A NOVEL GENETIC LINKAGE MAP OF CHOKECHERRY (*PRUNUS VIRGINIANA* L.) AND QTL MAPPING FOR X-DISEASE (*CANDIDATUS* PHYTOPLASMA PRUNI) RESISTANCE

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Ryan Rae Lenz

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Plant Sciences

June 2016

Fargo, North Dakota

North Dakota State University
Graduate School

Title

A NOVEL GENETIC LINKAGE MAP OF CHOKECHERRY (*PRUNUS VIRGINIANA* L.) AND QTL MAPPING FOR X-DISEASE (*CANDIDATUS* PHYTOPLASMA PRUNI) RESISTANCE

By

Ryan Rae Lenz

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Wenhao Dai
Chair

Dr. Ted Helms

Dr. G. Francois Marais

Dr. Xuehui Li

Dr. Stuart Haring

Approved:

7/8/2016
Date

Richard D. Horsley
Department Chair

ABSTRACT

A novel chokecherry genetic linkage map was constructed using 565 molecular markers and a previously published mapping population (n=101). Chokecherry (*Prunus virginiana* L.) is a potential model for genetic research of phytoplasmic diseases because of its natural resistance to X-disease (*Candidatus* Phytoplasma pruni). The novel chokecherry map was developed using JoinMap 4.0 and contains a complete set of 16 linkage groups. In total, the map spans a genetic distance of 2,172 cM with an average marker density of 3.97 cM. Three significant quantitative trait loci (QTL) associated with X-disease resistance were identified on linkage groups 15, 5, and 4 contributing to a total of 45.9% of the phenotypic variation. This novel genetic linkage map and the identified QTL linked to X-disease resistance will provide the framework needed to facilitate molecular genetics, genomics, and breeding research concerning X-disease in chokecherry and other *Prunus* species.

ACKNOWLEDGMENTS

I want to express my humble and sincerest gratitude to everyone who has supported me during my Master's experience at NDSU. I have learned a lot and owe so much of my progression to those that have helped me.

First, I want to thank my graduate advisor, Dr. Wenhao Dai for his guidance and positive mentorship. He has helped me progress through my thesis genetics work and other horticultural experiments which has given me a strong aptitude for becoming successful as a future Ph.D. student and professional plant scientist. Also, his support while writing manuscripts and my thesis is invaluable to me.

I would also like to thank NDSU faculty and support staff that have been instrumental in my academic success. I especially want to thank my graduate committee members, Drs. Ted Helms, Francois Marais, Xuehui Li, and Stuart Haring. Their support with my plan of study and thesis proposal has been beyond appreciated. I also want to thank my committee for taking the time commenting and improving my thesis.

Other people I want to thank for helping me with my thesis include Dr. Hongxia Wang, Dr. Jason Fiedler, and Dr. Yinghai Liang who shortened my learning curve associated with experimentation and genetic mapping software. Dr. Danqiong Huang did a wonderful job getting me familiar with the lab space and how to work with plant DNA. My undergraduate mentee, Peder Schmitz, did a great job helping me with tissue culture and running some molecular markers. Also, I want to thank Justin Hegstad for the extra thermocyclers and gel electrophoresis equipment.

I had insurmountable support from Vicki Magnusson whom I had the pleasure of working with on a daily basis. She has always given great advice and enjoyable conversations. I also want to thank her for helping me with side projects and field work.

Lastly, I need to convey my true appreciation to my whole family for being so supportive during my quest through school. My parents, Doug and Jeannie, are always encouraging me to push myself and work hard. Also, my wife Rebecca has been immensely supportive and very understanding of my graduate career. Her family has been great, and I appreciate everything I have shared with them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AFLP ..... Amplified fragment length polymorphism

AP ..... Apple proliferation

BN ..... Bois noir

Cho ..... Chokecherry

cM ..... centi-Morgan

COS ..... Conserved ortholog set

DNA ..... Deoxyribonucleic acid

EF x NY ..... 'Emperor Francis' by 'New York 54' sweet cherry

ESFY ..... European stone-fruit yellows

EST ..... Expressed sequence tag

GY ..... Grapevine yellows

IRAP ..... Inter-retrotransposon amplified polymorphism

ISBP ..... Insertion-site-based polymorphism

LG ..... Linkage group

LOD ..... Logarithm of the odds

LTR ..... Long terminal repeat

MAS ..... Marker assisted selection

Mt ..... Million tonnes

PaWB ..... Paulownia witches' broom

PCR ..... Polymerase chain reaction

QTL ..... Quantitative trait loci/locus

RBIP ..... Retrotransposon-based insertional polymorphism

RC ..... Resistant chokecherry

REMAP .....Retrotransposon-microsatellite amplified polymorphism

RFLP ..... Restriction fragment length polymorphism

RNA .....Ribonucleic acid

ROS..... Reactive oxygen species

SAR..... Systemic acquired resistance

SC..... Susceptible chokecherry

SCL .....Strongest cross link

SDRF..... Single dose restriction fragment

SNP ..... Single nucleotide polymorphism

SSR .....Simple sequence repeat

STS..... Sequence tagged sites

T x E.....'Texas' almond by ' Earlygold' peach

TE.....Transposable element

WGD.....Whole genome duplications

INTRODUCTION

Fruit and tree nut production contributes about $18 billion to the U.S. economy annually (Perez 2014). Although demand for fresh fruit is still growing, limiting factors such as plant disease reduce the yield potential in commercial production systems. X-disease is a major example which affects a variety of stone-fruit *Prunus* species such as peach, apricot, nectarine, cherry, plum, and chokecherry (Guo et al. 1998). Disease incidence as high as 60% and yield reductions ranging from 30% to 80% have been observed in Connecticut peach orchards (Douglas 1999) and can cause more than 50% mortality in orchards within three years post-infection (Peterson 1984). Current control measures for X-disease phytoplasma include pesticide treatment for leaf hopper vectors, elimination of infected trees, and antibiotic treatment as a last resort (Douglas 1999), but these methods have shown to be inefficient and expensive (Davis, 2013; Peterson, 1984). Chokecherry is the primary source of X-disease because it is a dominant reservoir of leafhoppers, by which the X-disease phytoplasma is vectored and transmitted.

It is commonly argued that X-disease resistant cultivars offer the best method of phytoplasma control (Olivier et al. 2009; Davis 2013; Peterson 1984); however, natural resistance to X-disease hasn't been documented in any plant host except for chokecherry (*Prunus virginiana* L.) (Guo et al. 1998; Wang et al. 2014). This makes chokecherry a potential model for genetic studies involving X-disease or any other phytoplasma-derived diseases. Previous work has developed a partial genetic linkage map for chokecherry which was used to identify a quantitative trait locus (QTL) associated with X-disease resistance (Wang et al. 2014) which accounted for only 26% of the total phenotypic variation. Until now, an inadequate repertoire of molecular genetics resources in chokecherry has prevented major advances in the genetic understanding of the host-pathogen interaction of X-disease.

The objectives of this research were to develop and utilize an improved chokecherry genetic linkage map for identifying genetic regions related to X-disease resistance. Simple sequence repeat (SSR), amplified fragment length polymorphism (AFLP), and long terminal repeat (LTR) markers were utilized in improving the genetic linkage map of chokecherry. Following the development of improved linkage maps, marker-assisted breeding and map-based cloning for X-disease resistance can be explored straightaway. Providing new information and resources for elucidating mechanisms involved with X-disease and other phytoplasma-derived diseases will advance future research regarding disease response. The thesis research presented will allow for continued exploration aimed at identifying specific genes associated with natural resistance mechanisms in chokecherry which can be used as a template for examining resistance in other *Prunus* species. Ultimately, this study could lead to more efficient fruit production worldwide and could address issues regarding food supply and affordability.

1. LITERATURE REVIEW

1.1. Phytoplasma and associated diseases

1.1.1. Phytoplasma biology

Phytoplasmas are cell wall-less, mycoplasma-like bacterial pathogens of plants that are transmitted by a variety of phloem-feeding insects, such as leaf-hoppers, plant hoppers, and psyllids or some parasite plants, such as dodder. The insect vectors belong to the taxonomic order Hemiptera (Weintraub and Beanland 2006). Phytoplasma can also be spread via direct contact between infected and healthy plant materials, such as grafting. Phytoplasma are largely associated with diseases commonly referred to as yellows diseases. These diseases were originally thought to be caused by virus until it was found that phloem sieve tube elements of yellows-diseased plants had numerous wall-less, pleomorphic bodies which resembled mycoplasmas (Marcone et al. 2014). Resultantly, phytoplasmas were historically classified as mycoplasma-like organisms (MLOs) (Doi et al. 1967), until it was replaced with phytoplasma in the 1990s. (International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of Mollicutes, 1993). Currently, phytoplasmas are placed in the class Mollicute, genus *Candidatus* (IRPCM Phytoplasma/Spiroplasma Working Team – Phytoplasma Taxonomy Group, 2004).

Phytoplasmas are obligate parasites which rely on their plant and insect hosts for dissemination. They are difficult to study because *in vitro* culturing has been proven unsuccessful. However, molecular methods have helped identify and elucidate the genetic diversity of these pathogens. For example, nested polymerase chain reaction (PCR) has been implemented to overcome the low titer of phytoplasma DNA being detected in plant tissues. Nested PCR begins with a preliminary amplification using universal phytoplasma 16S rRNA

primers and a subsequent amplification with primers designed within the previous amplicon. In this way, it is possible to identify low-titer phytoplasma infections with high specificity. In addition, the 16S rDNA sequences and the well-documented host-pathogen ecology have helped classify phytoplasma into their respective clades. Phytoplasma are typically differentiated on the 16S rRNA gene by means of restriction fragment length polymorphisms (RFLP) analysis (Lee et al., 1998; Bertaccini & Duduk, 2009). Ribosomal DNA is quite conserved within taxonomic clades, which facilitates the characterization of unknown phytoplasma with the RFLP patterns of known phytoplasma. Now, with improved affordability and utility of complete genome sequencing technologies, comparative genomics is underway for these plant pests (Kube et al. 2012).

Phytoplasmas have a small genome which can vary in size from 530 to 1350 kilobases (kb) (Marcone, 2014). In contrast to their small genomes, phytoplasma can infect/inhabit a wide range of plant hosts. It is hypothesized by Sugio and Hogenhout (2012) that the dynamic architecture of phytoplasma genomes, which is composed of a relatively high abundance of repetitive elements, may account for their adaptation to diverse environments of their plant and insect hosts. Phytoplasma genome sequences have been used to identify candidate genes that are likely to play major roles in phytoplasma–host interactions (Sugio and Hogenhout 2012). Among these are genes encoding surface membrane proteins and effector proteins. These effector proteins are ultimately the cause of phytoplasma pathogenicity and proliferation in plants.

1.1.2. Phytoplasma virulence

Phytoplasmas produce specific virulence proteins or effectors that modulate development of plant–pathogen interactions, leading to an increased transmission and improved evolutionary fitness of these organisms (Sugio and Hogenhout 2012). There are three well characterized

effector proteins produced by phytoplasma, TENGU, SAP11, and SAP54. It was reported that SAP11 and TENGU induced the production of vasculature tissue in plants, leading to witches' broom symptoms and a harbor for phytoplasma proliferation. SAP11 also reduces the production of jasmonic acid (JA) which is a plant hormone commonly used in stress signaling and defense (Turner et al. 2002). SAP54 was found to inhibit flowering in plants, which increases the vascular tissue and the proliferation window of the pathogen by preventing senescence in annual plant species (Sugio and Hogenhout 2012). These limited examples still exhibit how diverse phytoplasmas can be in their effective host range. In the future, it is expected that more phytoplasmic effector proteins will be characterized and associated to their particular symptomatic diseases in plants.

1.1.3. Phytoplasma symptoms

As described before, phytoplasmas reside in phloem sieve cells and can translocate via the pores of sieve plates separating cells. Infected plants exhibit various symptoms such as witches' brooms, chlorosis (leaf yellowing), virescence or phyllody (greening of flower organs), dieback, and altered volatile production (Mayer et al. 2008; Bertaccini & Duduk 2009; Sugio and Hogenhout 2012). Symptoms can appear within one week upon inoculation, but can take longer depending on temperature and plant species. It is known that the severity and combination of symptoms vary among phytoplasma strains and the infected plant host; however, management strategies for phytoplasma are quite limited. A few phytoplasma management strategies include rogueing or eliminating infected plants, insecticidal control of leaf hopper vectors, antibiotic treatment, and elimination of nearby wild plant hosts. Chemical control seems to be economical and practical; however, the efficacy of chemical control is far from complete with diseases continually spreading in many areas of the world (Firrao et al. 2007). It is also important to

consider the environmental and economic impact of continued pesticide use. The inefficiency of phytoplasma management has resulted in a collaborative effort to find natural resistance mechanisms in plant host germplasm.

1.1.4. Phytoplasma-derived diseases and resistance

It is known that phytoplasma diseases affect a wide range of plants. A comprehensive list of categorized phytoplasmic diseases can be found in a review by Marcone (2015). A few phytoplasmic diseases of particular interest include, Bois Noar (BN), apple proliferation (AP), European stone fruit yellows (ESFY), Paulownia Witches' Broom (PaWB) disease, and X-disease. They are particularly interesting because plant hosts of these diseases have documented levels of phytoplasma tolerance and/or resistance (Osler et al. 1999; Bertamini et al. 2002; Musetti et al. 2004 and 2005; Laimer et al. 2009; Albertazzi et al. 2009; Liu et al. 2013).

Grapevine (*Vitis vinifera* L.) cultivars in Europe are constantly assaulted by phytoplasma diseases including grapevine-yellows (GY), Flavescence dorée (FD), and Bois noir (BN), causing leaf rolling and curling, yellowing, weak canes and desiccated fruit clusters. This results in devastating production losses and poor fruit quality (Laimer et al. 2009). There are some cultivars that have exhibited a phenomenon known as recovery. Recovery is spontaneous symptom remission (Caudwell 1961) which may or may not involve the elimination of the pathogen from the host (Schmidt 1965). Recovery from phytoplasmic diseases was associated with an increase of hydrogen peroxide in phloem tissues. The reported hydrogen peroxide increase in recovered tissues was not found in diseased plants or in healthy control plants, suggesting a coordinated response mechanism has taken place (Romanazzi et al. 2009). Additional documentation of disease response in grapevine showed that gene expression profiles in infected plants are altered quite dramatically. Many of the differentially expressed genes seem

to be involved with various metabolic pathways (Albertazzi et al. 2009), nonetheless, identifying genes relating to direct disease resistance still needs to be uncovered.

Apple proliferation (AP) and European stone fruit yellows (ESFY) seen in apple (*Malus domestica* Borkh.) and apricot (*Prunus armeniaca* L.), respectively, have examples of tolerant responses to phytoplasma pathogens. Similarly to grapevines, recovery has been reported in both species (Osler et al. 1999; Musetti et al. 2004; Romanazzi et al. 2009). As mentioned before, recovery does not always mean elimination of the pathogen. For example, recovery from apple proliferation (AP) correlates with the disappearance of phytoplasma from the canopy but not from the roots (Loi et al. 2002). In apricot plants affected by ESFY, recovery does not appear to be correlated to disappearance of phytoplasma (Osler et al. 1999). It is important to note that recovery is not a common occurrence; however, it has been observed that recovered plants are less likely to become symptomatic in comparison to plants that were never infected. This indicates that systemic acquired resistance (SAR) is involved in inducing recovery (Osler et al. 1999 and Musetti et al. 2004), which means metabolites and reactive oxygen species (ROS) such as hydrogen peroxide have an active role in the recovery phenomenon.

Paulownia Witches' Broom (PaWB) is another example of a phytoplasma-derived disease having associated resistance studies. PaWB affects the hardwood *Paulownia* plant causing proliferation of branches (witches' broom), yellowing of leaves, and branch dieback (Liu et al. 2013). A forward genetics approach was used by Liu and colleagues (2013) to discover potential genetic loci associated with PaWB. They analyzed transcriptome data and expressed sequence tags (ESTs) related to biotic stress response and differential expression patterns induced by phytoplasma infection. Quantitative real-time PCR (qRT-PCR) amplified the ESTs of interest, which identified a relatively large group of candidate genes that could be related to

phytoplasma resistance. Studies like this will continue to increase not only in numbers but also in importance as it is becoming more apparent that resistant cultivars are needed to combat phytoplasmic diseases.

1.1.5. *Prunus* species and X-disease phytoplasma

Prunus species are categorized as stone-fruit tree species based on the characteristic of the pits found within the fruit. A few examples of well-known *Prunus* species include peach (*Prunus persica*), apricot (*Prunus armeniaca*), almond (*Prunus dulcis*), sweet cherry (*Prunus avium*), tart cherry (*Prunus cerasus*), plum (*Prunus domestica*), and chokecherry (*Prunus virginiana*). Stone fruits belong to the Rosaceae family, which holds a large global economic impact, estimated over 113 million tonnes (Mt) and 180 billion dollars in the year 2005 (Hummer and Janick 2009). *Prunus* species have a significant impact on these numbers with peach itself being produced at over 16 Mt. Although these production values are quite large, the demand for fresh fruit and the production limitations are both growing strong. The limitations mentioned include abiotic and biotic stresses, but one of particular significance is X-disease (Marcone 2015).

X-disease phytoplasma are categorized as '*Candidatus* Phytoplasma pruni,' belonging to subgroup 16SrIII-A (Davis et al. 2013). X-disease has historically been one of the major limiting factors in peach production in the United States (Stoddard et al., 1951). As an example, X-disease caused heavy losses in peach and cherry orchards in the United States during the 1990s (Bertaccini and Duduk 2009). The disease was first reported as 'X disease of peach' in 1933 in Connecticut. It was called this rather ambiguous name because of its unknown cause and mysterious nature (Stoddard 1938). X-disease is most problematic in orchards that are near wild chokecherry, because chokecherry is a natural host of the pathogen (Davis 1986). Being

vectored by over 12 reported leafhopper species (Olivier et al. 2009; Davis et al. 2013) and a host range in most *Prunus* species (Guo et al. 1998) makes this disease very troublesome, especially for peach and chokecherry (Rosenberger 1977; Peterson 1984).

The symptoms of X-disease include chlorosis (yellowing), tattered appearance of leaves (shothole), premature leaf drop, poor fruit quality, reduced yields, dieback of branches, and plant death (Stoddard et al. 1951, Douglas 1999). It has been reported that X-disease can cause more than 50% mortality in peach orchards within 3 years of infection (Peterson 1984). Application of insecticides in an effort to stem the spread of X-disease, destruction of nearby chokecherry, treatment of trees with tetracycline antibiotic, removal of symptomatic branches or full trees, replacement of trees that are killed by the disease, and an integrated approach to disease control have been practiced with varying degrees of success (Davis 2013; Peterson 1984). The most promising means of phytoplasma and X-disease control lies within plant resistance (Olivier et al. 2009). Unfortunately, natural disease resistance in plants for any phytoplasmic disease is limited, and that includes X-disease. Thankfully, chokecherry (*Prunus virginiana*) is a reported species that has varying levels of X-disease resistance and tolerance (USDA-NRCS 1993; Walla et al. 1996; Guo and Cheng, 1998), and recent advances in molecular genetics have provided a means to discover genetic loci (genes) associated with X-disease resistance in chokecherry (Wang et al. 2014).

1.2. Plant molecular genetics

1.2.1. Perceptions on molecular genetics

It is a well-documented phenomenon that plants have undergone multiple genomic duplications throughout their history (Van de Peer et al. 2009; Bowers et al. 2003; Jaillon et al. 2007; Vogel et al. 2009). Whole genome duplications (WGD) most commonly occur when a

diploid or unreduced gamete reproduces with another unreduced gamete (Van de Peer et al. 2009). If this occurs within the same species, it results in what is called an autotetraploid. If the diploid gametes came from different but related species, this results in an allotetraploid (Van de Peer et al. 2009). Once genome duplication occurs, the resulting offspring may undergo chromosome rearrangement, fusions, and deletions. These resultant consequences of WGD can lead to new species and ancestral lineages because of chromosome rearrangement and altered genes in terms of copy number, location, regulation, and function. Rearrangements occur in duplicated plants to address the costs associated with duplicate sets of genes and the need to efficiently produce bivalent chromosome pairing for reproduction. Many duplicate genes are eliminated or rearranged into altered genes. The changes in genes can result in new or similar functions and new phenotypes leading to the evolution of a species (Van de Peer et al. 2009).

Synteny is another result of the chromosomal changes from ancestral genome duplication (Tang et al. 2008). Synteny is described as similar chromosome blocks having collinear gene content and order among the species being compared. Because of synteny, it is possible to create molecular markers that are transferable among a taxomic group (Dirlewanger et al. 2004). This is because the similar chromosome blocks (or syntenic blocks) within closely related species will have similar DNA sequence. This provides a means to study a non-model species or a species that is not fully sequenced. Having this understanding reduces resource costs associated with identifying genetic loci of interest such as QTL involved with disease resistance like in chokecherry.

1.2.2. Basics of genetic mapping

Genetic mapping is the ordering of molecular markers in accordance with the relative genetic distances among them, and assigning them to their respective linkage groups based on

the recombination values from all the pairwise combinations (Jones et al. 1997). As a simplified example, if two loci are not linked (on different chromosomes), they will undergo independent segregation and will not be part of the same linkage group. However, if loci are linked and occur in close proximity on the same chromosome, the recombination frequency will be less and so will their genetic distance from one another. Recombination frequency is often expressed as an arbitrary map unit called centi-Morgan (cM). Markers that map together as one linkage group do so because they are all located in a single chromosome. The number of different linkage groups will correspond to the basic chromosome number of the species if there have been enough crossover events in the mapping population and if there have been enough markers employed (Jones et al. 1997). Although the markers can effectively map parallel to the number of chromosomes in a species, it is important to note that linkage groups are reported in genetic distances, not physical distances. In cases where crossovers are more abundant in certain clusters rather than being randomly distributed (i.e. telomeric versus centromeric regions), the genetic map will distort the physical distance. However, the advent and utility of DNA-based markers can provide anchor points for genetic maps to be integrated to its representative chromosomes (Jones et al. 1997). Generally, it is desirable to saturate a genetic map with as many markers as possible with a large population with many possible recombination events. Ultimately, map-based cloning projects can be conducted in a much more efficient manner with a robust map. Also, genome sequencing projects benefit from detailed genetic maps because the anchor points assist assembly of the sequence reads (Nelson 2012).

1.2.3. Short history of genetic mapping

Plant molecular genetics has continued to evolve since the discovery of Mendel's work with peas (Reid and Ross 2011). In today's world, we benefit from years of genetic research

through the use of molecular markers, genomics, and bioinformatics. Understanding and deciphering the genetic factors involved with traits of interest is a major goal of many researchers. If we look back through the development of plant molecular genetics tools and using tomato as a reference, we see that it begins with phenotype (Hedrick and Booth 1907).

Phenotypic or morphological markers were basically the beginning of mapping and associating traits of interest. In 1907, Hedrick and Booth provided a summary of phenotypic inheritance of simple tomato traits. This study provided a Mendelian prediction of a few qualitative characteristics but a genome-wide explanation was not available until an increasing repertoire of molecular markers has come to be. Isozymes were the first type of molecular marker developed that could assist in genetic mapping and phenotypic prediction. For example, the first complete genetic map which included all 12 chromosomes of tomato was published by Butler in 1952. Soon after was the development of DNA-based markers such as RFLPs, AFLPs, SSRs, and SNPs. These new systems allowed tomato to be one of the first species to be mapped in a 'high-density' fashion (Tanksley 1992). The progression of these marker systems in other species improved the approaches taken to conduct genetic mapping, QTL identification, and association mapping (Viruel et al. 1995; Abbott et al. 1998; Dettori et al. 2001; Kalendar et al. 2011; Salazar et al. 2014).

1.2.4. Combining molecular genetics and phenotypic variation

Molecular markers are 'neutral', meaning that their presence does not directly influence phenotype (besides genetic loci that are expressed genes). Markers rather provide a means of flagging regions of the genome that may be involved with a phenotype. The 'flags' or genetic loci are what constitute a genetic map (Asins 2002). Refinement of genetic maps is facilitated through the use of more individuals in a population (more recombination) and the more markers

you can use. Genetic maps rely on heterozygous populations with variations in phenotype. This is also necessary for QTL mapping or association mapping (Asins 2002).

QTL mapping arose as a means to identify markers that lie close to a genetic factor (gene) influencing phenotype (Asins 2002). If strong linkage is identified between a marker and a QTL, then breeders can use the marker to screen breeding populations for traits of interest, which is known as marker-assisted selection (MAS) (Lande and Thompson 1990). Association mapping is a similar approach to find genetic factors of interest, but requires identification of population structure and interactions between the subpopulations. Once population structure is accounted for, linkage disequilibrium (nonrandom association of alleles at different loci) attributes the association of a marker to the trait of interest (Flint-Garcia et al. 2003). If identified markers are validated in other populations, they can be used in MAS to help enhance selection efficiency and reduce resource costs in screening plants for beneficial genetics.

Association and QTL studies provide a means to identify genetic regions, but it is usually the goal to discover the physical location and the complete sequence of the genetic factors in question (Jander et al. 2002). DNA sequencing has provided this availability. Combining genetic linkage groups, QTL, and sequencing data allows for map-based cloning of candidate genes. Once a candidate gene is identified, its function needs to be tested and confirmed. Biotechnology approaches such as genetic engineering can be used to conduct complementation to confirm the function of the gene (Jander et al. 2002). Overall, molecular genetics is continually improving to provide researchers with all the necessary tools to discover genes that are involved with a species and traits of interest. This understanding is the basic driving force behind the molecular studies in plants including chokecherry and X-disease research.

1.3. Chokecherry: a model woody species for X-disease disease research

1.3.1. Chokecherry biology and ecology

Chokecherry (*P. virginiana*) is a small tree or shrub species in the Rosaceae family. Chokecherry has the same base chromosome number as other *Prunus* species (x = 8), but is one of the few tetraploids, having 32 chromosomes (2n = 4x = 32) (Dai, unpublished). It is native to North America with geographical origins stretching from northern Mexico to most of Canada and parts of Alaska. Leaves are simple, glabrous, and oval in shape with dark green color and serrated edges. They form perfect flowers that are aromatic and arranged in cylindrical racemes 3 to 6 inches long. Chokecherry fruits are dark red to black spherical drupes that are high in antioxidant and anthocyanin content. The common name, chokecherry, came from the bitter taste of the fruit and the relatively large seed within the pulp. (USDA-NRCS National Plant Data Team, http://plants.usda.gov/plantguide/pdf/cs_prvi.pdf)

The fruit was a staple for numerous Indian tribes across the North American continent, especially to tribes who lived on the plains and prairies. Chokecherries were routinely cooked before they were eaten or dried thoroughly to neutralize the naturally occurring hydrocyanic acid produced when cyanogenic glycosides in the leaves, stem, and stone are disturbed (USDA-NRCS). Chokecherry fruits are still collected today and used to make jellies, jams, pie-fillings, syrups, sauces, and wines. More recently, chokecherry is being implemented as a minor crop in the prairie provinces of Canada for juice production. Estimated fruit production potential is 15,000 pounds per acre from mature plants (USDA-NRCS). In the United States, chokecherry is a popular ornamental for its size, attractive white flowers, and autumn color (USDA-NRCS).

Chokecherry is tolerant to many stresses including drought, cold temperatures, and alkaline soil, which is attributed to its extensive root system. Chokecherry plantings are used

extensively in shelterbelts, windbreaks, wildlife habitat, and erosion control. Also, it is effectively grown on disturbed sites such as mined land reclamation, highway right-of-ways, and construction sites (USDA-NRCS National Plant Data Team, http://plants.usda.gov/plantguide/pdf/cs_prvi.pdf). Most recently, chokecherry is being proposed as a model system for X-disease research and resistance germplasm for *Prunus* and Rosaceous crop systems (Dai, personal communication).

