ABSTRACT

LIN, YING-CHEN. Improved and Novel Resources for Blueberry Mapping and Genomics. (Under the direction of Dr. Allan Brown.)

Blueberry (*Vaccinium* spp.) is an economically important small fruit crop (564.4 million pounds harvested in the U.S. with a value of $850.9 million in 2012), but relatively few genomic resources are available to researchers and breeders for the purposes of plant improvement. Traditional blueberry breeding is constrained by severe inbreeding depression, the complex nature of the polyploidy genome (2n=24, 4n=48 and 6n=72), and a relatively long period of juvenility. With the advent of next-generation sequencing (NGS) and the development of bioinformatics tools to interpret this information, blueberry breeding efforts are likely to benefit from these technologies.

With the exception of two much older RAPD-based maps (Rowland & Levi, 1994; Qu & Hancock, 1997), a map of diploid population [(*Vaccinium darrowii* Fla4B × *Vaccinium corymbosum* W85-20) F_1 #10 × *V. corymbosum* W85-23] is the only published genetic linkage map of blueberry. However, two other maps of a tetraploid blueberry and an interspecific blueberry × sparkleberry population, are currently under construction. Simple sequence repeat (SSR) and single nucleotide polymorphism (SNP) primer information from these genetic linkages maps, and one from cranberry, were used to position 43% of the current blueberry draft genome assembly. The draft assembly currently consists of over 9,000 scaffolds (greater than 2 K in length) with an N50 of 241 kb. The assembly was created through a combination of Roche 454 pyrosequencing, Illumina Hiseq and Sanger BAC-end sequencing. Multiple software programs were used to predict the genomic location of the markers and to provide alignment (and in some cases, orientation) of genomic scaffolding.

In addition to the alignment of scaffolding, new SSR markers were designed to ensure that the largest scaffolds and targeted genes in the anthocyanin biosynthesis pathway were positioned on the published linkage map. Sixty-six putative anthocyanin biosynthesis genes were identified from scaffolding and linked SSR markers were used to position several of these to the diploid map. This work has significantly improved the current blueberry genetic map, provided evidence for alignment between the current map and those in progress, and it has provided functionality to the maps by anchoring 43% of the genome (and tens of thousands of annotated genes) of blueberry. This research provides both an immediate resource to current blueberry researchers, and a roadmap for the completion and refinement of the sequenced blueberry genome.

Improved and Novel Resources for Blueberry Mapping and Genomics

by
Ying-Chen Lin

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Horticultural Science

Raleigh, North Carolina

2015

APPROVED BY:

_____	_____
Allan Brown, Committee Chair	Penelope Perkins-Veazie
_____	_____
Craig Yencho	Zhao-Bang Zeng
_____	
Jeannine L. Rowland	

BIOGRAPHY

Ying-Chen Lin is from Taiwan (R.O.C.) and grew up in Taipei city. A girl who is not good at taking care of plants but chose horticultural science as her major in National Taiwan University. As an undergraduate student, she participated in a summer research project in North Carolina State University and spent the summer in the Forestry Biotechnology Lab, which inspired her to study abroad and pursue molecular biology and biotechnology related graduate work. In 2012, she came to U.S.A. for her MS degree in Horticultural Science Department in NC State University. Currently she works on a blueberry genomic project in Dr. Allan Brown's lab in the Plants for Human Health Institute (PHHI), Kannapolis, NC.

ACKNOWLEDGMENTS

I would like to give my sincere gratitude to my advisor, Dr. Allan Brown, for introducing me to this project and for patiently guiding me throughout this research, and at this moment, staying up late at the last minute to correct my thesis. During my work with Dr. Brown, I had a chance to meet many blueberry breeders and scientists, who enlightened me further in my pursuit of knowledge in my blueberry genomic study.

I would like to thank other committee members, Dr. Penelope Perkins-Veazie, for offering favor and suggestions regardless of the academic or research aspect; Dr. Craig Yencho and Dr. Zhao-Bang Zeng for giving me advice and comments on this research and thesis; and finally, Dr. L. Jeannine Rowland, for generously supporting this project and providing DNA samples and data. Additional thanks go to Dr. Nahla Bassil, Dr. Susan McCallum, and Dr. James Olmstead for providing their linkage map data. I hope this thesis will be helpful for their research.

I'm also grateful to have Gad Yousef as a lab mentor, he has always encouraged me to be optimistic not matter good or bad things happen. Also, thanks to Robert Reid for taking care of the blueberry genome assembly and providing bioinformatic support for this project. Aswathy Thomas and Danielle Nonnemacher, without your full support and hard work, I wouldn't be able to finish on time. Also thanks to Ryan Park and Yoshinori Tanimura, they were the best labmates I've ever had. Thanks to Brooklyn Phillips and Elizabeth George, they did tremendously a great job in the lab helping with thousands of PCR and keeping the lab neat and clean. This research was supported by Plant Pathway Elucidation Project (P^2EP) and by a Specific Cooperative Agreement with the USDA-ARS. It was a great experience to be a part

of the P^2EP summer internship program, during these two summers for all the people and interns I worked with.

Finally, thanks to my parents and friends it would be really hard without their unconditional support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One Literature Review

Economic Importance of Blueberry

Blueberry (*Vaccinium* spp.) is an economically important small fruit crop worldwide, and ranks as the second most important berry crop in the United States after strawberry (USDA, 2015). The United States is the world's largest blueberry producer. In 2012, 564.4 million pounds of blueberries were harvested with an estimated value of over $850.9 million (NASS 2013). In 2013, 83% of cultivated blueberries were harvested in the USA from the states of Michigan, Georgia, New Jersey, Oregon, Washington and North Carolina (USDA 2013). Maine is the largest producer of unimproved/wild blueberry or lowbush blueberry (*V. angustifolium*). The U.S.A., Canada and Poland (first, second and third respectively) are the world's major blueberry producing countries and accounted for up to 80% of all blueberry production in 2012 (Brazelton, 2011; USDA-ERS report). In addition to increased worldwide blueberry production in recent years, consumer demand of blueberry also has increased around the world as blueberry has significantly increased in popularity; attributable in part to its reputation of possessing high levels of health-beneficial compounds.

As a native North American plant, blueberries have been consumed by indigenous people for thousands of years (Moerman, 1998). Native Americans not only gathered blueberries for food (fresh and preserved) but also used them for medicinal purpose, such as relieving fevers, headaches and persistent coughs (Trehane, 2004). However, despite this long involvement in dietary history, the domestication of blueberries did not occur until the early

20th century. Domestication of wild blueberries was initiated by Elizabeth White, a farmer's daughter from New Jersey, and Frederick Coville, a USDA botanist. After several years of collaboration, the first commercial blueberry cultivars were released in 1916 (Coville, 1921). Currently, many of these cultivars are still being used by public and private breeding programs around the world to develop superior F_1 hybrids with improved horticultural characteristics.

Blueberry Taxonomy

The genus *Vaccinium*, a member of the family *Ericaceae*, consists of approximately 450 species that can be grouped into ~30 sections (subgenera) (Stevens, 1970) (Table 1). Most *Vaccinium* species are distributed in the temperate and subtropical regions of the Northern Hemisphere, but a number can be found in regions as diverse as the Himalayas, New Guinea and South America (Luby et al., 2001). *Vaccinium* species are not found in Australia or Antarctica. Among *Vaccinium* berries, the most familiar are blueberries (*V. corymbosum*; *V. angustifolium*; *V. virgatum*), cranberries (*V. marcrocarpon*), lingonberries (*V. vitis-idaea*) and bilberry (*V. myrtillus*). As with blueberries, these other fruits are of recent domestication but they have also been traditionally consumed by indigenous peoples. *Vaccinium* species display variations in ploidy levels that include diploids (2n=24), tetraploids (4n=48) and hexaploids (6n=72). This phenomenon in *Vaccinium* is believed to be due to its autopolyploid in nature and results from the spontaneous doubling of the chromosomes or through the fusion of unreduced 2n gametes (Lyrene & Perry, 1982; Hokanson & Hancock, 1993; Costich et al., 1993; Qu et al., 1998). The triploid block is a problematic issue in blueberry hybridization.

Diploid plants crossed with tetraploid blueberries generally produce few viable offspring (Darrow et al., 1944). As with other species in the family *Ericaceae*, blueberry grows well on acidic, moist soils that are often not conducive for other plants (Song & Hancock, 2011).

Blueberries commercially grown in the United States belong to the section (or sub-genus) Cyanococcus of the genus *Vaccinium*, and can be categorized into three main varieties: highbush (*V. corymbosum* L. introgressed with other species) (4n=48), rabbiteye [*V. ashei* Reade (syn. *V. virgatum* Ait.)] (6n=72), and lowbush [*V. angustifolium* Ait. (4n=48) and *V. myrtilloides* (2n=24)] (Vander Kloet, 1988). Highbush and lowbush blueberry are the major commercial types, while rabbiteye is of regional importance in the southeastern United States (and increasingly, in the western United States). Highbush blueberries can be further classified as either northern or southern highbush depending on their chilling requirements and the degree of winter hardiness. They often have significant genomic introgression from other species (Rowland et al., 2011).

Differences in vernalization requirements and cold hardiness are evident between highbush and lowbush blueberry and determine their geographical distribution. Lowbush blueberry (generally unimproved clones), tolerate much colder conditions, have a longer vernalization period and are principally harvested in the state of Maine and in the provinces of Eastern and Central Canada (Strik & Yarborough, 2005). Highbush blueberry requires less chilling than lowbush and is more tolerant of high summer temperatures. Highbush varieties are grown in 37 American states, six Canadian provinces, as well as Australia, Chile, Argentina, New Zealand and several countries in Europe (Strik, 2005).

Southern highbush involves genomic introgression between *V. corymbosum* and one or more southern *Vaccinium* species (*V. darrowii*, *V. ellotii*, and *V. virgatum*). A single hybrid, US75, was the result of hybridization between the diploid *V. darrowii* selection Fla4B and tetraploid highbush cultivar 'Bluecrop' (*V. corymbosum*) and has become the primary source of low chilling requirement for southern highbush cultivars (Draper & Hancock, 2003). The *V. darrowii* parent Fla4B is an evergreen southern highbush blueberry that is more tolerant to higher summer temperature, thus, Fla4B is an important donor parent to several breeding programs that aim to extend the range of southern highbush blueberry production.

Rabbiteye blueberries, in addition to their traditional range in the Southeast, are currently grown in California, the Pacific Northwest and Chile, as they ripen later than highbush blueberry and are better suited to the long milder seasons in these regions. Rabbiteye have superior adaptation to heat, droughty, and soil with limited organic matter and have been used to introduce these traits into southern highbush. Hybrids of highbush and lowbush (half-highs) have been developed for the Upper Midwest (Minnesota and other states), and are more tolerant of severe winter weather than highbush. As the name implies, the average height of half-high blueberry is 0.5~1.0 meter, between that of lowbush and highbush. Although half-highs blueberries do not have a significant impact on the consumer market, they are widely used in gardens as ornamental plants (Rowland et al., 2011).

Molecular Markers and Linkage Mapping

Several types of molecular markers have been used in blueberry genome research (Semagn et al., 2006). Random Amplified Polymorphic DNA (RAPD) markers have been used to identify and evaluate genetic relationships among a limited number of blueberry cultivars (Levi & Rowland, 1997) and in the development of the first RAPD-based genetic linkage map (Levi & Rowland, 1994). However, RAPD markers are dominant in nature, which make them less informative in establishing linkage relationships with other markers and with linked quantitative trait loci (QTL). Reproducibility is also an issue as variation has been observed between laboratories that can be attributed to several factors including differences in equipment, reagents, and PCR reaction conditions (Penner et al., 1993; Jones et al., 1997; Schierwater & Ender, 1993).

Expressed sequence tags (ESTs) markers are short, single-pass nucleotide sequences (300~500 bp) that are developed from coding regions of expressed genes. These markers are often derived from cDNA libraries that represent hundreds to thousands of expressed DNA sequences under specific environmental or developmental conditions. A few hundred nucleotides of the 5' and/or 3' end of the expressed DNA are sequenced to create 5'-ESTs or 3'-ESTs. If these sequences vary in length due to insertions and deletion (indels) or other polymorphisms, they are easily visualized on agarose gels. The concept of ESTs was first proposed in 1970s, and term EST was coined in 1991 during work on human gene discovery research (Adams et al., 1991). Over the past 30 years, there has been an exponential growth in EST sequences deposited in public database including several thousand blueberry sequences

(Rowland et al., 2003a, 2003b; Boches et al., 2005; Bell et al., 2008; Rowland et al., 2012; Die & Rowland, 2013). ESTs are particularly useful in gene discovery as they represent the expressed portion of the gene sans introns. EST markers have the advantage of targeting desirable genes that are expressed in specific tissues or developmental stages and the cross-taxonomic conservation of these sequences may be higher among related species than non-expressed sequences (Boches et al., 2006a, 2006b; Ayeh, 2008). ESTs have been successively used in gene discovery, phylogenetic studies, gene structure identification, SNP characterization, genome map construction, and proteome analysis (Nagaraj & Ranganathan, 2007; Parkinson & Blaxter, 2009). However, many ESTs deposited in databases often have low sequence quality that increases the probabilities of errors. For blueberry genomic research, about 5,000 Sanger ESTs were initially generated from cDNA libraries created for cold acclimation research (Dhanaraj et al., 2004, 2007; Rowland et al., 2012, 2014). Expressed sequence tag-polymerase chain reaction (EST-PCRs) markers were developed from these blueberry cDNA libraries of floral buds undergoing cold acclimation and non-acclimation. EST-PCR markers were used to study cold tolerance and initial analysis of blueberry genetic relationships (Rowland et al., 2002, 2003b).

Genomic variations in plant DNA can be classified as single-nucleotide polymorphisms (SNPs), small insertion-deletion polymorphisms (indels), and copy number variation (CNV) (Żmieńko et al., 2014). Possible origins of these variations include double-stranded DNA recombination, replication slippage, mismatch of double strand break and repair, and retrotransposition (Wang et al., 2009). Copy number variation includes satellites, mini-

satellites, and microsatellites, which differ primarily by the copy size (Varshney et al., 2005). CNV of the shortest sizes are microsatellites or Simple Sequence Repeats (SSRs) and can be repeated as many as 100 times (Litt & Luty, 1989). There is no fixed definition of the base pair length of these short sequences, and there is considerable difference of opinion among researchers (1~6 bp or 2~8 bp) (Schlötterer, 1998; Armour et al., 1999). In this thesis, we will adopt a strict definition of 2 to 6 bp repeats as suggested by Chambers and MacAvoy (2000). SSRs commonly occur throughout the nuclear genome of eukaryotes in both coding and non-coding regions, but those in coding sequences show less polymorphisms (Tautz & Renz, 1984; Lagercrantz et al., 1993; Powell et al., 1996). Bian et al. (2014) also confirmed that in blueberry the frequency of dinucleotide SSRs identified were highest in 5' UTR (370.0 per Mb), followed by 3' UTR (143.6 per Mb), while the coding sequences contained fewer SSRs that were primarily trinucleotides and hexanucleotides. In addition, coding sequences were shown to have the shortest average length of SSRs and 5' UTR had the longest. SSR markers are designed by creating forward and reverse primers (~20 bp) that correspond to unique, conserved DNA sequence flanking the target repeats. SSRs fragments are then amplified using multiple cycles of thermo-cycling known as polymerase chain reaction (PCR) with an annealing temperature that is determined by primer melting temperature (Tm), primer quality, and length of PCR product. The general rule is to use annealing temperature that is 5° lower than the primer Tm. There are a number of benefits associated with SSR markers. The analysis generally requires relatively small amounts of DNA. The products of PCR (usually 100 to 500 bp long) can be viewed as co-dominant on agarose gels or with a high throughput platform such as the ABI DNA Analyzer 3730xl (if florescent labels have been incorporated). In diploid

genomes, SSR markers are easily scored as co-dominant but they can also be scored as dominant markers in polyploid genomes. In addition, SSRs are locus-specific, highly polymorphic, reproducible, and more adaptable to automation than previously used markers. SSRs have been widely applied to parental analysis, population genetic structure and relationship studies.

It has been suggested that SSR markers are species specific or that their use is restricted to closely related species (Jarne and Lagoda, 1996), but several *Vaccinium* researchers have been successful in amplifying homologous sequences across *Vaccinium* sections (sub-genera) (Georgi et al., 2013; Bassil et al., 2008; Rowland et al., 2012; Schlautman et al., 2015). An additional concern with SSRs is that the development of these markers is laborious and time-consuming. Before the advent of large-scale genomic sequencing projects, SSRs were isolated from libraries that were initially enriched with probes corresponding to various repeat sequences (Rassmann et al., 1991; Zane et al., 2002). Another concern with SSRs involves the scoring of null allele. Null alleles are alleles that fail to amplify to detectable levels in PCR assay (Senan et al., 2014) and arise from mutation in annealing sites due to sequence divergence, preferential amplification of short-length alleles, and PCR failure due to inconsistent DNA quality or low template quantity (Chapuis & Estoup, 2007; Dakin & Avise, 2004). The presence of null alleles complicates the analysis of allele calls and genotype frequency, which may lead to underestimation of heterozygosity called size homoplasy, which results in scoring two different alleles as the same one (Kumar et al., 2009; Pompanon et al., 2005).

The advent of high throughput transcriptomic and genomic sequencing projects has provided an inexpensive alternative method for discovery and mining of SSRs on a large-scale basis. SSRs are routinely mined from large EST transcriptome databases (Gupta et al., 2003; Bhat et al., 2005) and provide a number of benefits and disadvantages to genomic SSRs. EST-SSRs can be generated relatively inexpensively when compared to genomic SSRs (Thiel et al., 2003), and have been found to be more conserved and transferable among related species (Coulibaly et al., 2005; Varshnet et al., 2005; Pashley et al., 2006). Another advantage of EST-SSR markers is that they have a lower probability of generating null alleles (Leigh et al., 2003; Rungis et al., 2004). The drawbacks of EST-SSRs is that they can be restricted by the availability of high quality sequences, and are often less polymorphic than genomic SSRs due to inherent sequence conservation of the transcribed regions (Cho et al., 2000; Gupta et al., 2003; Chabane et al., 2005). The first two sets of blueberry EST-SSR markers were developed from microsatellite-enriched genomic libraries of *V. corymbosum* cv. Bluecrop, and two EST libraries that were derived from cold acclimation and non-acclimation floral buds (Boches et al., 2005). Blueberry EST-SSRs have been shown to be transferable across species in the Cyanococcus section, and are also useful in sections that include such as cranberry species (*V. marcrocarpon*, *V. oxycoccos*) (Bassil & Hummer, 2009).

The development of *de novo* draft sequences of many plant species has greatly accelerated the identification of genomic SSRs. Bian et al. (2014) identified 43,594 SSRs (2-6 bp repetitive elements) from the genomic scaffolding of the diploid blueberry accession W85-20 (*V. corymbosum*) using a PERL5 script MIcroSAtellite (MISA). The successful

amplification of these predicted SSRs by multiple labs has ranged from 80 to 87% (Allan Brown, personal communication). A subset of these genomic blueberry SSRs has been used to investigate the genetic diversity and population structure among multiple species of cultivated blueberry (Bian et al., 2014).

Single Nucleotide Polymorphisms (SNPs) have gained immense attention with the increasing number of next-generation sequencing (NGS) and genotyping by sequencing (GBS) projects in the past few years. SNPs reflect polymorphism at a single nucleotide between individuals of a population or paired chromosomes with a frequency greater than 1% in a given population. SNPs are the most abundant potential molecular markers and occur in both coding and non-coding regions. SNPs are estimated to occur at a rate of about one per 300 bp (Seeb et al., 2001), but in many plant species they are often more abundant (Ayeh, 2008). SNP markers are extensively used in phylogenic analysis, marker-assisted selection, genetic mapping of quantitative trait loci (QTL), bulked segregation analysis, genome selection, as well as genome-wide association studies (GWAS) (Seeb et al., 2011; Kumar et al., 2012).

