from other flavonoid biosynthesis genes in apple fruit skin. Plant Science 170:487 - 499

Tam SM, Mhiri C, Vogelaar A, Kerkveld M, Pearce SR, Grandbastien M-A (2005) Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. Theoretical and Applied Genetics 110 (5):819-831

Tanaka Y, Sasaki N, Ohmiya A (2008) Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant Journal 54 (4):733 - 749

Tanksley SD (1993) Mapping polygenes. Annual Review of Genetics 27:205-233. doi:10.1146/annurev.ge.27.120193.001225

Tanksley SD, Bernatzky R, Lapitan NL, Prince JP (1988) Conservation of gene repertoire but not gene order in pepper and tomato. Proceedings of the National Academy of Sciences of the United States of America 85 (17):6419-6423

Tanksley SD, Medina-Filho H, Rick CM (1982) Use of naturally-occurring enzyme variation to detect and map genes controlling quantitative traits in an interspecific backcross of tomato. Heredity 49 (1):11-25

Testolin R, Marrazzo M, Cipriani G, Quarta R (2000) Microsatellite DNA in peach (*Prunus persica* L. Batsch) and its use in fingerpinting and testing the genetic origin of cultivars. Genome 43 (3):512-520

Thoday JM (1961) Location of polygenes. Nature 191:368-370. doi:10.1038/191368a0

Thompson MM (1995a) Chromosome numbers of *Rubus* cultivars at the National Clonal Germplasm Repository. HortScience 30 (7):1453-1456

Thompson MM (1995b) Chromosome numbers of *Rubus* species at the National Clonal Germplasm Repository. HortScience 30 (7):1447-1452

Thomson MJ, Zhao K, Wright M, McNally KL, Leung H, McCouch SR (2010) Development and applification of 96- and 384-plex single nucleotide polymorphism (SNP) marker sets for diversity analysis, mapping and marker-assisted selection in rice. Paper presented at the Second Africa Rice Congress, Bamako, Mali, 22-26 March 2010

Tohge T, Nishiyama Y, Hirai M, Yano M, Nakajima J, Awazuhara M, Inoue E, Takahashi H, Goodenowe D, Kitayama M, Noji M, Yamazaki M, Saito K (2005) Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants over-expressing an MYB transcription factor. Plant Journal 42 (2):218 - 235

Torre LC, Barritt BH (1977) Quantitative evaluation of *Rubus* fruit anthocyanin pigments. Journal of Food Science 42 (2):488-490. doi:10.1111/j.1365-2621.1977.tb01528.x

Tulio AZ, Reese RN, Wyzgoski FJ, Rinaldi PL, Fu R, Scheerens JC, Miller AR (2008) Cyanidin 3-rutinoside and cyanidin 3-xylosylrutinoside as primary phenolic antioxidants in black raspberry. Journal of Agricultural and Food Chemistry 56 (6):1880-1888. doi:10.1021/jf072313k

Van Deynze A, Stoffel K, Buell CR, Kozik A, Liu J, van der Knaap E, Francis D (2007) Diversity in conserved genes in tomato. BMC Genomics 8:465

Van Deynze AE, Nelson JC, Yglesias ES, Harrington SE, Braga DP, McCouch SR, Sorrells ME (1995) Comparative mapping in grasses. Wheat relationships. Molecular and General Genetics MGG 248 (6):744-754

van Dyk MM, Koning G, Simayi Z, Booi S, Maharaj R, Selala MC, du Preez MG, Rees DJG, Labuschagné IF, Warnich L (2005) Molecular genetic studies on pears (*Pyrus* spp.) in the Western Cape. Acta Horticulturae (ISHS) 671:259-265

van Ooijen JW (1992) Accuracy of mapping quantitative trait loci in autogamous species. Theoretical and Applied Genetics 84 (7):803-811. doi:10.1007/bf00227388

van Ooijen JW (2004) MapQTL® 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma, B.V., Wageningen, Netherlands

van Ooijen JW, Voorrips RE (2001) JoinMap(R) 3.0, Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands

Vandersteen JG, Bayrak-Toydemir P, Palais RA, Wittwer CT (2007) Identifying common genetic variants by High-Resolution Melting. Clinical Chemistry 53 (7):1191-1198. doi:10.1373/clinchem.2007.085407

Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D, Salvi S, Pindo M, Baldi P, Castelletti S, Cavaiuolo M, Coppola G, Costa F, Cova V, Dal Ri A, Goremykin V, Komjanc M, Longhi S, Magnago P, Malacarne G, Malnoy M, Micheletti D, Moretto M, Perazzolli M, Si-Ammour A, Vezzulli S, Zini E, Eldredge G, Fitzgerald LM, Gutin N, Lanchbury J, Macalma T, Mitchell JT, Reid J, Wardell B, Kodira C, Chen Z, Desany B, Niazi F, Palmer M, Koepke T, Jiwan D, Schaeffer S, Krishnan V, Wu C, Chu VT, King ST, Vick J, Tao Q, Mraz A, Stormo A, Stormo K, Bogden R, Ederle D, Stella A, Vecchietti A, Kater MM, Masiero S, Lasserre P, Lespinasse Y, Allan AC, Bus V, Chagne D, Crowhurst RN, Gleave AP, Lavezzo E, Fawcett JA, Proost S, Rouze P, Sterck L, Toppo S, Lazzari B, Hellens RP, Durel C-E, Gutin A, Bumgarner RE, Gardiner SE, Skolnick M, Egholm M, Van de Peer Y, Salamini F, Viola R (2010) The genome of the domesticated apple (*Malus* x *domestica* Borkh.). Nature Genetics 42 (10):833-839. doi:http://www.nature.com/ng/journal/v42/n10/abs/ng.654.html#supplementary-information

Ventelon M, Deu M, Garsmeur O, Doligez A, Ghesquière A, Lorieux M, Rami JF, Glaszmann JC, Grivet L (2001) A direct comparison between the genetic maps of sorghum and rice. Theoretical and Applied Genetics 102 (2):379-386

Verde I, Lauria M, Dettori M, Vendramin E, Balconi C, Micali S, Wang Y, Marrazzo M, Cipriani G, Hartings H, Testolin R, Abbott A, Motto M, Quarta R (2005) Microsatellite and AFLP markers in the *Prunus persica* [L. (Batsch)] × *P. ferganensis* BC_1 linkage map: Saturation and coverage improvement. Theoretical and Applied Genetics 111 (6):1013-1021

Vilanova S, Sargent D, Arús P, Monfort A (2008) Synteny conservation between two distantly-related Rosaceae genomes: *Prunus* (the stone fruits) and *Fragaria* (the strawberry). BMC Plant Biology 8 (1):67

Viruel MA, Messeguer R, De Vicente MC, Garcia-Mas J, Puigdomenech P, Vargas F, Arús P (1995) A linkage map with RFLP and isozyme markers for almond. Theoretical and Applied Genetics 91:964-971

Vision TJ, Brown DG, Shmoys DB, Durrett RT, Tanksley SD (2000) Selective mapping: A strategy for optimizing the construction of high-density linkage maps. Genetics 155 (1):407-420

Voorrips RE (2002) MapChart: Software for the graphical presentation of linkage maps and QTLs. Journal of Heredity 93 (1):77-78

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23 (21):4407-4414

VSN International L (2010) GenStat 12.2.0.3717 edn.,

Wang D, Karle R, Brettin TS, Iezzoni AF (1998) Genetic linkage map in sour cherry using RFLP markers. Theoretical and Applied Genetics 97 (8):1217-1224

Warburton ML, Xianchun X, Crossa J, Franco J, Melchinger AE, Frisch M, Bohn M, Hoisington D (2002) Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. Crop Science 42 (6):1832-1840

Way RD, Aldwinckle HS, Lamb RC, Rejman A, Sansavini S, Shen T, Watkins R, Westwood MN, Yoshida Y (1991) Apples (*Malus*). Acta Horticulturae (ISHS) 209:3-46

Weebadde CK, Wang D, Finn CE, Lewers KS, Luby JJ, Bushakra J, Sjulin TM, Hancock JF (2008) Using a linkage mapping approach to identify QTL for day-neutrality in the octoploid strawberry. Plant Breeding 127 (1):94-101. doi:10.1111/j.1439-0523.2007.01430.x

Wikström N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proceedings of the Royal Society B 268 (1482):2211-2220. doi:10.1098/repb.201.1782

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18 (22):6531-6535

Wittwer CT, Reed GH, Gundry CN, Vandersteen JG, Pryor RJ (2003) High-Resolution Genotyping by Amplicon Melting Analysis using LCGreen. Clinical Chemistry 49 (6):853-860. doi:10.1373/49.6.853

Wolfe KH, Sharp PM, Li WH (1989) Mutation rates differ among regions of the mammalian genome. Nature 337:283-285

Woodhead M, McCallum S, Smith K, Cardle L, Mazzitelli L, Graham J (2008) Identification, characterisation and mapping of simple sequence repeat (SSR) markers from raspberry root and bud ESTs. Molecular Breeding 22 (4):555-563