1.3.2. Chokecherry and foundational X-disease research

Chokecherry is susceptible and a natural host to X-disease phytoplasma. However, its native range and genetic diversity have provided a few accessions that display resistance (or at least tolerance) to disease pressure. In 1983, a chokecherry seed source located in Bismarck, ND, was established by the United States Department of Agriculture-Natural Resources Conservation Service (USDA-NRCS) to examine potential X-disease resistant materials. Over 3,000 established germplasm from 179 seed sources collected from ND, MN and the surrounding region were planted and evaluated for X-disease symptoms. By 1994, 44% of the plants were dead and the remaining 1,792 plants still contained X-disease phytoplasma. Only 5% of the remaining plants displayed little to no X-disease damage. Walla et al. (1996) reported that the few plants with little damage and/or zero observable symptoms may be resistant or highly tolerant to X-disease phytoplasma. This chokecherry planting paved the way for recent molecular genetic studies in chokecherry (Wang et al. 2014) and the novel results discussed later in this thesis.

1.3.3. Genetic mapping in *Prunus*

The *Prunus* reference map was constructed using an F_2 population (n=75) from a cross between ‘Texas’ almond and ‘Earlygold’ peach (T x E) (Joobeur et al. 1998). The T x E map was

first reported having 246 markers (235 RFLPs and 11 isozymes) in the expected eight linkage groups (Joobeur et al., 1998). An updated map with an addition of 96 simple sequence repeats (SSRs) has been reported by Aranzana et al. (2003). Another 220 markers consisting of 89 SSRs, five sequence-tagged sites (STS), and 126 RFLPs were added by Dirlewanger et al. (2004). The current map has 562 markers, covering 519 cM with an average density of 0.92 cM per marker (Genome Database of Rosaceae, website).

The sweet cherry map, developed from a cross between the cultivar 'Emperor Francis' and a wild forest cherry called 'New York 54' (EF x NY) was constructed using primarily simple sequence repeat (SSR), amplified fragment length polymorphism (AFLP) and sequence tagged sites (STS) markers derived from the *Prunus* reference map. Only 26% of the SSR markers were successfully placed on the parental EF or NY maps (Olmstead et al. 2008). The study attributed this to a reduced transferability of non-cherry *Prunus* markers and a low level of polymorphism between the mapping parents. However, Olmstead and colleagues (2008) successfully developed new markers for sweet cherry. The increased marker density resulted in the expected eight linkage groups for both parental maps. The EF and NY maps were 711.1 and 565.8 cM, respectively, with the average distance between markers of 4.94 and 6.22 cM. It was reported that 82 markers were shared between the EF x NY and the *Prunus* reference map and the majority of the markers were ordered the same as the *Prunus* reference. This shows the homology and the linear nature of *Prunus* diploid genomes, but it can be expected that polyploid species such as tart cherry and chokecherry may have rearrangements relative to the shared loci. Chokecherry is a minor plant species and has far less genetic resources to employ compared to its significant *Prunus* cohorts like peach, cherry, and almond. Therefore, identification of genetic loci contributing to important traits is inhibited and the production of genetic resources is

warranted; however, the limitations of genetic mapping in tetraploid species need to be considered.

1.3.4. Mapping in tetraploid species

Complex inheritance has limited the advancement of genetic maps in tetraploids, especially for autotetraploids. Random combinations of bivalent paring and quadrivalent paring from four homologous chromosomes and genetic anomalies like double reduction are a few examples leading to the complexity. Nevertheless, autotetraploid species such as alfalfa, potato, blueberry, and cut roses have been successfully mapped (Bradshaw et al. 2008, Gar et al. 2011; Robins et al. 2008; McCallum et al. 2016). Many allotetraploid species such as tart (sour) cherry, durum wheat, cotton, and rapeseed (*Brassica napus*) have also been mapped successfully (Canli 2004; Rong et al. 2007; Marone et al. 2012; Cai et al. 2014). True allotetraploids have the same inheritance as diploids, and are much easier to genetically map; however, it is often thought that many tetraploid species have intermediate inheritance (Hickok 1978a,b; Stift et al. 2008; Koning-Boucoiran et al. 2012). This means that some chromosomes have diverged enough to preferentially pair to produce diploid inheritance (disomic), while others are similar enough to have levels of randomly pairing or produce quadrivalents during meiosis, leading to tetrasomic inheritance. Segregation analysis is the recommended way to determine allelic inheritance and the corresponding polyploid type (Krebs and Hancock 1989; Soltis and Rieseberg 1986). The aforementioned considerations may seem dire for genetic mapping of tetraploids, but a universal approach has been established and widely applied.

Wu et al. (1992) proposed the use of single dose restriction fragments (SDRF) to overcome the difficulty of mapping polyploids. This type of mapping strategy accounts for syntenic rearrangements of homeologous chromosomes. It has been also been proposed by Wu

et al. (1992) that a population size of at least 75 is needed to confidently map in tetraploid species. Molecular markers used for tetraploids are scored as 'single dose' alleles because this produces simplex by nulliplex arrangements that have recombination frequency estimates consistent across both types of inheritance. Ultimately this provides a means for software programs such as JoinMap (Van Ooijen 2006) and TetraploidMap (Hackett and Luo 2003; Hackett et al. 2007) to construct reliable genetic maps.

Potato is a good example of a tetraploid crop species mapped via JoinMap and TetraploidMap. In 2008, Bradshaw et al. was able to successfully develop genetic linkage groups in TetraploidMap by using 38 AFLPs and 514 SSR markers. The genetic maps were subsequently subjected to QTL analysis using the Interval Mapping function in the TetraploidMap program. This resulted in the identification of 16 QTL associated with agronomic traits, including yield. More recent applications of JoinMap and TetraploidMap in potato were reported by Hackett et al. (2013) and Massa et al. (2015). Both studies utilized the Infinium 8303 Potato SNP array in conjunction with allele dosage information to develop new genetic maps for their populations and successive QTL identification for certain agronomic traits.

Autotetraploid blueberry has also been successfully mapped using both TetraploidMap and JoinMap. McCallum et al. (2016) utilized SNPs and SSR markers to construct the first representative linkage groups in blueberry. They used TetraploidMap first to identify the groups, and then JoinMap was used to refine the maps one linkage group at a time. Additionally, JoinMap has the advantage of being able to combine parental maps based on the shared markers. This function was utilized for blueberry and resulted in an improved genetic map and a consensus framework to conduct future studies like QTL mapping.

Tart cherry was the first polyploid *Prunus* species mapped and has provided an outline for genetic mapping in other tetraploid *Prunus* species (i.e. chokecherry). Tart cherry is thought to be derived from natural hybridization between sweet cherry (*P. avium*) and ground cherry (*P. fruticosa*). This origin was first suggested by Olden and Nybom (1968), and was later confirmed with molecular biology and genetics approaches (Hancock and Iezzoni, 1988; Santi and Lemoine, 1990; Schuster and Schreiber, 2000). Genetic mapping of tart cherry utilized the single dose restriction fragments SDRF or the 'single dose allele' strategy of molecular marker genotyping even after previous allele segregation analysis determined disomic inheritance (Beaver et al. 1993, Canli 2004). JoinMap software constructed the final linkage groups. Further analysis of the linkage groups revealed 11 QTL linked to flower and fruit traits.

1.3.5. QTL mapping for disease resistance in *Prunus* species

Diseases in *Prunus* species are major contributors to economic losses and production limitations. QTL mapping for disease resistance using genetic maps has facilitated efforts to combat diseases in *Prunus* species. An early effort in disease QTL identification was to identify QTL related to plum pox virus (PPV) resistance in apricot (Lambert et al. 2007). A population of 220 F_1 progenies from a cross between a susceptible cultivar 'Polonais' and a resistant cultivar 'Stark Early Orange' were analyzed for resistance phenotypes and subjected to QTL analysis. Four genomic regions were identified as being involved in PPV resistance. In addition, candidate genes associated to these major QTL were found via sequencing and map-based cloning. More recently, Rubio et al. (2015) was able to identify gene expression patterns related to disease onset via RNA-seq of apricot.

Peach is the most significant fruit species in *Prunus*. Peach is also the most genetically studied and has a high quality reference genome (IPGI 2013). Employing all the tools available,

many QTL and genetic loci have been determined to be associated with phenotypic traits. Zhebentyayeva et al. (2013) utilized genetic maps and sequencing information to elucidate the genetic mechanisms involved with chilling requirement for floral development in peach cultivars. They found two genes that are associated with a major QTL located on linkage group 1 of the peach map. Disease resistance QTL have also been well documented in peach.

Powdery mildew, caused by *Sphaerotheca pannosa* is one of the most severe diseases seen in European peach orchards (Foulogne et al. 2003). Foulongne et al. (2003) identified 13 QTL associated with the disease, which were subsequently used for crop improvement via marker-assisted selection (MAS). In 2012, QTL contributing to resistance to phloem feeding aphids was documented (Sauge et al. 2012). Their results suggest a change in phloem characteristics contributes to aphid resistance. Another severe disease in peach called bacterial spot (caused by *Xanthomonas arboricola*) has been studied extensively. In a report by Yang et al. (2013), a total of 14 QTL with additive effects on bacterial spot resistance were identified. The availability of sequence information and high quality molecular markers (i.e. peach SNP arrays) help facilitate this field of research. High quality genetic maps collaborate with the sequence information to deduce the genes contributing to QTL effects. Molecular genetics and genomics tools in other *Prunus* species may not be as full-bodied as in peach, but utilizing synteny and genome sequence similarity has effectively provided related species the resources needed to conduct molecular genetics studies.

1.3.6. Transferable molecular markers for studying related species

A cost-effective way to conduct genetic mapping studies in inadequately studied species involves exploiting molecular markers that are transferable to closely related species. A great example of this approach has been seen in Rosaceae. A somewhat novel approach of developing

markers for intra-family genetic analysis was conducted by Cabrera et al. (2009). They developed a group of Conserved Ortholog Set (COS) markers specifically for rosaceous species. The effectiveness of these markers was demonstrated by determining the amplification frequencies as well as their polymorphic extent. Transferability of molecular markers across the Rosaceae family (primarily in the *Malus*, *Prunus*, and *Fragaria* genera) was verified, indicating their great value for genetic analyses within family due to the conserved nature of the markers.

It has also been shown that commonly used markers, such as simple sequence repeat (SSR) and single nucleotide polymorphisms (SNPs) are transferable. Peach SSR markers used in the T x E reference map have been used in many other *Prunus* species including sweet cherry, tart cherry, chokecherry, apricot, almond, and more with varying percentages of transferability (Canli 2004; Olmstead et al. 2008; Dangl et al. 2009; Wang et al. 2012; Lambert et al. 2007). Wide transferability of other rosaceous markers (i.e. apple and pear) to *Prunus* species has also been demonstrated (Cabrera et al. 2009; Canli et al. 2004 & 2008; Zhang et al. 2014).

Transferability of chokecherry SSR and retrotransposon-based molecular markers has been demonstrated (Wang et al. 2012; Liang et al. unpublished). Retrotransposons are uniformly distributed throughout plant genomes and their long terminal repeats (LTR) are quite conserved. The conserved LTR sequences are a result of copy and paste transpositions of these DNA elements (Havecker et al. 2004). Transposition/insertion of TEs on plant chromosomes generates unique junctions between TEs and their flanking sequences (Bennetzen 2000). Several molecular marker types (REMAP: retrotransposon-microsatellite amplified polymorphism, IRAP: inter-retrotransposon amplified polymorphism, RBIP: retrotransposon-based insertional polymorphism, ISBP: insertion-site-based polymorphism) have been developed based on LTR retrotransposons and have been employed for genetic and QTL mapping in many

plant species (You et al. 2010; Mazaheri et al. 2014; Monden et al. 2014; Sun et al. 2015). These LTR markers have been used in the mentioned studies to increase marker density on genetic linkage groups. Combining several types of molecular markers from related species is a proven concept that has been employed for QTL mapping in many plant species (Olmstead et al. 2008; Keyser et al. 2010; Lambert et al. 2007; Guajardo et al. 2015; Canli 2004; Dirlewanger et al. 2004).

1.3.7. Chokecherry genetic and QTL mapping for X-disease

Chokecherry is currently the only plant species to display natural resistance to X-disease phytoplasma. Until recently, there was no genetic information on chokecherry and X-disease. Wang et al. (2014) developed a genetic map (RC x SC) from a cross between an X-disease resistant chokecherry (RC) and susceptible chokecherry (SC). The mapping population was composed of 101 progenies with varying degrees of X-disease resistance. Commonly used markers, SSR and AFLP, were used based on the knowledge that they are transferable and polymorphic among species (Dirlewanger et al. 1998; Canli et al. 2004; Gisbert et al. 2009; Olmstead et al. 2008; Wang et al. 2012; and Yamamoto et al. 2005). More recently, retrotransposon-based LTR markers developed from chokecherry genome sequences have shown to be polymorphic in chokecherry and are being used for mapping (Liang et al. unpublished). The previous genetic mapping progress in chokecherry has identified a quantitative trait locus (QTL) associated with X-disease resistance, accounting for 26% of the phenotypic variation (Wang et al. 2014). Large gaps and unassigned linkage group segments of the previous maps warrant the improvement of the chokecherry linkage map, which will provide opportunities for more QTL identification, map-based cloning, and facilitated sequence assembly.

2. MATERIALS AND METHODS

2.1. Plant Materials

2.1.1. Mapping population

The chokecherry (*Prunus virginiana* L.) mapping population used in Wang et al. (2014) was used again to construct a novel genetic linkage map and to identify QTL associated with X-disease resistance. The mapping population consists of 101 progenies which derived from a cross between a susceptible (SC) and a resistant (RC) chokecherry lines. The two parental lines were selected from a large chokecherry germplasm collection that was established in 1983 by the USDA Natural Resources Conservation Service (NRCS). The female parent (R-II-2010-3, 'RC') was rated as highly resistant to X-disease phytoplasma, and the male parent (S-V-2007-3, 'SC') was rated highly susceptible.

2.1.2. Phenotyping of X-disease resistance in the mapping population

Chokecherry hybrid seedlings were inoculated with an aggressive X-disease phytoplasma strain using a side grafting method (Wang et al. 2014). In brief, scions consisting of fresh symptomatic chokecherry branches were collected from the source tree less than two hours before being grafted to the stem of the seedling. Non-inoculated seedlings were used as controls. After 2 weeks of growth, nested PCR was used to confirm X-disease infection. Nested PCR utilized universal phytoplasma primers (R16 F2- ACGACTGCTGCTAAGACTGG and R16 R2-TGACGGGCGGTGTGTACAAACCCCG) and X-disease specific primers (R16 (III) F2-AAGAGTGGAAAAACTCCC and R16 (III) R1-TCCGAACTGAGATTGA). For more information on nested PCR conditions see Wang et al. (2014).

Phenotypic data has been collected over a period of four years (2008, 2009, 2010, 2011) using the disease resistance scoring method described in Walla et al. (2004) and Wang et al. (2014). In short, X-disease severity was rated on a scale of 0–5 based on the level of X-disease symptoms

and tree vigor (0 = whole plant died; 1 = leaves were discolored, most shoots and leaves were very stunted, and very low growth vigor; 2 = leaves were discolored, most shoots and leaves were stunted, and low growth vigor; 3 = leaves were discolored, most shoots and leaves were moderately stunted, and moderate growth vigor; 4 = all or part of tree with slight symptoms and high growth vigor; and 5 = no symptoms and high growth vigor). The final disease severity rating used in QTL analysis was an average score from the last two years of evaluation (5–6-year-old trees).

2.1.3. DNA extraction

Leaf samples from all offspring and both parental lines were collected and stored in a -80^{o} C freezer. DNA was extracted from leaf tissues based on cetyltrimethylammonium bromide (CTAB) methods described in Lodhi et al. (1994) with some modifications. Additional steps included washing ethanol-precipitated DNA with 70% ethanol in a slow moving shaker for 2-3 h before being dissolved in tris-EDTA (TE) buffer and then digesting DNA solutions with RNase A (10 mg/ml) followed by Proteinase K (1mg/ml) for 60 and 15 min, respectively. DNA concentration was determined using a NanoDrop ND-1000 Spectrophotometer (Thermo Fisher Scientific Inc.) and stored at 4 °C until use.

2.2. Molecular markers used in chokecherry genetic mapping

2.2.1. Previous markers proven in chokecherry

Simple sequence repeat (SSR) and amplified fragment length polymorphism (AFLP) markers produced a total of 498 ‘single dose’ alleles analogous to single dose restriction fragments (SDRFs) suitable for tetraploid chokecherry mapping (Wang et al. 2014). The SSR markers originated from chokecherry and six other *Prunus* species including peach (*P. persica*), sweet cherry (*P. avium*), Japanese plum (*P. salicina*), apricot (*P. armeniaca*), almond (*P. amygdalus*), and tart (sour) cherry (*P. cerasus*). The information on SSR markers from other *Prunus* species were obtained from the Genome Database for Rosaceae (GDR) website

(www.rosaceae.org). All 498 qualified markers were re-tested for the usability in this research. All molecular markers were scored in the same fashion as single dose restriction fragments (SDRFs) and identified as single dose alleles based on the presence and absence of each band in the image from the parents and all 101 progeny.

2.2.2. Development of LTR retrotransposon markers

Long terminal repeat (LTR) retrotransposon markers were developed from a partial genome sequence in chokecherry. LTR development is explained in more detail in Liang et al. (unpublished). In brief, sequence contigs of chokecherry (Wang et al. 2012) were searched for LTRs using the online software 'RepeatMasker' (Smit, Hubly, and Green 2013). The function 'cross_match' search engine and '*Prunus*' was specified as the DNA source. Regions containing LTRs were used to design primers flanking unique repeat junctions. The primers were designed using Primer Premier 5.0 (http://www.premierbiosoft.com). A total of 78 polymorphic chokecherry LTR markers were scored as single dose alleles for the whole mapping population (Table 1). Polymerase chain reaction (PCR) conditions are described below in section 2.3.

2.2.3. Development of SSR markers using peach genome sequences

A total of 48 new SSR marker primer pairs were designed from the peach reference genome (*Prunus persisca* whole genome v1.0; Verde et al. 2013). The software 'RepeatMasker' was used as described in 2.2 above with the additional option 'Only mask simple…' to quickly locate simple sequence repeats from the scaffolds. Peach scaffolds two, four, and six were used in the simple repeat search. Primers were designed based on the flanking regions of the simple repeats using Primer Premier 5.0 (http://www.premierbiosoft.com). Primers were screened via BLASTn to examine off-target effects. If off-target effects were detected, primers were either discarded or re-designed. See Table 2 for a complete list of SSR markers designed. The 48

developed markers were examined in a subpopulation of chokecherry (n=8) to analyze amplification and polymorphic potential. PCR amplification of the SSR primers were done in 20 µl reactions consisting of 60 ng template DNA, 1.5 mM $MgCl_2$, 1× PCR buffer (Promega M791A), 200 µM dNTP, 0.2 µM of each primer, and 0.2 U *Taq* DNA polymerase. PCR amplifications were performed with the following procedure: 94 ^{o}C for 4 min, followed by 35 cycles of 94 ^{o}C for 30 s, 56 ^{o}C for 30 s, 72 ℃ for 60 s (8 min for the final extension cycle). The PCRs were conducted with a Programmable Thermal Controller PTC-100^{TM}. PCR products were separated in a non-denaturing 6% polyacrylamide gel (29:1 acrylamide: bis, J.T. Baker, Mallinckrodt Baker, Inc. NJ). The gel was prepared as following. A mixture of 7.5 ml 10× TBE buffer, 22.5 ml acrylamide (29:1), 119.9 ml ddH_2O, 106 mg APS (Ammonium Persulfate), and 110.7 µl TEMED (Tetramethylenediamine) was poured between two glass plates and left to polymerize for one hour. A pre-run of the resultant gel was done at 350 V in 0.5× TBE buffer for 1 h to allow ethidium bromide to migrate from the buffer into the gel. A total of 8µl of PCR products were loaded into wells of the gel and ran at 250 V for 3 h. The gel was visualized under UV lights and images were captured using software 'Alphaimager 2200'.

2.2.4. Newly adopted SSR markers from the Rosaceae family

Additional markers from recent publications were adopted for use in this study. Dettori et al. (2015) successfully developed 216 long core (tri, tetra, and penta) simple sequence repeats from the whole peach genome sequence (Peach v1.0; Verde et al. 2013). A group of 176 SSR markers that amplified products in sweet cherry (Dettori et al. 2015) were tested in chokecherry. Zhang et al. (2013) developed 194 SSR markers from assembled sequences of pear (*Pyrus*). Only 17 pear SSR markers were reported transferable in Rosaceae in which 11 were tested in chokecherry. Sun et al. (2015) developed 8 new highly polymorphic and multi-allelic LTR-

based markers and these 8 pear markers were all tested in chokecherry. All markers were tested in a subpopulation (n=8) of chokecherry to identify polymorphic banding patterns. Polymerase chain reaction was conducted via the same protocol found in 2.3; however, changes in annealing temperature were made according to the primer characteristics (Table 3). Polymorphic banding patterns within the chokecherry subpopulation (n=8) identified the adopted markers used for the full mapping population.

2.3. Genetic map construction and QTL identification

2.3.1. Segregation analysis

All of the molecular markers were scored for the presence/absence of individual marker alleles and tested in accordance to Koning-Boucoiran et al. (2012) and Beaver et al. (1993). The terminology 'phenotypic class' was used to describe when the marker genotype in terms of dosage of an allele could not be observed directly. Segregation of molecular markers with two (duplex) or three (triplex) unique single-dose alleles in one parent and nulliplex in the other parent was analyzed. The segregation ratios of the phenotypic classes of the progeny were used to reveal the allelic constitution of the gametes contributed by a parent, which allows the identification of the meiosis mode of inheritance for a particular locus. The hypotheses of segregation according to disomic and tetrasomic inheritance were both tested by Chi-square goodness-of-fit test at $\alpha = 0.01$. In the case of allotetraploids with a disomic inheritance, four phenotypic classes are expected with a frequency of 1/4 each. In the case of tetrasomic inheritance, random homeologous chromosome pairing and resultant six phenotypic classes in the progeny are expected with frequencies of 1/6 each. These six classes can be directly recognized if the parent has four alleles at a locus that segregate in the progeny. If a marker has only three different alleles segregating, the presence or absence of the fourth allele (null allele O)

can be inferred assuming a single-locus situation. Markers that produce only two unique single-dose alleles in one of the parents of the mapping population can also be used. In this situation, not all possible phenotypic classes can be distinguished for autotetraploids. Alternatively, four phenotypic classes are expected with frequencies 1/6, 2/6, 2/6, and 1/6. Disomic inheritance will have expected frequencies of 1/4, 1/4, 1/4, and 1/4.

2.3.2. Map construction using JoinMap

Linkage analysis was performed using JoinMap 4.0 (Van Ooijen, 2006) for cross-pollinated progeny. All molecular marker single dose alleles were coded as simplex or nulliplex in reference to the parental type (nn x np, lm x ll), because estimating recombination frequencies in autotetraploids and diploids are identical for simplex by nulliplex markers. Map constructions were performed following a 'Two-Step' strategy (Tavassolian et al. 2010; Klagges et al. 2013) that involved constructing parental maps separately before combining. The segregation pattern of markers was tested and distorted markers (chi-square threshold of 0.001) were eliminated from analysis. 'Suspect Linkage' and 'Genotype Probabilities' tabs were used to identify mis-grouped markers and double recombination, respectively. Regression mapping was used as the mapping algorithm with Kosambi's mapping function to convert recombination frequency into map distance. Logarithm of odds (LOD) of 3.0 was the minimum LOD score to establish linkage groups. The inspection of proper assignment of a marker to a group was done by calculating the Strongest Cross Link (SCL) parameter. Also, ungrouped markers were manually transferred into established groups by examining SCL and related LOD values.

The process of removal of unfit loci, reassigning groups, and mapping was done for each individual parental map until a limited number of markers could not be assigned to a linkage group. Final linkage groups were compared between each parental map to define a consensus

grouping based on homologous loci. Non-corresponding linkage groups were aligned with each consensus parental grouping to check for conflicting markers before combining the parental maps. Newly joined chokecherry maps were developed by using the 'Combine Groups for Map Integration' function of JoinMap. If combined groups showed markers that were not suited for regression mapping, MergeMap Online (Wu et al. 2008) was used to finish ordering the new groups. All combined chokecherry linkage groups were drawn using MapChart 2.30 (Voorrips 2002) and compared with previous chokecherry (RC x SC) maps (Wang et al. 2014).

2.3.3. QTL analysis using Qgene

A combined map of the redefined chokecherry parental maps (Wang et al. 2014) was used for the QTL analysis using the software QGene 4.3.10 (Joehanes and Nelson, 2008). Normality test of the phenotypic data was reported as a Kolmogorov-Smirnov (K-S) *p*-value. Composite Interval Mapping (CIM) was used to detect quantitative trait loci (QTL). Nearby loci with the highest LOD scores were selected as cofactors per the default parameters set in the program. Permutation tests with 1,000 iterations were used to determine significant LOD thresholds at the 95% and 99% confidence levels for the experiment-wise Type I error. An overall schematic of the process of starting with PCR products to establishing significant QTL is illustrated in Figure 1

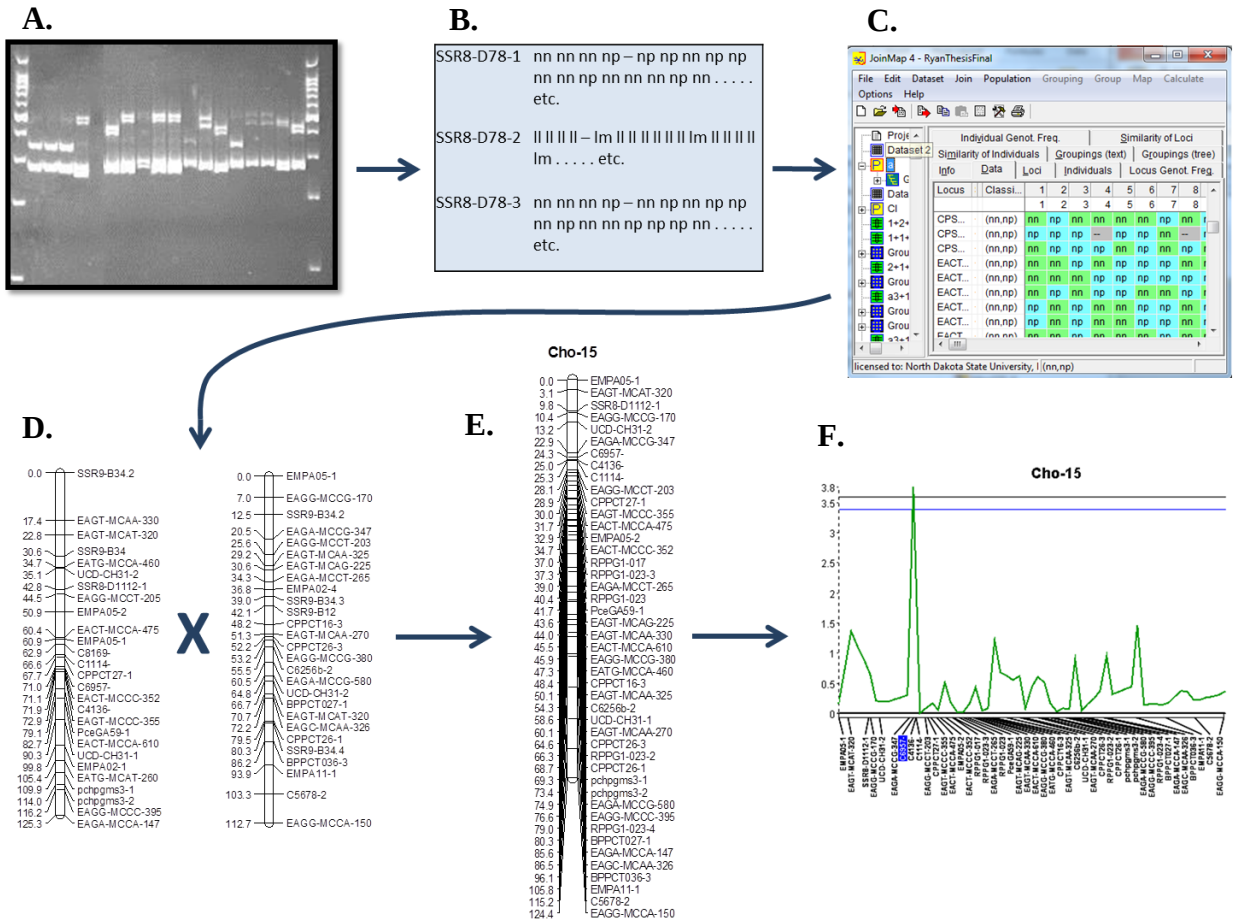

Fig. 1. A flowchart summarizing the steps taken to create a genetic linkage map and identify significant quantitative trait loci (QTL). A. Gel electrophoresis image of amplified alleles from a single PCR-based molecular marker (i.e. SSR); B. The amplicons presented in the gel image are recorded individually as 'single dose' alleles according to the coding requirements of the utilized software package; C. A collection of all the genotyping data is uploaded to a software package (i.e. JoinMap) and tests are run (i.e. segregation distortion) to eliminate molecular markers from the genetic mapping process. Genetic mapping is conducted within the software program to develop separate parental linkage groups; D. Parental linkage groups are compared via MapChart to identify shared genetic loci for integration; E. Homologous parental linkage groups as indicated by shared loci are combined to produce an integrated consensus map; and F. Progeny genotype data within each linkage group is associated with its corresponding phenotypic (i.e. disease resistance) score and analyzed via composite interval mapping and permutation within a QTL analysis software program (i.e. QGene).