There are several ways to identify SNPs. For known gene sequences, fragments can be amplified through PCR and sequenced on a small-scale. This can often be effectively applied to non-model organisms without reference genomes or with limited sequence information. Another method is large-scale, cost-efficient, and rapid NGS-based techniques such as GBS. NGS techniques have been applied in genetic marker discovery and genotyping for various organisms. With NGS, hundreds of thousands of SNPs can be generated in a relatively short period of time, which can remedy the drawbacks of time-consuming and laborious analysis

associated with other methods (Gupta et al., 2001; Seeb et al., 2001; Davey et al., 2011; Kumar et al., 2012).

Often, it is necessary to reduce the complexity of polyploid plant genomes that have large amounts of repetitive DNA through methods such as reduced representation sequencing and complexity reduction of polymorphic sequences (Garvin et al., 2010). Davey et al. (2001) summarized methods to improve SNP calling using NGS techniques into three categories: reduced representation sequencing (RRS) or reduced representation libraries (RRL) and complexity reduction of polymorphic sequences (CRoPS), restriction site associated DNA (RAD)-seq, and low coverage genotyping such as GBS. RRS, CRoPs, and RAD-seq reduce genome complexity by decreasing the sequenced proportion of a genome to targeted regions, and increasing the sequencing coverage in these targeted regions for detailed analysis.

Low coverage sequencing and genotyping such as GBS and multiplexed shotgun genotyping (MSG) produce large numbers of markers in limited numbers of reactions. They produce large numbers of markers, and work extremely well with populations that have defined parental genotypes. Microsatellites are widely used for mapping and other applications in non-model crops with limited genomic resources due to their simplicity of use, and highly polymorphic co-dominant nature; however, there are still some shortcomings when compared to SNPs. SNPs are extremely abundant, highly distributed throughout the whole genome and are particularly well adapted to automated genotyping when sufficient genomic resources are available. An example of this is the recent 60,000 SNP *Brassica napus* array used in our laboratory to produce highly saturated genetic linkage maps of *Brassica oleracea* (Brown et.

al. 2014; Brown et. al. 2015). The linkage map has came out with an estimate of about 96% coverage of the current reference genomic sequence of *B. oleracea* accession 'TO1000', including 547 markers that covered 948.1 cM across nine chromosomes.

Automated genotyping produces lower error rates when compared with allele calling of microsatellites made by even the best technicians (Ball et al., 2010, Seeb et al., 2011; DeFaveri et al., 2013). Low error rates are essential to constructing accurate, high resolution linkage maps; however, unlike multi-allelic microsatellites, SNPs can be limited in some applications (such as estimations of genetic diversity) because they provide less information per locus. Ball et al. (2010) and Seeb et al. (2011) suggested that instead of using only one type of markers, combinations of microsatellites and SNPs might produce better results: microsatellite markers can act as anchors in a linkage map, and large numbers of SNPs can then be fit into the map framework to provide higher density.

Next-Generation Sequencing and Plant Genome Study

The strategies of sequencing can be categorized into two groups: hierarchical sequencing (HS, or called clone-by-clone) and whole genome shotgun (WGS) sequencing. In the hierarchical method, a low-resolution physical map of the target genome is required prior to the sequencing step. To create a sequence library, the genome is broke down into many overlapping sequence fragments, that are then inserted into vectors such as bacterial artificial chromosome (BAC), yeast artificial chromosome (YAC) or P1-derived artificial chromosome (PAC) to be sequenced in parallel. These high-capacity vectors can take up insert fragments

with average size of several hundred kb in length, an average of 300 kb insert can be taken by BACs, while YACs can take up to a megabase of pair fragments (Green, 2001). Sequence-tagged site (STS), or BAC-end sequencing, acts as a tag on one or both ends to search neighboring clones. According to STS, order and orientation of vectors can be determined. To sequence the inserted DNA fragment, the fragment is purified and sheared randomly into smaller pieces. The overlapped sequences of these small pieces are then viewed as anchors for assembly work (Green, 1997; Green, 2001; Tammi, 2003).

The HS method provides a higher quality of sequencing reads and is suitable for large and complex genomes, but it is cost- and time-consuming. The whole genome shotgun method, in contrast, cleaves the genome sequences randomly, and the fragments then go through the sequencing process. Assembly is based on overlapping regions. This method relies more heavily on computational assembly techniques. To overcome issues associated with repetitive sequences found in many plants, clones with large inserts are used and/or nested primers are used to sequence gaps from both ends, called pairwise end sequencing, to avoid sequence gaps and provide higher quality information (Green, 2001; Brown, 2002). This method is widely applied in next-generation sequencing technologies.

The first DNA sequencing technique, Sanger sequencing, was described by Sanger et al. in 1977, and is now considered as first-generation sequencing. The concept is based on the chain-termination method, where labeled di-deoxynucleotidetriphosphates (ddNTPs) are incorporated into DNA synthesis by DNA polymerase. DNA synthesis is terminated at incorporation as these ddNTPs lack a 3'-OH preventing elongation, and results in various

lengths of terminated reads. The order of the four nucleotides is visualized and scored from a gel. Automated Sanger sequencing usescapillary electrophoresis, which enables larger sample numbers and is faster. In automated methods, tagged fluorescent labels can be either attached to primers or ddNTPs, and tagged nucleotides are then detected by laser camera for nucleotide calling. With this higher throughput method, the human genome was completed in 2001 (Metzker, 2005, 2010; Liu et al., 2012). Sanger sequencing can produce reads with an average length of 800 bp (up to 1000 bp) with 99.9% accuracy. Before the development of next-generation sequencing, Sanger sequencing was the dominant sequencing technique for about 25 years. However, the high cost and time-consuming features are overcome by next-generation sequencing techniques.

Next-generation sequencing (NGS) emerged in the late 1990s. The most well-known commercial NGS platforms include: pyrosequencing (454 Life Sciences, now owned by Roche Diagnostics), sequencing by synthesis (Illumina), sequencing by ligation (SOLiD, Applied Biosystems by Life Technologies), Ion semiconductor (Ion Torrent) sequencing, and most recently, single-molecule real-time (SMRT) sequencing (Pacific Bioscience) (Table 2). NGS is renowned for its massive parallel sequencing, high-throughput, and low cost features, and has started a new phase of genome sequencing. Pyrosequencing is based on the sequencing by synthesis principle, nucleotides are called by the real-time detection of pyrophosphate (PPi) release from DNA synthesis. Pyrosequencing is able to produce read lengths of 700 bp on average (up to 1,000 bp) with the most recent platform, GS FLX Titanium. However, the error rate of this platform is relatively high, particularly when it comes to a single repeated

nucleotide (homopolymer). Illumina, SOLiD, and Ion Torrent platforms, with the capability of generating higher throughput with less cost, were released later and are still crucial tools for genomic research. These sequencing platforms all generate short reads in large numbers. Illumina, for esxample, generates 30 million reads of ~100 bp; and SOLiD generates 100 millions reads of 50-75 bp in length; Ion Torrent can generate about 6 million reads with an average length of 345 bp. The latest platform, PacBio SMRT sequencing, can produce reads up to 10~20 kb, but the price is higher (Metzker, 2005, 2010; Mardis, 2008; Liu et al., 2012; van Dijk et al., 2014). One of the results of NGS technologies is the high demand for bioinformatics tools for data handling. Various novel alignment algorithms that calculate the exact location for each read have been developed to solve assembly issues, but in complex plant genomes, assembly is still a challenge. Another issue is the distribution of reads, which relates directly to the coverage of a genome. The random DNA shearing process may result in unequal frequency and coverage due to differences in shearing ability in repetitive regions, retrotransposons, or chromatin stage (Ahn, 2011).

The first complete plant genome, *Arabidopsis thaliana,* was published in 2000. A relative of the mustard family, *Aradibopsis* became the model organism for plant genomic studies due to its relatively small nuclear genome, short generation time and large number of offspring that can be produced (Arabidopsis Genome Initiative, 2000). Taking advantage of BAC, PAC and transformation-competent artificial chromosome (TAC) libraries, researchers were able to put the pieces of sequences together. Shortly after, the draft genome of rice (*Oryza sativa*) was released in 2002, which also utilized clone-by-clone sequencing and whole-

genome shot gun sequencing method (Goff et al., 2002; Yu et al., 2002). The development of high-throughput sequencing technologies facilitated various plant genome sequencing projects including poplar (*Populus trichocarpa*; Tuskan et al., 2006), maize (*Zea mays*; Schnable et al., 2009), papaya (*Carica papaya*; Ming et al., 2008), and more. The advent of next-generation sequencing has provided more handy sequencing methods that are less costly tan older methods like Sanger sequencing and Maxam-Gilbert sequencing, and are able to conduct parallel sequencing and generate millions or billions of sequence data within a relatively short period of time. However, large genome size, high copy number, polyploidy, and presence of transposable elements have been challenges for plant genome sequencing. Also, in most eukaryotes, centromeres are difficult to sequence because of highly repetitive elements such as retroelements, transposons, and microsatellites reside in these regions, thus, sequencing information is often lacking of. Therefore, the study of plant genomes still needs enormous experimental and computational efforts to improve the-state-of-art technologies.

Blueberry Breeding in the Advent of Biotechnology

Traditional blueberry breeding is complicated by the polyploid nature of the plant, a relatively long juvenile period (~3 years), and by severe inbreeding depression that prevents the development of inbred lines (Die & Rowland, 2013). As with other plants, phenotypic selection of blueberry is complicated by incomplete expression of traits through a combination of environmental factors such as climate, soil type and abiotic and biotic stress that cause differential morphologies (Collard et al. 2005; Winter and Kahl, 1995). Blueberry breeding

(and blueberry research in general) would benefit substantially from the development of genomic resources and the use of these resources in marker-assisted selection (MAS) programs, particularly in regards to early generational plant selection. One of the advantages of MAS is that selection based on molecular markers (observable variation at DNA level) is not affected by the environment or by the developmental stage of plants, and is transferrable from one generation to the next. Markers are a form of indirect selection for linked genes of interest and they provide confident criteria for screening large numbers of progeny (Collard et al. 2005; Winter and Kahl, 1995). Screening for markers linked to desirable traits in juvenile blueberries, could lead to a reduction in the cost of labor and field space, and shorten the time required to develop new varieties (Collard & Mackill, 2008).

Traditional blueberry breeding programs have focused on improving desirable fruit characteristics, such as fruit size and flavor, expanding the production range and harvest season, providing insect and disease resistance, and broadening the adaptation of blueberries to less acidic soils (Rowland et al., 2011). Molecular techniques, including tissue culture screening for tolerance to higher pH, in-vitro chromosome doubling, and genetic transformation have been successfully incorporated into blueberry improvement programs (Finn et al., 1993; Lyrene & Perry, 1982; Song & Sink, 2004; Bassil, 2012), but the scarcity of genomic resources has been a hindrance. Until recently, only limited numbers of molecular markers have been available for genotyping of populations (Anjos et al., 2008; Brevis et al., 2008; McCallum et al., 2012), constructing molecular maps, and examining the genetic relationships and distance among various blueberry accessions (Rowland & Levi, 1994; Aruna et al., 1993, 1995; Levi,

& Rowland, 1997; Burgher et al., 2002; Bell et al., 2008; Debnath, 2009). Molecular markers reveal heritable differences in the DNA sequence that are referred to as polymorphisms. These polymorphisms can be identified by hybridization with radiolabeled probes, through polymerase chain reactions, or with arrays or assays designed to detect single base pair changes. When polymorphisms revealed by molecular markers occur within (or are closely linked to) genes that affect desirable traits, selection can be made based on the marker which is often more advantageous or easier than selection based on the trait itself.

The haploid blueberry genome was estimated to be 500-608 Mb, which is about five times larger than the 125-Mb Arabidopsis thaliana genome (Costich et al., 1993). Different species of blueberry all have the same basic number of chromosome, n=12, with natural polyploid complements ranging from the diploid to the hexaploid. Given the diversity among the blueberry sections and the difficulty in making crosses among the sections, the size of the genome is likely to vary slightly but this has not been systematically studied.

Diploid Blueberry Linkage Map

Rowland et al. (2014) recently published the most saturated diploid blueberry genetic linkage map to date using a combination of RAPD, EST-PCRs and genomic and EST- SSR markers. The interspecific diploid population [(*Vaccinium darrowii* Fla4B × *Vaccinium corymbosum* W85-20) F_1 #10 x *V. corymbosum* W85-23] was constructed specifically to identify genes segregating for cold hardiness and chilling requirement. The population is pseudo-backcross population. Instead of using W85-20 in the backcross, a similar *V.*

corymbosum line W85-23 was crossed to the F_1. Genetic analysis of W85-20 and W85-23 have shown that these lines are genetically similar but not identical (Bian et al., 2014). The *V. darrowii* selection Fla4B is an evergreen native plant of the southeastern United States, and has provided an important source of low chilling requirement genes for southern highbush blueberries. W85-20 and W85-23 are deciduous selections from New Jersey, need higher chilling requirements, and are cold hardy.

The current diploid blueberry map consists of 265 markers across 12 linkage groups that correspond to the base chromosome number of blueberry. The total length of the map is 1,740 cM (Haldane mapping function) with the average distance between markers of 7.2 cM, and the estimated map coverage is 89.9%. The mapped markers included SSRs, EST-PCRs, RAPDs and a limited number of single nucleotide polymorphisms (SNPs). EST-PCR markers designed initially with a focus on cold-responsive genes expressed in flower buds, than later also included genes expressed in fruit. PCR primer sequences were designed from assembled transcriptome data generated by Sanger sequencing (Dhanaraj et al., 2004, 2007) and 454 transcriptome sequencing (Rowland et al., 2012). SSR markers were from several sources including 454-sequenced ESTs, Sanger sequenced ESTs in EST database, and genomic SSRs (designed KAN) obtained from the initial draft of the whole genome assembly of diploid blueberry selection W85-20, and from EST-SSRs obtained from New Zealand Institute for Plant and Food Research Ltd (Auckland, New Zealand). In the published genetic map, two linkage groups were comparatively short with only five to eight markers. The lack of markers in these regions may be due to chance, or reduced rates of polymorphism or recombination in

these regions. Screening additional genomic SSRs (with a higher predicted rate of polymorphism over EST-SSRs) could potentially remedy this situation. One of the objectives of the current study was to provide additional markers to further saturate this map and improve the efficacy of this important resource. Two additional blueberry molecular maps, of a tetraploid and interspecific hybrid population, are currently under construction (utilizing different sets of markers) and a single map of cranberry has been published.

Tetraploid Blueberry Linkage Map

The first tetraploid blueberry map is currently being constructed in a joint venture that includes researchers from the USDA (Beltsville, MD and Corvalis, OR), Michigan State University, The University of Florida, The New Zealand Institute for Plant & Food Research Ltd., The James Hutton Institute (Scotland) and North Carolina State University (Brown, personal communication). The population associated with the map was created by crossing the northern highbush cultivar 'Draper' with the southern highbush cultivar 'Jewel' (tetraploid *V. corymbosum*), at Michigan State University. The cross created an initial population of 105 F_1 progenies that has since been expanded to a population of ~200, that segregates for key phenotypic traits such as fruit size, fruit shelf life, and other important horticultural characteristics.

Two tetraploid linkage maps are being created from markers that segregate from each parent ('Drape'r and 'Jewel') as well as a consensus map of both parents. With the aid of Illumina HiSeq2000 and genotyping-by-sequencing (GBS), 109,044 SNPs were identified and

after trimming sequences and removing those with missing values, 17,846 SNPs were used for further linkage analysis. SNPs that showed polymorphism in either parent were selected for screening, and 1,288 exhibited a 1:1 ratio with the segregation of one allele. These SNPs, along with other markers such as SSRs and AFLPs, were then incorporated into what will be referred to as 'Draper' and 'Jewel' linkage maps (or jointly as the tetraploid blueberry map).

The authors of this work (which include A. Brown from North Carolina State University) have provided us with a working copy of this map, which includes 1,251 total markers (1056 SNPs, 53 co-dominant SSR markers and 142 dominant SSR markers) grouped into 12 linkage groups and analyzed with JoinMap 4.0 with a LOD score of 4.0. The linkage maps cover 1,613 cM in length and the average number of markers in each group was 60.9. To compare with the diploid blueberry linkage map that was published in 2014 by Rowland et al., 15 markers from the diploid map were screened and nine were mapped in the tetraploid Draper linkage map. Linkage group (LG) 1 in the tetraploid had two markers, CA344 and VCBC09467, that were placed in LG 5 in diploid blueberry; another LG pair was LG 8 in the tetraploid and LG 7 in the diploid that had two markers in common (personal communication). This number of markers, however is insufficient to determine unambiguously the relationships between the diploid and the tetraploid maps. It should be noted that at the time of this thesis, the linkage map is still under review and the information reported here could change before publication.

The Interspecific Hybrid Linkage Map

Intersectional crosses in *Vaccinium* are difficult but partially fertile hybrids do occur. One example includes crosses between tetraploid blueberry (*V. corymbosum*) and bog berry (*V. uliginosum*, or northern bilberry). Hiirsalmi (1977) made crosses between the tetraploid F_1 hybrids of *V. uliginosum* × *V. corymbosum* and *V. corymbosum*. The progenies turned out to be viable and fertile with considerable variation in the BC_1 progeny in several characteristics such as berry size, color and flavor.

Sparkleberry (*Vaccinium arboreum* Marshall), from the section Batodendron, has been used extensively in the University of Florida's breeding program. Sparkleberry is a diploid (2n=24) that can grow up to 5 m tall, and is native to the southeastern United States (Camp, 1945; Vander, 1988). Although the fruits are inedible, several traits of sparkleberry are favorable to blueberry breeding programs. Compared to blueberry, sparkleberry is more drought and heat tolerant, and can grow on up-land soils with low organic matter and a pH above 6.0. Sparkleberry also flowers and ripens later than highbush blueberry, which could prevent freeze damage in blueberry and allow for the development of late-ripening blueberry cultivars.

Sparkleberry has been particularly important to the Florida highbush blueberry breeding program, as they hope to broaden the adaptation of southern highbush blueberry. Direct introgression of genes from sparkleberry to tetraploid blueberry is difficult due to a strong triploid block in *Vaccinium*. *V. darrowii* is often used as a bridge species to introduce

genes of sparkleberry to tetraploid highbush blueberry. Sparkleberry seeds have also been treated with aqueous colchicine (0.1%~0.2%) to induce chromosome doubling to tetraploid level (Lyrene, 2011; Lyrene & Olmstead, 2012). At the University of Florida, the first cross between sparkleberry and blueberry was made in 1981 using evergreen *V. darrowii* as a bridge (Lyrene, 1991). Crosses were easy to make and hybrids were vigorous, but low in fertility. These F1 hybrids were then able to open-pollinate with other blueberry species at the research farm in Gainesville. The progenies from open-pollination are variable in morphology and fertility. The fertile seedlings (termed MIK, mother is known) are tetraploid or near tetraploid, which are believed to originate from unreduced megagametes of F_1 hybrids and other tetraploid southern highbush (Lyrene & Brooks, 1996). The vigorous progenies from MIK backcrossed with southern highbush cultivars had great variation in fertility, fruit ripening time, and fruit size (only 20~30% of the seedlings met cultivar standards). Plants were more evergreen than conventional southern highbush cultivars (Lyrene, 1997). Selected high-quality progenies were backcrossed with southern highbush cultivars again for two selection cycles, and these seedlings were incorporated into southern highbush breeding germplasm at University of Florida (Olmstead et al., 2013). These seedlings were maintained in the field and flowered heavily each year. Some of these plants never set fruit, while some produced small fruits annually (Brooks & Lyrene, 1998a, 1998b).