Woodhead M, Weir A, Smith K, McCallum S, MacKenzie K, Graham J (2010) Functional markers for red raspberry. Journal of the American Society for Horticultural Science 135 (5):418-427

Wu FN, Eannetta NT, Xu YM, Durrett R, Mazourek M, Jahn MM, Tanksley SD (2009a) A COSII genetic map of the pepper genome provides a detailed picture of synteny with tomato and new insights into recent chromosome evolution in the genus Capsicum. Theoretical and Applied Genetics 118 (7):1279-1293. doi:10.1007/s00122-009-0980-y

Wu FN, Eannetta NT, Xu YM, Tanksley SD (2009b) A detailed synteny map of the eggplant genome based on conserved ortholog set II (COSII) markers. Theoretical and Applied Genetics 118 (5):927-935. doi:10.1007/s00122-008-0950-9

Wu FN, Tanksley SD (2010) Chromosomal evolution in the plant family Solanaceae. BMC Genomics 11. doi:10.1186/1471-2164-11-182

Wu S-B, Tavassolian I, Rabiei G, Hunt P, Wirthensohn M, Gibson J, Ford C, Sedgley M (2009c) Mapping SNP-anchored genes using high-resolution melting analysis in almond. Molecular Genetics and Genomics 282 (3):273-281

Xu Y, Zhu L, Xiao J, Huang N, McCouch SR (1997) Chromosomal regions associated with segregation distortion of molecular markers in F_2, backcross, doubled haploid, and recombinant inbred populations in rice (*Oryza sativa* L.). Molecular and General Genetics 253 (5):535-545. doi:10.1007/s004380050355

Yamamoto T, Kimura T, Saito T, Kotobuki K, Matsuta N, Liebhard R, Gessler C, van de Weg WE, Hayashi T (2004a) Genetic linkage maps of Japanese and European pears aligned to the apple consensus map. Acta Horticulturae (ISHS) 663:51-56

Yamamoto T, Kimura T, Sawamura Y, Kotobuki K, Ban Y, Hayashi T, Matsuta N (2001a) SSRs isolated from apple can identify polymorphism and genetic diversity in pear. Theoretical and Applied Genetics 102 (6):865-870

Yamamoto T, Kimura T, Shoda M, Ban Y, Hayashi T, Matsuta N (2002a) Development of microsatellite markers in the Japanese pear (*Pyrus pyrifolia* Nakai). Molecular Ecology Notes 2 (1):14-16

Yamamoto T, Kimura T, Shoda M, Imai T, Saito T, Sawamura Y, Kotobuki K, Hayashi T, Matsuta N (2002b) Genetic linkage maps constructed by using an interspecific cross between Japanese and European pears. Theoretical and Applied Genetics 106 (1):9-18

Yamamoto T, Kimura T, Soejima J, Sanada T, Ban Y, Hayashi T (2004b) Identification of quince varieties using SSR markers developed from pear and apple. Breeding Science 54 (3):239-244

Yamamoto T, Kimura T, Terakami S, Nishitani C, Sawamura Y, Saito T, Kotobuki K, Hayashi T (2007) Integrated reference genetic linkage maps of pear based on SSR and AFLP markers. Breeding Science 57 (4):321-329

Yamamoto T, Shimada T, Imai T, Yaegaki H, Haji T, Matsuta N, Yamaguchi M, Hayashi T (2001b) Characterization of morphological traits based on a genetic linkage map in peach Breeding Science 51 (4):271-278

Young ND (2001) Constructing a plant genetic linkage map with DNA markers. In: Phillips RL, Vasil IK (eds) DNA-Based Markers in Plants vol 6. Advances in Cellular and Molecular Biology of Plants, 2nd edn. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 31-47

Zorrilla-Fontanesi Y, Cabeza A, Domínguez P, Medina J, Valpuesta V, Denoyes-Rothan B, Sánchez-Sevilla J, Amaya I (2011a) Quantitative trait loci and underlying candidate genes controlling agronomical and fruit quality traits in octoploid strawberry (*Fragaria* × *ananassa*). Theoretical and Applied Genetics 123 (5):755-778. doi:10.1007/s00122-011-1624-6

Zorrilla-Fontanesi Y, Cabeza A, Torres A, Botella M, Valpuesta V, Monfort A, Sánchez-Sevilla J, Amaya I (2011b) Development and bin mapping of strawberry genic-SSRs in diploid *Fragaria* and their transferability across the Rosoideae subfamily. Molecular Breeding 27 (2):137-156. doi:10.1007/s11032-010-9417-1

APPENDIX: Reprints

Thesis publication reprints