2.3.4. Comparative analysis of *Prunus* genetic maps

Synteny analysis with other *Prunus* maps was conducted. The linkage groups from the refined chokecherry map (linkage groups Cho-1 to 16) were compared to the *Prunus* reference

map (T x E) (Joobeur et al. 1998; Aranzana et al. 2003; Dirlewanger et al. 2004) and the sweet cherry linkage map (EF x NY) (Olmstead et al. 2008). Homologous loci within the T x E and EF x NY maps are reported. The new chokecherry map produced herein was also compared to the previous chokecherry (RC x SC) map (Wang et al. 2014). MapChart 2.30 (Voorrips 2002) was used to visualize how the new map combined the parental maps and the small chromosome segments.

Table 1. Chokecherry retrotransposon-based long terminal repeat (LTR) markers that produced polymorphisms in chokecherry

Marker name	Marker type[a]	Source species	Forward sequence	Reverse sequence	Ta[b]
LTR1-A12	RBIP	Chokecherry	GATGTAAAGACAGCCTTCCTCC	CTACACCCAATCCGCAGCATA	56
LTR1-B56	RBIP	Chokecherry	ATGTTACCTCCTCAATCCG	AGATAACGCATAATCCAGTAC	56
LTR1-B78	IRAP	Chokecherry	CTAAAGCAGGCACCTCGAC	GATCTCCTGCCATGTCTGC	56
LTR1-C12	RBIP	Chokecherry	TTCCTACAACCGTCTCACT	CCATCATCAACCACTCCGT	56
LTR1-C56	RBIP	Chokecherry	TTCGGTTTCAAGCCTCCAT	AGCCTTTGACCTTCTGGAC	56
LTR1-D1112	RBIP	Chokecherry	TGGCGGAAAGAAAGTTAGGT	CAGTATTGCTGCTGAGGTG	56
LTR1-F78	RBIP	Chokecherry	ACAACCCCTTATGAGCCT	TCAAGTACTGCTGGAAATTC	56
LTR1-G12	RBIP	Chokecherry	ATACCCTTTACTGGTTTCAC	TATTGTGGTTGAGGTGGT	56
LTR1-G78	RBIP	Chokecherry	TTATGGGACAAAGAAGGAAC	CGCTACCTCGGGTGTCTC	56
LTR2-A34	RBIP	Chokecherry	CAGCCCACTTGGTGAAATAG	TCCGAGTCAACTTGGATGG	56
LTR2-B1112	RBIP	Chokecherry	TCTCGTCGTGGACAAGTAAC	GATTCGGGATAGAGGCTTG	56
LTR2-C12	RBIP	Chokecherry	TGTAAAACTCCCTGGGTCTTG	CAGATTCTACAGAGTCTATGG	56
LTR2-C56	RBIP	Chokecherry	ACGCCCAGCAAGTGTCTAT	CAAGGGTCGTCCTGTCTCC	56
LTR2-C78	RBIP	Chokecherry	GCACCATTCTAACTCCTCA	GCCACCGACTTCTTCACG	56
LTR2-C1112	RBIP	Chokecherry	ACCCTATTGAACCAGAACC	TTTGACACTTGAACTCGGT	56
LTR2-D34	RBIP	Chokecherry	CATGGCTCACTTACCTCAA	GTCATCTACTACGCATCCC	56
LTR2-D78	RBIP	Chokecherry	CTCTGCTTCGTTGTTCTGG	GGGTTCAAATCCTCCACTTC	56
LTR2-D910	IRAP	Chokecherry	CCCCTAAGTCCGATCCATT	CTCAGATTCTTGCTGGTCC	56
LTR2-E12	IRAP	Chokecherry	TGGAGGTCTAACAACAGAAG	TGCCTGCTGATTTTGTAGTGT	56
LTR2-G34	RBIP	Chokecherry	GAAACCTCTTCCGTTAGTG	AACCGTGAACCTACTGATG	56
LTR2-G56	RBIP	Chokecherry	GTTGGTGAGGGTTGGTTTC	GTGAGAAGTATGATGTGCTG	56
LTR2-G78	RBIP	Chokecherry	GAACTAATACTGTCAAGGGAG	CTGCGAGATTTAGACCACC	56
LTR2-G1112	RBIP	Chokecherry	GGATTTATGTGCGGCTCAAG	TGTCGAACCCTGTGCTGTC	56
LTR2-H12	RBIP	Chokecherry	GAAACTTACCTTGTGCGTGCTG	CGAGTCAAGCCCTATCCGT	56
LTR2-H78	ISBP	Chokecherry	TCCACTTCAGCAGCCTTAG	CTGTCAGTCACACTATGTAAGC	56
LTR3-A12	RBIP	Chokecherry	TTAACCGCAGGGTCCGTCT	CTATGTCTCCGAGGCTGAAG	56
LTR3-A56	RBIP	Chokecherry	GCCCTCCGATGCTAAAGTT	CAAGTCCCCAACCAAGAAG	56

Table 1. Chokecherry retrotransposon-based long terminal repeat (LTR) markers that produced polymorphisms in chokecherry (continued)

Marker name	Marker type[a]	Source species	Forward sequence	Reverse sequence	Ta[b]
LTR3-B34	RBIP	Chokecherry	CCGAAAGCACAACAACCAT	TCGCTATCTCAAGGGGTCT	56
LTR3-B1112	RBIP	Chokecherry	GTCAGCACCATCATAGTTACG	GGTCGGGTCCTGTCAACTT	56
LTR3-C12	RBIP	Chokecherry	TAGAAACCGTTAGGAATAGG	CCTCAAGATGCCTAAGACC	56
LTR3-E12	RBIP	Chokecherry	ACTAACTCGGAAAGTGCTC	CTTTTCCTATCTTTCGTTCC	56
LTR3-E1112	RBIP	Chokecherry	TTGGACTACCTACCACTCT	TAATCTCAGCATGTAAGTCG	56
LTR3-H78	RBIP	Chokecherry	AGATTGCTCTGTATCGCCCAT	CACCGAATCCTGCTGAAGTT	56
LTR4-A34	IRAP	Chokecherry	GAGGACGTGTCACCTATGGT	TGACAGGGATCAAGCCACAC	56
LTR4-A910	IRAP	Chokecherry	CACAGGCTTATAGGCTGGAG	CACCAGACGACCCAAGAGAT	56
LTR4-A1112	RBIP	Chokecherry	GCAACCTTGTAGGGATTAG	AACTTTCCAGAACCAGCGT	56
LTR4-B12	RBIP	Chokecherry	TATGCTTTGTGCTCTGTCT	CTTCATAGAGGGCATTTTC	56
LTR4-B1112	RBIP	Chokecherry	AGGACCCAGGGAGTTTTAC	GGTGGCGACTTGGCTTTC	56
LTR4-C34	RBIP	Chokecherry	GGAAACTATTCTAGGGATGT	GACAACAAAGCCGCAATAG	56
LTR4-C78	RBIP	Chokecherry	GGTGAAGCCCTGATGACTG	CCAATGGAGGAGGAACTGG	56
LTR4-C1112	RBIP	Chokecherry	TCCTGGCATTTCTTTGTACG	ATCAGCCCCACGCAAATCG	56
LTR4-D34	RBIP	Chokecherry	GGTGATGGGCTTTTAGGGT	GAAGAAGCATTTGGGAACTC	56
LTR4-E34	RBIP	Chokecherry	ACCAAATGCTCACTACCTTC	GAATGAAGCTGACTTTGATGGT	56
LTR4-E56	RBIP	Chokecherry	TTATCCTCCCTGCTGACTT	TGGAAGGAACGAGTTTGGT	56
LTR4-E78	IRAP	Chokecherry	AAGTTTGGGTGTTGTCTCC	GATTGAACTCGCAAAGCAG	56
LTR4-F12	RBIP	Chokecherry	CTTTTGGCTAGGGCTTTCC	TGGATACATCAGCACAACAC	56
LTR4-F1112	RBIP	Chokecherry	ATTCCACAAATCTGCAAAGAC	ATCCCTGGAGCAGAAATGG	56
LTR4-G56	RBIP	Chokecherry	AGGCAAGTCGTTTGAAGAAG	TCTACTCTTGGGGAGTGAG	56
LTR4-G78	IRAP	Chokecherry	CAAAGCCTGCTCTGATACC	AAGTTGGGTGTTCTTGACAT	56
LTR4-H910	RBIP	Chokecherry	CGGTGGGGTTAGTATCTTG	ATCTCCTTGCCTTCCTGAC	56
LTR5-B56	RBIP	Chokecherry	TGGGAAATAGAGCGATACT	CAACTCAAGTATGCGTATG	56
LTR5-D1112	RBIP	Chokecherry	CTGGAGAAAGGGCAATAGT	CGAAACATACCCGAGAAAC	56
LTR5-E12	RBIP	Chokecherry	CATTCTCCAGGTGCTATCT	ATCCTCACCACAACAGCAT	56
LTR5-F12	IRAP	Chokecherry	ATGTGGTACAAGCACTTCGGT	CCCAAACAAGGAACTCACG	56

Table 1. Chokecherry retrotransposon-based long terminal repeat (LTR) markers that produced polymorphisms in chokecherry (continued)

Marker name	Marker type[a]	Source species	Forward sequence	Reverse sequence	Ta[b]
LTR5-F78	RBIP	Chokecherry	CAAAGAATCGCTCTACCTC	ACCCGAACAATGATAGGAC	56
LTR5-H78	RBIP	Chokecherry	TATGCTGGCGTTAAGTGGT	CCTCATTTCGTGGAGTGCT	56
LTR6-A78	RBIP	Chokecherry	AATCGGCAGGGGCTGTAAT	TATTCCTCAGATGGCAGAC	56
LTR6-B34	RBIP	Chokecherry	CCTTGATTTGTTGGTGCTG	TTTGGGATGAGCCTTGTGG	56
LTR6-B910	RBIP	Chokecherry	GACACCTCCCTTTGATACTG	CTCCGATGCCTAAGTCAGT	56
LTR6-D12	IRAP	Chokecherry	TTCACTGGACATGGCTTGC	TTCCTTATGAAGTCTGGGTG	56
LTR6-F56	RBIP	Chokecherry	CTACAACCTTAACGTGGAC	CAGATTTCAGTCTAACCCAG	56
LTR7-B78	RBIP	Chokecherry	TCGGTATTTTCAGAGTAGGT	AGATACGCAACTTCGGTCC	56
LTR7-B910	RBIP	Chokecherry	GATGACCTCAGACGCCTTGT	CAACCAAGAAGGGCAAAGAG	56
LTR7-C12	RBIP	Chokecherry	AGACCGACACGCTTAGGGAC	AGTGTGTTGAGTCTCCATTGCT	56
LTR7-C56	RBIP	Chokecherry	TGGGCGATACAGAGCAATC	TGGAACGGTGGGATTTACT	56
LTR7-C78	RBIP	Chokecherry	CTCCGAGCGAGTTGCTATC	TTCACGAGCCTCTGATACG	56
LTR7-C1112	RBIP	Chokecherry	TTCCAGACAAGTCGTTCCAC	CCAGTTCCTAATCCGCAC	56
LTR7-D12	RBIP	Chokecherry	CGAATCCTGTGGTTTTCCG	ATTGGGTTTTCGTGGTTGT	56
LTR7-D34	RBIP	Chokecherry	TAAGGCATTCAACTCACTCT	CATATTGCTCTTGCTGTCTC	56
LTR7-D56	RBIP	Chokecherry	GGAATGTCTATCACCCCTT	TTGGTCTGCGATGGGTTCT	56
LTR7-D910	RBIP	Chokecherry	TGATACCAAGTTGACACGAC	TTTCGGCAGAAGTCAAACC	56
LTR7-D1112	RBIP	Chokecherry	CATACTAACCCCACCGTTTC	GTGTGACTGTAGACCCTTG	56
LTR7-E12	RBIP	Chokecherry	GTTAGAGCCTTCCCAGATG	TACAAGCCAAAGCCTAAGTG	56
LTR7-E1112	RBIP	Chokecherry	CAACCAAGAACCCCAACG	GAAGGGCATGAGTACATAG	56
LTR7-G56	REMAP	Chokecherry	CCTTCCTTGCTTGTATTGTC	GTACTAACCCCACCGTTTC	56
LTR7-G78	RBIP	Chokecherry	CCATTGTTGGTGCTGTTGC	ATCCTGAAAAGAGCAATGTG	56
LTR7-H56	RBIP	Chokecherry	TCTTCCCCTTTGGTAACTG	GCCTCCACAGCAACAAGAT	56

[a]Marker types - REMAP: retrotransposon-microsatellite amplified polymorphism, IRAP: inter-retrotransposon amplified polymorphism, RBIP: retrotransposon-based insertional polymorphism, ISBP: insertion-site-based polymorphism

[b] Annealing temperature (°C) for polymerase chain reaction (PCR)

Table 2. Complete list of newly designed simple sequence repeat (SSR) markers from the peach reference genome

Marker name	Marker type	Scaffold	Forward sequence	Reverse sequence	Ta[a]
SSR8-A12	SSR	2	TGGCGTCTGAACTCCGATTG	GAACAACTTCCTTCCTCCATCT	56
SSR8-A34	SSR	2	TCTGAGCCCGTCACCTTTC	GAGAGAAACCGATTAGCACC	56
SSR8-A56	SSR	2	CAAGGCAGAAAATTGTGAGT	TTCTTATTTGTCGGTGAGG	56
SSR8-A78	SSR	2	ATTGACACTGGCACTCTTG	TCAACACCGATTCAAGCAC	56
SSR8-A910	SSR	2	ACGGGCTGATTTTGCCTTGG	CAACTGCGAAGTGCTTGACG	56
SSR8-A1112	SSR	2	GCCTAAGGTTATTAGGTGC	TTTCCAGTGCTGCTCTATC	56
SSR8-B12	SSR	2	GTGGAAGGAGTACAATGGATG	CCCGTCACTTTTGACTACTTG	56
SSR8-B34	SSR	2	CAAAACAGAAATTGTGAGTGG	ATCTCGGAAAGAGGTTGGT	56
SSR8-B56	SSR	2	TTTTGTGGCTCTTTACCCT	CCTTTCTAACCATCCCTCT	56
SSR8-B78	SSR	2	CCCTTGGGATACCGAATAG	AACAGTCCATCAATGCTCC	56
SSR8-B910	SSR	2	TGGAAAGTTACCCAAATGAC	TGAAAGGTTGGAACATAGAG	56
SSR8-B1112	SSR	2	TCAGGGTCTGGAGAACGGAG	TAAAATCGGGTAGGTGAGGAAC	56
SSR8-C12	SSR	2	TTTCTGGGGCAACAGTGAC	TTGAGCCTGTTTGGGATTG	56
SSR8-C34	SSR	2	CAGATGTGGTGTTTGCCGACT	GATCAACGGAAGTATAGTGGC	56
SSR8-C56	SSR	2	AAGGATTGGCAATGCGTGT	AAGACAAAGGTGAGAAAAGC	56
SSR8-C78	SSR	2	TGGTCAAAAGAAACCCAGTGG	TTTGGCACCAGTCCCGTCT	56
SSR8-C910	SSR	2	GATTTCAACCCGATTCTAC	GATAAATCCCACTACCCAT	56
SSR8-C1112	SSR	2	TCCAATCTCACAAAATCAGG	ACTTACTGAATCACCTTTGC	56
SSR8-D12	SSR	2	GTGTGGGCAAGAACTTATCAT	CCATTCCCAAAGTCACCGAT	56
SSR8-D34	SSR	4	ACTGCCCGCATTCTTGGTT	CTCTTCAGGCTTGACATACT	56
SSR8-D56	SSR	4	AAGCAGCATCTTCCTCTTC	TCAATACCGTCAAGATAACC	56
SSR8-D78	SSR	4	GTGATGATTGAATCTGCCAT	GACAAAGCTAATTCCACATC	56
SSR8-D910	SSR	4	ATTGCTGGGAAATCTGGAAG	ATCAAGCCAAACGGCAGAG	56
SSR8-D1112	SSR	4	TGTTGTACATGACCGTTAGC	CATTTTCACCAAGGCTTTACC	56
SSR8-E12	SSR	4	AGTCAGTGAGCCAGCCTATG	GGGAATGGGAACAGGAAACG	56
SSR8-E34	SSR	4	GAGATAAGGGCATTTTGGTCAT	CCCAACCATCAGCTAGAGC	56
SSR8-E56	SSR	4	CGCAAGGTTCATATTCTTC	CACCATCTGTCCAAATCAT	56

Table 2. Complete list of newly designed simple sequence repeat (SSR) markers from the peach reference genome (continued)

Marker name	Marker type	Scaffold	Forward sequence	Reverse sequence	Ta[a]
SSR8-E78	SSR	4	GCTTTGCCTTGTAGACGCTT	GTGGAGCATAAGCATACCTC	56
SSR8-E910	SSR	4	TGTTCTCATCCAAAGCAGC	TTTTCTTGAAGCACCGTGT	56
SSR8-E1112	SSR	4	ATGCCTGAATGAATGCTCT	AGTGGGAGATTGAGTTTGT	56
SSR8-F12	SSR	4	GCCTTGTAGACACTTCCCTG	TGTGGAACATAAGCGAACCTC	56
SSR8-F34	SSR	4	AGAACCTTGAACGATTGAC	TGTCTGGATCTTTGTTTGG	56
SSR8-F56	SSR	4	TGGGAGGTAATTTGGTGAC	AAATGGGAATCCTTGGTGT	56
SSR8-F78	SSR	4	GTTGTTGTTCTCATCCGAAAC	CTACCTTTGTTGTCCTCTGC	56
SSR8-F910	SSR	4	GTGTAGGAGTAGCCCTGGTT	AGAGGCAAAAGAGGAACAGTG	56
SSR8-F1112	SSR	4	ATTAGAGTTTAGGGTCGGTT	GCATTCCACCAGTACAACT	56
SSR8-G12	SSR	4	GTGACCGTTGGAGGCTGTAT	TACACCCATTTTGCTCACCC	56
SSR8-G34	SSR	4	CAGTGATACCTGCTACGAT	CTATCTGCTCCGATTCCTC	56
SSR8-G56	SSR	4	CTTTTGAAAGCAGACAGATC	GCCAGTCTAACCTTCTCAG	56
SSR8-G78	SSR	6	TTGCTAACAGTGGCACGCTC	AGAAAGAAGCAGTCGTCAAGC	56
SSR8-G910	SSR	6	CTTGGCAACAAAGTAAGAAC	CACTTAGAGCAATACCACTT	56
SSR8-G1112	SSR	6	CTTCTCGGGTGATGTGGTC	CGGTTGGTTGGTCAAGACAT	56
SSR8-H12	SSR	6	AACTCACAAGTTCCCAACC	GATTAGGTTTGGCAACTGAG	56
SSR8-H34	SSR	6	TTTGCTAGGTGTCCGTCTC	CGTTGATTGTAACTTCTTCG	56
SSR8-H56	SSR	6	CCTTTAGATGGGTAGTTTGC	GCACCCGTTCAAGAATCAC	56
SSR8-H78	SSR	6	ATTGCTTACTTCCCTGGTT	AGCGAGGCTTGTGATTGG	56
SSR8-H910	SSR	6	TGGAAAGGGTATAGCTGTG	CAAGATGACAACTGAGGCT	56
SSR8-H1112	SSR	6	GGTGTACGCTTCCACTATG	CGGATTATCGGACAAAGTG	56

[a] Annealing temperature (°C) for polymerase chain reaction (PCR)

Table 3. Simple sequence repeat (SSR) markers from (Dettori et al. 2015) and (Zhang et al. 2014) that were successfully anchored to the novel chokecherry genetic map

Marker name	Marker type	Source species	Forward sequence	Reverse sequence	Ta[a]
RPPG1-017	SSR	Peach	GCTCATCAAAACTCTCAACCA	CCCTTTCTTCAATCCCATC	56
RPPG1-023	SSR	Peach	GGCCTTTGTTTTCTTTCCTT	GGATTTCAGTTGACCCATTT	56
RPPG1-025	SSR	Peach	GATTTGATTCCTGTGGCATT	TGGGCATTCTTTTTCTCTTC	56
RPPG1-026	SSR	Peach	CTTCTGGCACTCTTCCATTT	GTTCCCAAGTTTTCCTCTCA	51
RPPG1-029	SSR	Peach	TCACTCCAGCATTTGAACC	AGCACTGAAAACACCACAGA	56
RPPG1-041	SSR	Peach	TGTTGTAATGGATGGTGTCTTC	CTTGGTCTTGGTTTCATTCA	56
RPPG2-007	SSR	Peach	GCATCAGAAGTCCCAATCA	GCGGTGGTGTGAAACTAAA	51
RPPG2-019	SSR	Peach	TTACGTGCTTTTCCCATGA	CGCCTTATCCCCTGACTAT	49
RPPG3-030	SSR	Peach	AAACTGCCCAAAACAAAGAC	GCAACCAACAAAGATGACAA	56
RPPG3-031	SSR	Peach	AGCGGAGAGAGAATGAGATG	GCAACAATACGAACAGCAAG	58
RPPG3-038	SSR	Peach	GTTTCCCATCCCATACCTC	CAACACAAGAAGCAAGCAAG	53
RPPG3-039	SSR	Peach	CAACACGTTATTGCCCATT	GTGAGCCACATTTACTATTGAGAG	49
RPPG3-041	SSR	Peach	TGCCATTCAACAACAAACAC	TCAAGGGAACAGGGATGA	54
RPPG4-074	SSR	Peach	AGTGGCTGTTCTGGTTTGAG	GTTTGGGGTTTGGAGAGAG	58
RPPG4-076	SSR	Peach	TGCCAACTATGCTCCTATTTAC	GGATTTGAATTGCCGAACT	54
RPPG4-097	SSR	Peach	GGCATGTGAAAGCAAAAGT	CTTCCTGAAAACCCCATTC	49
RPPG5-008	SSR	Peach	CCTGAATGGCTCTCTCTTTC	TGTTGGTGGGACTAATGATG	58
RPPG5-020	SSR	Peach	CAAGAATTTGGCTTGGAACT	GTGTATCATGGACAGCTTGC	56
RPPG5-024	SSR	Peach	TTAGAAAACGGGACAAGCAC	CAACGACACCATTGAAAACT	56
RPPG6-009	SSR	Peach	GGGCTTGGCTGATAAAATAA	TGGTAAAATAGAAGAGCGAGAAG	53
RPPG6-010	SSR	Peach	ACTTGACGTAGAGAGCATACCTAA	ATTATGGGCAGAAATGGTTG	51
RPPG6-014	SSR	Peach	ACCCAATACACAAGATTGACC	CTTTGGAAGCAGGGATTAGA	53
RPPG6-018	SSR	Peach	TCTGCTATCTGTTTGGTGGA	GACTACAGTGGGGGATGAAC	58
RPPG6-024	SSR	Peach	CTTGGAGATTGGGGCATA	CACAAGATGGACTAGGCAAA	49
RPPG6-025	SSR	Peach	GATAAAAGGGTAGGTAGGTCCA	AGTCCCATGTGCTTGTTTCT	53
RPPG6-030	SSR	Peach	GATGACACCGAGTTTCGATT	CAGATCGGGTTTACGCTACT	58
RPPG6-033	SSR	Peach	CATTATCAAACCACGACCAA	AAAGCTCAACAGCGACTTCT	56

Table 3. Simple sequence repeat (SSR) markers from (Dettori et al. 2015) and (Zhang et al. 2014) that were successfully anchored to the novel chokecherry genetic map (continued)

Marker name	Marker type	Source species	Forward sequence	Reverse sequence	Ta[a]
RPPG6-036	SSR	Peach	GCTATTTTCTCCACCAGCTC	GCCATAGTTGACTGCATTGAT	55
RPPG6-038	SSR	Peach	GCATAGGGTGTGTTCCTCA	CCAGTGACATCTAGCCCATT	53
RPPG7-018	SSR	Peach	TTGTCATCAGGTCGTTCATC	TCCTCCCACTCTGTATTTGG	58
RPPG7-020	SSR	Peach	GATCCAACTTCCACCACACT	CAGGGCACCATCTCTTAAAC	55
RPPG7-023	SSR	Peach	TTTAGCCATTTACCCATTTTG	CATTCCTGTTCCCTTTTTGT	56
RPPG7-026	SSR	Peach	TTTGGTGAGTGGGCTCTATT	CTATCGTTCGCTGGTCTTCT	53
RPPG7-029	SSR	Peach	CGAAGTGGAAACAGAAGATGA	GAGGTTGAAGACGGAAGATG	55
RPPG7-032	SSR	Peach	AAGGGAGGAGGATTGTGAA	TGGTAGACGGGTAGATGTTG	53
RPPG8-007	SSR	Peach	ACCACCACCTCTTCCAATC	ACCTCAAAGTGTCCCAGAAA	53
RPPG8-011	SSR	Peach	GCTTCTTCTTTGCTTGGAGT	CCGTTCATCATCTACCTTCC	53
RPPG8-014	SSR	Peach	ACTTGAATGGGCTAAAACGA	GAGAAGAAAAGAGCGTGGAG	56
RPPG8-017	SSR	Peach	AAACTATGCCTTGCTTGAGAAC	GCGGCGTTTCTTTCCTTT	53
RPPG8-020	SSR	Peach	CTGATCTGACAAAAGCACCA	TGAAGGCAACAAGAACGTAG	58
RPPG8-030	SSR	Peach	GCAAGTCAAACCACACAAGA	TGAAAGTGAAACCAACGAGA	56
RPPG8-031	SSR	Peach	ATCATGTCCTTTGGGCTCT	GGGCAAATCGAAGTTGTG	54
NAUpy_E603	SSR	Pear	GAAAGTCCTTTTATCTAATTGGAATCCT	CAGGGCAAAGCTTTCTCTTATTTT	58

[a] Annealing temperature (°C) for polymerase chain reaction (PCR)

3. RESULTS

3.1. Molecular markers successfully mapped to the new chokecherry linkage groups

3.1.1. LTR molecular markers

A previous study tested a total of 336 chokecherry long terminal repeat based (LTR) primer pairs for amplification and polymorphisms in eight chokecherry lines and ten representative Rosaceous species (Liang et al. unpublished). The results revealed 78 primer pairs that produced polymorphic alleles in chokecherry. All 78 polymorphic primers were evaluated in the full mapping population of chokecherry, resulting in the identification of 59 qualified single dose alleles based on expected segregation ratios of 1:1 or 3:1 ($\chi^2 = 0.001$). Of the qualified LTR markers, 20 were successfully anchored to the chokecherry genetic map (Table 4). Additionally, eight pear LTR markers retrieved from Sun et al. (2015) were tested in the subpopulation (n=8) of chokecherry. No pear LTR markers produced amplicons in chokecherry and were not further examined in the full mapping population. Figure 2 provides an example of successful amplification of polymorphic alleles in LTR marker analysis.