An interspecific population of 342 individuals was created from cross of 'FL08-467' (*V. corymbosum* × *V. arboreum*) and 'Southern Belle' (*V. corymbosum*). 'FL08-467' was generated from a cross of the highbush cultivar 'Primadonna' and 'FL06-753' (*V. arboretum*).

To generate 'FL06-753', seeds of the sparkleberry were soaked in colchicine, and individuals that exhibited tetraploid-like morphology were selected (James Olmstead, personal communication). Currently, the marker data of the mapping population includes SNPs generated from GBS and SSRs. The linkage map for the parent 'FL08-467' consists of 12 linkage groups with 332 markers in total. As this research is still in progress, the intact, detailed map will be published later and the content might be slightly different from the information used in this publication.

Cranberry Linkage Map

The large-fruited or American cranberry [*Vaccinium marcrocarpon*, section Oxycoccus (Hill)] is a diploid (2n=2x=24) perennial species, with a haploid genome size of 570~608 Mbp (Costich et al., 1993). Unlike blueberry, cranberry is self-fertile and provides advantages in creating inbred lines. Belonging to the same genus as blueberry, cranberry shares numerous genetic characteristics and also provides abundant potential health benefits. The close relationship between these two crops allows for an important comparison model in terms of genetic mapping, sequencing and gene regulation.

Cranberry next-generation sequencing data has been generated from an inbred clone using the ABI short read [sequencing by oligonucleotide ligation and detection (SOLiD) mate-paired] sequencing platform in 2012 (Georgi et al., 2012). After assembly, the total estimated length of the sequence was 566.7 Mbp, and the assembly contained 441,159 contigs in 68,496 scaffolds larger than 300 bp, with N50 scaffold length of 26,335 bp. From this de novo

assembled sequence data, the first set of cranberry specific SSR markers were developed and used in genetic mapping. In addition to genome sequence data, the sequence data of transcriptome was completed by Polashock et al. (2014) via Illumina GAIIx sequencing, which provided further information to investigate the cranberry genome.

Georgi et al. (2013) published the first genetic map of cranberry. The linkage map consisted of 138 markers that mapped into 14 linkage groups with a total distance of 879.9 cM. These markers include, blueberry-derived SSR markers, cranberry SSR markers that were developed from either assembled SOLiD or Illumina GAIIx sequencing data mined for large scaffolds, putative conserved orthologous set (COS) sequences, defense-related or flavonoid biosynthesis pathway genes, and cranberry-derived sequence-characterized amplified region (SCAR) markers. Four mapping populations (CNJ98-153, CNJ98-154, CNJ98-164, and CNJ97-86) were used in creating the map, and these populations were investigated for field fruit-rot resistance.

Diploid Blueberry Genome Assembly

The blueberry genomic assembly was initiated by NCSU in 2009 and utilized sequence reads from multiple sequencing platforms, including 454-pyrosequencing, Illumina GA and HiSeq. W85-20 was chosen as the target of sequencing because it is a diploid and a parent of the only (at the time) blueberry population with an existing genetic linkage map. Assembly of the sequence was conducted utilizing multiple software programs including Newbler, GARM (Genome Assembler, Reconcilation and Merging), Masurca (Maryland Super Read Cabog

Assembler), and SSPACE (SSAKE-based Scaffolding of Pre-Assembled Contigs after Extension). The long sequences generated from pyrosequencing laid the foundation of assembly, while sequences of Illumina were fragmented but able to be aligned to some longer pieces or fit into gaps. The previous reported assembly, was generated from Newbler assembly, and consisted of 13,757 sequences with an N50 of 145 kb.

Recently, two BAC libraries of W85-20, were constructed by Amplicon Express (Pullman, WA) using two restriction enzymes (Table 3). The Hind*III* BAC library (HVC) and BamH*I* BAC libraries (BVC) have 39,168 clones in total (18,432 in HVAC and 20,736 in BVC), with 163 clones lacking an insert. The average insert size of HVC and BVC were 120 kb and 110 kb, respectively, based on the manufacturer's report. BAC-end sequencing (BES) of 10,000 clones (5,000 from each library) was also conducted by Amplicon Express. After trimming non-paired BES, the average insert size was 90 kb based on the alignment results (Table 4). The combined BAC libraries represent as 5.49 fold coverage of the blueberry genome of 608 Mbp (Costich et al., 1993).

The resulting assembly of NGS and BES using SSPACE resulted in improvements in all parameters of the assembly, increased confidence against misalignments of scaffolds and resulted in a greater number of larger scaffolds with the maximum length of 1,796,319 bp (Table 5). The N50 of the assembly increased from 145 kb to 241 kb, and the coverage was estimated to be more than 150 fold (Allan Brown, personal communication). The average length of the new assembly was significantly lower using the Newbler assembly – only 4627 bp long. The reason for the low number was that with the Newbler assembler, sequences less

than 2 kb were trimmed while SSPACE kept all the sequences. Multiple annotations have been run on the blueberry assembly included Augustus, GeneMark.hmm, Glimmer, Cufflink, Stringtie, Interproscan, and Maker (Allan Brown, personal communication; Gupta et al. 2015). The annotation results were deposited in the Integrated Genome Browser (IGB) developed by Dr. Ann Loraine at UNC-Charlotte.

Anthocyanin Biosynthesis in Blueberry

Blueberry has gained considerable attention in recent years due to the reported health benefits associated with a unique phytochemical profile that includes anthocyanins, phenolic acids, procyanidins, flavonols, stilbeles and other related compounds. Blueberry contains amongst the highest concentration of anthocyanins in commonly consumed fruits and vegetables (Prior et al., 1998). Unlike other fruits and vegetables, blueberry also accumulates significant concentrations of five of the six-anthocyanin aglycones commonly found in foods (malavidin, dephinidin, cyanidin, peonidin, and petunidin) (We et al., 2006; Prior & We, 2006). The only one of the six-anthocyanin aglycones not found in blueberry is the least common one found in plants, pelargonidin. In a recent analysis by North Carolina State University of over 1300 wild and cultivated blueberry commercial cultivars and advanced breeding selections, the total anthocyanin concentration in blueberry has been observed to rage from 74 to 537 mg/100 g in replicated analyses over multiple years (Yousef et al., 2014). Generally, it has been observed that fruits of lowbush blueberry contain higher levels of these compounds than highbush or rabbiteye (Moyer et al., 2002). However, the measurement of anthocyanin content

varies between studies due to extraction methods, analysis and by location. Also, due to differences in vernalization and cold hardiness it is difficult to grow lowbush in the same environment as rabbiteye so these comparisons are often done in different environments. Anthocyanin concentration in rabbiteye, in general tends to be significantly higher than most southern highbush blueberry cultivars grown in the same environment but there are cultivar-specific exceptions to this (Yousef et al., 2014).

In addition to anthocyanins, blueberry contains additional phenolic and polyphenolic compounds such as proanthocyanidins, flavonols, stilbenes and others. Anthocyanin and proanthocyanidin concentrations are particularly high in blueberry when compared to other fruits and vegetables (Jaakola et al., 2002; Connor et al., 2002; We et al., 2006; Kalt et al., 2008). The concentrations of these compounds depend on the species, cultivar, fruit maturity, and postharvest processing methods (Lee et al., 2004). While concerns still exist as to the bioavailability of many of these compounds, mounting evidence suggests links between the consumption of these compounds and reduced incidences of diabetes, cognitive function disorders associated with aging, cancer, heart disease, and impaired night vision (Cho et al., 2004; Kalt et al., 2007; Zafra-Stone et al., 2007; He & Guisti, 2010). In many plants, anthocyanins are the primary source of red and purple color. In the petals, the vivid color attracts pollinators and in the fruit, it plays an important role in seed dispersal. In addition, anthocyanins may function in the repellence of herbivores and parasites and prevent damage from UV-B irradiation (Holton & Cornish, 1995; Willson & Whelan, 1990; Solovchenko & Schmitz- Eiberger, 2003).

The anthocyanin and flavonoid biosynthetic pathway has been well established in a number of model crops including maize (*Zea mays*), snapdragon (*Antirrhinum majus*), petunia (*Petunia hybrida*), and grape (*Vitis vinifera*). A large number of anthocyanin and flavonoid structural and regulatory genes have been identified in multiple plant species. In petunia, over 35 genes have been described that affect flower color (Wiering & De Vlaming, 1984; Tanaka & Ohmiya, 2008; Koes et al., 2005). The study of anthocyanin biosynthesis in flowers has had a great impact on floriculture research, particularly in the generation of new or modified colors (Gould et al., 2008). The biosynthetic pathway of anthocyanins involves both a generalized flavonoid pathway and specific genes associated with anthocyanin production, regulation and modification (He, et al., 2010). Figure 2 provides a generalized depiction of these pathways. Anthocyanin substrates initially originate from the Calvin cycle and then proceed through the pentose phosphate pathway, before feeding into the shikimate pathway, which ultimately produces phenylalanine. The first committed steps of flavonoid biosynthesis include PAL (phenyalanine ammonia lyase), cinnamate 4-hydroxylase (C4H) and 4-coumarate-CoA: ligase (4CL). These three genes convert phenylalanine into 4-coumaroyl-CoA. Substrates are then catalyzed by CHS (chalcone synthase) and CHI (chalcone isomerase) to produce naringenin. Naringenin is converted to dihydrokaempferol (DHK) by flavanone 3-hydroxylase (F3H). DHK can be hydroxylated by flavonoid 3'-hydroxylase (F3'H) to produce dihydroquercetin (DHQ), or converted to dihydromyricetin (DHM) by flavonoid 3'5'-hydroxylase (F3'5'H), or it can be reduced to leuocopolargonidin by DFR (dihydroflavonol 4-reductase). DHK and DHQ can either be further reduced to produce kaempferol and myricetin by FLS (flavonol synthase), or reduced by DFR and form leucoanthocyanidinds. The next enzyme involved is ANS

(anthocyanidin synthase). ANS (sometimes referred as LDOX or leucoanthocyanidin dioxygenase) reduced leucoanthocyanidinds to unstable anthocyanidins such as cyanidin and delphinidin. Anhocyanidins are then rapidly catalyzed by UFGT (UPD glucose: flavonoid 3-0-glucosyltransferas) to form stable anthocyanins (Holton & Cornish, 1995; Springob et al., 2003; He, et al., 2010). Another enzyme, anthocyanin permease (ANP), is not involved in anthocyanin biosynthesis but is important in anthocyanin accumulation. The function of ANP is to transport synthesized anthocyanin into the vacuole, where water-soluble pigments are stored (Gonzali et al., 2009; Aza-González et al., 2013).

Two broad categories of genes are involved in anthocyanin biosynthesis: structural genes that directly synthesize intermediate compounds, and regulatory genes that influence the transcription of the structural genes. In plants, transcription factors (TFs) that commonly influence biosynthesis include members of three families: the R2R3 Myb and Myc, basic helix-loop-helix (bHLH), and WD-repeat proteins (Lepiniec et al., 2006). Myb transcription factors are of particular importance in the biosynthesis of anthocyanins in several fruit crops, particular in grapes (*Vitis vinifera*). The principle Myb-related gene identified in *Vitis labruscana*, *VlmybA1-1*, was shown to regulate anthocyanin biosynthesis through the expression of UFGT (Kobayashi et al., 2002, 2005; Azuma et al., 2007). Later, additional homologs of *VlmybA1-1* were found in related species of grape. *VvmybA1*, a homolog isolated from *Vitis vinifera*, was studied by Kobayashi et al. (2005) and Azuma et al. (2007, 2008) who demonstrated that transcripts of *VvmybA1* were absent from white-skinned fruits and that the cause of this was related to the presence of a retrotransposon in the *VvmybA1* promoter (Kobayashi et al., 2004).

Research Objectives

There are four objectives to the current research. First, to add additional SSR markers to the existing diploid genetic linkage map created by Rowland et al (2014). Primers were designed from sequence information generated from the on-going blueberry genomic sequencing of W85-20 (one of the parents of the population). This objective will improve the saturation of the existing map and provide anchors to the genomic sequence.

The second objective was to use the SSR primer information associated with markers on the existing genetic linkage map (referred to in this thesis as the 'diploid map') to identify the corresponding genomic scaffolds from which they originated. We were also provided SSR and SNP sequence information from two additional blueberry genetic linkage maps that are currently under construction. This information was provided confidentially by Dr. Nahla Bassil at the USDA-National Clonal Germplasm Repository at Corvallis, Oregon, Dr. Susan McCallun at The James Hutton Institute in Scotland, and Dr. James Olmstead and his PhD student Hilda Patricia Rodriguez-Armenta at the Horticultural Science Department, University of Florida. The current blueberry assembly contains thousands of genomic scaffolds that range in size from 1.7 Mbp to a few hundred base pairs (N50 = 2.4 kb). The objective was to assign as many of these scaffolds as possible to these three maps based on the published and provided sequence information. This will provide functionality to the maps by linking markers to thousands of annotated genes on the scaffolding while also significantly contributing to the genomic sequence assembly by providing order and orientation to genomic scaffolds.

The third objective was to use scaffolds assigned to all blueberry maps as common anchors for the purpose of comparing the current published map to those under construction. These maps are built largely with unrelated marker sets and it is currently difficult to make meaningful comparisons. Providing these anchors would provide a framework for the eventual construction of a consensus map of *Vaccinium*.

The final objective was to utilize sequence homology of known genes (from model crops) in the flavonoid and anthocyanin biosynthesis pathways to identify orthologous sequences in the blueberry genomic assembly. Currently very little is known about how blueberries produce such high (and unique) concentrations of anthocyanins. Identifying these sequences and their genomic location will provide important information to researchers interested in health benefits associated with blueberry.

Chapter Two
Improved and novel resources for blueberry mapping and genomics

Introduction

Blueberry is an economically important small fruit, increasingly popular in the United States and worldwide. This popularity is due in part to the publicity arising with the reputed health benefits of blueberries which are associated with a high concentration of putative health promoting compounds such as anthocyanins, related flavanoids and phenolic acids. The past decade has seen several research programs generating blueberry genomic tools (genomic and RNA libraries, molecular markers and maps, and gene annotations) to address some of the problems associated with current breeding strategies and also to gain a better understanding of the blueberry genome.

The only published genetic linkage map of blueberry that is not based solely on RAPD markers was generated from an interspecific diploid blueberry population created from a cross between *V. darrowii* and *V. corymbosum*. The population has been used to research cold hardiness and chilling requirement, and is being used to map a number of fruit quality traits. (Rowland et al., 2003, 2014). The linkage map is comprised of 265 molecular markers spanning 12 linkage groups but considerable gaps remain in the map and two of the linkage groups are comparatively short with only five to eight markers. (Rowland et al., 2014). In addition, one of the parents of the mapping population 'W85-20' is the target for next generation shotgun and BAC-assisted genomic sequencing of blueberry, but currently no anchors exist to

link this map to the current draft assembly. Anchoring the assembly to the linkage map could benefit both breeders (by providing access to putative gene candidates and additional molecular markers in the vicinity of the QTL) and genomicists (by providing anchors to align and orient genomic scaffolding). Furthermore, there are two additional genetic linkage maps of blueberry currently being constrcuted. These maps utilized, for the most part, different sets of molecular markers so it is difficult to compare these maps on the basis of common markers. However, if individual markers from different marker sets can be mapped to the same scaffolds, these scaffolds can then be used to align linkage groups from different maps.

The objectives of this study were to 1) enhance the existing genetic linkage map by adding additional SSR markers designed from larger scaffolds, 2) identify the scaffolding location of all SSR primers and GBS sequences associated with new and existing maps, and 3) integrate as much of the genomic scaffolding as possible to the framework of the existing blueberry map and the two maps currently under construction.

Materials and Methods

Plant material

The diploid F1 population, [(*V. darrowii* selection Fla4B × *V. corymbosum* selection W85-20) F_1 #10 × *V. corymbosum* W85-23], segregates for cold tolerance, chilling requirement and a number of fruit quality traits. The population was developed by Dr. Rowland and plants are maintained in the greenhouse at the USDA-ARS, Beltsville Agricultural Research Center-

West, Beltsville, MD (Rowland et al., 2014). One of the parents, ‘Fla4B’ is the primary source of the low chilling requirement trait in many modern southern highbush blueberry cultivars and is therefore of significant economic value. A second parent, 'W85-20' is the target of the on-going genomic sequencing project at NCSU.

The population is a modified backcross, which was created to compensate for the severe inbreeding depression in diploid blueberry that prevents the development of a true backcross or F_2 population. The parents 'Fla4B' and 'W85-20' were crossed to produce the F_1 designated #10. The F_1 was then crossed to 'W85-23', a genetically similar line of 'W85-20'. A recent diversity study showed that while these lines were genetically similar, there were distinct at a number of genetic loci Bian et al. (2014). After examining the marker segregation patterns in the progeny of this population, Rowland et al. (2014), determined it would be most appropriate score these markers as a testcross population, concentrating on those markers that were polymorphic between Fla4B and W85-20, present in #10 but absent in W85-23.

DNA from 91 progenies, and four parental genotypes was extracted from young leaves by Dr. Rowland and shipped for use to The Plants for Human Health Institute, Kannapolis, NC. The concentration of DNA received was quantified using a NanoDrop 1000 (Thermo Scientific, Waltham, MA) and diluted to 7.5 ng/µl for PCR use.

SSR marker selection

SSR primers were designed from sequence information of the assembled 13,757 scaffolds (Newbler assembly) and a list of mined SSRs available on the Genome Database for

Vaccinium (GDV) website (http://www.vaccinium.org/content/ssr-markers). SSR markers were selected from the 600 largest scaffolds, scaffolds containing putative genes associated with anthocyanin biosynthesis (see chapter three for more detail) and from published cranberry SSR markers (Georgi et al., 2013). The selection criteria of SSRs were tri-nucleotide repeats with an expected PCR product size of 200~300 bp. One primer pair each was selected from the 600 largest scaffolds and six additional SSRs were selected from the 10 largest scaffolds. A total of 27 SSR markers from the cranberry genetic linkage map were also screened. Additional SSR markers were screened from a previous genetic diversity and population structure study conducted by our lab (Bian et al., 2014). The primer sequences were then aligned to the current genomic assembly for scaffold assignment.

SSR marker screening and PCR

The 20-bp primers were obtained from Integrated DNA Technologies (IDT, Coralville, IA). Primers were designed with 18 bp FAM-labeled M13 tail for fluorescent extension (5'-TGTAAAACGACGGCCAGT -3') on the 5' end of the forward primer in each pair, to permit labeling of fragments by PCR with a fluorescently tagged M13 primer (Oetting et al, 1995; Schuelke, 2000) for further allele analysis.

Primer screening panels were established and results examined before testing markers on whole population. Each screening panel consisted of the four parents (Fla4B, W85-20, W85-23, and F_1 #10) and 12 randomly chosen progenies. Markers that displayed polymorphism in the screening panel were selected to genotype the entire population. Each

PCR reaction contained 15 ng of genomic DNA template, 1x buffer, 2 mM MgCl2, 200 μM of dNTP, 0.1 μM M13 (-21) tailed forward primer, 0.2 μM reverse primer, 0.2 μM FAM-labeled M13 (-21) universal primer and 0.25 units of Taq DNA polymerase (Bioline Inc., Taunton, MA) in a total volume of 10 μl. ABI GeneAmp 9700 Thermocycler™ (Applied Biosystems, Foster City, CA) was used for all the PCR reactions. The modified touchdown PCR (TD-PCR) protocol (Bian et al., 2014) was used and the procedure is described as following: 94°C for 5 min; 19 cycles of 94°C for 40s, 60°C for 40s, and 72°C for 40s, annealing temperature reduced 0.5°C per cycle, until finally reach 50°C; 20 cycles of 40s, 50°C for 40s, and 72°C for 40s; and final extension at 72°C for 15 min; the last step was keeping PCR product at 4°C for further analysis.

Fragment analysis and genotyping

From a total volume of 10 μl PCR product, 9 μls was run on a 2% agarose gel for a PCR quality check and as a comparison with capillary electrophoresis result. The remaining 1 μl was genotyped using a capillary fluorescent sequencer, ABI 3730xl DNA Analyzer™ (Applied Biosystems), with POP-7™ Polymer and a 50-cm capillary array. Binning and scoring of the alleles was done with GeneMapper V4.0 (Applied Biosystems). Criteria of marker scoring followed the description of Rowland et al. (2014) for presence or absence of an allele in every individual.

Anchor SSRs onto diploid linkage map

SSR markers were examined for segregation distortion using a chi-square goodness-of-fit test at p=0.05. Markers were tested for 1:1 and 1:2:1 segregation ratios. Markers that did not fit either the ratio were removed. Scoring criteria followed Rowland et al. (2014) to ensure consistency. Markers were added to existing linkage groups using JoinMap 4.0 (Kyazma; Wageningen, Netherlands) and mapping distance was calculated using the Haldane mapping function, the rest of parameters followed default setting. With the original data file, which was kindly provided by Dr. Rowland, new marker data could be added directly into the file. Instead of processing all the markers at once, we added markers to each linkage group one at a time. The original markers from each linkage group were added first and then all new markers were added. The highest grouping LOD for each new marker was used to determine its appropriate group. Using a minimum LOD of 4.0 or higher, grouped markers (old and new) were used to construct a new linkage group. These linkage groups were further analyzed by reconstructing them one marker at a time from both ends of the linkage group (new and old markers). If JoinMap could not unambiguously map the markers in a single round of analysis or the addition of the marker grossly distorted the alignment, then the marker was removed. All maps of the final alignment were drawn from results using MapChart 2.2 (Voorrips, 2002) and GIMP (the GNU Image Manipulation Program).

Marker sequence alignment

SSR primer and GBS sequences were aligned to maps [the diploid map constructed here; the preliminary tetraploid map provided by Dr. Nahla Bassil and Dr. Susan McCaullum; the preliminary interspecific hybrid map provided Dr. James Olmstead; and the published cranberry map (Georgi et al., 2013)] via the Bowtie alignment tool (Lagnmead et al, 2009). Default options were used except for –f for multifasta, -S for SAM format and –p 12 for 12 processors. These alignments required a percent identify > 95%. The quality of alignments were then visualized and examined using the Geneious software platform (version R8, Genious, Auckland, New Zealand). For alignment of SSR primers, both forward and reverse primers had to be on the same scaffold, they could be no further than 1000 bp apart, and they had to be in complementary direction in the sequence (5' to 3' and 3' to 5').

Results and Discussion

New SSR markers and diploid map improvement

A total of 1,073 SSR markers were screened for this work. These markers included 386 SSRs derived from the 600 largest Newbler scaffolds of the blueberry genome assembly, 180 SSRs related to genes of the anthocyanin biosynthesis pathway (chapter 3), 432 SSRs from previous SSR mining work from genome assembly (Bian et al., 2014) and 27 SSR primer sequences from the published cranberry genetic map (Georgi et al., 2013). Among all markers screened, 281 (26.18%) revealed polymorphisms between the parents and 45 were

monomorphic. These 281 polymorphic markers, including nine from cranberry, allowed us to score 173 alleles that fit a 1:1 ratio of presence/absence phenotypes for linkage analysis in the population.

In addition to 1:1 ratio, a 1:2:1 ratio was also examined and only three markers exhibited 1:2:1 ratio (scaf00113, scaf00125, scaf00281). Markers that exhibited skewed segregation were removed ($p \leq 0.05$) before adding to the map, although in the original linkage map (Rowland et al., 2014), 42 markers with distorted ratio were mapped and clustered at the ends of LG 1, LG 3 and LG 5. The causes of segregation distortion could not only be genetic, but also physiological and environmental factors can affect allele frequency (Taylor & Ingvarsson, 2003).

The most common repeats in plant genomes are mono-, di-, and tri-nucleotide, but there are also small proportions of tetra-, penta- and hexa-nucleotide repeats. Di-nucleotide repeats were the most abundant in the blueberry genome except for in gene coding regions and among the di-nucleotide repeats, AG/CT motifs were the most abundant while CG/CG were less common (Bian et al., 2014). The less frequent appearance of CG-content repeat motifs was also observed in tri-nucleotide, CCG/CGG was the least frequent across the whole blueberry genome. From 96 SSR markers with polymorphisms, none of them were CG repeats, whereas 20 out of 35 di-nucleotide SSRs were AG/CT, which coincided with previous research. In coding regions, tri-nucleotide repeats were more common not only in blueberry, also in other species such as monocots like rice (*Oryza sativa*) and sorghum (*Sorghum bicolor*), and dicots like Arabidopsis (*Arabidopsis thaliana*) and poplar (*Populus tremula*) (Sonah et al., 2011). The

triplet nature of the codon might explain the abundance of tri-nucleotide in coding regions since repeats other than tri-nucleotide may result in frameshift mutation. The variation of repeat length is also thought to be as an evolutionary force when SSRs reside within a gene, the length variance can lead to strengthening or loss of gene function.

After analyzing 173 new markers in the diploid population, the number of mapped markers increased to a total of 318 (from 265 in the original map). These new mapped markers represent 92 (28.9%) of the 318 total markers on the current map, while 39 of the original markers were removed due to low LOD score (less than 4.0) or (Fig. 3). In the new diploid linkage map, the total map length is 1,659 cM and the average distance between markers is 5.45 cM. The largest LG is DI02 (198.7 cM) while the smallest LG is DI11 (43.56) cM in length. The average number of markers per linkage group is 26.5 and the number of markers per linkage group ranged from 8 to 37. The largest gap is located on LG 3 (30.54 cM) (Table 6). Most of the large gaps are at the ends of the linkage groups. These gaps may result from differences in recombination frequency in these areas.

In addition to adding more SSR markers to the existing map, a number of significant changes to existing markers and linkage groups was indicted based on the new marker information. Previously, linkage group 1 was the largest linkage group in the published map. When we attempted to reconstruct this linkage group one marker at a time beginning from the end of the LG, the markers formed two groups that could not be grouped above LOD level 2.0. When linkage groups of the two separate groups were mapped, no marker from one group

could be added to the other group at a grouping LOD above 2.0. They are currently designated as DI01a and DI01b in this map version.

We were unable to add any new markers to LG 12 (which had only five markers in the original map). After examining the original LG we found we could only group four of these markers above an LOD of 2.0 and considerably large gaps existed between the them. We are suggesting that for the time being, this linkage group be considered a fragment. We could not link any of the markers from LG 12 to any other linkage group. We also attempted to see if LG 11 (also a short LG) would link to any other existing linkage groups, and it did link at an LOD 4.0 but was separated at an LOD of 5.0. The results show that LG 11 clustered at the end of LG 2, however, the total length of the LG increased significantly to 243.1 cM, which made the new linkage group suspiciously long, possible due to the misplacement of markers (Fig. 4). At this point, the decision was to leave these as separate LGs.

Aligning scaffolds to linkage maps

In total 696 scaffolds from the genome assembly were assigned to locations on one or more of the genetic linkage maps of blueberry. These scaffolds represent 214,984,227 bp or 44.3% of the current assembly (Appendix A-D). Assigned scaffolds ranged from 1,794,986 bp to 334 bp in length. Not surprisingly, the likelihood of assigning a scaffold to a linkage group was a function of the size of the scaffold. We were able to assign 83 of the largest 100 scaffolds. These 83 scaffolds had a mean length of 775 kb and represented 64,393,139 bp (Appendix A-

D). Of the 1000 largest scaffolds, 510 could be assigned which represents 202,569,426 bp with an average size of 397 kb.

The alignment of GBS markers to unique scaffolds was more successful than the alignment of the SSRs. When aligning SSR primers to the blueberry genomic scaffolds, it was common to have a single forward or reverse primer align to many scaffolds or not appear as a pair on any one scaffold. There are a number of possible explanations for this, the GBS sequences were longer than the combined length of the forward and reverse SSRs and likely consisted of unique and highly conserved regions flanking the polymorphism. Additionally, many of the SSR markers were derived from EST data and the flanking region of these repeats are coding regions of genes which could have several paralogs or could represent conserved domains within genes. Some evidence for this is presented by the results of the primer screening where multiple bands (loci) where amplified with the same primer sequence. Also these SSR repeats could represent regions that the genomic assembly software might not have been able to successfully resolve resulting in either a break between scaffolds or a gap within a scaffold.

The Draper and Jewel linkage maps were the most useful for assigning scaffolds. Ten of the 15 largest scaffolds could be unambiguously assigned to the same linkage groups in both maps (Table 7). These ten linkage groups represent 47 GBS markers, 17 Mbp and 1,892 annotated genes. Using these ten scaffolds to estimate GBS spacing, it can be estimated that one GBS marker occurs every 375 kb on this map. All linkage groups in this population

contained multiple markers that mapped to the same scaffold (Appendix B and C). This markers were not always contiguous but they tended to cluster in the same region.

The assignment of the markers to scaffolds may also provide some clues to help resolve various issues that mapping groups are having constructing the linkage groups. An example is shown in Figure 5. While linkage group 11 in Draper appears as a single linkage group, the same linkage group in Jewel is currently separated into two groups (J11a and J11b). According to scaffold assignments D11 and J11b are perfectly aligned with both suggesting that scaffold 10 appears near the end of the linkage groups, followed by scaffold 329, scaffold 2, scaffold 3, and scaffold 21. In a similar manner, scaffold00148, 1107, 118, 110 and 3147 are co-linear between D11 and J11a. Some discrepancy appears to exist at the proximal end of these two linkage groups with the scaffold positions showing an inverse linear relationship of 5 common scaffolds with scaffold 1938 appearing near the end of D11 and scaffold 1138 appearing near the end of Jewel 11a. This would suggest that perhaps one or more markers from Jewel 11a may have resulted in an inversion in their order and that this marker(s) may also be preventing the two halves of Jewel LG 11 from coming together. In addition to linkage group 11, more than half of the linkage groups in Jewel are currently fragmented when compared to Draper and these included LG 3 (J03a, J03b, J03c, J03d), LG 4 (J04a, J04b), LG 6 (J06a, J06b), LG 8 (J08a, J08b), J10 (J10a, J10b, J10c), and LG 12 (J12a, J12b). It is our expectation that the work we have done here will ultimately help identify potential trouble spots on these linkage groups as well and lead to a successful resolution of the Jewel map.

Despite these problems, the tetraploid linkage maps of Draper and Jewel, were the most effective of all maps in assigning scaffold locations (Table 9). These maps are composed primarily of different sets of GBS markers, and the alignment shows strikingly similar positions and often represent similar distances in cM. Multiple markers from scaffold00002, for example, mapped about 8 cM apart in both the Draper and Jewel map near the end of LG 11; scaffold00008 was linked to Draper, Jewel, and interspecific hybrid with the distances of 12 cM, 10 cM, and 7 cM, respectively (Table 10).

The scaffold assignments provide a benefit to linkage map construction but also provide a benefit to the continuedd sequencing and assembly of W85-20. The latest assembly of W85-20 contains integrated BAC end sequences that are assigned to specific scaffolds. A number of these sequences are unpaired (one end of the BAC end occurs within a given scaffold and the other doesn't match). Sequencing of specific BACs could be used to expand scaffold length and to resolve gaps within the scaffolds.

Consensus linkage groups

Identifying consensus between all of the linkage groups of the different maps was not possible at the present time. In part, this is attributable to the continuedd evolution of the tetraploid blueberry and the Florida blueberry maps. These maps have been revised several times in the past couple of months and our estimations of linkage group alignment have often changed with the new maps. As we have demonstrated in Figure 5, the alignment between Draper and Jewel has (in most cases) been as expected, but we have not found sufficient

agreement between these maps and the other maps to make definitive alignments of all linkage groups.

On the basis of multiple scaffolds, we can align Jewel LG1, Draper LG 1, interspecific hybrid LG 7, and Diploid LG 5 (Fig. 6). The other example is alignment of Jewel LG 12 and Draper LG 12 with interspecific hybrid LG 10 and Diploid LG 10 (Fig. 7). Multiple scaffolds occur in consecutive order throughout the length of these linkage groups.

Conclusion

The addition of new markers to the diploid genetic linkage map has improved the saturation of the map and helped to identify an assembly issue with LG01. We have provided the mapping file to the original author of this publication (Rowland et al., 2014) who has verified our analysis that LG01 is actually two separate linkage groups that we have designated DI01a and DI01b. Further, we suspect that LG12 may be misaligned, a fragment or a statistical anomaly as the five markers on this linkage group are only joined together (over a distance of 59 cM) by a single RAPD marker. We could not add any additional markers to this group or join these markers to any other group.

We have assigned 696 scaffolds to one or more genetic linkages maps on the basis of primer and GBS sequences. These scaffolds represent 214,984,227 bp or 44.3% of the current assembly of blueberry. As these scaffolds have been annotated and the sequences and annotations are available through the database *vaccinium.or*g, they provide functionality to the genetic linkage maps by allowing researchers to access genomic information in the vicinity of

specific markers. The website also contains information on over 40,000 SSR markers associated with the assembly which will allow researchers to add additional markers to maps for resolving linkage issues and for purposes such as fine mapping. The scaffold assignments are also of considerable importance to the improvement of the draft genomic sequence of blueberry. The recent addition of BES to the genome provides a blueprint for additional sequencing and refinement of the sequence through the selective sequencing of individual BAC clones.

We were not as successful as we had hoped in providing alignments between linkage groups of different maps. In part, this can be attributed to the delays in finalizing the map order of the tetraploid blueberry and Florida maps. We expect the finalized Draper and Jewel map to be released shortly and the Florida map should be finalized before the end of the year. We also attribute some of the difficulty we had to locating unique scaffold assignments to the blueberry EST-SSRs primers, which often mapped to multiple scaffolds across the genome. This was in contrast to the GBS sequence, which rarely mapped to more than a single unique locus. The alignment of the scaffolds is consistent between Draper and Jewel and may help to resolve some issues associated with the Jewel map (Fig. 7).

Chapter Three
Identification of genes in anthocyanin biosynthesis in blueberry

Introduction

Anthocyanins are important compounds that give flowers the colors of red, blue and purple and also protect plants from UV-light, insects and herbivores (Hatier & Gould, 2009). In addition, anthocyanins are increasingly thought to be significantly beneficial to human health (Kalt et al., 2007; Zafra-Stone et al., 2007; He & Giusti, 2010). The value of studying anthocyanins in blueberry is that blueberry contains surprisingly high concentrations compared to other crops (Wu et al., 2006). Given that the demand and production of blueberry has been grown exponentially for the past few years, increased production is expected to continued at least for the near future.

Anthocyanin studies in blueberry have primarily focused on biochemical analysisand compound profiling (Yosef et al., 2014), and the biological effect on human health (Prior & Wu, 2006; Zafra−Stone et al., 2007; He et al., 2010). Little is known, however, regarding the genetic mechanism of anthocyanin biosynthesis (Routray & Orsat, 2011). Several genes involved in the pathway have been deposited separately in The National Center for Biotechnology Information (NCBI) such as chalcone synthase (AFA53723), dihydroflavonol 4-reductase (AHK23093), anthocyanin synthase (AFA53722), and R2R3 Myb transcription factor (AEV21970) (Zifkin et al., 2012), however, there is no information regarding copy number or genomic location of these genes but gene expression has been studied in both

blueberry and bilberry by Zifkin et al. (2012). The pathway has been elucidated in many model crops including: petunia (*Petunia* × *hybrida*), maize (*Zea mays*), grape (*Vitis vinifera*), and Arabidopsis (*Arabidopsis thaliana*). Anthocyanin distribution in these plants is tissue specific, anthocyanins accumulate in the petal of petunias, in the seeds of maize, and in the skin of grapes and berries, while anthocyanins were found in all organs of Aradidopsis. Among all these well-studied plants, grape might be the best model for the study of anthocyanins in blueberry because grape accumulates anthocyanins in the skin of fruits, is a semi-perennial woody plant, and produces an anthocyanin profile that is similar to what is found in blueberry.

The objective of this chapter was to use available protein sequences of grapes to identify homologs in the blueberry genomic sequence. Further, we report on the estimated copy numbers of these genes and examine, through phylogenetic analysis, the relationship between these putative sequences and known sequences of other plants.

Materials and Methods

BLAST for genes related to anthocyanin biosynthesis

To identify homologous sequences in the blueberry genomic draft assembly, known anthocyanin and flavonoid genes, amino acid sequences of these genes were acquired from the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database for flavonoid (entry: map00941) and anthocyanin biosynthesis (entry: map00942) using grape (*Vitis vinifera*) as reference. A previous publication providing annotation to the draft genomic sequence of

blueberry has noted that grape sequences are more closely related to blueberry than any other model organism (Gupta et al., 2015). Amino acid sequences of nine key genes in the anthocyanin/flavonoid pathway were obtained including: phenyalanine ammonia lyase (PAL), 4-coumarate-CoA:ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), dihydroflavonol 4-reductase (DFR), anthocyanidin synthase [ANS, or LDOX (leucoanyhocyanidin dioxygenase)], flavanone 3-hydroxylase (F3H), flavonoid 3'-hydroxylase (F3'H), flavonoid 3'5'-hydroxylase (F3'5'H), and UPD-glucose: flavonoid 3-0-glucosyltransferas (UFGT).

Additionally, the sequence of a known anthocyanin transcription factor, Myb transcription factor *VvmybA1-1*, was obtained from the DNA Data Bank of Japan (DDBJ) with accession number AB427165, as reported by Azuma et al. (2008). The amino acid sequences from grape were then searched against the blueberry scaffold database (*V. corymbosum* 454-Scaffolds) on the Genome Database of *Vaccinium* (GDV) website (http://dev.vaccinium.org/tools/blast) to identify which genomic scaffolds contained putative candidates to these sequences. Searches were conducted with tblastn which uses protein sequences to search nucleotide databases for matches in all six reading frames. Sequences were then visualized in the Integrate Genome Browser (IGB), which contains annotated blueberry genome assembly sequences (Gupta et al., 2015) to validate and visualize the results and to determine if the automated annotation programs used by the site had correctly annotated the sequence. The number of putative candidates for each gene in blueberry was determined by an E-value equal or greater than E-20.

Phylogenetic analysis

To compare the putative genes identified in blueberry with known sequences of unrelated species, phylogenetic trees were constructed with ClustalW2, an online program for multiple sequences alignment (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The phylogenetic trees were generated with EvolView, an online tool for visualizing phylogenetic trees (Zhang et al., 2012) (http://www.evolgenius.info/evolview/#login). For the analysis, known DNA sequences of the ten genes previously described were obtained through the NCBI database for the species: petunia (*Petunia* × *hybrida*), Arabidopsis (*Arabidopsis thaliana*), barley (*Hordeum vulgare* L.), broccoli (*Brassica oleracea*), grape (*Vitis vinifera*), strawberry (*Fragaria* × *ananassa*), and cranberry (*V. marcrocarpon*). Not all sequences were available for all seven species.