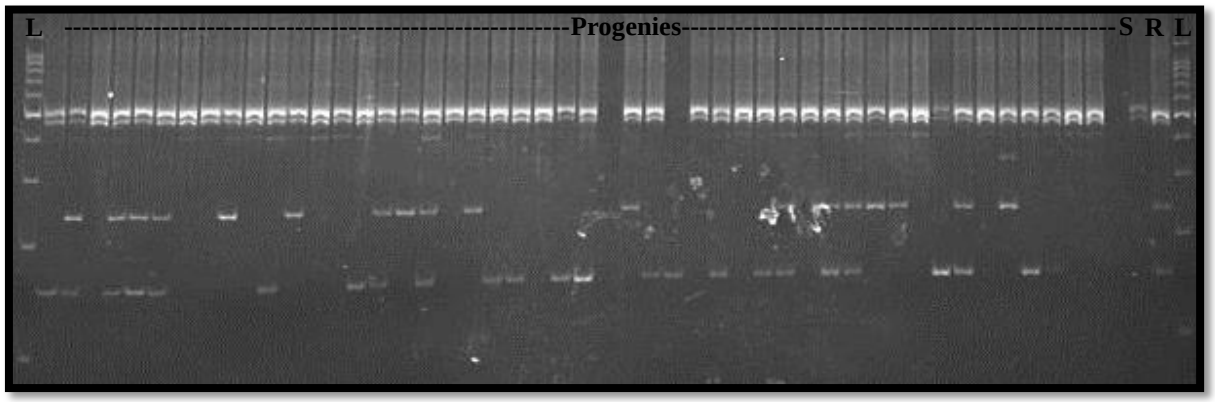

Fig. 2. Polyacrylamide gel electrophoresis showing amplification patterns of chokecherry LTR marker (LTR-D1112) in a chokecherry subsample population (n=49) and the two parental types. From left to right: 100bp DNA ladder (L), 1-49 chokecherry progenies, susceptible chokecherry parent (S), resistant chokecherry parent (R), 100bp DNA ladder (L).

3.1.2. SSR markers

A total of 257 simple sequence repeat (SSR) markers used previously for construction of the RC x SC maps (Wang et al. 2014) were re-analyzed in this study. The scoring data by Wang et al. (2014) was converted to the JoinMap format and subjected to the same statistical tests of the other molecular markers utilized. The results showed that the majority (97.7%) of the SSR markers were successfully anchored to the new chokecherry linkage groups (Table 4).

A total of 48 newly developed SSR primer pairs were tested in a subpopulation (n=8) of chokecherry, one peach genotype, and one sour cherry genotype. The results showed 100% amplification of the primer pairs in peach, while only 27 and 11 pairs produced amplicons in chokecherry and sour cherry, respectively. The amplified bands in chokecherry produced a total of 19 polymorphic markers (Table 5), with 11 being successfully anchored to the chokecherry linkage map (Table 4).

Additional SSR markers were adopted from Dettori et al. (2015) and Zhang et al. (2014) with a total of 176 peach and 11 pear primer pairs being tested in a chokecherry subpopulation (n=8). Peach SSRs were more transferable with 130 primer pairs (73.9%) producing amplicons; only four pear SSRs were amplifiable (36.4%) in chokecherry. A total of 117 polymorphic SSR markers (116 peach, one pear) were examined in the full mapping population. A total of 55 SSR markers (54 peach, one pear) were successfully anchored to the chokecherry genetic map (Table 4). Details on the distribution of all anchored markers are summarized in Table 6 ‘Marker distribution and map statistics for the integrated chokecherry (Cho) map’. Figure 3 provides an example of successful amplification of polymorphic alleles in SSR marker analysis

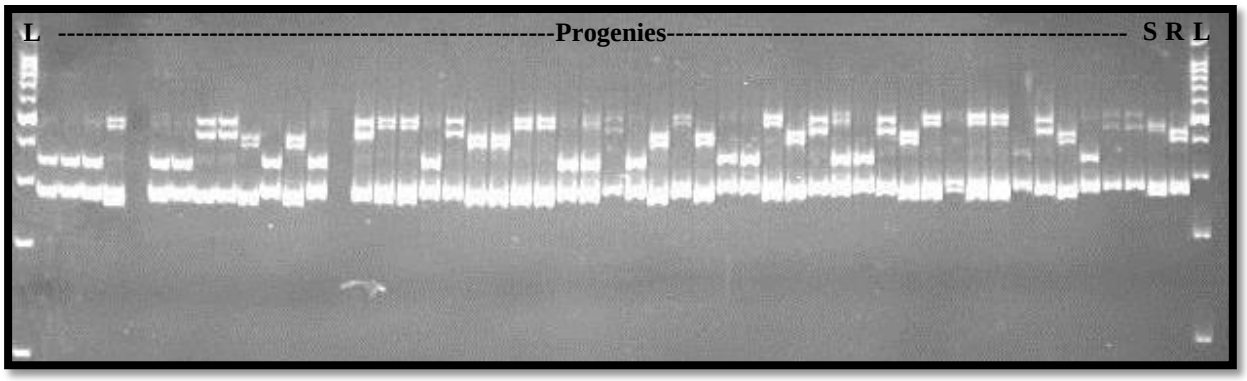

Fig. 3. Polyacrylamide gel electrophoresis showing amplification patterns of peach SSR marker (SSR8-D78) in a chokecherry subsample population (n=49) and the two parental types. From left to right: 100bp DNA ladder (L), 1-49 chokecherry progenies, susceptible chokecherry parent (S), resistant chokecherry parent (R), 100bp DNA ladder (L).

3.1.3. AFLP markers

There were no new amplified fragment length polymorphism (AFLP) markers tested in this study; however, a total of 241 AFLP markers that were previously proven qualified for genetic mapping in chokecherry (Wang et al. 2014) were analyzed for their suitability in constructing the new chokecherry map and QTL mapping in this study. A total of 228 of the 241 qualified markers (94.6%) were successfully mapped to the new chokecherry genetic map (Table 4).

Table 4. Origin and overall performance of molecular markers tested in genetic mapping of the novel chokecherry map

Marker type	Source species	Tested	Amplified	Polymorphic	Qualified[c]	Mapped	% Mapped[d]	Reference
SSR	Peach	176	130	116	85	54	30.7%	Dettori et al. 2015
	Peach	48	27	19	17	11	22.9%	Present study
	Pear	11	4	1	1	1	9.1%	Zhang et al. 2014
LTR	Chokecherry	336	283	78	59	20	6.0%	Liang, unpublished
	Pear	8	0	0	0	0	0.0%	Sun et al. 2015
SSR	*Prunus*[a]	257	257	257	257	251	97.7%	Wang et al. 2014
AFLP	Chokecherry[b]	241	241	241	241	228	94.6%	Wang et al. 2014
TOTAL		1077	942	712	660	565	52.5%	

[a] *Prunus* SSR markers originate from chokecherry, peach, sweet cherry, black cherry, Japanese plum, apricot, almond, and sour cherry
[b] Chokecherry AFLP markers are described in more detail in Wang et al. 2014
[c] Markers were considered qualified if segregation distortion ratios did not exceed Chi-Square's test at (p<0.001)
[d] Percentage of markers that were mapped of the total markers tested

Table 5. Newly designed short simple repeat (SSR) markers that produced polymorphisms in chokecherry

Marker name	Marker type	Source Species	Forward sequence	Reverse sequence	Ta[a]
SSR8-A12	SSR	Peach	TGGCGTCTGAACTCCGATTG	GAACAACTTCCTTCCTCCATCT	56
SSR8-A910	SSR	Peach	ACGGGCTGATTTTGCCTTGG	CAACTGCGAAGTGCTTGACG	56
SSR8-D1112	SSR	Peach	TGTTGTACATGACCGTTAGC	CATTTTCACCAAGGCTTTACC	56
SSR8-D12	SSR	Peach	GTGTGGGCAAGAACTTATCAT	CCATTCCCAAAGTCACCGAT	56
SSR8-D34	SSR	Peach	ACTGCCCGCATTCTTGGTT	CTCTTCAGGCTTGACATACT	56
SSR8-D56	SSR	Peach	AAGCAGCATCTTCCTCTTC	TCAATACCGTCAAGATAACC	56
SSR8-D78	SSR	Peach	GTGATGATTGAATCTGCCAT	GACAAAGCTAATTCCACATC	56
SSR8-E1112	SSR	Peach	ATGCCTGAATGAATGCTCT	AGTGGGAGATTGAGTTTGT	56
SSR8-E12	SSR	Peach	AGTCAGTGAGCCAGCCTATG	GGGAATGGGAACAGGAAACG	56
SSR8-E34	SSR	Peach	GAGATAAGGGCATTTTGGTCAT	CCCAACCATCAGCTAGAGC	56
SSR8-E78	SSR	Peach	GCTTTGCCTTGTAGACGCTT	GTGGAGCATAAGCATACCTC	56
SSR8-G34	SSR	Peach	CAGTGATACCTGCTACGAT	CTATCTGCTCCGATTCCTC	56
SSR8-F78	SSR	Peach	GTTGTTGTTCTCATCCGAAAC	CTACCTTTGTTGTCCTCTGC	56
SSR8-F910	SSR	Peach	GTGTAGGAGTAGCCCTGGTT	AGAGGCAAAAGAGGAACAGTG	56
SSR8-G12	SSR	Peach	GTGACCGTTGGAGGCTGTAT	TACACCCATTTTGCTCACCC	56
SSR8-G56	SSR	Peach	CTTTTGAAAGCAGACAGATC	GCCAGTCTAACCTTCTCAG	56
SSR8-G78	SSR	Peach	TTGCTAACAGTGGCACGCTC	AGAAAGAAGCAGTCGTCAAGC	56
SSR8-H1112	SSR	Peach	GGTGTACGCTTCCACTATG	CGGATTATCGGACAAAGTG	56
SSR8-H910	SSR	Peach	TGGAAAGGGTATAGCTGTG	CAAGATGACAACTGAGGCT	56

[a] Annealing temperature (°C) for polymerase chain reaction (PCR)

Table 6. Marker distribution and map statistics for the integrated chokecherry (Cho) genetic map

Linkage group	SSR	AFLP	LTR	TOTAL	Length	Average distance[a]	Longest gap[b]
					(cM)	(cM)	(cM)
Cho-1	12	7	1	20	129.7	6.5	19.4
Cho-2	20	12	0	32	127.8	4.0	18.3
Cho-3	17	13	1	31	149.2	4.8	16.6
Cho-4	38	9	0	47	142.6	3.0	9.9
Cho-5	32	3	0	35	142.5	4.1	11.7
Cho-6	20	9	2	31	98.3	3.2	8.7
Cho-7	18	21	0	39	126.9	3.3	13.1
Cho-8	10	25	0	35	145.6	4.2	21.2
Cho-9	7	23	0	30	156.8	5.2	13.2
Cho-10	20	23	0	43	161.6	3.8	11.3
Cho-11	17	0	15	32	98.2	3.1	10.7
Cho-12	18	15	0	33	168.6	5.1	15.2
Cho-13	25	14	0	39	171.4	4.4	10.1
Cho-14	15	21	1	37	135.0	3.6	16.5
Cho-15	25	20	0	45	124.4	2.8	9.4
Cho-16	23	13	0	36	93.6	2.6	9.1
TOTAL	317	228	20	565	2172.1	3.97	-

[a] Average distance in centi-Morgans (cM) between markers per linkage group
[b] Largest gap between markers per linkage group

3.2. Construction of a novel chokecherry genetic map

3.2.1. Segregation analysis

The hypothesis of preferential pairing of chromosomes in allotetraploids (disomic inheritance) was investigated using markers for which one single parent has a single dose for two or three marker alleles. Six phenotypic classes are expected for markers with three single-dose alleles in the case of autotetraploids (tetrasomic inheritance) if they belong to the same locus. Eight are possible in the case of disomic inheritance. The hypothesis of equal 1/8 or 1/6 segregation ratios for disomic or tetrasomic inheritance, respectively, was tested with a Chi-square goodness-of-fit test ($\alpha = 0.01$) for the three-allelic or triplex by nulliplex marker, C6363. Progeny for this marker exhibited segregation patterns in a ratio not significantly different from

the expectations for tetrasomic or disomic inheritance (Table 7). More molecular markers showing segregation of two single dose alleles derived from one parent were also tested to provide further insight into the mode of inheritance. In this situation, only four marker classes are possible, but it is possible to test the inheritance pattern. The corresponding phenotypic classes have expected frequencies of 1/6, 2/6, 2/6, and 1/6 for tetrasomic inheritance due to the random pairing of bivalents or quadrivalent formation. Disomic inheritance phenotypic classes are expected to segregate at 1/4 each. Frequencies of the phenotypic classes of the progeny varied between the hypothesis of disomic and tetrasomic inheritance (Table 8). For example, markers C11508, LTR5-F78, PS7a2, UCD-CH17, and UDP409 fit the ratios of both disomic and tetrasomic inheritance. The majority of the duplex markers (9 of 18) rejected the goodness-of-fit of tetrasomic inheritance while not rejecting disomic inheritance. The remaining 4 markers rejected the disomic inheritance pattern but were statistically similar to tetrasomic segregation ratios. The 565 molecular markers mapped to chokecherry linkage groups were scored as single dose alleles to circumvent the mathematical differences between disomic and tetrasomic genetic linkage mapping.

3.2.2. Combining parental maps

Two parental genetic maps were created and analyzed against each other to find homologous loci shared by any linkage groups. All linkage groups between the parents had a homologous group based on having at least 2 shared loci (common markers). Before combining the groups in JoinMap, all markers were double checked to make sure none were shared in any other groups. A total of 12 new linkage groups were created by joining and re-ordering parental groups via the 'Combine Groups for Map Integration' function of JoinMap. The remaining four pairs of linkage groups had the required shared loci; however, there were some markers that did

not order properly once combined. To overcome this problem, MergeMap Online (Wu et al. 2008) was utilized to create the newly formed chokecherry linkage maps. All four pairs were successful joined and ordered to finish the development of sixteen newly formed linkage groups, representative of the haploid chromosome number (n = 2x = 16) of chokecherry (2n = 4x = 32) (Wang et al. 2014; Liang et al. unpublished).

3.2.3. Marker distribution and map statistics

A total of 16 novel chokecherry linkage groups have been created to represent the 16 chromosome pairs found in this tetraploid *Prunus* species (Fig. 4). Chokecherry linkage group 1 (Cho-1) had the fewest markers anchored to it; twelve SSR, seven AFLP, and one LTR were distributed across a map distance of 129.7 centi-Morgans (cM). The chokecherry linkage group with the most markers was Cho-4, having a total of 47 anchored markers in which 38 markers are SSRs and nine AFLPs markers, spanning a total map length of 142.6 cM. The linkage group with the longest map distance was Cho-13, spanning 171.4 cM, whereas Cho-16 was the shortest at 93.6 cM. The longest gap between any two markers in all linkage groups was 21.2 cM near the end of Cho-8. Overall marker density for all linkage groups was 3.97 cM per molecular marker. The total genetic length of the collective chokecherry linkage groups is 2172.1 cM (Table 6).

Table 7. Distribution of gametic contribution to chokecherry progeny for the triplex SSR marker, C6363 (ABCO x OOOO)

Marker	Gametic contribution									χ^2 probability of inheritance[a]		Chokecherry
	AO	BO	CO	AB	AC	BC	OO[b]	ABC[b]	Unknown	Disomic	Tetrasomic	linkage groups
C6363	16	11	14	9	10	11	18	11	1	0.58715	0.06225	14, 9

[a]Chi-square goodness-of-fit test (p<0.01) to assess segregation different from disomic (1:1:1:1:1:1:1:1) or tetrasomic (1:1:1:1:1:1) inheritance for gametes AO, BO, CO, AB, AC, BC, OO, and ABC (O being a null allele)
[b]Gametes OO and ABC are expected and included only for disomic phenotypes

Table 8. Distribution of gametic contribution to chokecherry progeny for various duplex by nulliplex markers (ABOO x OOOO)

Marker	Gametic contribution					χ^2 probability of inheritance[a]		Chokecherry linkage groups
	AB	AO	BO	OO	Unknown	Disomic	Tetrasomic	
BPCT028	12	33	27	29	0	0.02	0.00	8
C11508	29	29	24	19	0	0.44	0.20	12, 14
C2109	21	22	27	26	5	0.78	0.00	2, 13
C6256b	26	13	30	29	3	0.06	0.00	15
CPPCT22	15	31	24	28	3	0.12	0.00	2, 16
CPPCT26	28	41	18	13	1	0.00	0.32	15
CPSCT006	32	15	23	28	3	0.09	0.00	1
EMPA02	38	28	24	11	0	0.00	0.09	3
EMpaS10	25	18	32	24	2	0.26	0.00	4, 12
LTR5-F78	22	16	22	13	28	0.34	0.01	6
LTR7-D1112	21	21	26	32	1	0.35	0.00	1, 16
PMS3	28	21	25	25	2	0.80	0.00	4
PS7a2	28	38	19	15	1	0.01	0.57	8
RPPG4-084	18	17	29	31	6	0.08	0.00	-
SSR8-E34	35	24	11	19	12	0.00	0.24	3, 9
UCD-CH17	28	28	22	20	3	0.56	0.25	12
UDP409	27	29	20	23	2	0.58	0.18	13
UDP98-406	46	34	7	13	1	0.00	0.01	14

[a]Chi-square goodness-of-fit test ($p<0.01$) to assess segregation different from disomic (1:1:1:1) or tetrasomic (1:2:2:1) inheritance for gametes AB, AO, BO, and CO (O being a null allele)

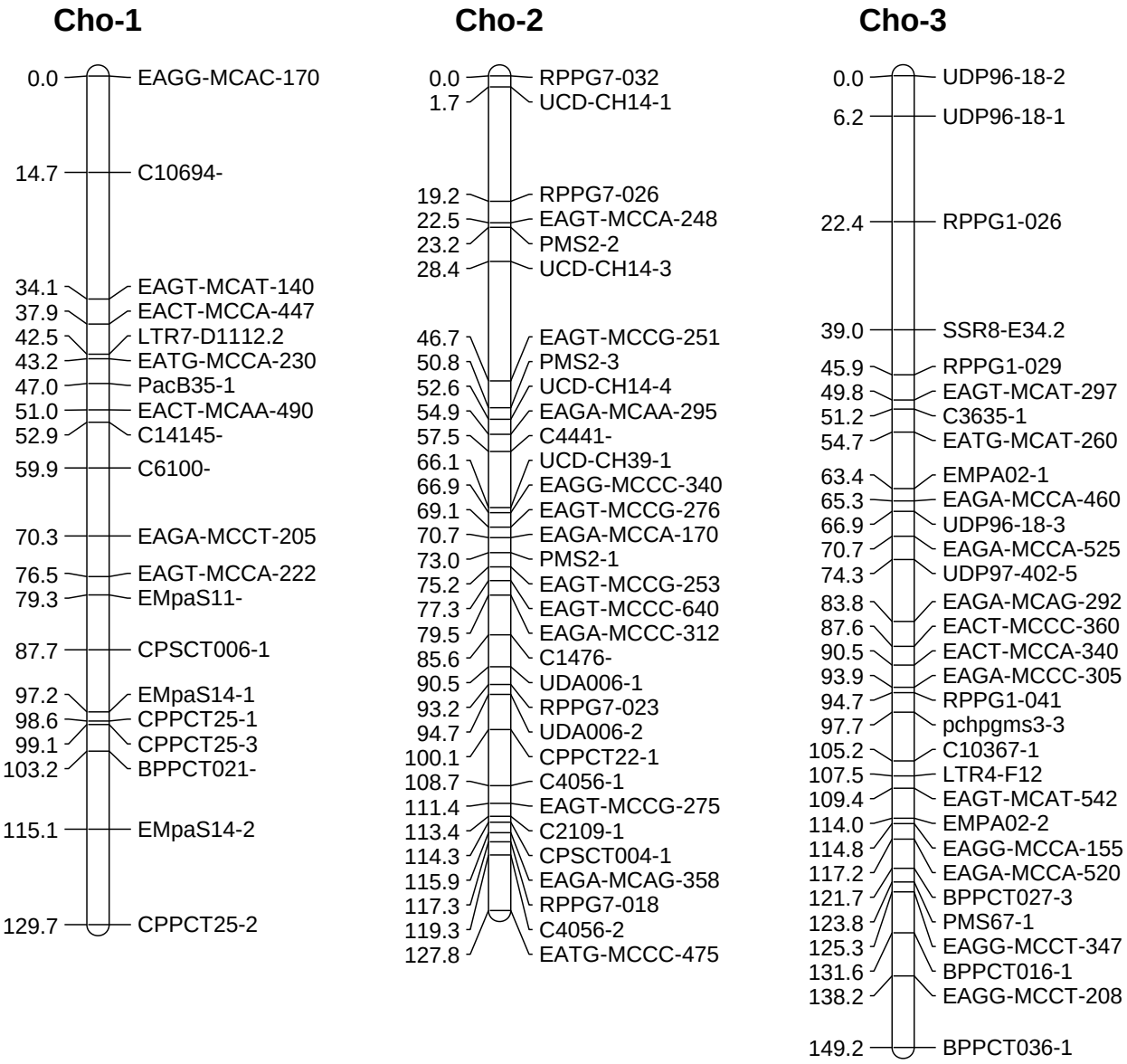

Fig. 4. Linkage groups 1-16 of the novel chokecherry (Cho) genetic map

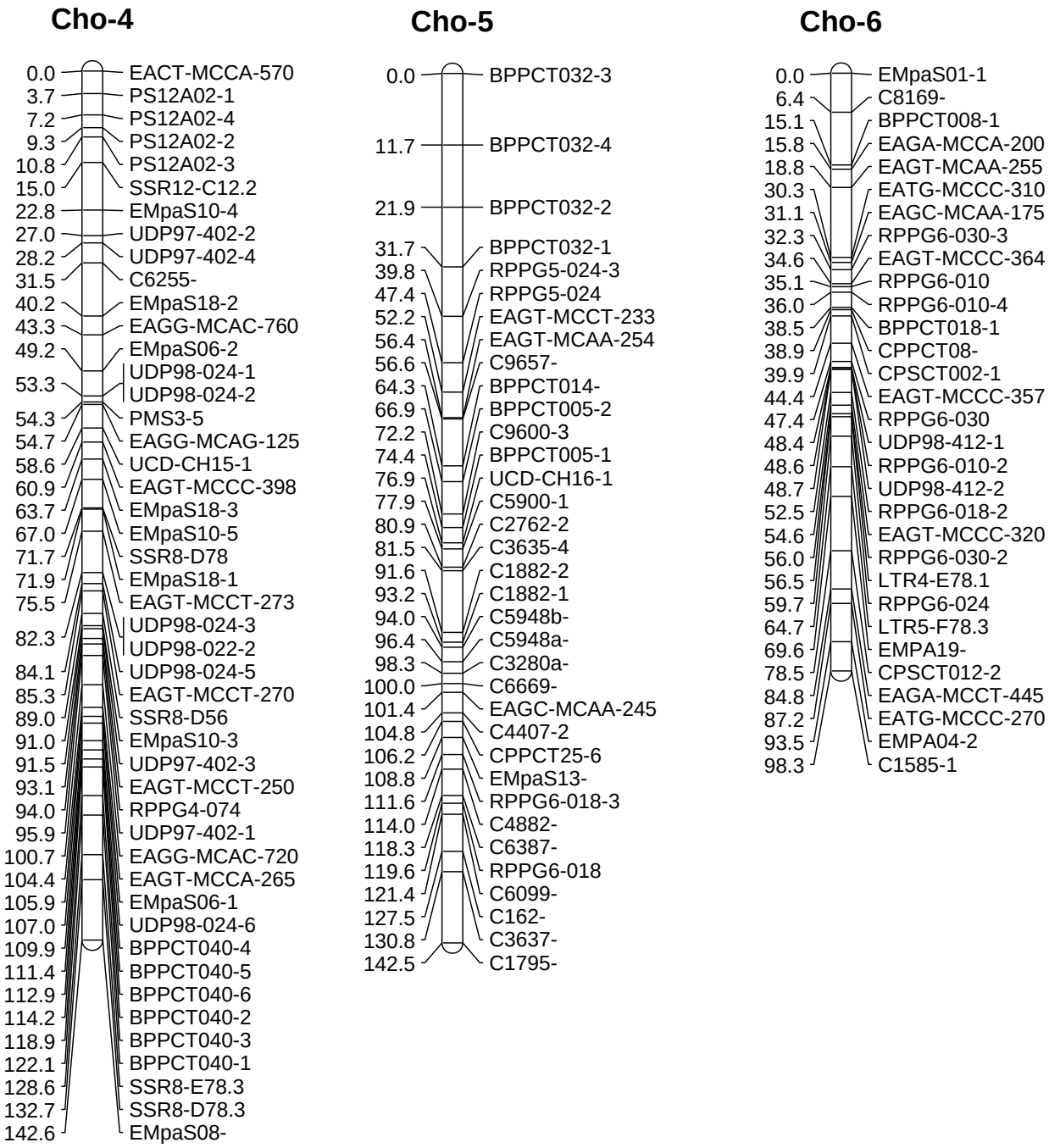

Fig. 4. Linkage groups 1-16 of the novel chokecherry (Cho) genetic map (continued)

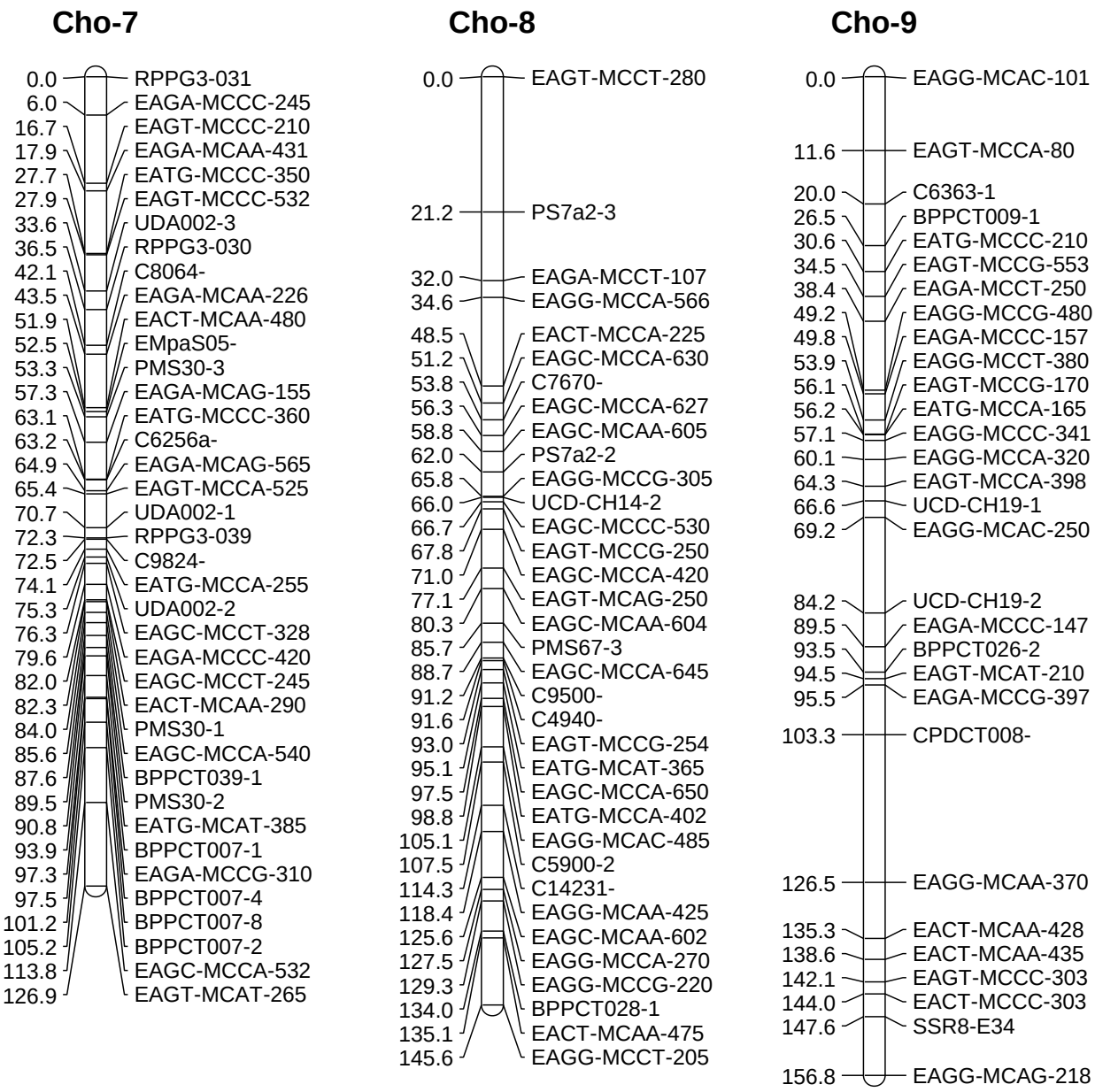

Fig. 4. Linkage groups 1-16 of the novel chokecherry (Cho) genetic map (continued)