Results and Discussion

The use of grape gene sequences to identify orthologous sequences in blueberry was based on the observations of Gupta et al. (2015), who performed an annotation of an earlier version of the blueberry genomic assembly and observed that, in general, the sequences were more closely related to grape (*Vitis vinifera*) than any other model system. The anthocyanin biosynthetic pathway has been studied extensively in grape due to the economic value associated with color quality of wine. The availability of these sequences provided ample resources for searching blueberry.

The results of the gene searches (E-values, predicted gene length, percent nucleotide identity matches) are presented in Table 11 and resulted in 63 gene copies in total. Sequence identity scores ranged from 89% (flavanone 3-hydroxylase) to 29% for flavonoid 3' hydroxylase. It should be noted that sequence homology is only one way to compare homologous sequences and that genes that perform similar functions often share little homology outside of functional or targeting domains. If gene annotations were provided by IGB they are also included in Table 11. With scaffold number, start, stop location obtained from BLAST results, corresponding annotated genes can be searched in IGB, which houses annotation results generated by four software programs: GeneMark, Augustus, and GlimmerHMM annotated using diploid blueberry genomic sequences while Cufflinks annotated RNA-seq data from Dr. Ann Loraine at Bioinformatics, UNC-Charlotte. Also, the relationship of these putative genes with five linkage maps investigated (diploid blueberry, Jewel and Draper of tetraploid blueberry, interspecific hybrid of sparkleberry and blueberry, and cranberry) was shown in Table 12.

Several of these putative genes were not annotated by all four gene-finding programs in IGB (chalcone isomerase on scaffold04077, flavonoid 3' hydroxylase on scaffold00412, UDP-glucose: flavonoid-3-O-glucosyltransferase on scaffold00427, and anthocyanin permease on scaffold00396). These putative genes might be highly specific to blueberry, they may be pseudo genes or they may simply have been missed by the automated gene prediction programs. All gene prediction programs have associated error rates that generally require additional manual annotation.

Two full length sequences (1748 bp) of PAL were identified on scaffolds00017 (CUFF.1944) and scaffold03505 (CUFF.51381.1). Another copy was identified on scaffold02351 (gene.g25340.t1) and identified by GeneMark, Augustus, and Glimmer but not validated by Cufflinks. Blasts of sequence to NCBI did suggest it is a third blueberry copy of PAL. One of these sequences (gene.g25340.t1) clustered closely with the grape sequence while the other two appear more closely related to the sequence from bilberry (*V. myrtillus*) (Fig. 8).

Six putative copies of CHS were identified on scaffolds 14, 52, 491, 519, 814, and 1392. Four copies of CHS clustered with grape (CUFF.39764 CUFF.30053, CUFF.22324, gene.g12175.t1) , but two (CUFF.2188 and CUFF.4712) appear to be more distant and perhaps more closer in structure to CHS in strawberry. All but gene.g12175.t1 were correctly identified by all annotation programs. In a similar manner to PAL and CHS, three copies were found for CHI, five copies of DFR, five copies of F3H, 10 copies of F3'H, five copies of F3'5H, five copies of AMOT, four copies of UFGT, five copies of ANP, and five copies of 4CL. In most plants these genes generally represent small gene families, so the numbers are expected. The cluster analysis provides some evidence for prioritizing the eventual analysis of these gene sequences. For example, UFGT is a glycotransferases and a member of a large gene family. The copies of the genes on scaffold01120 and scaffold00304 were predicted by annotation software to be anthocyanidin 3-O-glucosyltransferases (others were generalized glucosyltransferases) and these two sequences clustered more tightly with known anthocyanidin 3-O-glucosyltransferases than the other four identified sequences. In general, the peptide sequences of grape were more similar to the predicted peptide sequence of the

blueberry genes than sequences from Arabidopsis, Brassica, or barley, but where available, sequences from bilberry or cranberry were much more conserved. Considerable variation has been observed in the accumulation of anthocyanins in blueberry, and while much of this variation is attributable to environmental conditions, a significant portion is genetic in nature (Yousef et al., 2014). Understanding how blueberries accumulate such high levels of health related compounds like anthocyanins will help to eventually develop cultivars that could be even more effective than ones currently available in reducing risks of chronic disease such as diabetes and cancer. Genomic studies in *Vaccinium* are at an early but exciting stage. Recently, the draft genomic sequence of Kiwifruit (*Actinidia chinensis*) became the first sequenced species in the order *Ericale* (Huang et al., 2013) and as other closely related plants in this order (tea) are used to generate additional resources (Shi et al., 2011; Tan et al., 2013), this information should provide exciting opportunities for comparative analysis of anthocyanin gene sequences in these fruits that are often associated with beneficial health effects. The identification of anthocyanin related genes in blueberries reported here is an important first step toward that goal.

Another consideration is evolutionary divergence. The species selected can be traced back to the same angiosperm ancestor, which diverged from gymnosperm around 245-202 million years ago, and the earliest flowering plant discovered so far appeared about 130-140 million years ago (Soltis et al., 2008). After hundreds of millions of years of evolution, numerous species of plants today are adapted to a multitude of various climates and environments. Thus, the study of compararive genomics has become a crucial strategy to

identify homologs in non-model organisms, especially in those lacking sequence information. The forces of genome evolution usually are considered to be genome duplication, genome fusion, or speciation (Koonin, 2005). According to the forces of evolution, homologs can be further categorized into two groups: orthologues, which are driven by speciation; and paralogues, which result from genome duplication events (Koonin, 2005; Gabaldón & Koonin, 2013). Although orthologues that form a clade are monophyletic, by strict definition of orthology, the same family can be separated into different groups depends on the desired level of resolution, since the relatedness of an orthologous is compared with the last common ancestral species (Gabaldón & Koonin, 2013). Besides natural selection, artificial selection can also influence the genetic drift with different breeding goals. By selecting individuals with favorable phenotypes, the frequency of genes linked to particular traits is then increased in the next breeding cycle, and by repeating the inbreeding process, this can lead to the creation of a specific plant or crop population that is based on human needs. Thus, even the same species can have different genetic compositions regarding a particular gene or trait based on breeder's needs; and vice versa, different crops can have similar genetic information after selecting for the same characteristic in both crops for an extended period of time.

Conclusion

The integration of blueberry genomic scaffolding and anthocyanin gene identification provides a foundation for studies in the anthocyanin biosynthesis pathway. The results not only show the locations of the genes which will be informative in marker-assisted breeding,

but they also can be helpful in studying, gene interaction and evolution of paralogous genes. For example, among ten genes investigated, F3'H and F3'5'H showed were consistently identified on the same scaffolds. This might suggest these tandem genes evolved from a common ancestor.

The understanding of genetic mechanisms in blueberry anthocyanin biosynthesis is still in its infancy, and the results reported here should provide a roadmap for more detailed research.

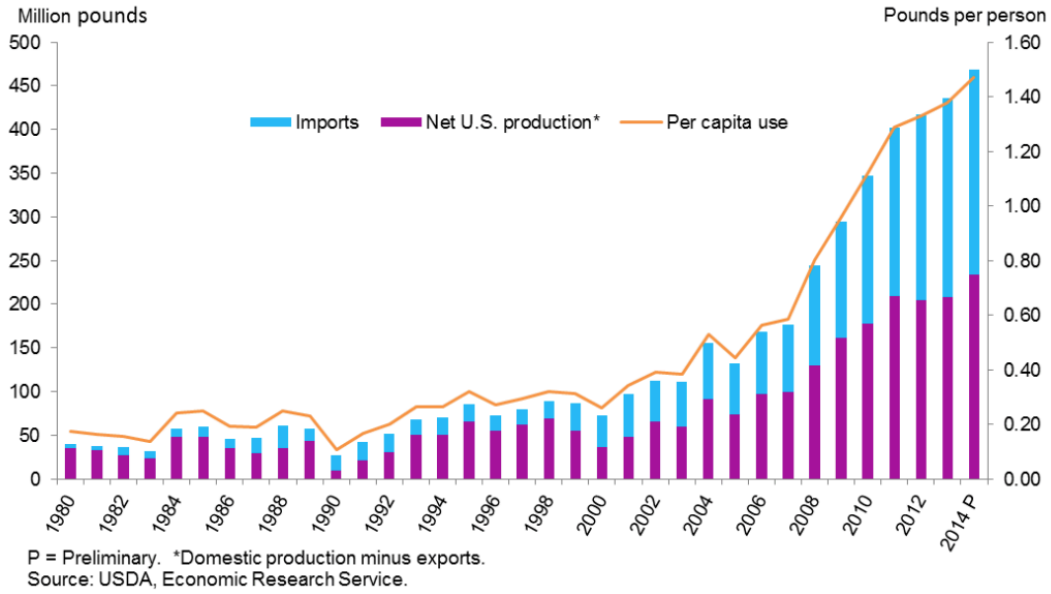


Fig. 1. The demand and production of blueberry since 1980. Purple bar indicates the growth of US blueberry production, blue bar indicates the import of blueberry while the orange line indicates the demand of blueberry in the U.S.A. (reference: USDA, Economic Research Service, Fruit and Tree Nuts Outlook, March, 2015)

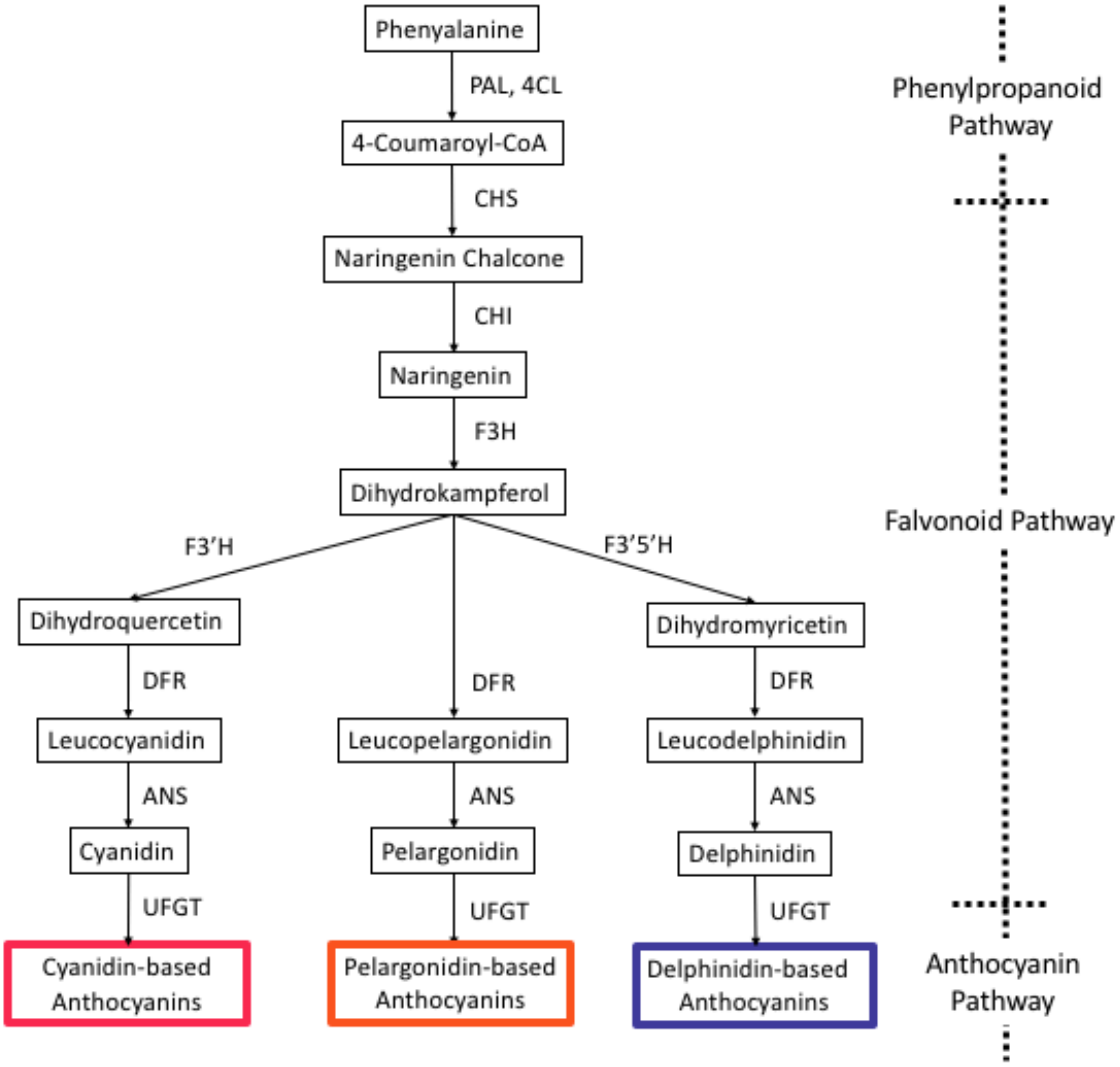
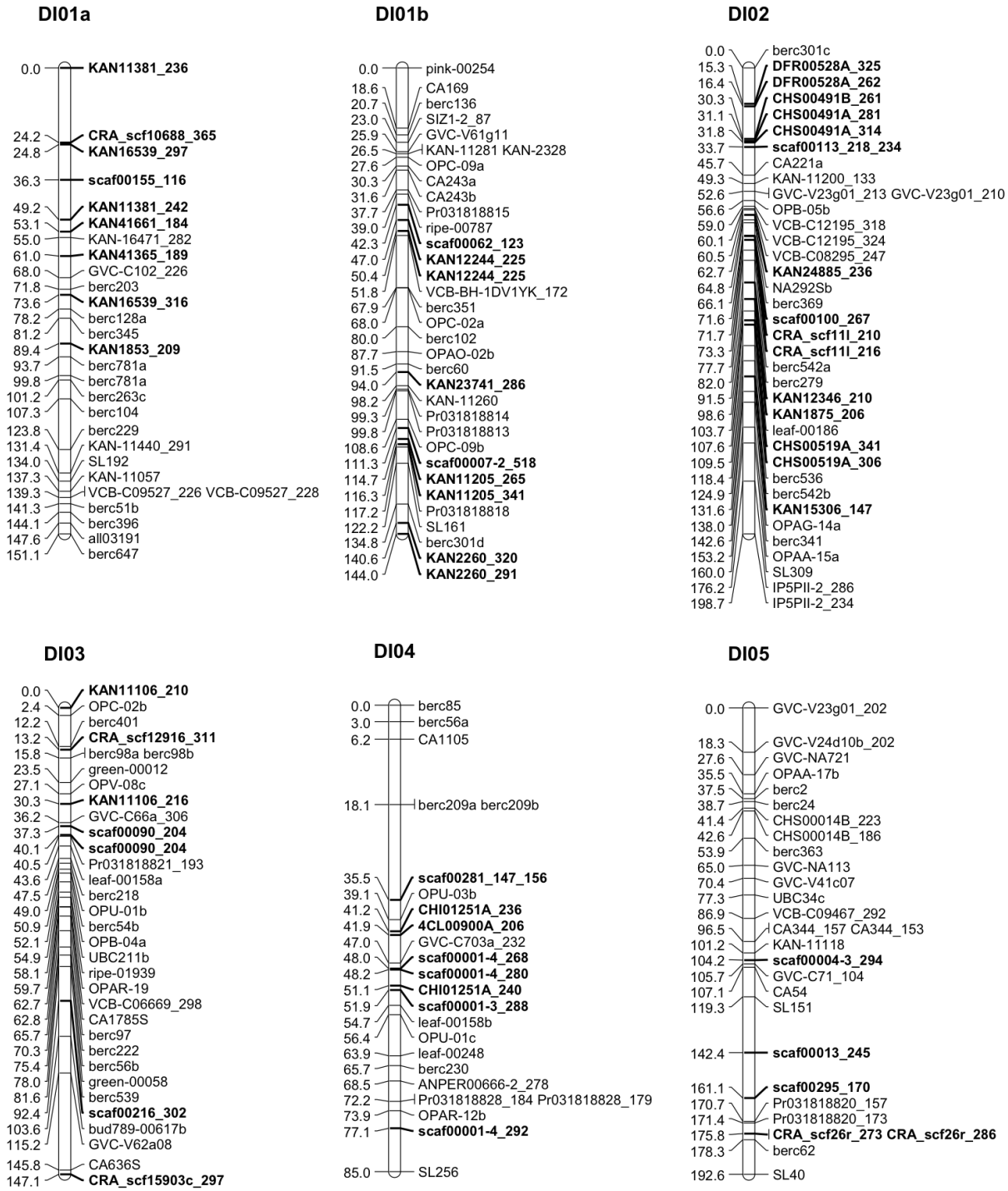
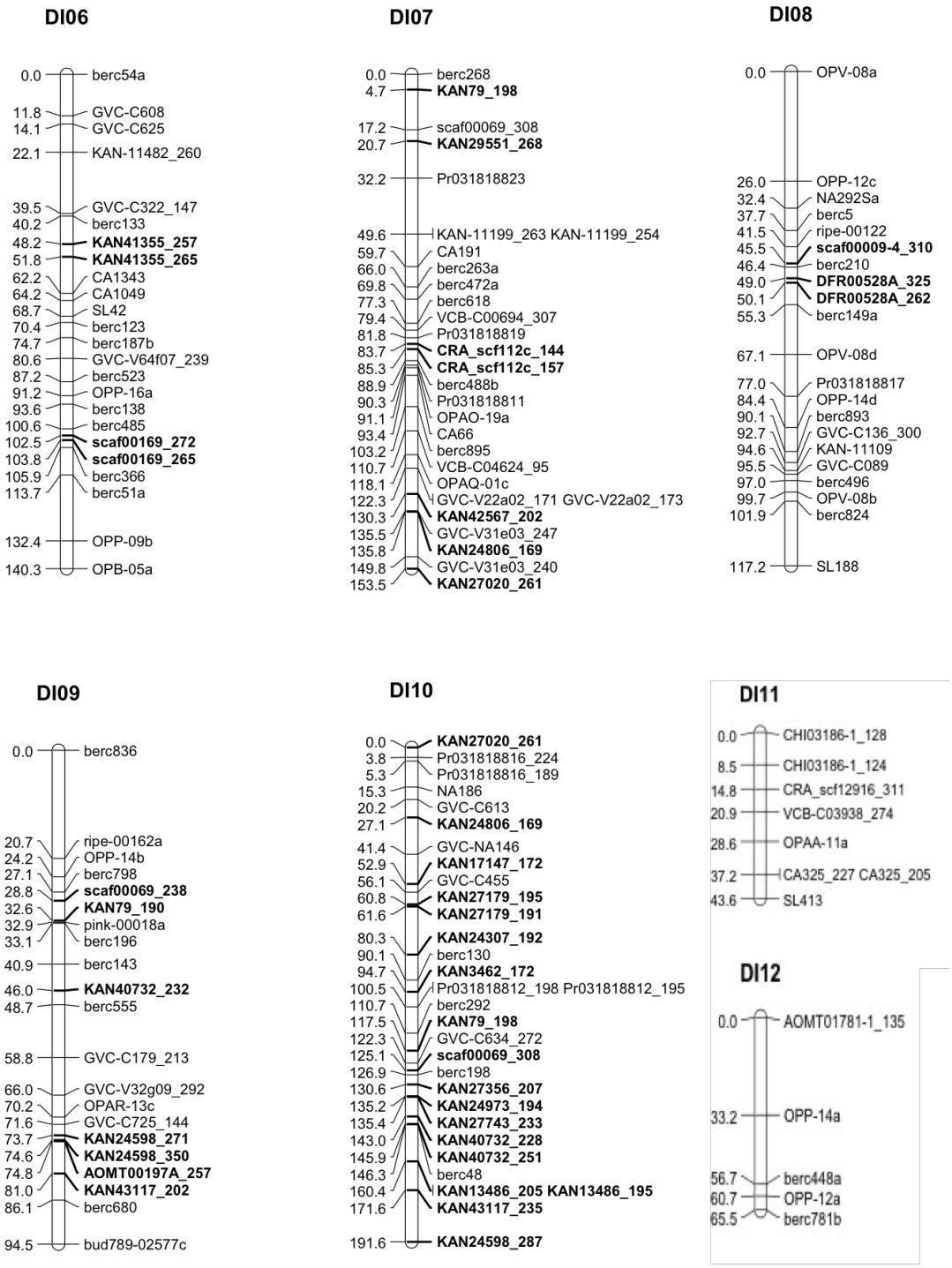


Fig. 2. Simplified anthocyanin biosynthesis pathway. The complete pathway includes phenylpropanoid pathway, flavonoid pathway, and specific anthocyanin pathway. Genes involved in the pathway are PAL (phenyalanine ammonia lyase), 4CL (4-coumarate-CoA: ligase), CHS (chalcone synthase), CHI (chalcone isomerase), F3H (flavanone 3-hydroxylase), F3'H (flavonoid 3'-hydroxylase), F3'5'H (flavonoid 3'5'-hydroxylase), DFR (dihydroflavonol-4-reductase), ANS (anthocyanidin synthase), UFGT (UPD glucose: flavonoid 3-0-glucosyltransferas). Cyanidin-based and delphinidin-based anthocyanins are found in blueberry but not pelargonidin-based blueberry.