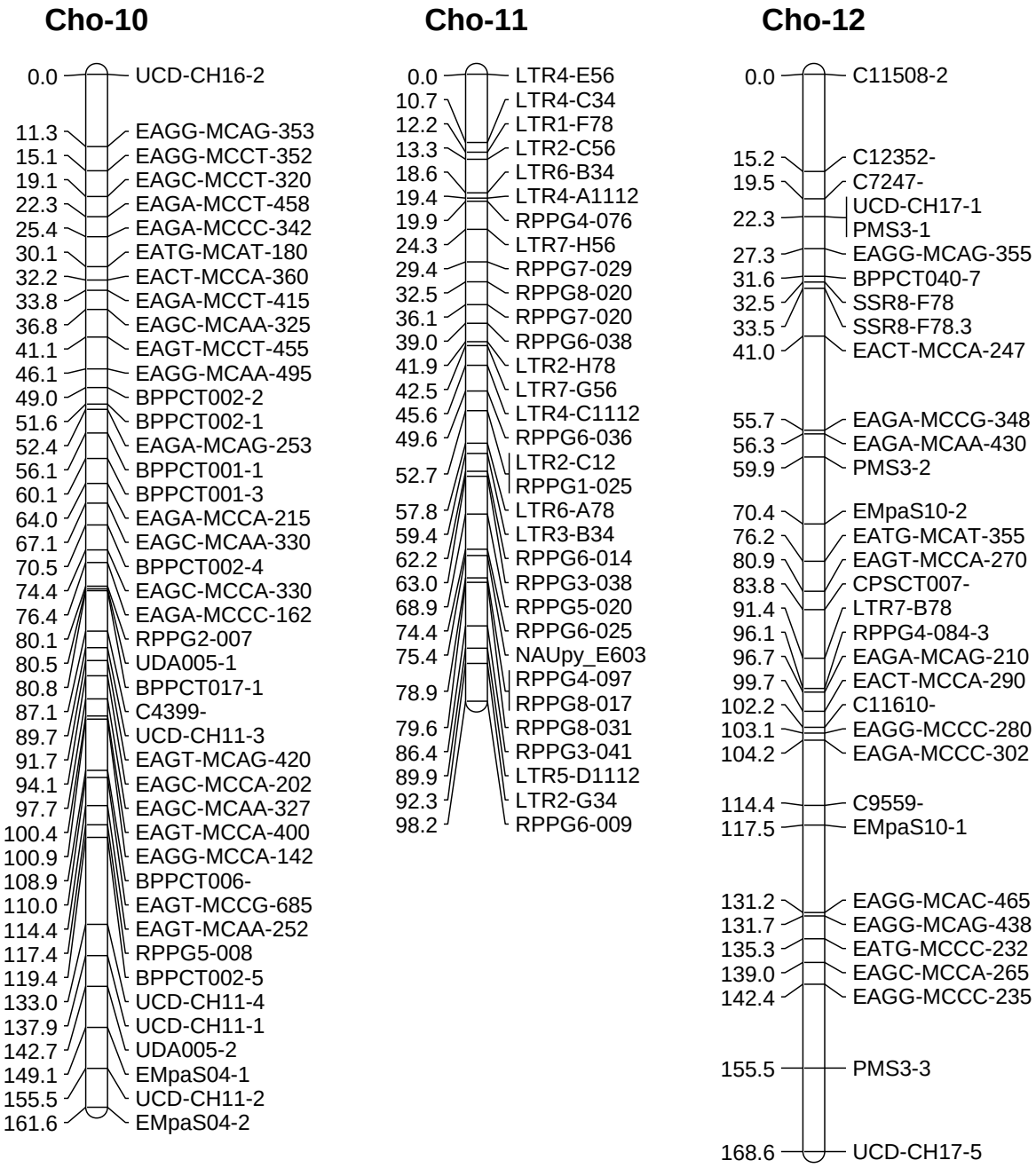

Fig. 4. Linkage groups 1-16 of the novel chokecherry (Cho) genetic map (continued)

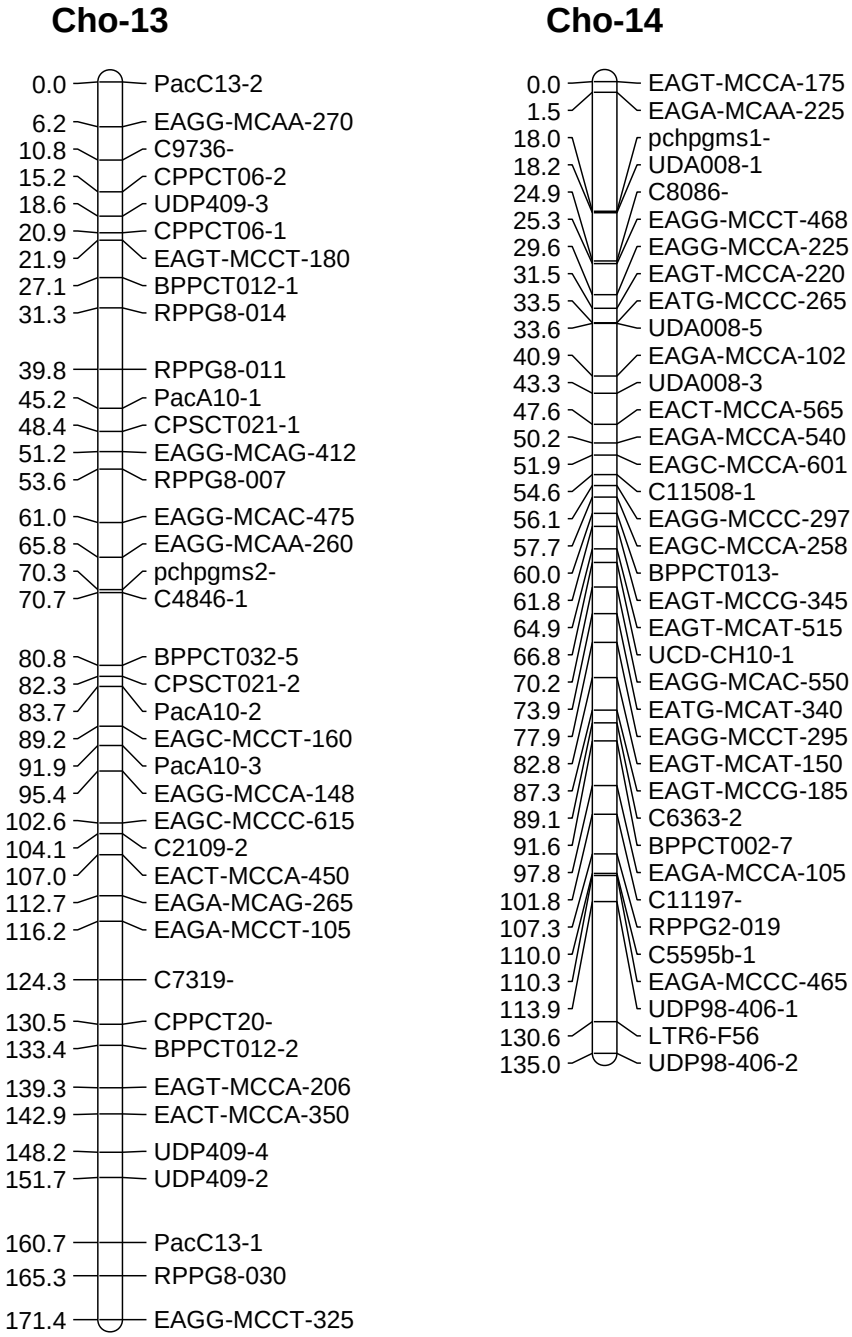

Fig. 4. Linkage groups 1-16 of the novel chokecherry (Cho) genetic map (continued)

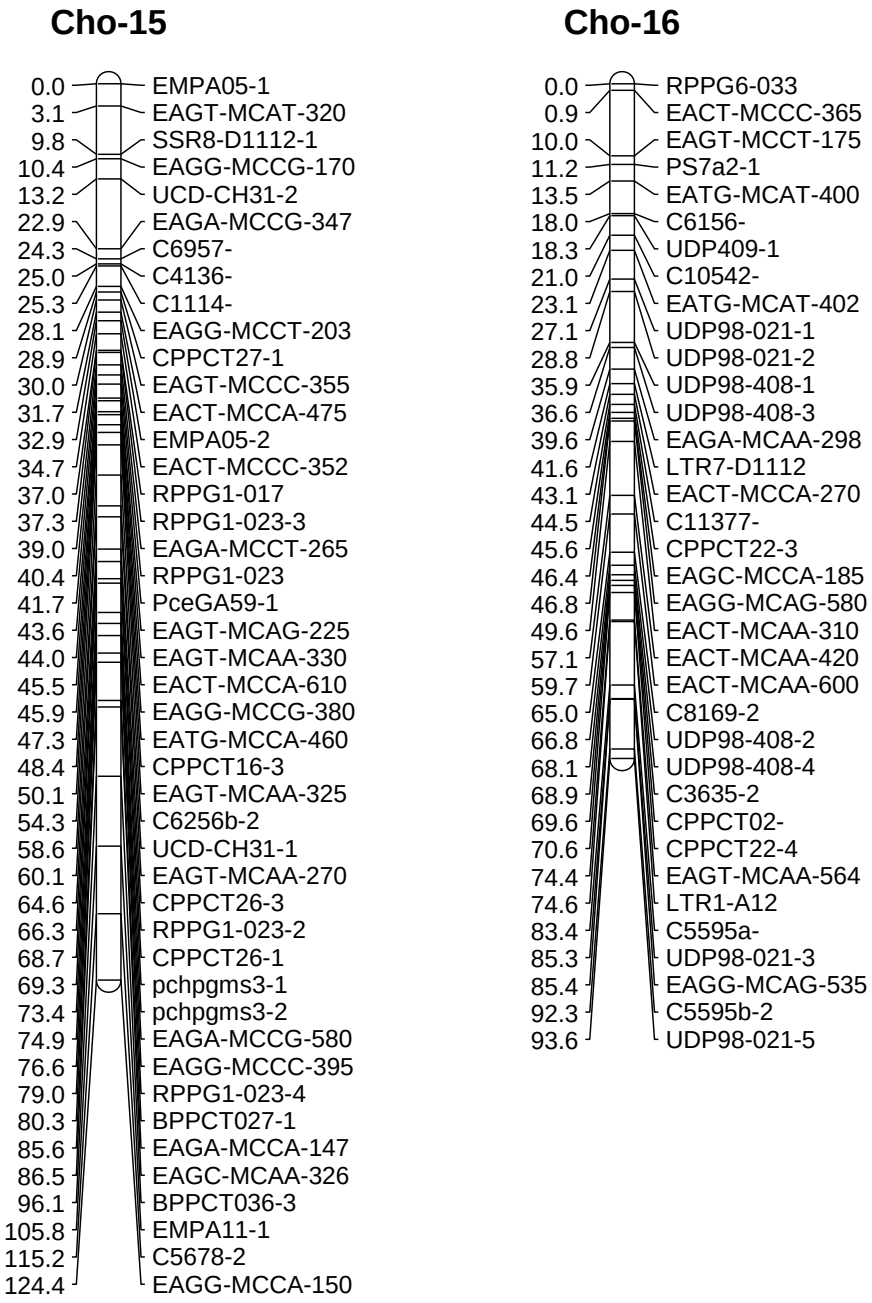

Fig. 4. Linkage groups 1-16 of the novel chokecherry (Cho) genetic map (continued)

3.2.3. Comparative analysis of chokecherry maps

The newly developed chokecherry linkage groups have been compared to the previously published RC x SC chokecherry map (Wang et al. 2014). The previous map had separate linkage groups for the two parents (resistant chokecherry, RC and susceptible chokecherry, SC) and a few segmented groups due to large gaps spanning 30 cM or more. The present chokecherry map produced herein was labeled according to the homologous relationship to previous chokecherry linkage group numbers. Most new linkage groups corresponded to unique previous linkage group pairs, but a few groups combined different combinations of previous linkage group segments or none at all. Linkage group Cho-1 corresponded best with SC-1 and RC-11 in the RC x SC map (Wang et al. 2014) (Fig. 5). Linkage group Cho-2 corresponded best with RC-2 and SC-10 (Fig. 6). Cho-10 was mostly a combination of RC-10 and the segmented group SC-2b (Fig. 7). Lastly, Cho-4 corresponded with the segmented groups RC-12a and SC-4e (Fig. 8). These new combinations of linkage groups eliminated previous linkage groups that were labeled 12 and 11. To compensate, the two new chokecherry groups with the least homology with previous linkage groups were assigned arbitrarily as Cho-11 and Cho-12. Group Cho-11 showed zero homology with previous maps, while Cho-12 showed homology to segment 14b of the RC linkage map and the top half of SC-5 (Fig. 9).

The overall genetic length of the chokecherry maps was changed as well. The revised chokecherry linkage groups spanned a total of 2172.1 cM compared to the 2089.0 cM and 1562.0 cM for the previous RC and SC maps, respectively (Table 9). This represents a 104% and 139% increase in genetic length for the new chokecherry map. Nevertheless, the marker density was increased substantially, going from 6.9 cM (RC) and 6.0 cM (SC) to 3.97 cM per marker.

Table 9. Size comparison in centi-Morgans (cM) of the revised chokecherry map to reference maps in *Prunus*

	RC[a]	SC[b]	TxE[c]	EFxNY[d]
	(cM)	(cM)	(cM)	(cM)
Chokecherry map	2172.1	2172.1	2172.1	2172.1
Reference map	2089.0	1562.0	621.2	638.5
Difference	83.05	610.1	1550.9	1533.6
Percent difference	104%	139%	350%	340%

[a] Resistant chokecherry (RC) parent map (Wang et al. 2014)
[b] Susceptible chokecherry (SC) parent map (Wang et al. 2014)
[c] 'Texas' almond x ' Earlygold' peach (TxE) reference map for *Prunus* species (Dirlewanger et al. 2004; Aranzana et al. 2003; Joobeur et al. 1998; Horn et al. 2005; Lalli et al. 2005)
[d] 'Emperor Francis' x 'New York 54' (EFxNY) sweet cherry map (Olmstead et al. 2008)

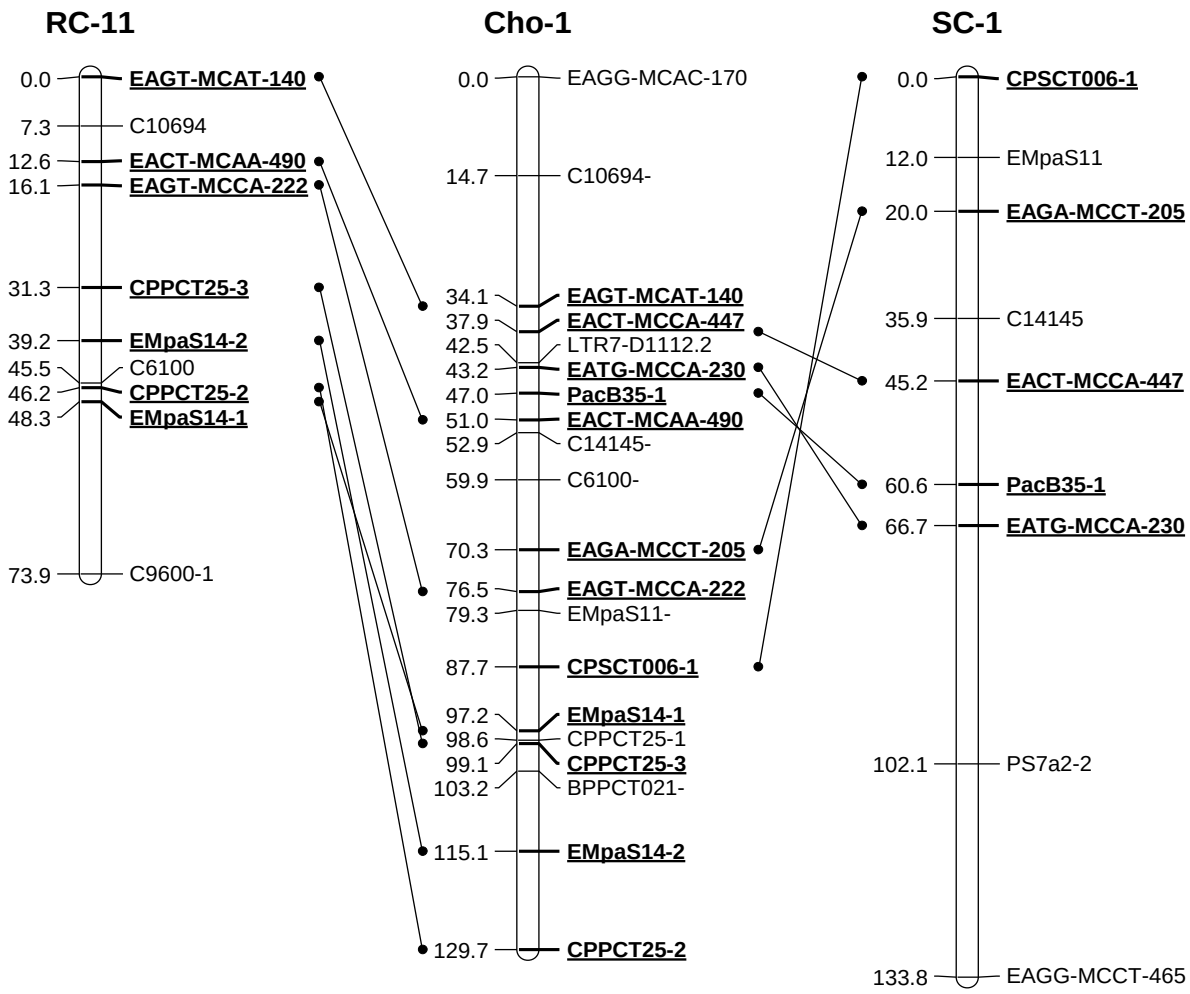

Fig. 5. Integrated chokecherry linkage group 1 (Cho-1) joins resistant chokecherry (RC-11) and susceptible chokecherry (SC-1) linkage groups from Wang et al. 2014

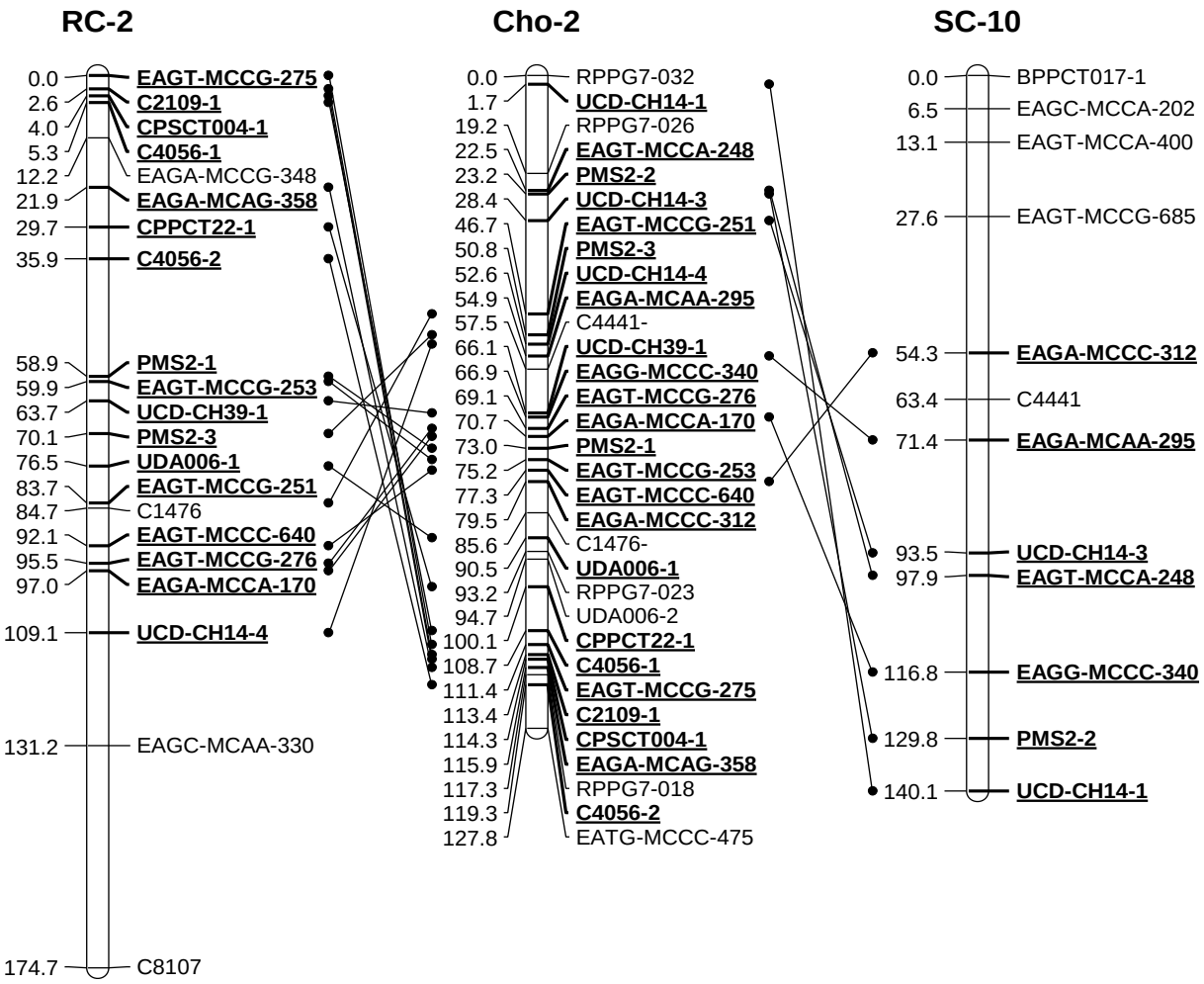

Fig. 6. Integrated chokecherry linkage group 2 (Cho-2) joins resistant chokecherry (RC-2) and susceptible chokecherry (SC-10) linkage groups from Wang et al. 2014

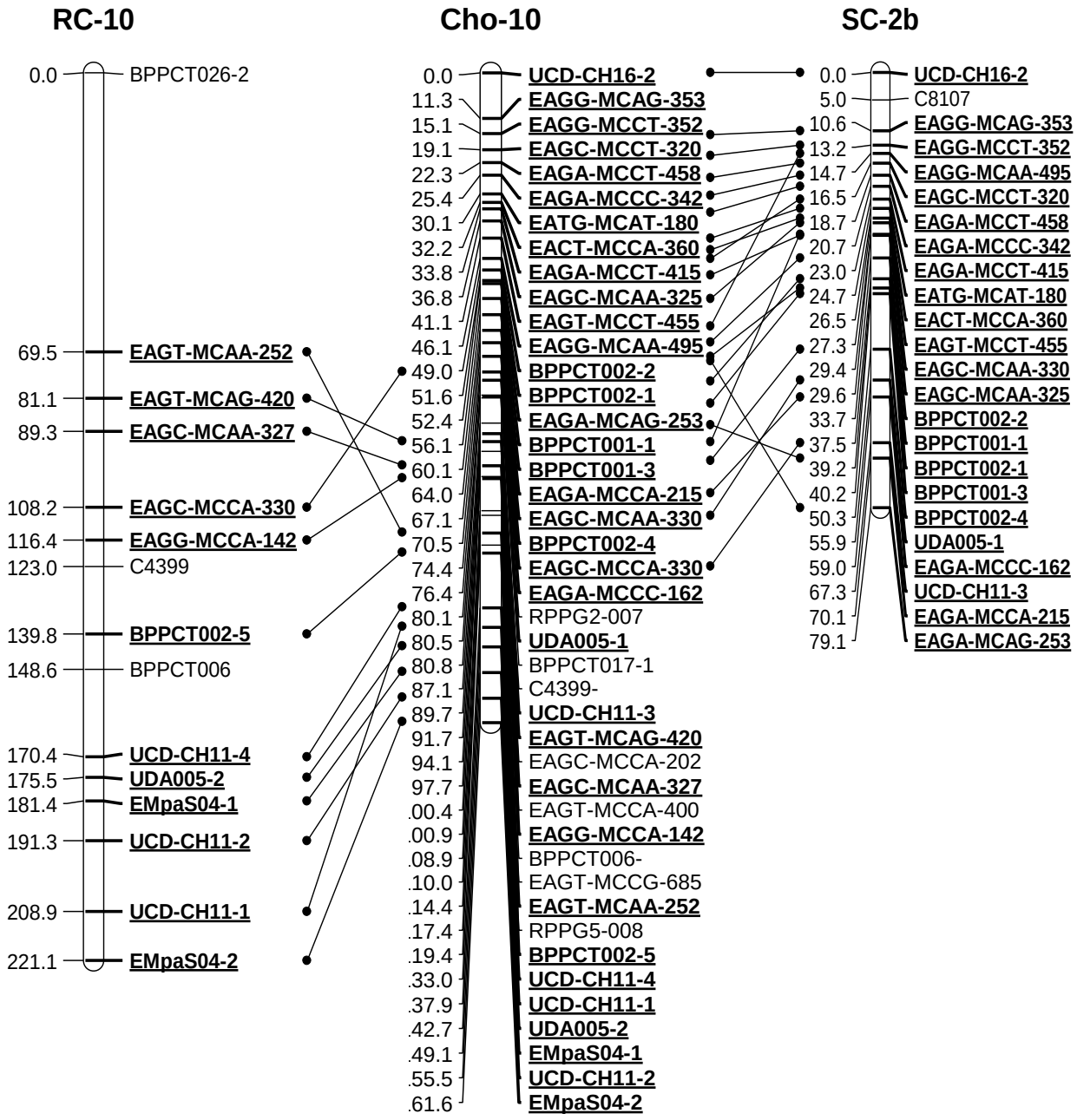

Fig. 7. Integrated chokecherry linkage group 10 (Cho-10) joins resistant chokecherry (RC) linkage group 10 and susceptible chokecherry (SC) segment 2b from Wang et al. 2014