Fig. 3. Genetic linkage map of diploid blueberry population [(Fla4B × W85-20) F_1 #10 × W85-23]. The markers emphasized in bold are new added SSR (simple sequence repeats) markers, the detail information of these markers are describe in Appendix F.


DI01a
0.0 KAN11381_236
24.2 CRA_scf10688_365
24.8 KAN16539_297
36.3 scaf00155_116
49.2 KAN11381_242
53.1 KAN41661_184
55.0 KAN-16471_282
61.0 KAN41365_189
68.0 GVC-C102_226
71.8 berc203
73.6 KAN16539_316
78.2 berc128a
81.2 berc345
89.4 KAN1853_209
93.7 berc781a
99.8 berc781a
101.2 berc263c
107.3 berc104
123.8 berc229
131.4 KAN-11440_291
134.0 SL192
137.3 KAN-11057
139.3 VCB-C09527_226 VCB-C09527_228
141.3 berc51b
144.1 berc396
147.6 all03191
151.1 berc647
DI01b
0.0 pink-00254
18.6 CA169
20.7 berc136
23.0 SIZ1-2_87
25.9 GVC-V61g11
26.5 KAN-11281 KAN-2328
27.6 OPC-09a
30.3 CA243a
31.6 CA243b
37.7 Pr031818815
39.0 ripe-00787
42.3 scaf00062_123
47.0 KAN12244_225
50.4 KAN12244_225
51.8 VCB-BH-1DV1YK_172
67.9 berc351
68.0 OPC-02a
80.0 berc102
87.7 OPAO-02b
91.5 berc60
94.0 KAN23741_286
98.2 KAN-11260
99.3 Pr031818814
99.8 Pr031818813
108.6 OPC-09b
111.3 scaf00007-2_518
114.7 KAN11205_265
116.3 KAN11205_341
117.2 Pr031818818
122.2 SL161
134.8 berc301d
140.6 KAN2260_320
144.0 KAN2260_291
DI02
0.0 berc301c
15.3 DFR00528A_325
16.4 DFR00528A_262
30.3 CHS00491B_261
31.1 CHS00491A_281
31.8 CHS00491A_314
33.7 scaf00113_218_234
45.7 CA221a
49.3 KAN-11200_133
52.6 GVC-V23g01_213 GVC-V23g01_210
56.6 OPB-05b
59.0 VCB-C12195_318
60.1 VCB-C12195_324
60.5 VCB-C08295_247
62.7 KAN24885_236
64.8 NA292Sb
66.1 berc369
71.6 scaf00100_267
71.7 CRA_scf11l_210
73.3 CRA_scf11l_216
77.7 berc542a
82.0 berc279
91.5 KAN12346_210
98.6 KAN1875_206
103.7 leaf-00186
107.6 CHS00519A_341
109.5 CHS00519A_306
118.4 berc536
124.9 berc542b
131.6 KAN15306_147
138.0 OPAG-14a
142.6 berc341
153.2 OPAA-15a
160.0 SL309
176.2 IP5PII-2_286
198.7 IP5PII-2_234
DI03
0.0 KAN11106_210
2.4 OPC-02b
12.2 berc401
13.2 CRA_scf12916_311
15.8 berc98a berc98b
23.5 green-00012
27.1 OPV-08c
30.3 KAN11106_216
36.2 GVC-C66a_306
37.3 scaf00090_204
40.1 scaf00090_204
40.5 Pr031818821_193
43.6 leaf-00158a
47.5 berc218
49.0 OPU-01b
50.9 berc54b
52.1 OPB-04a
54.9 UBC211b
58.1 ripe-01939
59.7 OPAR-19
62.7 VCB-C06669_298
62.8 CA1785S
65.7 berc97
70.3 berc222
75.4 berc56b
78.0 green-00058
81.6 berc539
92.4 scaf00216_302
103.6 bud789-00617b
115.2 GVC-V62a08
145.8 CA636S
147.1 CRA_scf15903c_297
DI04
0.0 berc85
3.0 berc56a
6.2 CA1105
18.1 berc209a berc209b
35.5 scaf00281_147_156
39.1 OPU-03b
41.2 CHI01251A_236
41.9 4CL00900A_206
47.0 GVC-C703a_232
48.0 scaf00001-4_268
48.2 scaf00001-4_280
51.1 CHI01251A_240
51.9 scaf00001-3_288
54.7 leaf-00158b
56.4 OPU-01c
63.9 leaf-00248
65.7 berc230
68.5 ANPER00666-2_278
72.2 Pr031818828_184 Pr031818828_179
73.9 OPAR-12b
77.1 scaf00001-4_292
85.0 SL256
DI05
0.0 GVC-V23g01_202
18.3 GVC-V24d10b_202
27.6 GVC-NA721
35.5 OPAA-17b
37.5 berc2
38.7 berc24
41.4 CHS00014B_223
42.6 CHS00014B_186
53.9 berc363
65.0 GVC-NA113
70.4 GVC-V41c07
77.3 UBC34c
86.9 VCB-C09467_292
96.5 CA344_157 CA344_153
101.2 KAN-11118
104.2 scaf00004-3_294
105.7 GVC-C71_104
107.1 CA54
119.3 SL151
142.4 scaf00013_245
161.1 scaf00295_170
170.7 Pr031818820_157
171.4 Pr031818820_173
175.8 CRA_scf26r_273 CRA_scf26r_286
178.3 berc62
192.6 SL40


DI06
0.0 berc54a
11.8 GVC-C608
14.1 GVC-C625
22.1 KAN-11482_260
39.5 GVC-C322_147
40.2 berc133
48.2 KAN41355_257
51.8 KAN41355_265
62.2 CA1343
64.2 CA1049
68.7 SL42
70.4 berc123
74.7 berc187b
80.6 GVC-V64f07_239
87.2 berc523
91.2 OPP-16a
93.6 berc138
100.6 berc485
102.5 scaf00169_272
103.8 scaf00169_265
105.9 berc366
113.7 berc51a
132.4 OPP-09b
140.3 OPB-05a
DI07
0.0 berc268
4.7 KAN79_198
17.2 scaf00069_308
20.7 KAN29551_268
32.2 Pr031818823
49.6 KAN-11199_263 KAN-11199_254
59.7 CA191
66.0 berc263a
69.8 berc472a
77.3 berc618
79.4 VCB-C00694_307
81.8 Pr031818819
83.7 CRA_scf112c_144
85.3 CRA_scf112c_157
88.9 berc488b
90.3 Pr031818811
91.1 OPAO-19a
93.4 CA66
103.2 berc895
110.7 VCB-C04624_95
118.1 OPAQ-01c
122.3 GVC-V22a02_171 GVC-V22a02_173
130.3 KAN42567_202
135.5 GVC-V31e03_247
135.8 KAN24806_169
149.8 GVC-V31e03_240
153.5 KAN27020_261
DI08
0.0 OPV-08a
26.0 OPP-12c
32.4 NA292Sa
37.7 berc5
41.5 ripe-00122
45.5 scaf00009-4_310
46.4 berc210
49.0 DFR00528A_325
50.1 DFR00528A_262
55.3 berc149a
67.1 OPV-08d
77.0 Pr031818817
84.4 OPP-14d
90.1 berc893
92.7 GVC-C136_300
94.6 KAN-11109
95.5 GVC-C089
97.0 berc496
99.7 OPV-08b
101.9 berc824
117.2 SL188
DI09
0.0 berc836
20.7 ripe-00162a
24.2 OPP-14b
27.1 berc798
28.8 scaf00069_238
32.6 KAN79_190
32.9 pink-00018a
33.1 berc196
40.9 berc143
46.0 KAN40732_232
48.7 berc555
58.8 GVC-C179_213
66.0 GVC-V32g09_292
70.2 OPAR-13c
71.6 GVC-C725_144
73.7 KAN24598_271
74.6 KAN24598_350
74.8 AOMT00197A_257
81.0 KAN43117_202
86.1 berc680
94.5 bud789-02577c
DI10
0.0 KAN27020_261
3.8 Pr031818816_224
5.3 Pr031818816_189
15.3 NA186
20.2 GVC-C613
27.1 KAN24806_169
41.4 GVC-NA146
52.9 KAN17147_172
56.1 GVC-C455
60.8 KAN27179_195
61.6 KAN27179_191
80.3 KAN24307_192
90.1 berc130
94.7 KAN3462_172
100.5 Pr031818812_198 Pr031818812_195
110.7 berc292
117.5 KAN79_198
122.3 GVC-C634_272
125.1 scaf00069_308
126.9 berc198
130.6 KAN27356_207
135.2 KAN24973_194
135.4 KAN27743_233
143.0 KAN40732_228
145.9 KAN40732_251
146.3 berc48
160.4 KAN13486_205 KAN13486_195
171.6 KAN43117_235
191.6 KAN24598_287
DI11
0.0 CHI03186-1_128
8.5 CHI03186-1_124
14.8 CRA_scf12916_311
20.9 VCB-C03938_274
28.6 OPAA-11a
37.2 CA325_227 CA325_205
43.6 SL413
DI12
0.0 AOMT01781-1_135
33.2 OPP-14a
56.7 berc448a
60.7 OPP-12a
65.5 berc781b

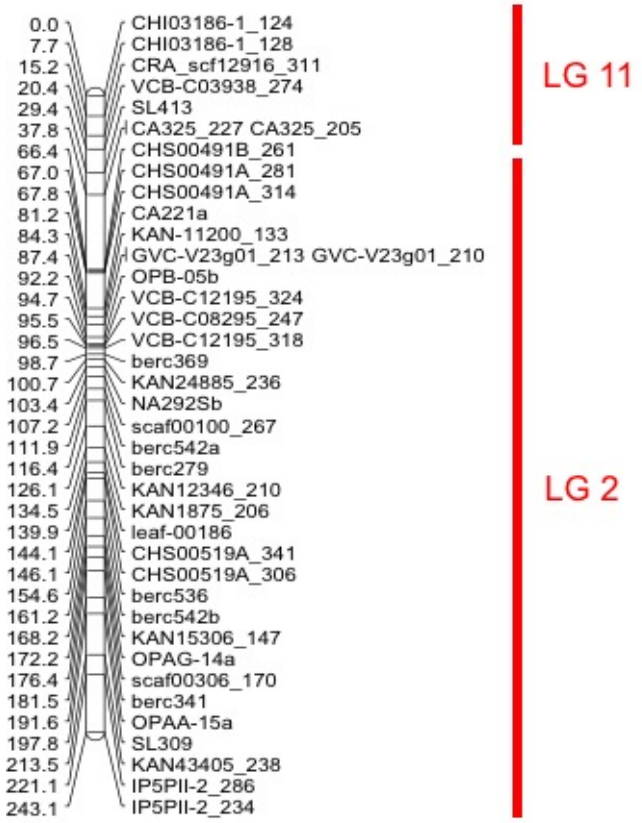


Fig. 4. The integration of LG 2 and LG 11 of diploid blueberry linkage map.

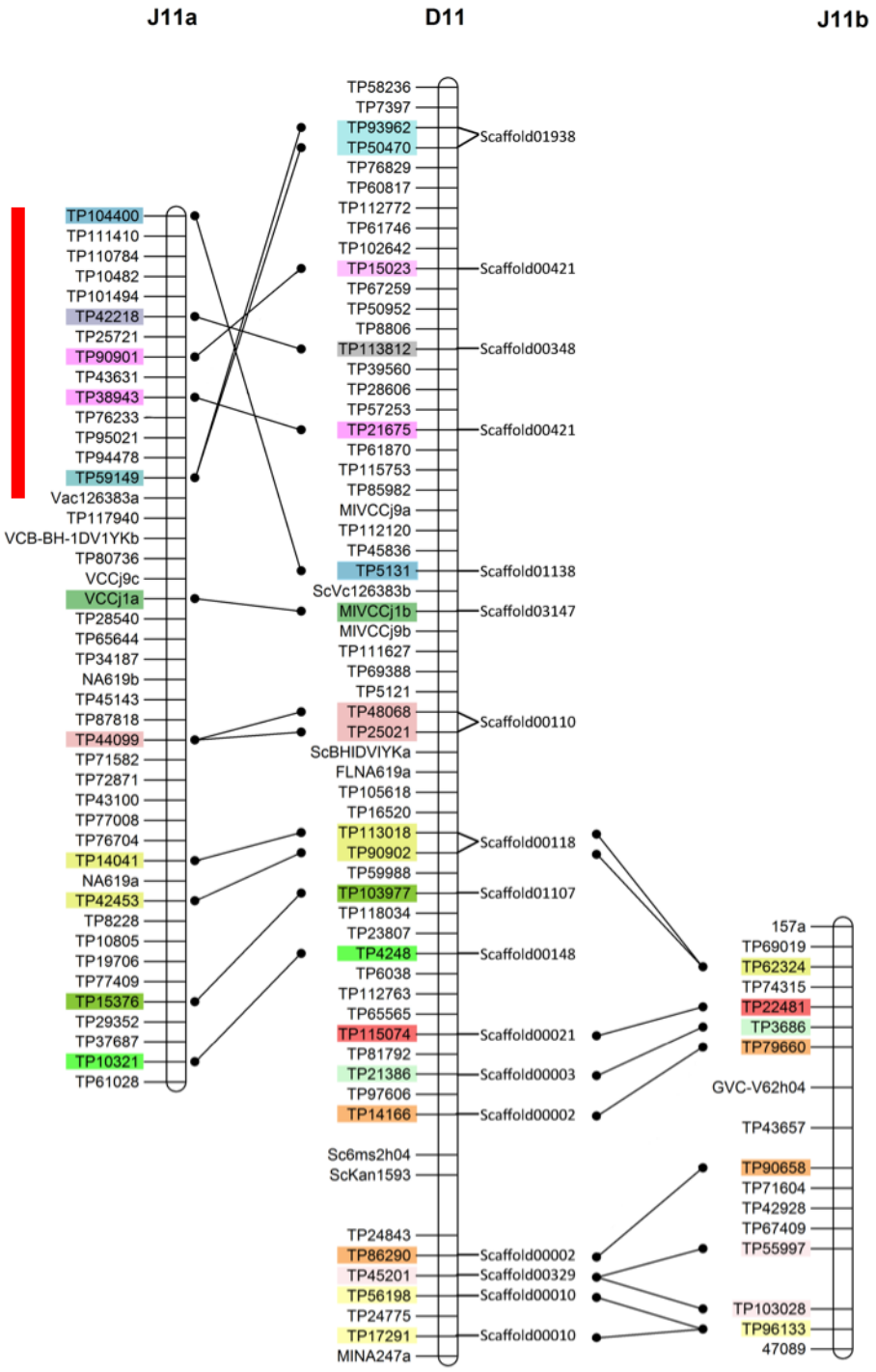


Fig. 5. Scaffold alignment comparison between Draper linkage group 11 (D11) and two Jewel linkage group 11 (J11a and J11b). The red bar indicates the regions of the inverse scaffold order between J11a and D11.

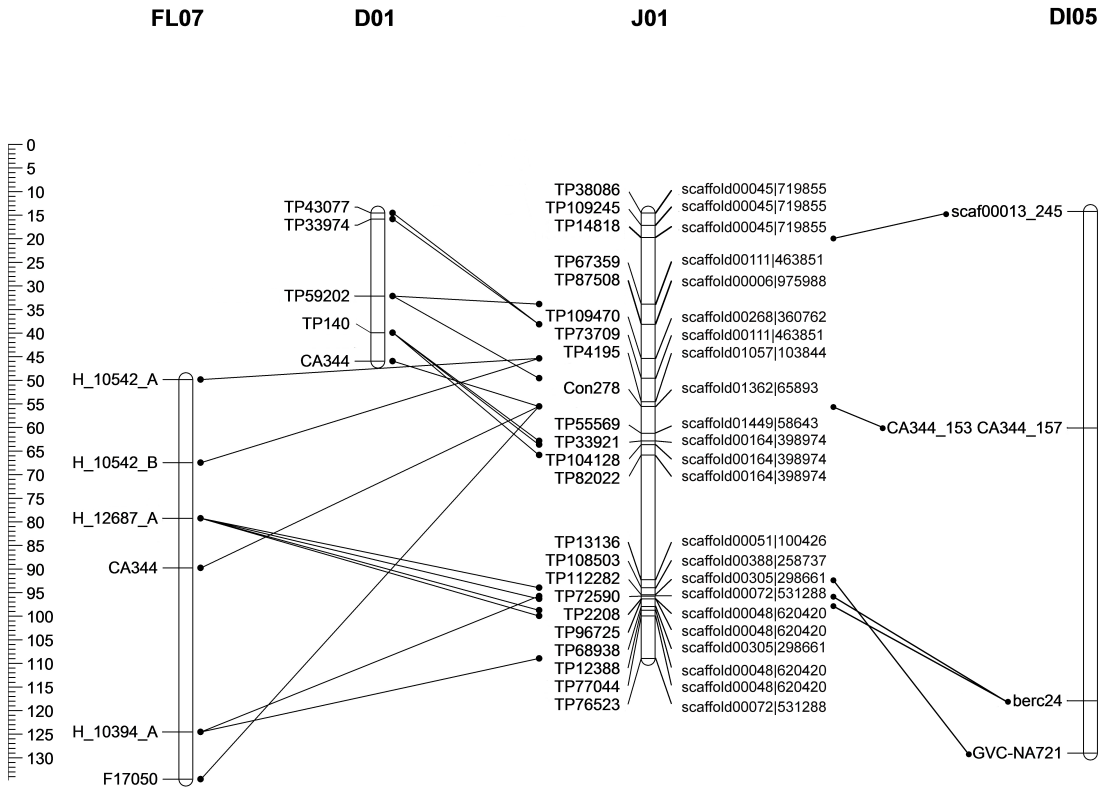


Fig. 6. Consensus linkage maps consisted of diploid blueberry linkage group 5 (DI05), Jewel linkage group 1 (J01), Draper linkage group 1 (D01) and Florida interspecific hybrid linkage group 7 (FL07). Each line that linked to two markers indicates that both markers were assigned to the same scaffold. The number followed scaffold number is the size (bp) of the scaffold.