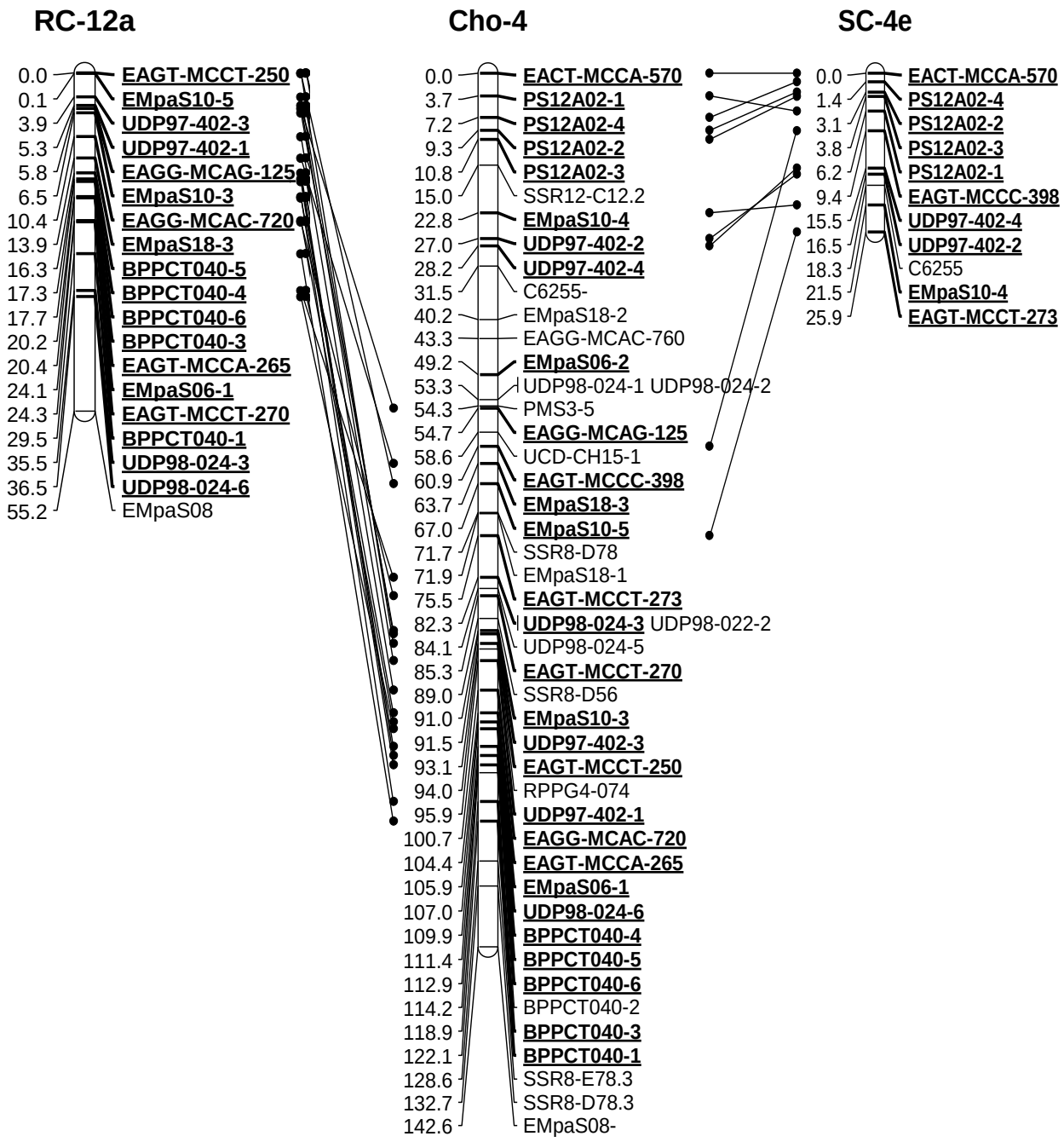

Fig. 8. Integrated chokecherry linkage group (Cho-4) joins resistant chokecherry (RC) segment 12a and susceptible chokecherry (SC) segment 4e from Wang et al. 2014

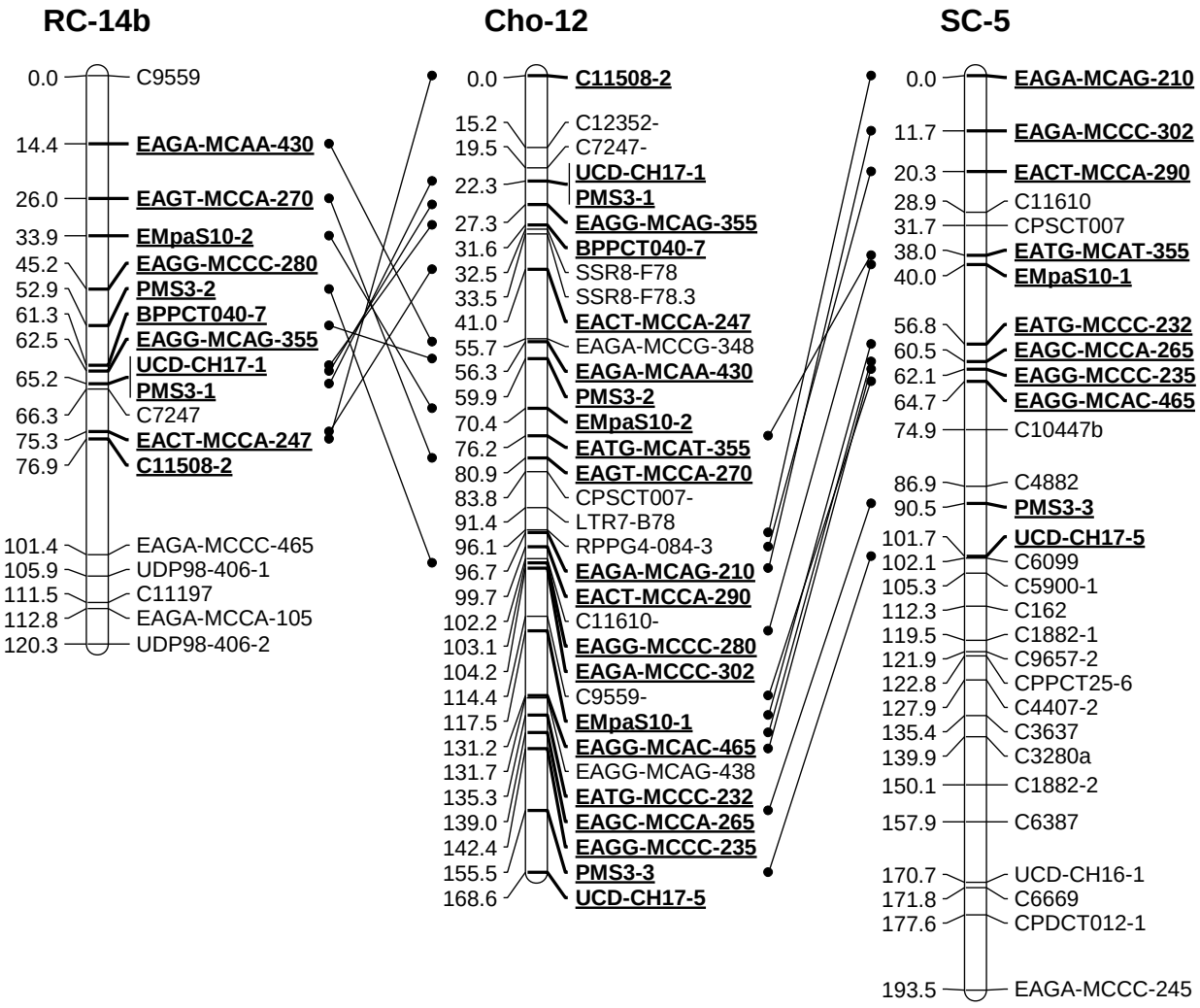

Fig. 9. Integrated chokecherry linkage group (Cho-12) joins the first half of resistant chokecherry (RC) linkage group 5 and susceptible chokecherry (SC) segment 14b from Wang et al. 2014

3.2.4. Comparative analysis of chokecherry and the *Prunus* reference maps

Chokecherry synteny with the *Prunus* reference map 'Texas' almond x 'Earlygold' peach (T x E) (Dirlewanger et al. 2004) was examined. It was discovered that 32 total loci are orthologous between chokecherry and the T x E map (Table 10). The strongest homology is seen between chokecherry group 4 (Cho-4) and T x E-4 with a total of seven orthologous loci. Other groups showing a strong relationship by having at least three orthologous loci include Cho-2 and T x E-7, Cho-3 and T x E-1, Cho-5 and T x E-5, Cho-7 and T x E-3, Cho-10 and T x E-2, and Cho-12 and T x E-4. All other chokecherry groups have only one or two orthologous loci except for Cho-16 and Cho-11 which have no corresponding loci found in the T x E map.

3.2.5. Comparative analysis of chokecherry and sweet cherry maps

Chokecherry linkage maps were compared with the 'Emperor Francis' x 'New York 54' (EF x NY) sweet cherry maps for syntenic relationships. A total of 23 marker loci were shared between chokecherry and sweet cherry. The largest number of orthologous loci is seen between Cho-4 and EF x NY-4 (Table 11) in which six loci were shared. Linkage groups Cho-2, Cho-7, Cho-10, and Cho-12 have at least three loci shared with corresponding sweet cherry linkage groups. Most of the shared loci between chokecherry and sweet cherry are seen in the same linkage groups as peach (Table 12).

Table 10. Orthologous loci between chokecherry and *Prunus* reference T x E linkage groups

	TxE-1[b]	TxE-2	TxE-3	TxE-4	TxE-5	TxE-6	TxE-7	TxE-8
Cho-1[c]					1			
Cho-2							4	
Cho-3	3[a]			1				
Cho-4				7				
Cho-5					5			
Cho-6						2		
Cho-7			5					
Cho-8	2							
Cho-9			1		1	1		
Cho-10		6			1			1
Cho-11								
Cho-12				4				
Cho-13		2			1			2
Cho-14		2						
Cho-15	2			1				
Cho-16								

[a] Number of orthologous loci shared between linkage groups
[b] 'Texas' almond x 'Earlygold' peach (TxE) linkage groups
[c] Chokecherry (Cho) linkage groups

Table 11. Orthologous loci between chokecherry and sweet cherry linkage groups

	EFxNY-1[b]	EFxNY-2	EFxNY-3	EFxNY-4	EFxNY-5	EFxNY-6	EFxNY-7	EFxNY-8
Cho-1[c]					1			
Cho-2							3	
Cho-3	1[a]							
Cho-4				6				
Cho-5				2	1			
Cho-6						2		
Cho-7			4					
Cho-8	1							
Cho-9					1	1		
Cho-10		5				1		
Cho-11								
Cho-12				4				
Cho-13		2						
Cho-14		1		1				
Cho-15	1							
Cho-16								

[a] Number of orthologous loci shared between linkage groups
[b] 'Emperor Francis' x 'New York 54' (EFxNY) sweet cherry linkage groups
[c] Chokecherry (Cho) linkage groups

Table 12. List of homologous loci and their corresponding linkage groups

Marker		Linkage group		
Name	Type	Cho[a]	TxE[b]	EFxNY[c]
EMPaS11	SSR	1	5	-
CPSCT006	SSR	1	-	5
PMS2	SSR	2	7	7
CPSCT004	SSR	2	7	-
PMS67	SSR	3	-	1
BPPCT027	SSR	3	1	-
BPPCT016	SSR	3	1	-
BPPCT036	SSR	3	4	-
PMS3	SSR	4	-	4
BPPCT040	SSR	4	-	4
BPPCT014	SSR	5	5	5
BPPCT005	SSR	5	-	4
BPPCT032	SSR	5	5	-
EMPaS01	SSR	6	6	6
BPPCT008	SSR	6	6	6
CPSCT012	SSR	6	6	-
EAT-MCCC-350	AFLP	7	-	3
PMS30	SSR	7	-	3
BPPCT039	SSR	7	3	-
BPPCT007	SSR	7	3	-
PMS67	SSR	8	1	1
BPPCT028	SSR	8	1	-
BPPCT026	SSR	9	5	5
BPPCT009	SSR	9	-	6
CPDCT008	SSR	9	3	-
BPPCT002	SSR	10	2	2
BPPCT006	SSR	10	8	2,6
BPPCT002	SSR	10	2	2
BPPCT001	SSR	10	1	-
BPPCT017	SSR	10	5	-
PMS3	SSR	12	4	4
BPPCT040	SSR	12	4	4
CPSCT021	SSR	13	2	2
BPPCT012	SSR	13	8	-
BPPCT032	SSR	13	5	-
BPPCT013-	SSR	14	2	4

a Chokecherry (Cho) linkage groups
b ‘Texas’ almond x ‘Earlygold’ peach (TxE) linkage groups
c ‘Emperor Francis’ x ‘New York 54’ (EFxNY) sweet cherry linkage groups

Table 12. List of homologous loci and their corresponding linkage groups (continued)

Marker		Linkage groups		
Name	Type	Cho[a]	TxE[b]	EFxNY[c]
BPPCT002	SSR	14	2	2
PceGA59	SSR	15	1	1
BPPCT027	SSR	15	1	-
BPPCT036	SSR	15	4	-

[a] Chokecherry (Cho) linkage groups
[b] 'Texas' almond x 'Earlygold' peach (TxE) linkage groups
[c] 'Emperor Francis' x 'New York 54' (EFxNY) sweet cherry linkage groups

3.3. QTL related to X-disease resistance

3.3.1. Data distribution

The distribution of phenotypic data for X-disease resistance in the chokecherry mapping population (101 progenies) was tested for normality using the Kolmogorov-Smimov normality test. This test is conducted using the 'Trait analysis' function in QGene 4.0 (Joehanes and Nelson 2008). Analysis produced a normality coefficient at 0.171, showing that the data are distributed normally and are suitable for quantitative trait loci (QTL) analysis.

3.3.2. Identified QTL and their significance

Three significant quantitative trait loci (QTL) were identified from the novel chokecherry linkage groups (Table 13). The composite interval mapping (CIM) function of QGene was used to detect the likelihood value of a QTL, expressed as an experiment-wise logarithm of the odds (LOD) score. The most likely position of the QTL was estimated by peak LOD score. The percentage of phenotypic variation explained (R^2) was estimated for all three QTL. A QTL accounting for the greatest contribution of X-disease phenotypic variation was identified on linkage group Cho-15. This particular locus spanned a distance of 2.1 cM, accounted for 18.4% of the phenotypic variation, had an additive effect of 0.71, and was significant at both the 95% and 99% levels of confidence (Fig. 10). Chokecherry linkage group Cho-5 also had a significant

QTL. This particular locus explained 14.6% of the phenotypic variation, had an additive effect of 0.42, was significant at the 99% confidence level, and spanned a genetic distance of 11.5 cM (Fig. 11). Linkage group Cho-4 had the third detected QTL, accounting for 12.9% of the phenotypic variation and an additive effect of 0.66. This locus was only significant at the 95% confidence level and spanned a distance of 6.9 cM (Fig.12).

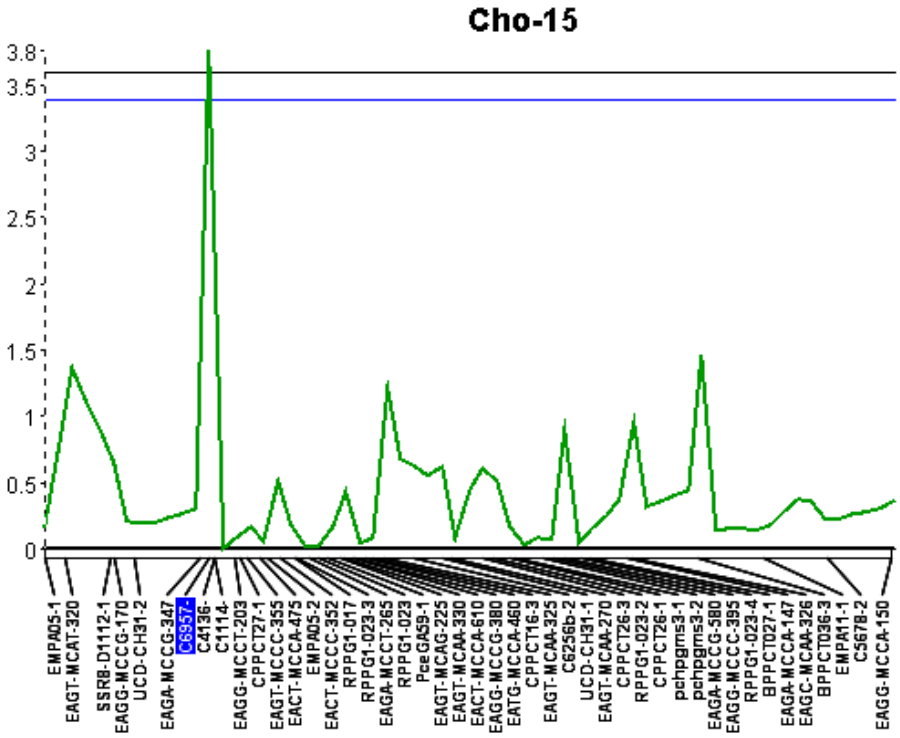

Fig. 10. Quantitative trait locus identified on linkage group 15 (Cho-15) with an LOD score of 3.8. Upper and lower line represent 1% and 5% significance thresholds respectively

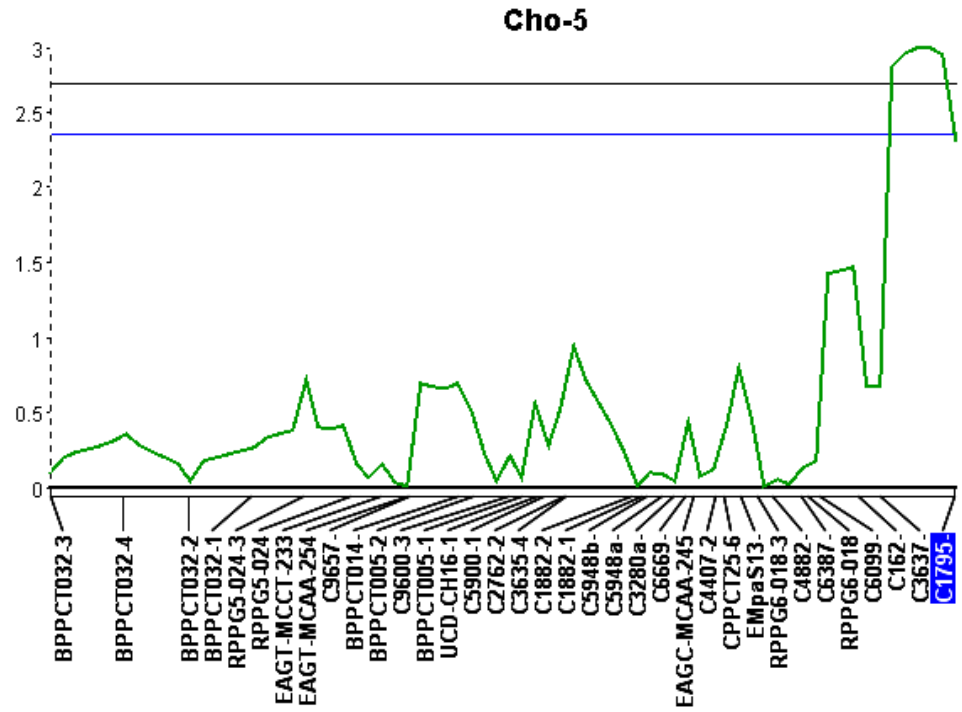

Fig. 11. Quantitative trait locus identified on linkage group 5 (Cho-5) with an LOD score of 3.0. Upper and lower line represent 1% and 5% significance thresholds respectively

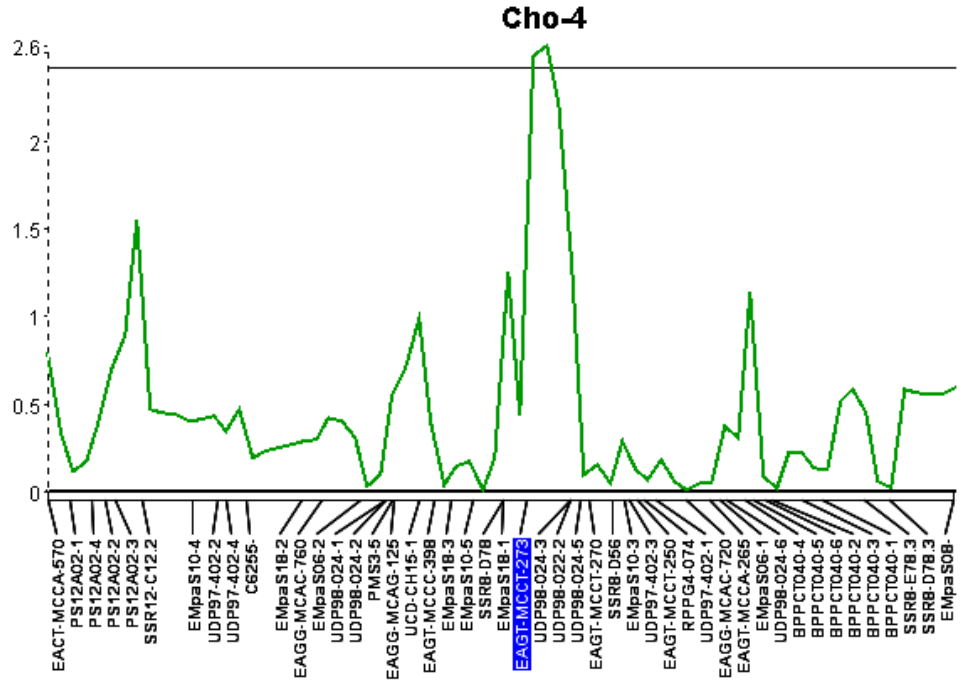

Fig. 12. Quantitative trait locus identified on linkage group 4 (Cho-4) with an LOD score of 2.6. Horizontal line represents a significance threshold of 5%

Table 13. Significant quantitative trait loci (QTL) statistics and associated markers distances in centi-Morgans (cM)

QTL	Linkage group	Position (cM)	Peak (LOD)	Permutation significance[a]	Additive effect[b]	Phenotypic Variance[c]	Flanking marker 1	Flanking marker 2	Interval (cM)[d]
1	Cho-15	24	3.8	**	0.71	18.4%	EAGA-MCCG-347	C4136	2.1
2	Cho-5	138	3.0	**	0.42	14.6%	C3637	C1795	11.5
3	Cho-4	78	2.6	*	0.66	12.9%	EAGT-MCCT-273	UPD98-024-3	6.9

[a] Significance thresholds were set after 1000 permutation iterations: * Significant at ($\alpha = 0.05$), ** Significant at ($\alpha = 0.01$)
[b] Additive effect represents the phenotypic score change due to QTL presence
[c] Phenotypic variance represents the R^2 value produced by the QTL
[d] Interval is the genetic distance between the flanking markers in which the QTL resides

4. DISCUSSION

4.1. Adoption and development of molecular markers for chokecherry genetic mapping

The main objective of the present work was to provide legitimate resources for X-disease research and management in *Prunus* species. X-disease is a major limiting factor of stone fruit production and current management strategies are inefficient (Douglas et al. 1999; Davis et al. 2013; Peterson 1984). Chokecherry has a documented ability to overcome X-disease phytoplasma and could provide valuable resources for biological, genetic, and molecular research on the host-pathogen interaction of X-disease. An inadequate repertoire of molecular genetics resources in chokecherry has prevented major advances in understanding of this potential germplasm source.

4.1.1. Transferability of SSR markers used in chokecherry mapping

Genetic studies and molecular breeding approaches require basic genomic resources, such as molecular markers and linkage maps. In chokecherry, previous studies have demonstrated the utility of simple sequence repeat (SSR), amplified fragment length polymorphism (AFLP), and long terminal repeat (LTR) markers (Wang et al. 2014; Liang unpublished). Many of these markers are derived from DNA sequences other than chokecherry, meaning they are transferable among species with closely related ancestry. Transferability of markers is best described as the percentage of markers that produce polymorphisms across genera/ species (Mnejja et al. 2010). Many studies showed the inter-specific and inter-generic transferability of rosaceous SSR markers (Dondini et al. 2007; Gasic et al. 2009; Mnejja et al. 2010; Wang et al. 2012; Zhang et al. 2014; Dettori et al. 2015). In this study, 60.3% of peach SSRs were polymorphic and transferable to chokecherry, which is slightly better than the 55% transferability observed by Wang et al. (2014). This may be explained by using a relatively high

number of long-core SSR markers that may be more unique and phylogenetically conserved; however, this idea would need to be confirmed with a larger number of markers tested in multiple chokecherry populations. The newly developed peach SSR markers (Table 3) were amplified 100% in peach, 29.2% in cherry, and 56.2% in chokecherry. It is interesting that 51.9% of the amplified SSR markers were multi-allelic (amplifying 2-6 bands) in tetraploid chokecherry and only 4.2% and 0.0% were multi-allelic in peach and cherry (diploid), respectively. It could be explored whether or not the polyploid nature of chokecherry plays any role. Additionally, this data supports the seemingly higher percentage of genomic synteny between chokecherry and peach versus chokecherry and other *Prunus* species (Wang et al. 2012; 2014).

4.1.2. Transferability of LTR markers used in chokecherry mapping

It is known that long terminal repeats (LTRs), produced by 'copy and paste' transposition of retrotransposons are well dispersed in plant genomes. The abundance, uniqueness, and non-sequence-biased transposition of LTR-retrotransposons (Bennetzen 2000) make them valuable resources for marker identification and gene mapping. Also, sequences of retrotransposons are similar within plant species and related genera (Lou and Chen 2007; Kalendar et al. 2011); however, Liang et al. (unpublished) found that the transferability rate of chokecherry LTR primers was lower than that of chokecherry SSR primers. For example, an average of 32.0% of chokecherry LTR primer pairs amplified bands in other *Prunus* species (cherry, peach, plum, and apricot), whereas 61.0% of chokecherry SSR primer pairs were amplified in *Prunus* (Liang et al. (unpublished; Wang et al. 2012). Furthermore, only 10% of chokecherry LTR primer pairs versus 47.2% SSR primer pairs amplified bands from non-*Prunus* rosaceous species (Liang et al. unpublished; Wang et al. 2012). This indicates that the transferability of LTR markers to more

distantly related species is sharply reduced. Nevertheless, LTR markers have proven utility in genetic mapping (You et al. 2010; Mazaheri et al. 2014; Monden et al. 2014; Sun et al. 2015) and have successfully been anchored to the novel chokecherry linkage map (Liang et al. unpublished).

4.2. Development of a novel chokecherry linkage map

4.2.1. Segregation analysis and genetic mapping

Molecular markers were scored for the absence or presence of individual marker amplicons (alleles), and those segregating at a ratio of 1:1 in the progeny were used for mapping. Scoring markers as single dose alleles overcomes the discrepancy of the mathematical process of calculating the recombination frequencies for autotetraploid (tetrasomic) inheritance or allotetraploid (disomic) inheritance, because this type of mapping strategy accounts for syntenic rearrangements and homeologous chromosome loci (Wu et al. 1992; Canli 2006; Koning-Boucoiran et al. 2012). It is also important to understand that many polyploids have intermediate inheritance patterns (Hickok 1978a,b; Stift et al. 2008; Koning-Boucoiran et al. 2012). This means that some chromosomes have diverged enough to preferentially pair to produce diploid inheritance (disomic), while others are similar enough to have levels of randomly pairing or produce quadrivalents during meiosis, leading to tetrasomic inheritance. Segregation analysis is the recommended way to determine allelic inheritance and the corresponding polyploid type (Krebs and Hancock 1989; Soltis and Rieseberg 1986).