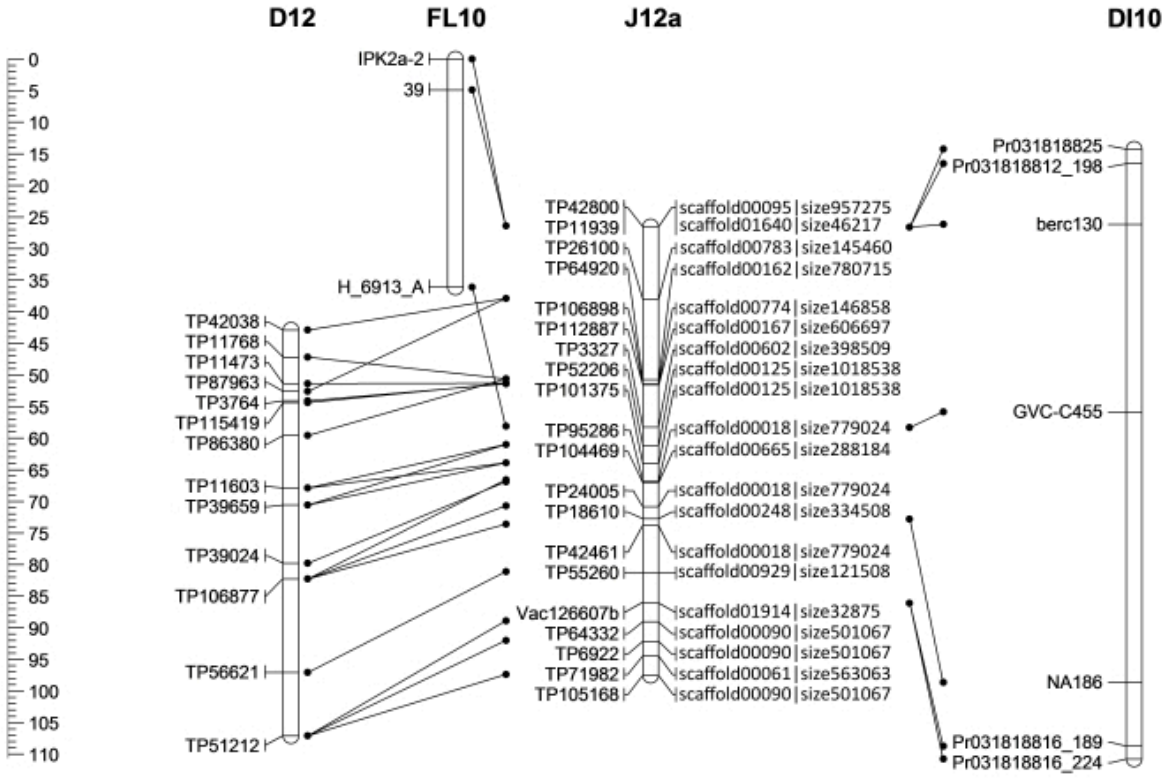
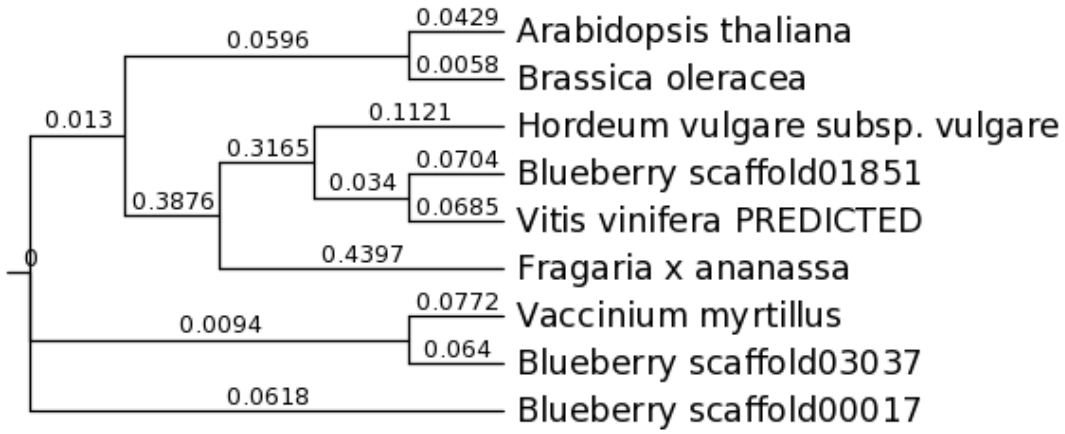
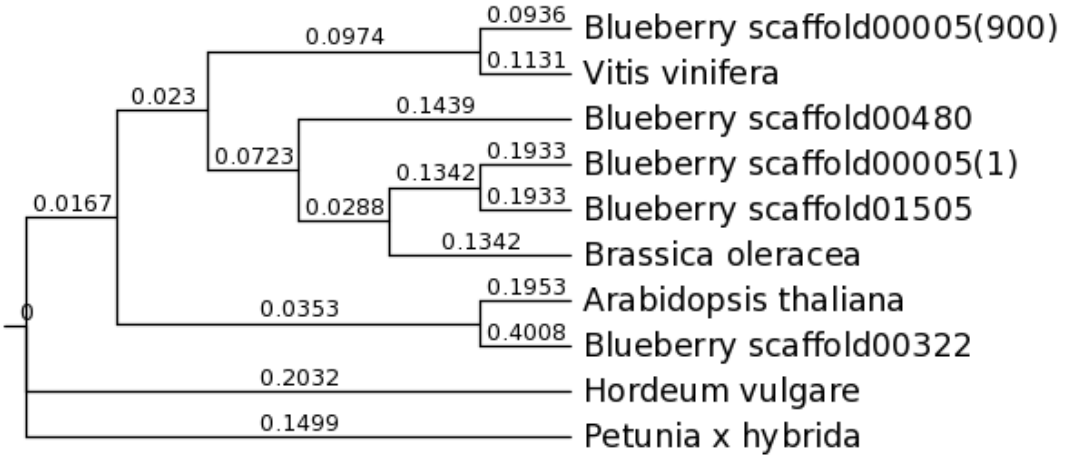


Fig. 7. Consensus linkage maps that consisted of Draper linkage group 12 (D12), interspecific hybrid linkage group 10 (FL 10), Jewel linkage group 12a (J12a), and Diploid linkage group 10 (DI10). Each line that linked to two markers indicates that both markers were assigned to the same scaffold. The number followed scaffold number is the size (bp) of the scaffold.

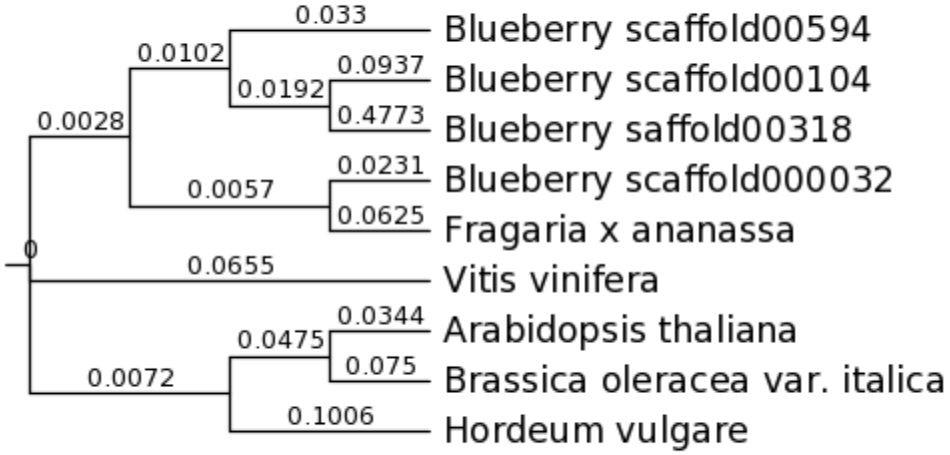
Fig. 8. Phylogenetic tree analysis of genes involved in anthocyanin biosynthesis pathway. Species used in the analysis included putative blueberry candidates (named after the scaffold number), grapevine (*Vitis vinifera*), Aradibopsis (*Arabidopsis thaliana*), barley (*Hordeum vulgare*), broccoli (*Brassica oleracea*), strawberry (*Fragaria* × *ananassa*), wild strawverry (*Fragaria vesca*), maize (*Zea mays*), petunia (*Petunia* × *hybrida*), rapeseed (*Brassica napus*), cranberry (*Vaccinium macrocarpon*), and bilberry (*Vaccinium myrtillus*). (a) phenyalanine ammonia lyase (PAL); (b) 4-coumarate-CoA: ligase(4CL); (c) chalcone synthase (CHS); (d) chalcone isomerase (CHI); (e) flavanone 3-hydroxylase (F3H); (f) flavonoid 3'-hydroxylase (F3'H); (g) flavonoid 3'5'-hydroxylase (F3'5'H); (h) dihydroflavonol-4-reductase (DFR); (i) anthocyanidin synthase (ANS) or leucoanyhocyanidin dioxygenase (LDOX); (j) UPD glucose: flavonoid 3-0-glucosyltransferas (UFGT).


0.0429
Arabidopsis thaliana
0.0596
0.0058
Brassica oleracea
0.013
0.1121
Hordeum vulgare subsp. vulgare
0.3165
0.0704
Blueberry scaffold01851
0.034
0.3876
0.0685
Vitis vinifera PREDICTED
0.4397
0
Fragaria x ananassa
0.0772
Vaccinium myrtillus
0.0094
0.064
Blueberry scaffold03037
0.0618
Blueberry scaffold00017

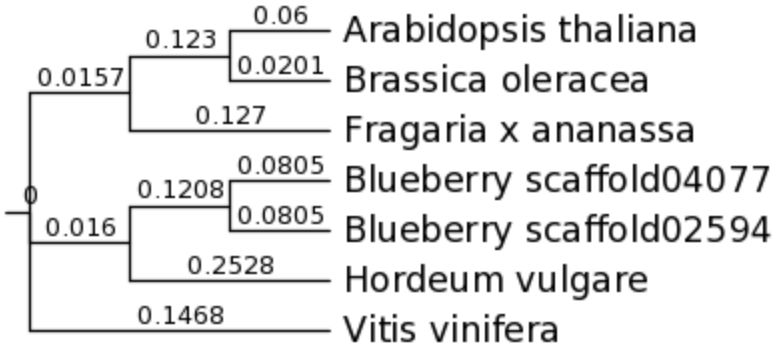
(a)


0.0936
Blueberry scaffold00005(900)
0.0974
0.1131
Vitis vinifera
0.023
0.1439
Blueberry scaffold00480
0.0723
0.1933
Blueberry scaffold00005(1)
0.1342
0.0167
0.1933
Blueberry scaffold01505
0.0288
0.1342
Brassica oleracea
0.1953
Arabidopsis thaliana
0
0.0353
0.4008
Blueberry scaffold00322
0.2032
Hordeum vulgare
0.1499
Petunia x hybrida

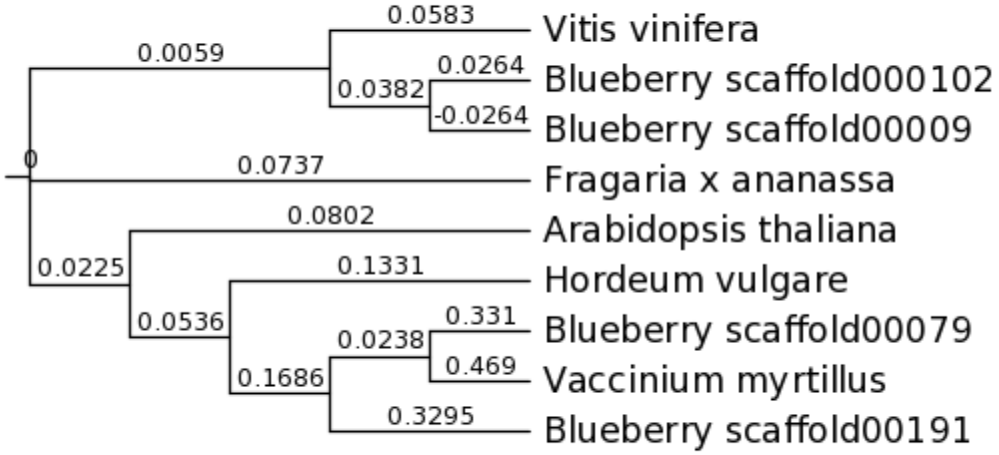
(b)


0.033
Blueberry scaffold00594
0.0102
0.0937
Blueberry scaffold00104
0.0192
0.4773
Blueberry saffold00318
0.0028
0.0231
Blueberry scaffold000032
0.0057
0.0625
Fragaria x ananassa
0
0.0655
Vitis vinifera
0.0344
Arabidopsis thaliana
0.0475
0.075
Brassica oleracea var. italica
0.0072
0.1006
Hordeum vulgare

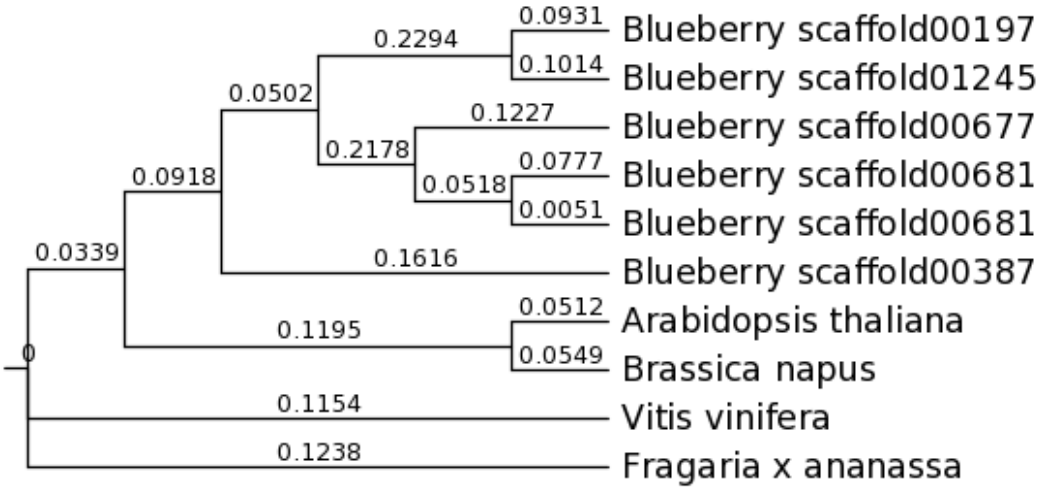
(c)


0.06
Arabidopsis thaliana
0.123
0.0201
Brassica oleracea
0.0157
0.127
Fragaria x ananassa
0.0805
Blueberry scaffold04077
0.1208
0
0.0805
Blueberry scaffold02594
0.016
0.2528
Hordeum vulgare
0.1468
Vitis vinifera

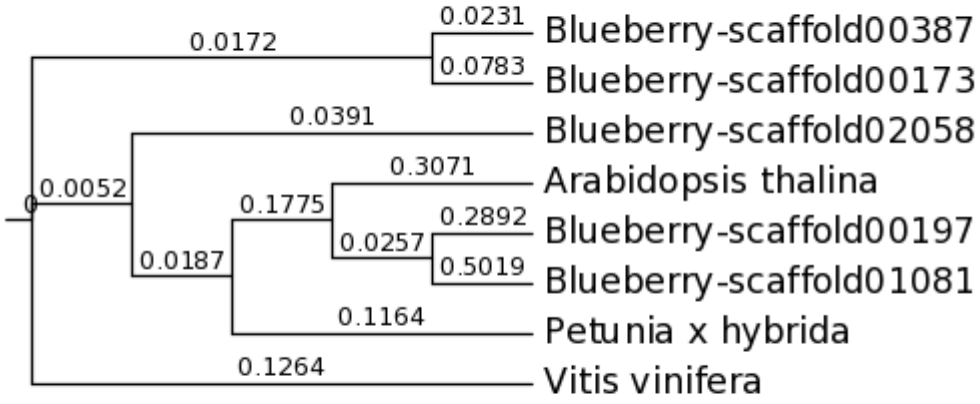
(d)


0.0583
Vitis vinifera
0.0059
0.0264
Blueberry scaffold000102
0.0382
-0.0264
Blueberry scaffold00009
0
0.0737
Fragaria x ananassa
0.0802
Arabidopsis thaliana
0.0225
0.1331
Hordeum vulgare
0.0536
0.331
Blueberry scaffold00079
0.0238
0.469
Vaccinium myrtillus
0.1686
0.3295
Blueberry scaffold00191

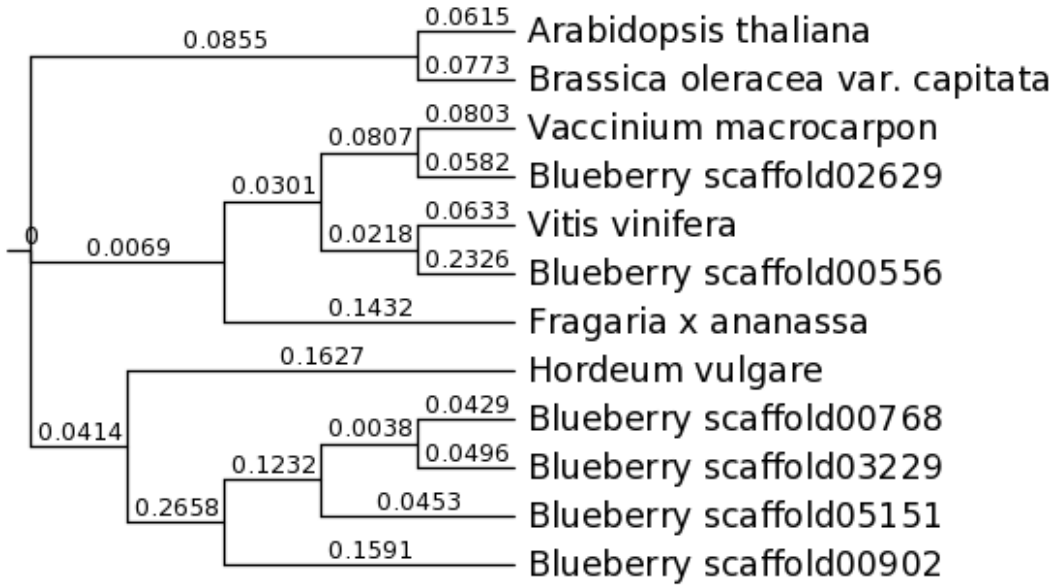
(e)


0.0931
0.2294
Blueberry scaffold00197
0.1014
Blueberry scaffold01245
0.0502
0.1227
Blueberry scaffold00677
0.2178
0.0777
Blueberry scaffold00681
0.0918
0.0518
0.0051
Blueberry scaffold00681
0.0339
0.1616
Blueberry scaffold00387
0.0512
Arabidopsis thaliana
0.1195
0.0549
Brassica napus
0
0.1154
Vitis vinifera
0.1238
Fragaria x ananassa

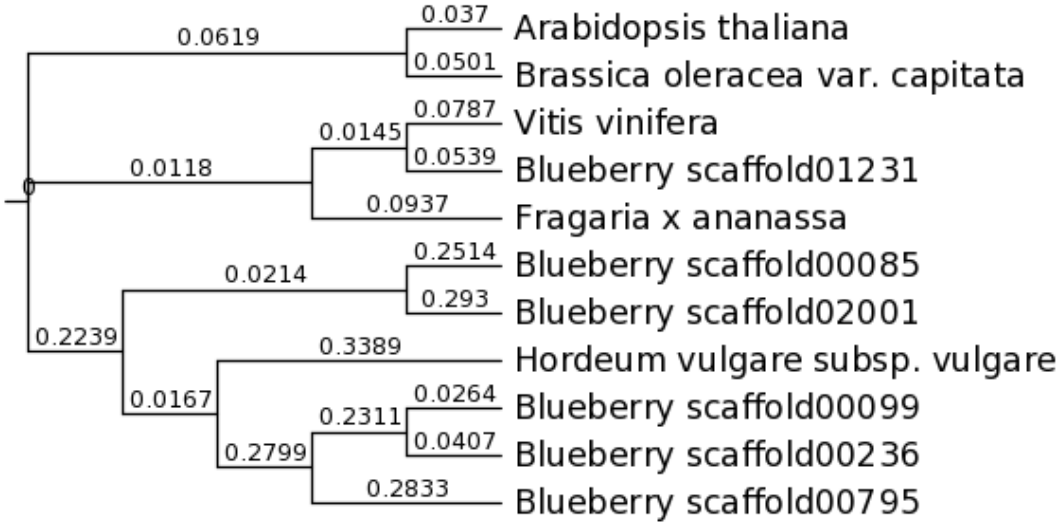
(f)