Segregation analysis of markers with two or three unique single dose alleles in one of the parents was analyzed in detail using the distribution of the number of individuals over the different phenotypic classes encountered in the progeny. The results showed that about half of the markers (9/19) were strictly compatible with disomic inheritance, yet 4 were strictly

compatible with tetrasomic inheritance and 6 did not reject either inheritance model (Table 7 and 8). This suggests an intermediate mode of inheritance of chokecherry. A total of 11 chokecherry linkage groups are described by the markers used for inheritance determination, but only 4 are strictly disomic and 1 is strictly tetrasomic. The remaining linkage groups have supporting evidence of both tetrasomic and disomic inheritance patterns. More molecular markers need to be developed and utilized to study the tetraploid inheritance in more detail; however, this determination does not affect the genetic mapping process other than providing a means to explore individual chromosomes more unambiguously.

Tart cherry was the first polyploid *Prunus* species mapped and has provided an outline for genetic mapping in other tetraploid *Prunus* species like chokecherry. Genetic mapping of tart cherry utilized the single dose restriction fragments SDRF or the 'single dose allele' strategy of molecular marker genotyping even after previous allele segregation analysis determined it was an allotetraploid with disomic inheritance (Beaver et al. 1993, Canli 2004). This is because homologous loci may be present among non-pairing homeologous chromosomes. Overall, the tart cherry example of allotetraploid mapping and many autotetraploid examples mentioned in the literature review support the rationale that single dose alleles of molecular markers can be used in the intermediate tetraploid chokecherry population for genetic mapping and QTL identification. Also, Figures 13 and 14 provide examples of the collinearity of genetic maps produced from diploid mapping software (JoinMap) and autotetraploid mapping software (TetraploidMap).

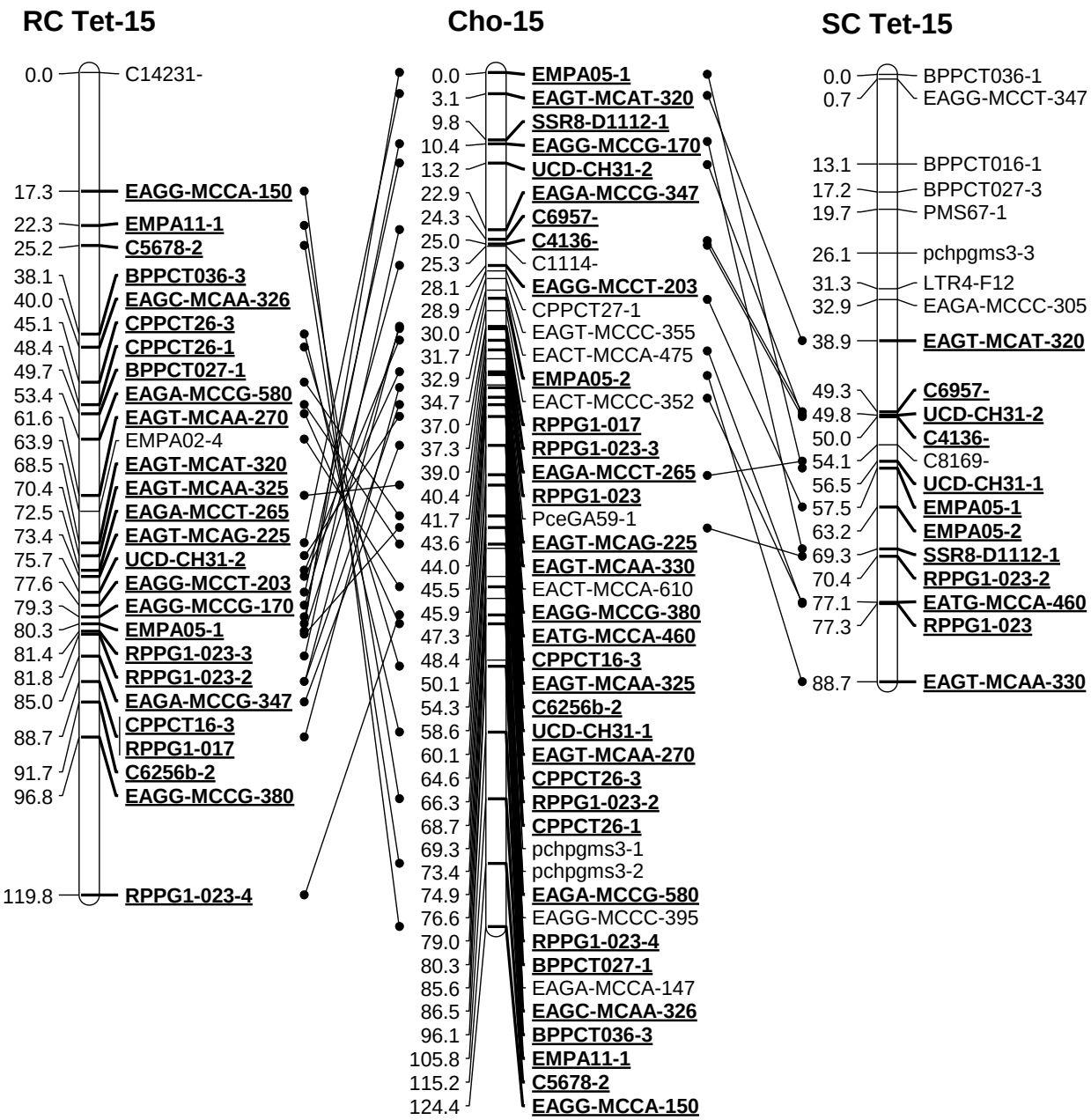

Fig. 13. Integrated chokecherry linkage group (Cho-15) and its colinearity with TetraploidMap-developed parental linkage groups: Resistant chokecherry (RC Tet-15) and susceptible chokecherry (SC Tet-15)

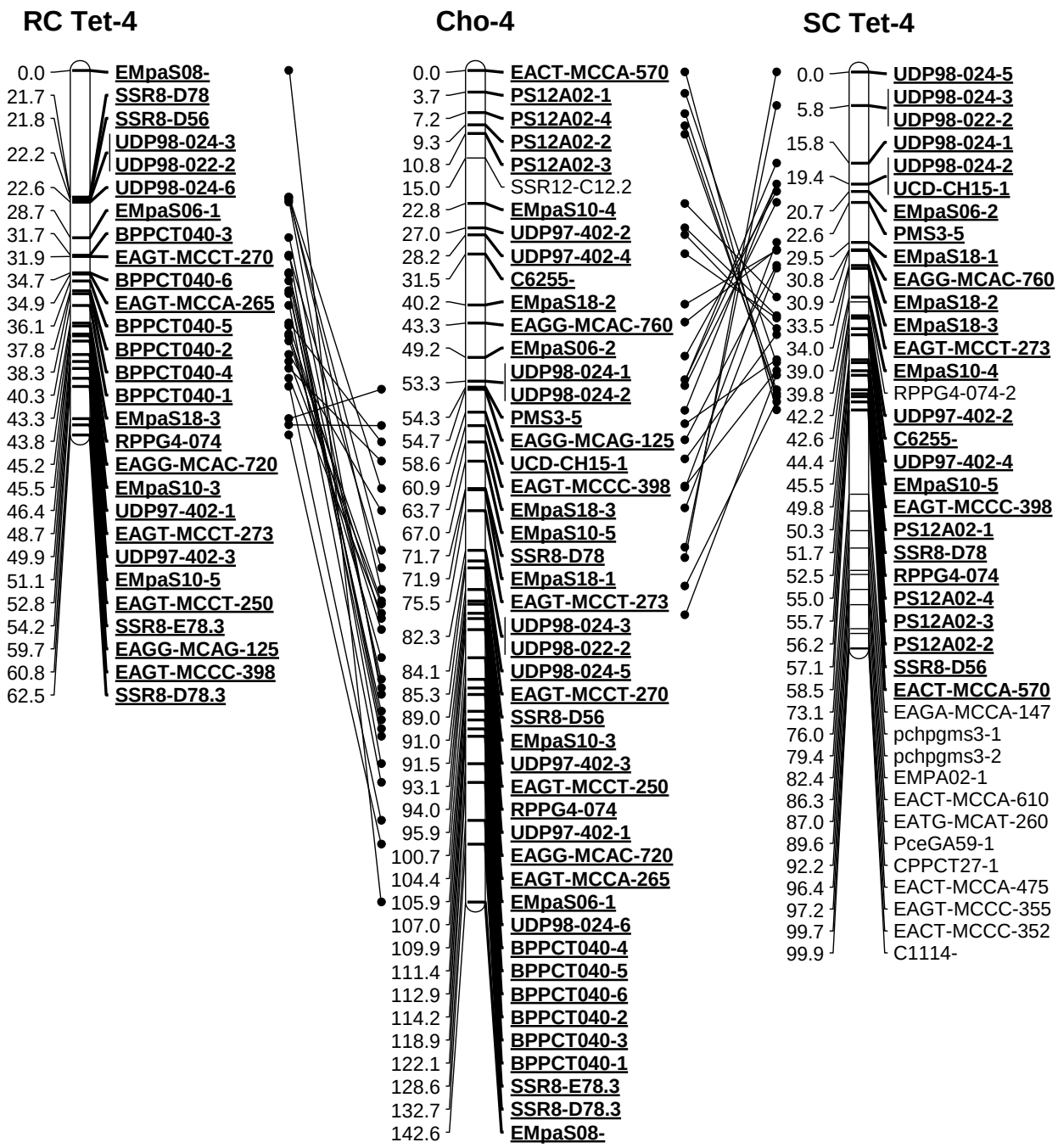

Fig. 14. Integrated chokecherry linkage group (Cho-4) and its colinearity with TetraploidMap-developed parental linkage groups: Resistant chokecherry (RC Tet-4) and susceptible chokecherry (SC Tet-4)

4.2.1. Genetic mapping in cross-pollinating species

Genetic maps produced from reliable molecular markers and heterogeneous populations are the basis for forward genetics, comparative genomics, and QTL identification. Chokecherry, being the only plant species reported as resistant to X-disease, will benefit greatly from a reference genetic map; however, a few limitations besides having limited genetic resources needed to be addressed. Chokecherry trees are obligate outcrossing species, have a long juvenile period, and complicated genetic make-up; therefore, alternatives to development of true F_2 or backcross populations for mapping are needed. Weeden et al. (1994) described appropriate alternatives such as simple hybridization, resulting in a so called pseudo-F_2 population. In open-pollinating species, particularly in those that are self-incompatible woody species, seeds derived from a single tree may represent similar haploid genotypes and so can be considered as pseudo-F_1 plants; thus, a pseudo-F_2 population can be developed by crossing two pseudo-F_1 trees (Weeden et al. 1994). This strategy was recently used for genetic linkage map construction of flowering dogwood (Wang et al. 2009) and other woody species including many *Prunus* species (Canli, 2004; Lambert et al. 2007; Olmstead et al. 2008; Klagges et al. 2013; Wu et al. 2014). Our results also showed that managed hybridization in a cross-pollinating species provides sufficient recombination for genetic mapping.

4.2.2. Integration of chokecherry parental maps

Wang et al. (2014) developed the first chokecherry genetic linkage map (RC x SC) consisting of individual parental linkage groups using TetraploidMap (Hackett et al. 2007). The RC (female parent) map consisted of 14 linkage groups in which four were composed of multiple segments. The SC (male parent) map had 16 linkage groups and four consisted of segments. The

linkage group segments resulted from large gaps (>30 cM). Combining parental maps increases marker density and can fill some gaps. Ultimately, combined maps have an advantage in QTL mapping because of the complete parental allelic consideration during QTL analysis and the decreased QTL interval (Keyser et al. 2010; Klagges et al. 2013; Wu et al. 2014). In this study a well-described 'psuedo-F_2' segregating population derived from a cross between two parents with contrasting X-disease resistance was used (Wang et al. 2014). JoinMap 4.0 (Van Ooijen, 2006) was selected to construct a single map for a cross pollinating (CP) population. Regression mapping in JoinMap was first used for each parental map separately. Regression mapping permits the construction of linkage groups by adding loci one at a time starting from the most informative pair of loci (Van Ooijen, 2006). The best position of the most informative markers is searched by comparing the goodness-of-fit of the calculated map for each tested position. When the goodness-of-fit measure decreases sharply for a locus, it is removed and the process is continued until a framework map is produced (Van Ooijen, 2006; Wang et al. 2011). Two more rounds of goodness-of-fit position appropriate marker loci that were previously removed. After linkage groups for each parent was established, joining homologous groups was successfully implemented.

Integration of maps by regression mapping in JoinMap is based on mean recombination frequencies and combined logarithm of the odds (LOD) scores. Applying the regression mapping algorithm requires at least two common markers to provide relative map distances. A total of 11 linkage groups from each parental map were successfully joined together. The remaining pairs of linkage groups were unstable meaning that a few conflicting markers had insufficient linkage leading to incomplete regression mapping. In an attempt to overcome this issue, markers showing segregation distortion or missing data were excluded from the parental

maps. Nevertheless, five out of 16 groups remained unstable and needed to be ordered with another program called MergeMap (Wu et al. 2008). MergeMap relies on graph theory (Yap et al. 2003; Jackson et al. 2005) and uses directed acyclic graphs (DAGs) to represent maps from individual populations and to resolve conflicts between maps. Although MergeMap does not make use of genotype data, simulations have shown that MergeMap can outperform JoinMap in terms of ordering accuracy and running time (Wu et al. 2008; Wang et al. 2011). It is important to understand that that MergeMap relies solely on the linear arrangement of molecular markers from each paired map and does not use the genotypic data to perform the map re-calculation. As a result, JoinMap tends to produce more accurate estimates of genetic distances. However, JoinMap has limited utility when a low number of shared markers are found between individual maps. JoinMap resolves marker order in the integrated map based on mean recombination frequencies and combined LOD scores (Wang et al. 2011). MergeMap resolves conflicts by identifying and eliminating a small number of markers that are of questionable value from the maps. MergeMap only requires the marker order and cM distances of the component maps, rather than the original genotypic data. Therefore, it is crucial that the original parental maps are a reliable representation of marker order.

The primary limitation of MergeMap is an overestimation of genetic length of the integrated maps (Wang et al. 2011). Overall, when accurate estimates of genetic distances are not the priority, MergeMap provides a rapid and relatively reliable solution (Wang et al. 2011). Indeed, JoinMap and MergeMap can generate integrated maps with good consistency in marker order, so both have been used to construct the novel chokecherry linkage groups. Furthermore, the increase in marker density produced by combining parental maps have improved QTL

mapping and have provided a resource for future examination of genetic and physical positions (i.e. map-based cloning, comparative genomics, and genome sequencing).

4.2.3. Syntenic relationship of chokecherry and other *Prunus* maps

Synteny, as described before, is the product of shared chromosomal segments with the same genetic order of molecular markers between closely related species (Tang et al. 2008; Dirlewanger et al. 2004; Cabrera et al. 2009). Transferability of molecular markers relies on synteny; however, these same transferable markers can provide a means of determining synteny between two species. The reliability of the peach genome sequence and genetic maps have been utilized in studies to confirm synteny and collinearity of peach and *Prunus* species (Arus et al. 2012; Shulaev et al. 2008; Zhebentyaveva et al. 2008; Klagges et al. 2013). In this study, the new chokecherry linkage maps did not contain enough shared markers with the *Prunus* reference map (T x E) or the sweet cherry genetic map (EF x NY) to conduct a thorough study of their synteny; however, certain linkage groups show homology to a few representative *Prunus* chromosomes. Tables 8, 9, and 10 show which chokecherry linkage groups are homologous to *Prunus* linkage groups. Since chokecherry is a tetraploid, it may be expected that rearrangements and duplicated loci have resulted in non-collinearity to other *Prunus* species. The tables aforementioned could elude to this postulation, however, it seems that more shared markers and/or genome sequencing will help deduce the evolutionary relationship of chokecherry and *Prunus* species.

4.3. QTL mapping for X-disease resistance in chokecherry

The joint analysis of genotype marker segregation in genetic maps and phenotypic data from individuals enables the detection of loci affecting quantitative traits. Crossing two parents having dissimilar phenotype allows different but linked loci to be co-segregating,

consequentially leading to the discovery of QTL and the inferred genetic location relative to linked molecular markers. The most obvious applications of QTL analysis seem to be marker-assisted selection (MAS) for breeding and QTL map-based cloning. More specifically, QTL characterization can help deduce the genetic mechanisms of plant–pathogen interaction, plant evolutionary genomics, DNA regulatory elements, and germplasm enhancement. The success in these applications depends on the reliability and accuracy of the QTL analysis and the underlying genetic linkage maps being used. It is also important to consider the limitations of QTL discovery. In 1994, Beavis reported some of these limitations in a mapping population (n=400) that were analyzed as four separate populations of 100. Results showed that none of the identified QTL were shared among the four groups. This profound documentation indicates that careful consideration needs to be done while using QTL information. Nonetheless, QTL mapping has proven its utility in a multitude of plant species and is still a useful tool used for genetic research.

In this study, new genetic linkage maps were grouped for chokecherry with the aim to provide a framework for future studies and to identify additional QTL located near molecular markers. The new linkage groups presented have a higher number of molecular markers and increased marker density. Three significant QTL associated with X-disease were identified. Although it is impossible to unveil all possible QTL using a single segregating population, our population was large enough to identify QTL that explained 45.9% of the phenotypic variation and had a cumulative additive effect of 1.79 to the phenotype scores. The QTL located on linkage group Cho-15 contributed the most to the overall phenotypic effect and also had the shortest genetic interval of 2.1 cM. This particular QTL was the only one previously identified by Wang et al. (2014). That study mapped the QTL within a ten-fold longer interval of 21.4 cM.

Additional markers and the integration of the parental maps resulted in an increase in marker density which explains the shorter genetic interval. Interestingly, the increased marker density and shorter QTL interval were associated with smaller phenotypic variation than previously reported at this locus. Wang et al. (2014) reported 26.6% phenotypic variation was explained by the locus, while present results show 18.4% of the phenotypic variation is explained. This may be attributed to the discovery of additional QTL contributing to the phenotype and the possibility of more molecular markers causing background interference during analysis.

The other two QTL located on groups Cho-5 and Cho-4 respectively, span 11.5 and 6.9 cM distances and explain 14.6 and 12.9% of the total phenotypic variation. It is important to consider the physical distance between the QTL flanking markers. Note that 10 cM equates to 300 kilo base pairs (kbp) in *Arabidopsis* and 6,000 kbp in wheat (Asins 2002). Also, genetic distance changes with chromosomal regions and the corresponding recombination frequencies. Until more resources are utilized in chokecherry, such as genomic sequencing and SNP markers, it will be difficult to determine how the genetic distances correlate with X-disease QTL. Although this issue is overcome by closely linked markers in MAS, map-based cloning and genome walking rely on markers being a short physical distance from the QTL or marker. Nevertheless, the present chokecherry linkage map is an improvement from previously developed maps. Map-based cloning can still be explored for all three QTL. Sequence information of the nearest molecular markers will allow for map-based cloning beginning with genome walking. In this way the DNA sequence information within the QTL location will produce additional markers to hone in on the QTL. Eventually, sequencing across the precise QTL position will provide a list of candidate genes that can be confirmed with complementation studies.

4.4. Future applications

In spite of its economic importance, few molecular genetics and genomic research studies have targeted X-disease phytoplasma. The paucity of genetic information on chokecherry is a big obstacle for studying the mechanisms of its genetic resistance to X-disease (*Candidatus* Phytoplasma pruni). The linkage maps constructed in this study have provided a basis to identify QTL and will help discover genes relevant to the X-disease resistance/susceptibility response. Also, the genetic maps will provide a framework for marker assisted selection (MAS) and genomic analysis. Future whole genome sequencing and/or transcriptome expression profiles will provide relevant data essential to expanding the genetic understanding of X-disease and other phytoplasmic 'yellowing' diseases.

Single nucleotide polymorphism (SNP) markers are an example of a valuable molecular tool used in genetic mapping. SNP markers are abundant and polymorphic, but SNP assay development is time consuming and cost-intensive; therefore, SNPs were not a practical resource option in chokecherry as of yet. Recently, it has been reported that new rosaceous and *Prunus* SNP arrays are somewhat transferable and could be applied to chokecherry in the future (Peace et al. 2012; Verde et al. 2012). Also, the availability of the peach reference genome and the partial sweet cherry genome (International Peach Genome Initiative 2013; Guajardo et al. 2015) will enable the alignment of SNPs used in future mapping studies of chokecherry.

Advances in next-generation sequencing (NGS) technology have provided a wealth of genomic data that has been a valuable resource for molecular genetic research of many plant species. For example, a large number of molecular markers, such as SSRs and SNPs, can be identified from NGS data (Cavagnaro et al. 2010; Zalapa et al. 2012; Dettori et al. 2015; Peace et al. 2012; Guajardo et al. 2015; Verde et al. 2012); however, due to a lack of genomic resources

aliquoted to non-model species, relatively few molecular markers have been discovered. As the efficiency, affordability, and accessibility continue to increase for DNA sequencing we can expect non-model species (i.e. chokecherry) to have the full capacity of genomic resources available (i.e. SNP markers).

The understanding of molecular mechanisms involved in symptom development and interaction between phytoplasmas and their hosts is quite limited; however, map-based cloning and RNA-seq are becoming the forefront of discovering candidate genes involved in the host-pathogen interaction such as seen for Paulownia witches' broom (PaWB) phytoplasma and Paulownia tree species (Liu et al. 2013). In congruence, biochemistry studies have linked gene expression data to deduce the host-pathogen interactions of phytoplasma diseases such as Bois Noir (BN), European Stone Fruit Yellows (ESFY), and Apple Proliferation (AP) (Bertamini et al. 2002; Musetti et al. 2004 and 2005). In chokecherry, X-disease QTL and gene candidates will support the previous phytoplasma research. As mentioned in the literature review, there are multiple types of phytoplasma and different associated diseases. Uncovering the genetic mechanisms of host resistance will be crucial for future understanding of phytoplasma and management strategies. After the resources are provided for chokecherry research, applications of RNA-seq and candidate gene cloning will be readily available and highly applicable.

CONCLUSION

To conclude, the novel chokecherry genetic map constructed in this study represents a high quality framework that can be used for the elucidation of X-disease ('*Candidatus* Phytoplasma pruni') response in woody plant species, especially in the *Prunus* genus. The present maps have been instrumental in current QTL analysis, and will be a reliable reference for genetic and genomic applications such as marker assisted selection (MAS), DNA sequencing,

RNA-seq, and candidate gene cloning. Improving the chokecherry germplasm with the proposed technology could offer great benefits for the fruit tree industry. Chokecherry is a troublesome host of X-disease, so providing disease resistance would greatly inhibit the spread of X-disease phytoplasma to other hosts including orchard production systems. This would directly benefit important *Prunus* crop species such as peach, apricot, plum, cherry, and nectarine by reducing production costs associated with X-disease management. As another use, *Prunus* breeding programs could utilize disease resistant chokecherry lines as germplasm sources for inter-specific crosses or gene integration via biotechnology. Lastly, natural resistance mechanisms in chokecherry could be used as a template to examine natural resistance in other *Prunus* species.

REFERENCES

Abbott AG, Rajapakse S, Sosinski B, Lu ZX, Sossey-Alaoui K, Gannavarapu M, Scorza R, Reighard G, Ballard RE, Callahan A, Baird WV (1998) Construction of saturated linkage maps of peach crosses segregating for characters controlling fruit quality, tree architecture and pest resistance. Acta Hort 465:141-150

Alberazzi G, Milc J, Caffagni A, Francia E, Roncaglia E, Ferrari F, Tagliafico E, Stefani E, Pecchioni N (2009) Gene expression in grapevine cultivars in response to Bois Noir phytoplasma infection. J Plant Sci 176(6): 792-804. doi:10.1016/j.plantsci.2009.03.001

Aru´ s P, Verde I, Sosinski B, Zhebentyayeva T, Abbott A (2012) The peach genome. Tree Genet Genomes 8: 531–547

Asins MJ (2002) Present and future of quantitative trait locus analysis in plant breeding. Plant Breed 121: 281−291

Audic S, Claverie JM (1997) The significance of digital gene expression profiles. Genome Research **7:** 986-995

Barakat A, Staton M, Cheng CH, Park J, Yassin NBM, Ficklin S, Yeh CC, Hebard F, Baier K, Powell W, Schuster SC, Wheeler N, Abbort A, Carlson JE, Sederoff R (2012) Chestnut resistance to the blight disease: insights from transcriptome analysis. BMC Plant Biology 12:38 doi:10.1186/1471-2229-12-38

Beaver JA and Iezzoni AF (1993) Allozyme inheritance in tetraploid sour cherry (*Prunus cerasus* L.). J Amer Soc Hort Sci 118(6): 873-877

Beavis WD (1994) The power and deceit of QTL experiments: Lessons from comparative QTL studies. Proceedings of the 49th Annual Corn and Sorghum Research Conference, American Seed Trade Association. 49:250-265

Belthoff LE, Ballard R, Abbott A, Baird WV, Morgens P, Callahan A, Scorza R, Monet R (1993) Development of a saturated linkage map of *Prunus persica* using molecular based marker systems. Acta Hort 336:51-56

Bennetzen J.L (2000) Transposable element contributions to plant gene and genome evolution. Plant Mol Biol 42:251-269. doi:10.1023/A:1006344508454

Bertaccini A & Duduk, B (2009) Phytoplasma and phytoplasma diseases: a review of recent research. Phytopathol Mediterr 48: 355–378

Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422, 433-438. doi: 10.1038/nature01521

Bradshaw JE, Hackett CA, Pande B, Waugh R, Bryan GJ (2008) QTL mapping of yield, agronomic and quality traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). Theor Appl Genet 116(2): 193-211

Butler L (1952) The linkage map of the tomato. J Hered 43: 25-35

Cai G, Yang Q, Yi B, Fan C, Edwards D, Batley J, et al. (2014) A complex recombination pattern in the genome of allotetraploid *Brassica napus* as revealed by a high-density genetic map. PLoS ONE 9(10): e109910. doi:10.1371/journal.pone.0109910

Canli FA (2004) Development of a second generation genetic linkage map for sour cherry using SSR markers. Pak J Biol Sci 7:1676-1683

Canli FA (2008) Progress in genetic mapping of *Prunus* species. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi 24: 414-424

Cavagnaro PF, Senalik DA, Yang L, Simon PW, Harkins TT, Kodria CD, Huang S, Weng Y (2010) Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). BMC Genomics 11: 569-586

Chen T, Lv Y, Zhao T, Li N, Yang Y, Yu W, He X, Liu T, Zhang B (2013) Comparative transcriptome profiling of a resistant *vs* susceptible tomato (*Solanum iycopersicum*) cultivar in response to infection by tomato yellow leaf curl virus. PLoS ONE 8(110: e80816.doi: 10.1371/journal.pone.0086816

Christensen N, Axelsen K, Nicolaisen M, Schulz A (2005) Phytoplasmas and their interactions with hosts. TRENDS in Plant Science 10: 526–535

Davis RE, Zhao Y, Dally EL, Lee IM, Jomantiene R, Douglas SM (2013) '*Candidatus* Phytoplasma pruni', a novel taxon associated with X-disease of stone fruits, *Prunus* spp.: multilocus characterization based on 16S rRNA, secY, and ribosomal protein genes. Int J Syst Evol Microbiol 63(Pt 2): 766-76. doi: 10.1099/ijs.0.041202-0

Dettori MT, Micali S, Giovinazzi J, Scalabrin S, Verde I, Cipriani G (2015) Mining microsatellites in the peach genome: development of new long-core SSR markers for genetic analyses in five *Prunus* species. Springerplus 4: 337. doi: 10.1186/s40064-015-1098-0

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

Dirlewanger E, Graziano E, Joobeur T, Garriga-Caldere F, Cosson P, Howad W, Arus P (2004) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. PNAS 101: 26. doi: 10.1073/pnas.0307937101

Dirlewanger E, Moing A, Rothan C, Svanella L, Pronier V, Guye A, Plomion C, Monet R (1999) Mapping QTLs controlling fruit quality in peach (*Prunus persica* (L.) Batsch). Theor Appl Genet 98:18-31

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

Doi Y, Teranaka M, Yora K, Asuyama H (1967) Mycoplasma or PTL-group-like microorganisms found in the phloem elements of plants infected with mulberry dwarf, potato witches'-broom, aster yellows, or paulownia witches'-broom. Ann. Phytopathol. Soc. Jpn. 33:259-266

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

Douglas SM (1986) Detection of mycoplasma-like organisms in peach and chokecherry with X-disease by fluorescence microscopy. Phytopathology 76: 784–787

Douglas SM (1999) Management of X-disease in stone fruits. The Connecticut Agricultural Experiment Station Extension Publ. PP062. Available from: http://www.ct.gov/caes/cwp/view.asp?a=2815&q=376862

Flint-Garcia SA, Thorsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu. Rev. Plant Biol. 54: 357-374. doi: 10.1146/annurev.arplant.54.031902.134907

Foulongne M, Pascal T, Pfeiffer F, Kervella J. (2003) QTLs for powdery mildew resistance in peach × *Prunus davidiana* crosses: consistency across generations and environments. Molecular Breeding 12(1): 33-50

Gar O, Sargent DJ, Tsai C-J, Pleban T, Shalev G, Byrne DH, Zamir D (2011) An autotetraploid linkage map of rose (*Rosa hybrida*) validated using the strawberry (*Fragaria vesca*) genome sequence. PLoS ONE 6:e20463

Gasic K, Han Y, Kerbundit S, Shulaev V, Iezzoni AF, Stover EW (2009) Characteristics and transferability of new apple EST-derived SSRs to other Rosaceae species. Mol Breed 23: 397–411

Gisbert AD, Martinez-Calvo J, Llacer G, Badenes ML, Romero C (2009) Development of two loquat [*Eriobotrya japonica* (Thunb.) Lindl.] linkage maps based on AFLPs and SSR markers from different Rosaceae species. Mol Breeding 23: 523-538. doi: 10.1007/s11032-008-9253-8.

Guajardo V, Solís S, Sagredo B, Gainza F, Muñoz C, Gasic K (2015) Construction of high density sweet cherry (*Prunus avium* L.) linkage maps using microsatellite markers and SNPs detected by genotyping-by-sequencing (GBS). PLoS ONE 10(5): e0127750. doi:10.1371/journal.pone.0127750

Guo Y, Cheng Z, Walla J, Zhang Z (1998). Large-scale screening for X-disease phytoplasma infection in chokecherry. HortScience 33(2): 293-295.

Hackett CA, Luo Z (2003) TetraploidMap: construction of a linkage map in autotetraploid species. J Heredity 94: 358-359

Hackett CA, McLean K, Bryan GL (2013) Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. PLoS One 8(5): e63939. doi: 10.1371/journal.pone.0063939

Hackett CA, Milne I, Bradshaw JE, Luo Z (2007) TetraploidMap for Windows Linkage map construction and QTL mapping in autotetraploid species. J Heredity 98: 727-729

Hancock AM, Iezzoni AF (1988) Malate dehydrogenase isozyme patterns in seven Prunus species. Hortscience 23: 381–383

Havecker ER, Gao X, Voytas DF (2004) The diversity of LTR retrotransposons. Genome Biology 5: 225. doi: 10.1186/gb-2004-5-6-225

Hickok LG, 1978a Homoeologous chromosome pairing and restricted segregation in the fern *Ceratopteris*. Am. J. Bot. 65: 516– 521.

Hickok LG, (1978b) Homoeologous chromosome pairing: frequency differences in inbred and intraspecific hybrid polyploid ferns. Science 202: 982–984.

Hedrick UP and Booth NO (1907) Mendelian characters in tomatoes. Proc Soc Hort Sci 5: 19-24

Huang D, Walla J, Dai W (2014) Quantitative phenotyping of X-disease resistance in chokecherry using real-time PCR. J Microbiol Meth 98: 1-7

Hummer KE, & Janick J (2009) Rosaceae: Taxonomy, economic importance, genomics. In: Plant genetics and genomics: Crops and models. Folta KM, Gardiner SE (eds.) doi: 10.1007/978-0-387-77491-6 1

Hurtado MA, Romero C, Vilanova S, Abbott AG, Llacer G, Badenes ML (2002) Genetic linkage maps of two apricot cultivars (*Prunus armeniaca* L.) and mapping of PPV (sharka) resistance. Theor Appl Gent 105: 182-191

International Committee on Systematic Bacteriology Subcommittee on the Taxonomy of *Mollicutes* (1993) Minutes of the Interim Meetings, August 1-2, 1992. Ames, Iowa. International Journal of Systematic Bacteriology 43: 394–397.

IRPCM Phytoplasma/Spiroplasma Working Team – Phytoplasma Taxonomy Group (2004) '*Candidatus* Phytoplasma', a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. International Journal of Systematic and Evolutionary Microbiology, 54, 1243–1255.

Jabs T, Tschope M, Colling C, Hahlbrock K, Scheel D (1997) Elicitor stimulated ion fluxes and O2- from the oxidative burst are essential components in triggering defense gene activation and phytoalexin synthesis in parsley. Proc Natl Acad Sci USA 94: 4800–4805

Jackson BN, Aluru S, Schnable PS: Consensus genetic maps (2005) A graph theoretic approach. IEEE Computational Systems Bioinformatics Conference, Proceedings 2005: 35-43

Jaillon et al. (2007) The French–Italian Public Consortium for Grapevine Genome Characterization. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449: 463-467. doi: 10.1038/nature06148

Jander G, Norris SR, Rounsley SD, Bush DF, Levin IM, Last RL (2002) *Arabidopsis* map-based cloning in the post-genome era. Plant Physiology 129(2): 440-450. doi: http://dx.doi.org/10.1104/pp.003533

Joehanes R, Nelson JC (2008) QGene 4.0, an extensible Java QTL-analysis platform. Bioinformatics 24:2788-2789

Jones AL, Sutton TB (1996) Diseases of Tree Fruits in the East. Michigan State University, East Lansing, MI

Joobeur T, Periam N, de Vicente MC, King GJ, Arús P (2000) Development of a second generation linkage map for almond using RAPD and SSR markers. Genome 43:649–655

Joobeur T, Viruel MA, de Vicete MC, Jauregui 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 almond × peach F_2 progeny. Theor Appl Genet 97: 1034-1041

Kalendar R, Flavell AJ, Ellis THN, Sjakst T, Moisy C, Schulman AH (2011) Analysis of plant diversity with retrotransposon-based molecular markers. J Heredity 106: 520-530. doi:10.1038/hdy.2010.93

Keyser ED, Shu QY, Van Bockstaele E, De Riek J (2010) Multipoint-likelihoot maximization mapping on 4 segregating populations to achieve an integrated framework map for QTL analysis in pot azalea (*Rhododendron simsii* hybrids). BMC Molecular Biology 11: 1

Kirkpatrick BC (1989) Strategies for characterizing plant pathogenic mycoplasma-like organisms and their effects on plants. In: Kosuge, T. and Nester, E. W. (Eds) Plant-Microbe Interactions-Molecular and Genetic Perspectives (pp. 241-294). McGraw-Hill Publishing Co., New York, NY

Kirkpatrick, B. C., J. K. Uyemoto, and A. H. Purcell. 1995. X-disease. Pp. 57-59. In J.M. Ogawa, E.I. Zehr, G.W. Bird, D.F. Ritchie, K. Uriu, and Uyemoto, J.K. (Eds.) Compendium of Stone Fruit Diseases. APS Press, St. Paul, MN.

Klagges C, Campoy JA, Quero-Garcia J, Guzman A, Mansur L, Gratacos E, Silva H, Rosyara UR, Iezzoni A, Meisel LA, Dirlewanger E (2013) Construction and comparative analyses of highly dense linkage maps of two sweet cherry intra-specific progenies of commercial

cultivars. PLoS ONE 9(1): 10. doi: 10.1371/annotation/80cd091d-31b2-473a-b028-ea9b211b5d9d

Koning-Boucoiran CFS, Gitonga VW, Yan Z, Dolstra O, van der Linden CG, van der Schoot J, Uenk GE, Verlinden K, Smulders MJM, Krens FA, Maliepaard C (2012) The mode of inheritance in tetraploid cut roses. Theor Appl Genet 125: 591-607. doi:10.1007/s00122-012-1855-1

Krebs SL and Hancock JF (1989) Tetrasomic inheritance of isoenzyme markers in the highbush blueberry, *Vaccinium corymbosum* L. Heredity 63: 11-18.

Kube M, Mitrovic J, Duduk B, Rabus R, Seemuller E (2012) Current view on phytoplasma genomes and encoded metabolism. Scientific World Journal 2012: Article ID 185942. doi: 10.1100/2012/185942

Laimer M, Lamaire O, Herrbach E, Goldschmidt V, Minafra A, Bianco P, Wetzel T (2009) Resistance to virsuses, phytoplasmas and their vectors in the grapevine in Europe: A review. Journal of Plant Pathology 91(1): 7-23. doi: http://dx.doi.org/10.4454/jpp.v91i1.620

Lambert P, Dicenta F, Rubio M, Audergon JM (2007) QTL analysis of resistance to sharka disease in the apricot (*Prunus armeniaca* L.) 'Polonais' × 'Stark Early Orange' F1 progeny. Tree Genetics & Genomes 3(4): 299-309

Lande R & Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124(3): 743-756

Lee IM, Davis RE (1992) Mycoplasmas that infect plants and insects. In: Maniloff, J., McElhaney, R. N., Finch, R. N., and Baseman, J. B. (Eds.) Mycoplasmas: Molecular Biology and Pathogenesis (pp.379-390). Amer Soc Microbiol, Washington, D.C.

Lee IM, Davis RE, Gundersen-Rindal DE (2000) Phytoplasma: phytopathogenic mollicutes. Annu Rev Microbiol 54:221-255

Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, *et al*: *De novo* assembly of human genomes with massively parallel short read sequencing. Genome Research 20: 265-272

Liu R, Dong Y, Fan G, Zhao Z, Deng M, et al. (2013) Discovery of Genes Related to Witches Broom Disease in *Paulownia tomentosa* × *Paulownia fortunei* by a *De Novo* Assembled Transcriptome. PLoS ONE 8(11): e80238. doi:10.1371/journal.pone.0080238.

Liu, Z., Wang, Y., Xiao, J., Zhao, J., Mengjun, L. (2014) Identification of genes associated with phytoplasma resistance through suppressive subtraction hybridization in Chinese jujube. Physiological and Molecular Plant Pathology. 86:43-48

Loi N, Ermacora P, Carraro L, Osler R, and Chen, TA (2002) Production of monoclonal antibodies against apple proliferation phytoplasma and their use in serological detection. Eur. J. Plant Pathol 108: 81-86

Lu ZX, 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: 199-207

Maliepaard C, Jansen JJ, van Ooijen (1997) Linkage analysis in a full-sib family of an outbreeding plant species: overview and consequences for applications. Genetical Research 70: 237-250

Marcone C (2015) Current status of phytoplasma diseases of peach. Acta Hortic. 1084: 569-578. doi:10.17660/ActaHortic.2015.1084.77. http://dx.doi.org/10.17660/ActaHortic.2015.1084.77

Marone D, Laido` G, Gadaleta A, Colasuonno P, Ficco DB, Giancaspro A, Giove S, Panio G, Russo MA, De Vita P, Cattivelli L, Papa R, Blanco A, Mastrangelo AM (2012) A high-density consensus map of A and B wheat genomes. Theor Appl Genet 125(8): 1619-1638. doi: 10.1007/s00122-012-1939-y

Massa AN, Manrique-Carpintero NC, Coombs JJ, Zarka DG, Boone AE, Kirk WW, Hackett CA, Bryan GJ, Douches DS (2015) Genetic linkage mapping of economically important traits in cultivated tetraploid potato (*Solanum tuberosum* L.). G3 5(11): 2357-2364. doi: 10.1534/g3.115.019646

Maust BE, Espadas F, Talavera C, Aguilar M, Santamaria JM, et al. (2003) Changes in carbohydrate metabolism in coconut palms infected with the lethal yellowing phytoplasma. Phytopathology 93: 976–981

Mayer CJ, Vilcinskas A, Gross J (2008) Phytopathogen lures its insect vector by altering host plant odor. J Chem Ecol 34: 1045-1049.

Mazaheri M, Kianian P, Mergoum M, Valentini GL, Seetan R, Pirseyedi S, . . . Kianian S. (2014) Transposable element junctions in marker development and genomic characterization of barley. The Plant Genome 7(1)

McCallum S, Graham J, Jorgensen L, Rowland LJ, Bassil NV, Hancock JF, Wheeler EJ, Vining K, Poland JA, Olmstead JW, Buck E, Wiedow C, Jackson E, Brown A, Hackett CA (2016) Construction of a SNP and SSR linkage map in autotetraploid blueberry using genotyping by sequencing. Mol Breeding 36: 41. doi: 10.1007/s11032-016-0443-5

McCarthy DJ, Chen y, Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 40: 4288-4297

Mnejja M, Garcia-Mas J, Audergon JM, Aru´s P (2010) *Prunus* microsatellite marker transferability across rosaceous crops. Tree Genet Genomes 6: 689–700

Monden Y, Fujii N, Yamaguchi K, Ikeo K, Yoshiko N, Waki T, Hirashima K, Uchimura Y, Tahara M (2014) Efficient screening of long terminal repeat retrotransposons that show high insertion polymorphism via high-throughput sequencing of the primer binding site. Genome 57: 245-252

Mou HQ, Lu J, Zhu SF, Lin AL, Tian GZ, Xu X, Zhao WJ (2013) Transcriptomic analysis of *Paulownia* infected by *Paulownia* witches'-broom phytoplasma. PLOS ONE 8(10): e77217. doi: 10.1371/journal.pone.0077217

Musetti R, Sanita di Toppi L, Ermacora P, Favali MA (2004) Recovery in apple trees infected with the apple proliferation phytoplasma: An ultrastructural and biochemical study. Phytopathology 94(2):203-208. doi: 10.1094/PHYTO.2004.94.2.203

Musetti R, Sanita di Toppi L, Martini M, Ferrini F, Loschi A, Favali MA, Osler R (2005) Hydrogen peroxide localization and antioxidant status in the recovery of apricot plants from European Stone Fruit Yellows. European Journal of Plant Pathogology 112(1): 53-61. doi: 10.1007/s10658-004-8233-z

Nelson DO (2012) Improving DNA sequencing accuracy and throughput. In: Genetic mapping and DNA sequencing 81: 185-201. Speed TT and Waterman MS (eds.). Springer Science and Business Media. doi: 10.1007/978-1-4612-0751-1

Olden EJ, Nybom N (1968) On the origin of *Prunus cerasus* L. Hereditas 70: 3321–3323.

Olivier CY, Lowery DT, Stobbs LW (2009) Phytoplasma diseases and their relationships with insect and plant hosts in Canadian horticultural and field crops. Can Entomolog 141, 425-426

Olmstead JW, Sebolt AM, Cabrera A, Sooriyapathirana SS, Hammar S, Iriarte G, et al (2008) Construction of an intra-specific sweet cherry (*Prunus avium* L.) genetic linkage map and synteny analysis with the *Prunus* reference map. Tree Genet Genomes 4:897-910

Osler R, Loi N, Carraro L, Ermacora P, and Refatti E. (1999) Recovery in plants affected by phytoplasmas. In 5th Proc. Congr. Eur. Found. Plant Pathol: 589-592

Perez A (ed.) (2014) Fruit and tree nuts. Retrieved from USDA Economic Research Service: http://www.ers.usda.gov/topics/crops/fruit-tree-nuts/background.aspx.

Peterson GW (1984) Spread and damage of western X-disease of chokecherry in eastern Nebraska plantings. Plant Disease 68:103-104

Rehder A (1940) Manual of cultivated trees and shrubs, 2nd ed. Collier MacMillan Ltd: 452–481

Reid JB and Ross JJ (2011) Mendel's genes: Toward a full molecular characterization. Genetics 189(1): 3–10. doi: 10.1534/genetics.111.132118

Robins JG, Hansen JL, Viands DR, Brummer EC (2008) Genetic mapping of persistence in tetraploid alfalfa. Crop Sci 48: 1780-1786

Rong J, Feltus FA, Waghmare VN, Pierce GJ, Chee PW, Draye X, Saranga Y, Wright RJ, Wilkins TA, May OL, Smith CW, Gannaway JR, Wendel JF, Paterson AH (2007) Meta-analysis of polyploidy cotton QTL shows unequal contributions of sub-genomes to a complex network of genes and gene clusters implicated in lint fiber development. Genetics 176: 2577-2588

Rosenberger DA, Jones AL (1977) Spread of X-disease in Michigan peach orchards. Plant Dis Rep 61:830-834

Rosenberger DA, Jones AL (1978) Leafhopper vectors of the peach X-disease pathogen and its seasonal transmission from chokecherry. Phytopathology 68:782-790

Rubio M, Ballester AR, Olivares PM, Castro de Moura M, Dicenta F, Martínez-Gómez P (2015) Gene expression analysis of plum pox virus (Sharka) susceptibility/resistance in Apricot (*Prunus armeniaca* L.). PLoS ONE 10(12): e0144670. doi: 10.1371/journal.pone.0144670

Salazar JA, Ruiz D, Campoy JA, Sánchez-Pérez R, Crisosto CH, Martínez-García PJ, Blenda A, Jung S, Main D, Martínez-Gómez P, Rubio M (2014) Quantitative trait loci (QTL) and Mendelian trait loci (MTL) analysis in *Prunus*: A breeding perspective and beyond. Plant Mol Biol Rep 32: 1-18. doi: 10.1007/s11105-013-0643-7

Salvi S, Belloti M, Conti S, Frascalori E, Giulani S, Landi P, Maccaferri M, Natoli V, Sanguinetti MC, Sponza G, Talamè V, Tuberosa R (2005) The art and science of cloning QTLs in plants. In: In the wake of the double helix: From the green revolution to the gene revolution. Tuberosa R, Phillips RL, Gale M (eds.) Avenue Media: 327−345

Santi F, Lemoine M (1990) Genetic markers for *Prunus avium* L.: Clonal identifications and discrimination from *P. cerasus* and *P. cerasus* x *P. avium*. Ann for Sci 47: 219–227

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: 292-301. doi: 10.1038/hdy.2011.74

Schuster M, Schreiber H (2000) Genome investigation in sour cherry, *P. cerasus* L. Acta Horticulture 538: 375–379

Shulaev V, Korban SS, Sosinski B, Abbott AG, Aldwinckle HS (2008) Multiple models for Rosaceae genomics. Plant Physiol 147: 985–1003.

Sinclair WA, Lyon HH (2005) Diseases of Trees and Shrubs. Comstock Publ. Assoc., Ithaca and London. Second Edition

Sinha RC, Chiykowski LN (1980) Transmission and morphological features of mycoplasmalike bodies associated with peach X-disease. Can J Plant Path 2:119-124

Smit AFA, Hubley R & Green, P. (2013) *RepeatMasker Open-4.0*. Available at http://www.repeatmasker.org

Socquet-Juglard D, Kamber T, Pothier JF, Christen D, Gessler C, et al. (2013) Comparative RNA-Seq Analysis of Early-Infected Peach Leaves by the Invasive Phytopathogen *Xanthomonas arboricola* pv. *pruni*. PLoS ONE 8(1): e54196. doi:10.1371/journal.pone.0054196

Soltis DE and Rieseberg LH (1986) Autopolyploidy in *Tolmiea menziesii* (Saxifragaceae): Genetic insights from enzyme electrophoresis. Amer. J. Bot. 73:310-318.

Stift M, Berenos C, Kuperus P, van Tienderen PH (2008) Segregation models for disomic, tetrasomic, and intermediate inheritance in tetraploids: A general procedure applied to Rorippa (yellow cress) microsatellite data. Genetics 179: 2113-2123. doi:10.1534/genetics.107.085027

Stoddard EM, Hildebrand EM, Palmiter DH, Parker KG (1951) In virus diseases and other disorders with viruslike symptoms of stone fruits in North America. United States Department of Agriculture. Agriculture Handbook 10: 37–42

Stoddard, E M (1938) The ‘‘X disease’’ of peach. Connecticut Agricultural Experiment Station Circular 122: 54–60

Sun J, Yin H, Li L, Song Y, Fan L, Zhang S, Wu J (2015) Evaluation of new IRAP markers of pear and their potential application in differentiating bud sports and other Rosaceae species. Tree Genetics & Genomes 11:25. doi: 10.1007/s11295-015-0849-y

Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH (2008). Synteny and collinearity in plant genomes. Science, 320(5875): 486–488. doi: 10.1126/science.1153917

Tavassolian I, Rabiei G, Gregory D, Mnejja M, Wirthensohn MG (2010) Construction of an almond linkage map in an Australian population 'Nonpareil' x 'Lauranne'. BMC Genomics 11: 551

The International Peach Genome Initiative (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet Mar 24. doi: 10.1038/ng.2586. [Epub ahead of print]

Tobutt KR, Boskovic R, Cerovic R, Sonneveld T, Ruzic D (2004) Identification of incompatibility alleles in the tetraploid species sour cherry. Theor Appl Genet 108:775-785

Turner JG, Ellis C, Devoto A (2002) The jasmonate signal pathway. The Plant Cell 2002(14). doi: http://dx.doi.org/10.1105/tpc.000679

USDA NRCS National Plant Data Team. Chokecherry: *Prunus virginiana* L. USDA NRCS. Available at: http://plants.usda.gov/core/profile?symbol=PRVI

Van de Peer Y, Fawcett JA, Proost S, Sterck L, Vandepoele K (2009) The flowering world: a tale of duplications. Trends Plant Sci 14(12): 680-688. doi: 10.1016/j.tplants.2009.09.001

van Ooijen JW (2006) JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV

Verde I, Bassil N, Scalabrin S, Gilmore B, Lawley CT, et al. (2012) Development and evaluation of a 9K SNP array for peach by internationally coordinated SNP detection and validation in breeding germplasm. PLoS ONE 7(6): 10.1371/annotation/33f1ba92-c304-4757-91aa-555de64a0768

Viruel MA, Messeguer R, de Vicente MC, Garcia-Mas J, Puigdomenech P, Vargas FJ, Arus P (1995) A linkage map with RFLP and isozyme markers for almond. Theor Appl Genet 91: 964-971

Vogel et al. (2010) International Brachypodium Initiative. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463: 763-768. doi: 10.1038/nature08747

Walla JA, Cheng ZM, Guo YH, Knudson MJ (1996) Field resistance of chokecherry to X-disease. Phytopathology 86:87

Walla JA, Cheng ZM, Guo YH, Knudson MJ, Dai W (2005) Monitoring for X-disease tolerance in chokecherry (*Prunus virginiana*). Phytopathology 95: S108

Wan J, Dunning FM, Bent AF (2002) Probing plant-pathogen interactions and downstream defense signaling using DNA microarrays. Funct Integr Genomics 2: 259–273

Wang H, Walla JA, Magnusson VA, Zhong S, Dai W (2014) Construction of a linkage map and QTL mapping for X-disease resistance in tetraploid chokecherry (*Prunus virginiana* L.) using SSR and AFLP markers. Mol Breed (doi:10.1007/s11032-014-0025-3)

Wang H, Walla JA, Zhong S, Huang D, Dai, W (2012) Development and cross-species/genera transferability of microsatellite markers discovered using 454 genome sequencing in chokecherry (*Prunus virginiana* L.). Plant Cell Rep 31:2047-2055

Wang J, Lydiate DJ, Parkin IAP, Flentin C, Delourme R, Carion PWC, King GJ (2011) Integration of linkage maps for the amphidiploid *Brassica napus* and comparative mapping with *Arabidopsis* and *Brassica rapa*. BMC Genomics 12: 10. doi: 10.1186/1471-2164-12-101

Wang X, Wadl PA, Rinehart TA, Scheffler BE, Windham MT, Spiers JM, Johnson DH, Trigiano RN (2009) A linkage map for flowering dogwood (Cornus florida L.) based on microsatellite markers. Euphytica. 165(1):165-175

Wang Z, Gerstain M, Snyder GK (2009) RNA-seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63

Wanjugi H, Coleman-Derr D, Huo N, Kianian SF, Luo M, Wu J, Anderson O, Gu YQ (2009) Rapid development of PCR-based genome-specific repetitive DNA junction markers in wheat. Genome. 52: 576-587. doi: 10.1139/G09-033

Ward JA, Ponnala L, Weber CA (2012) Strategies for transcriptome analysis in nonmodel plants. American J Botany 99: 267-276

Weeden NF, Hemmatt M, Lawson DM, Lodhi M, Manganaris AG, Reisch BI, Brown SK, Ye GN (1994) Development and application of molecular marker linkage maps in woody fruit crops. Euphytica. 77: 71–75

Weintraub PG, Beanland L (2006) Insect vectors of phytoplasmas. Annual Review of Entomology 51: 91–111.

Weintraub PG, Jones P (2010) Phytoplasmas. genomes, plant hosts and vectors. Wallingford, UK: CABI

Wu J, Li L, Li M, Khan MA, Li X, Chen H, Yin H, Zhang S (2014) High-density genetic linkage map construction and identification of fruit-related QTLs in pear using SNP and SSR markers. Journal of Experimental Botany 65(20): 5771-5781. doi: 10.1093/jxb/eru311

Wu J, Zhang Y, Zhang H, Huang H, Folta KM, Liu J (2010) whole genome wide expression profiles of *Vitis amurensis* grape responding to downy mildew by using solexa sequencing technology. BMC Plant Biol 10: 234

Wu KK, Burnquist W, Sorrells ME, Tew TL, Moore PH, Tanksley SD (1992) The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theor Appl Genet 83: 294–300

Wu Y, Bhat P, Close TJ, Lonardi S (2008) On the accurate construction of consensus genetic maps. Comput Syst Bioinformatics Conf 7: 285-296

Wu Y, Close TJ, Lonardi S (2008) On the accurate construction of consensus genetic maps. Comput Syst Bioinformatics Conf 7: 285-296

Yamamoto T, Yamaguchi M, Hayashi T (2005) An integrated genetic linkage map of peach by SSR, STS, AFLP and RAPD. J Japan Soc Hort Sci. 74(3): 204-213

Yang N, Reighard G, Ritchie D, Okie W, Gasic K (2013) Mapping quantitative trait loci associated with resistance to bacterial spot (*Xanthomonas arboricola* pv. *pruni*) in peach. Tree Genetics and Genomes 9(2): 573-586. doi: 10.1007/s11295-012-0580-x

Yap IV, Schneider D, Kleinberg J, Matthews D, Cartinhour S, McCouch SR (2003) A graph-theoretic approach to comparing and integrating genetic, physical and sequence-based maps. Genetics 165(4): 2235-2247

You FM, Wanjugi H, Huo N, Lazo GR, Luo M, Anderson OD, Dvorak J, Gu YQ (2010) RJPrimers: unique transposable element insertion junction discovery and PCR primer design for marker development. Nucleic Acids Research. 38: 313-320. doi:10.1093/nar/gkq425

Zalapa JE, Cuevas H, Zhu H, Steffan S, Senalik D, Zeldin E, McCown B, Harbut R, Simon P (2012) Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant science. Am J Bot 99: 193-208

Zhang M, Fan L, Liu Q, Song Y, Wei S, Zhang S, Wu J (2014) A novel set of EST-derived SSR markers for pear and cross-species transferability in Rosaceae. Plant Mol Biol Rep 32: 290-302. doi: 10.1007/s11105-013-0638-4

Zhang Z, Dai W, Cheng ZM, Walla JA (2000) A shoot-tip culture micropropagation system for chokecherry. J Environ Hort 18: 234-237

Zhao Y, Wei W, Lee IM, Shao J, Suo X, Davis RE (2009) Construction of an interactive online phytoplasma classification tool, *i*PhyClassifier, and its application in analysis of the peach X-disease phytoplasma group (16SrIII). Int J Syst Evol Microbiol 59: 2582-2593

Zhebentyayeva T, Swire-Clark G, Georgi L, Garay L, Jung S (2008) A framework physical map for peach, a model Rosaceae species. Tree Genet Genomes 4: 745–756