0.0231
0.0172
Blueberry-scaffold00387
0.0783
Blueberry-scaffold00173
0.0391
Blueberry-scaffold02058
0.3071
Arabidopsis thalina
0.0052
0.1775
0.2892
Blueberry-scaffold00197
0
0.0257
0.0187
0.5019
Blueberry-scaffold01081
0.1164
Petunia x hybrida
0.1264
Vitis vinifera

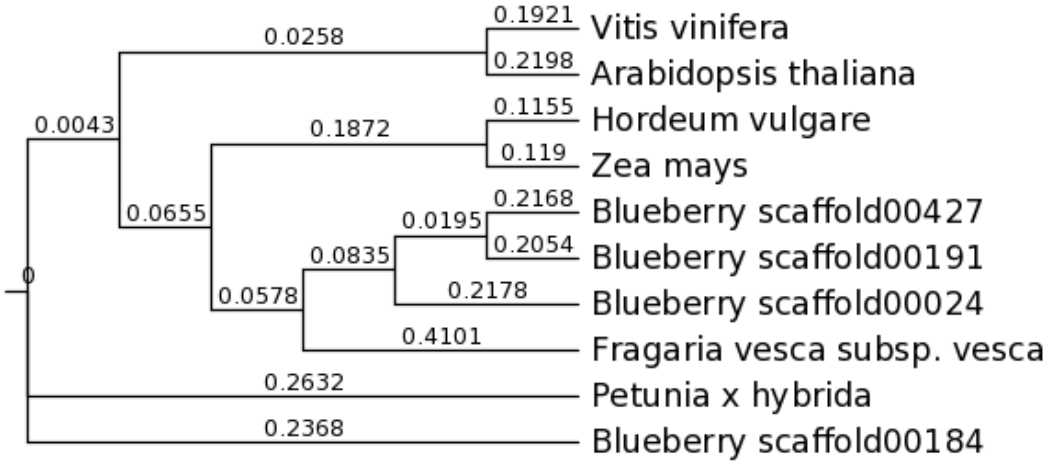
(g)


0.0855
0.0615
0.0773
Arabidopsis thaliana
Brassica oleracea var. capitata
0.0807
0.0803
0.0582
Vaccinium macrocarpon
Blueberry scaffold02629
0.0301
0.0218
0.0633
0.2326
Vitis vinifera
Blueberry scaffold00556
0
0.0069
0.1432
Fragaria x ananassa
0.1627
Hordeum vulgare
0.0414
0.0038
0.0429
0.0496
Blueberry scaffold00768
Blueberry scaffold03229
0.1232
0.2658
0.0453
Blueberry scaffold05151
0.1591
Blueberry scaffold00902

(h)


0.0619
0.037
0.0501
Arabidopsis thaliana
Brassica oleracea var. capitata
0.0145
0.0787
0.0539
Vitis vinifera
Blueberry scaffold01231
0.0118
0
0.0937
Fragaria x ananassa
0.0214
0.2514
0.293
Blueberry scaffold00085
Blueberry scaffold02001
0.2239
0.3389
Hordeum vulgare subsp. vulgare
0.0167
0.0264
0.2311
0.0407
Blueberry scaffold00099
Blueberry scaffold00236
0.2799
0.2833
Blueberry scaffold00795

(i)


0.1921
Vitis vinifera
0.0258
0.2198
Arabidopsis thaliana
0.1155
0.0043
Hordeum vulgare
0.1872
0.119
Zea mays
0.2168
0.0655
Blueberry scaffold00427
0.0195
0.2054
Blueberry scaffold00191
0.0835
0
0.2178
0.0578
Blueberry scaffold00024
0.4101
Fragaria vesca subsp. vesca
0.2632
Petunia x hybrida
0.2368
Blueberry scaffold00184

(j)

Table 1. Taxonomy of 30 sections of genus *Vaccinium* with representative species of each section. Modified from Vander Kloet & Dickinson (2009).

Section	Species	Common Name
Aethopus	*V. paucicrenatum Sleumer* (Himalayas)	
Baccula-nigra sectio nova	*V. fragile* Franch. (Yunnan)	
Barandanum	*V. barandanum* Vidal (Luzon)	
Batodendron (Nutt.)	*V. arboreum Marshall* (SE North America)	Sparkleberry
	Vaccinium crassifolium	Creeping Blueberry
Bracteata	*V. bracteatum* Thunb. (Japan)	
Calcicolus	*V. gaultheriifolium* (Griff.) Hook.f. (Himalayas)	
Ciliata	*V. ciliatum* Thunb. (Japan)	
Cinctosandra (Klotzsch)	*V. madagascariense* (Thouars)	
Conchophyllum	*V. conchophyllum Rehder* (Szechuan)	
	Vaccinium moupinense	Himalayan Blueberry
Cyanococcus A. Gray	*V. corymbosum* L.	
Cyanophthalmos	*V. modestum* W.W.Sm. (Alpine regions of the Himalayas)	
Eococcus	*V. sprengelii* (G. Don) Sleumer (Eastern Asia)	
Epigynium	*V. vacciniaceum* (Roxb.) Sleumer	
Euepigynium	*V. filipes Schltr.* (New Guinea)	
Galeopetalum (J.J. Sm.)	*V. dialypetalum* J.J. Sm. (Sumatra and Malaysia)	
Hemimyrtillus	*V. arctostaphylos* L. (Asia minor, Caucasus)	
Herpothamnus (Small)	*V. crassifolium* Andrews (Carolina)	
Myrtillus	*V. myrtillus* L. (circumboreal)	Common Bilberry, Blue Whortleberry
Neojunghuhnia (Koorders)	*V. insigne* (Koorders) J.J. Sm. (New Guinea)	
Nesococcus	*V. philippinense* Warb. (Luzon)	
Neurodesia (Klotzsch)	*V. crenatum* (D. Don)	
Oarianthe	*V. finisterrae* Schltr. (New Guinea)	

Table 1 continued

Oxycoccoides	*V. erythrocarpum* Michx	
Oxycoccus (Pers.)	*Vaccinium macrocarpon* *V. oxycoccus* L. (circumboreal)	American Cranberry, Common Cranberry
Polycodium	*V. stamineum* L. (E. North America and Central Mexico)	Deerberry
Praestantia	*V. praestans* Lamb. (NE Asia)	
Pyxothamnus (Nutt.)	*V. ovatum Pursh*	
Rigiolepis (Hook.f.)	*V. acuminatissimum* Miq	
Vaccinium L.	*V. uliginosum* L. (circumboreal)	Northern (or Bog) Bilberry (or Blueberry)
Vitis-idaea	*V. vitis-idaea* L.	Partridgeberry, Cowberry, Lingonberry

Table 2. Comparison of next-generation sequencing technology. Modified from Ahn (2001) and Lui et al. (2012).

Platform	Roche 454 GS FLX+	Illumina HiSeq X	SOLiD 5500xl	Ion Torrent PGM	PacBio RS II
Sequencing method	sequencing by synthesis	sequencing by synthesis	sequencing by ligation	sequencing by synthesis	SMRT sequencing
Amplification	Emulsion PCR	Bridge PCR	Emulsion PCR	Emulsion PCR	-
Read length (bp)	700	2 × 150	50 or 60	200 or 400	10-20k
Output (per run)	100 Mb	1.6-1.8 Tb (dual flow cell)	Up to 180 Gb	30 Mb – 2 Gb	500 Mb-1 Gb/cell
Run time	23 hrs	<3 d	6-10 d	2-7 hrs	30 mins – 4 hrs
Pros	Long reads, high accuracy	High throughput	Low error rate	Low cost	Long reads, no amplification needed, fast
Cons	High cost, low throughput	Short reads	Short reads	High cost	High cost, moderate throughput
Primary error	Homopolymer errors, insertion deletions	Single nucleotide substitutions	Substitutions (AT bias)	Homopolymer errors, short insertion deletions	CG deletions

Table 3. Comparison and statistics of two diploid blueberry BAC libraries.

	HVC[a]	BVC[b]	HVC-BVC
Digestion method	Hind *III*	BamH*I*	-
Vector	pCC1BAC	pCC1BAC	-
Number of clones	18,432	20,736	39,168
Average insert size (kb)	120[c]	110[c]	115
Average insert size (kb)	-	-	90[d]
Estimated coverage	3.67×	3.78×	5.85×

[a] BAC library digested by Hind*III* restriction enzyme
[b] BAC library digested by BamH*I* restriction enzyme
[c] the numbers were provided by manufacture
[d] the average insert size was obtained through software calculation

Table 4. Comparison and statistics of BAC-end sequences of two BAC libraries.

	HVC[a]	BVC[b]	HVC-BVC
Number of basepairs	13,776,067	13,689,236	27,464,303
Number of paired sequences	20,301	20,311	40,612
Average length (bp)	678	673	676
Maximum length (bp)	922	843	-
Minimum length (bp)	19	19	-
N50	714	700	-

[a] BAC library digested by Hind*III* restriction enzyme
[b] BAC library digested by BamH*I* restriction enzyme

Table 5. Statistics and comparison of diploid blueberry section W85-23 assembly, Newbler and SSPACE.

	Newbler	SSPACE
Number of raw reads	8,404,249	116.2 billion
Number of basepairs	393,161,046	484,516,305
Number of sequences	13,757	104,711
Maximum sequence length (bp)	1,030,549	1,796,319
Minimum sequence length (bp)	1,935	64
Average sequence length (bp)	28,578	4627.18
N50	145,047	241,296
Estimated coverage	-	>150×[a]

[a] based on personal communication with Dr. Robert Reid

Table 6. Summary of new diploid blueberry linkage map.

Linkage group	Number of markers	Number of new markers	Total length (cM)	Largest gap (cM)	Average distance between marker (cM)
DI01a	28	9	151.144	24.214	5.40
DI01b	34	9	144.015	18.562	4.24
DI02	37	15	198.706	22.464	5.37
DI03	33	7	147.088	30.54	4.46
DI04	24	8	84.953	17.406	3.54
DI05	28	5	192.62	23.114	6.88
DI06	24	4	140.281	18.627	5.85
DI07	29	7	153.53	17.464	5.29
DI08	21	3	117.188	26.046	5.58
DI09	21	7	94.462	20.728	4.50
DI10	31	18	191.593	19.98	6.18
DI11	8	0	43.596	7.744	5.45
Total	318	92	1,659.159	-	-
Average (per LG)	26.5	7.6	138.26	20.57	5.23

Table 7. Statistics of the markers from Jewel and Draper linkage maps located on the 15 largest scaffolds

Scaffold No.	Scaffold size (bp)	Number of Jewel markers	Number of Draper markers	Total of marker	Linkage group	Number of annotated gene[a]
scaffold00001	1,275,046	4	0	4	10	195
scaffold00002	1,526,805	5	7	12	11	248
scaffold00003	1,081,109	1	1	2	11	222
scaffold00004	1,245,574	2	2	4	10	240
scaffold00006	975,988	1	2	3	3	139
scaffold00007	907,789	2	1	3	1	149
scaffold00008	1,199,225	2	4	6	12	201
scaffold00010	843,235	1	2	3	11	178
scaffold00014	887,145	2	4	6	10	147
scaffold00015	808,311	2	2	4	8	143
Total	17,617,494	22	25	47	-	1,862

[a] The genes were obtained through Augustus for automated annotation.

Table 8. Statics of diploid blueberry genomic scaffolds assigned to five blueberry/cranberry genetic linkage maps among top 100, 300, 500, and 1000 scaffolds

	Number of scaffolds assigned to one or more maps	Number of base pairs	Average size of scaffold assigned to map (bp)
Top 100 scaffolds	83	64,393,139	775,824
Top 300 scaffolds	222	128, 972,974	580, 959
Top 500 scaffolds	342	162,111,378	474, 009
Top 1000 scaffolds	510	202,569,426	397,194

Table 9. Diploid blueberry genomic scaffolds of W85-23 assigned to each linkage map.

Linkage Map	Number of markers	Number of scaffolds	Total size (bp)
Tetraploid – Jewel	689	358	121,530,818
Tetraploid – Draper	576	328	112,427,224
Interspecific hybrid	322	190	74,069,152
Diploid	318	153	56,781,319
Cranberry	138	40	15,934,975

Table 10. Comparison of linkage groups that aligned to the same scaffold. Using scaffold00002, scaffold00008, and scaffold00118 as examples.

LG[a]	Marker	Position	Marker distance	Aligned scaffold	Scaffold size (bp)
D11	TP14166	135.5	8 cM	scaffold00002	1526805
	TP41775	137.5			
	TP59962	140.1			
	TP74568	141.7			
	TP86290	143.6			
J11b	TP90658	41.367	8 cM		
	TP15600	42.161			
	TP3322	45.099			
	TP21283	48.238			
	TP79660	50.4			
	TP20971	unmapped			
	TP7285	unmapped			
D12	TP109482	32.7	12 cM	scaffold00008	1199225
	TP64641	37.6			
	TP75932	39.9			
	MIVCCi2b	48.2			
J12b	VCCi2b	24.313	10 cM		
	1099	34.548			
FL12	H2656A	38.81	7 cM		
	H2034B	45.467			
D11	TP113018	70.7	3 cM	scaffold00118	551576
	TP90902	72.8			
J11a	TP14041	36.678	3 cM		
	TP42453	39.606			

[a] The letter stands for linkage map (D for Draper, J for Jewel, and FL for interspecific hybrid), and the number indicates the linkage group.

Table 11. BLAST search results of genes involved in anthocyanin biosynthesis pathway, using grapevine (*Vitis vinifera*) as reference and align against diploid blueberry genomic sequences of W85-23.

Scaffold No.	E-value	Length	Annotation	Locus Name	Entry	Organism
PAL (phenylalanine ammonia lyase)						
Grape Gene ID	100233012					
scaffold01851	6.00E-66	431	71156_t	-	-	-
			model.g25340.t1	-	-	-
scaffold00017	0	1748	CUFF.1944.1	Potri.008G038200.1	EC-4.3.1.24	*Populus trichocarpa*
scaffold03037	0	1748	CUFF.51381.1	Potri.010G224100.1	EC-4.3.1.24	*Populus trichocarpa*
4 CL (4-coumarate:CoA ligase)						
Grape Gene ID	100245991					
scaffold00005a[a]	E-141	1004	CUFF.31671.1	Glyma11g01240.1	EC-6.2.1.12	*Glycine max*
scaffold00480	6.00E-65	653	CUFF.16581.1	cassava4.1_027178m	EC-6.2.1.12	*Manihot esculenta*
scaffold00322	1.00E-64	1232	15863_t	-	-	-
			model.g5996.t1	-	-	-
scaffold00005b[b]	7.00E-48	197	CUFF.4.1	-	gi|297737319	*Vitis vinifera*
scaffold01505	2.00E-44	895	CUFF.46453.1	-	EC-6.2.1 3	*Carica papaya*
CHI (chalcone Isomerase)						
Grape Gene ID	100233078					
scaffold02594	2.00E-29	224	CUFF.50629.1	GSVIVT01032619001	EC-5.5.1.6	*Vitis vinifera*
scaffold04077	1.00E-11	164	-	-	-	-
scaffold01181	2.00E-05	156	CUFF.37529.1	GSVIVT01032685001	EC-5.5.1.6	*Vitis vinifera*
CHS (chalcone synthase/stilbene synthase)						
Grape Gene ID	100217471					
scaffold00032	e-137	964	CUFF.2188.1	cassava4.1_009402m	EC-2.3.1	*Manihot esculenta*

Table 11 continued

scaffold00104	e-136	989	gene.g12175.t1.1	Potri.003G176900.1	EC-2.3.1	*Populus trichocarpa*
scaffold00594	e-132	1767	CUFF.39764.1	cassava4.1_009402m	EC-2.3.1	*Manihot esculenta*
scaffold00318	5.00E-70	950	CUFF.30053.1	GSVIVT01018219001	EC-2.3.1.74	*Vitis vinifera*
DFR (dihydroflavonol 4-reductas)						
Grape Gene ID	100233141					
scaffold02629	4.00E-50	193	CUFF.50634.1	cassava4.1_010685m	EC-1.1.1.219	
scaffold05151	3.00E-29	518	85729_t	-	-	-
			model.g29841.t1	-	-	-
scaffold00768	3.00E-37	696	40172_g.1	-	gi|147769576	*Vitis vinifera*
scaffold03229	3.00E-36	578	gene.scaffold05042.path1.gene1.1		gi|225453895	*Vitis vinifera*
scaffold00902	2.00E-34	541	CUFF.30624.1	cassava4.1_011708m		*Manihot esculenta*
scaffold00556	5.00E-29	195	CUFF.23279.1		gi|225440266	*Vitis vinifera*
			CUFF.23278.1	GSVIVT01009743001	EC-1.1.1.219	*Vitis vinifera*
F3H (flavanone 3-hydroxylase)						
Grape Gene ID	100233079					
scaffold00293	4.00E-76	2758	CUFF.11841.1	cassava4.1_010212m	EC-1.14.11.9	*Manihot esculenta*
scaffold00009	2.00E-74	3468	CUFF.25304.1	AT3G51240.1	EC-1.14.11.9	*Arabidopsis thaliana*
scaffold00102	4.00E-32	200	4982_g.1	-	-	-
scaffold00191	3.00E-29	317	gene.g12745.t1.1	Potri.017G048700.1	EC-1.14.11.9	*Populus trichocarpa*
scaffold00079	2.00E-28	644	CUFF.8120.1	Potri.010G107500.1	-	*Populus trichocarpa*
F3'H (flavonoid 3' hydroxylase)						
Grape Gene ID	100232999					
scaffold00197	2.00E-83	1682	CUFF.28145.1	Potri.016G137400.1	-	*Populus trichocarpa*

Table 11 continued

scaffold00677	4.00E-78	1562	CUFF.42952.1	-	gi|255575503	*Ricinus communis*
			CUFF.42954.1	-	gi|225441018	*Vitis vinifera*
			CUFF.42953.1	-	gi|225441018	*Vitis vinifera*
scaffold00681	7.00E-77	626	CUFF.48276.1	-	gi|255575491	*Ricinus communis*
scaffold00681	1.00E-74	1298	31473_t	-	-	-
			model.g11591.t1	-	-	-
scaffold01245	3.00E-72	1457	CUFF.37869.1	Potri.016G137600.1	-	*Populus trichocarpa*
scaffold00387	1.00E-69	623	gene.g10884.t1.1	cassava4.1_005974m	-	*Manihot esculenta*
scaffold01081	2.00E-64	641	CUFF.37512.1	GSVIVT01033633001	-	*Vitis vinifera*
scaffold00173	6.00E-64	623	gene.g4106.t1.1	-	gi|256574658	*Rhododendron x*
scaffold00412	9.00E-62	524				
F3'5'H (flavonoid-3,5'-hydroxylase)						
Grape Gene ID	100232896		100261319			
scaffold00387	E-100	626				
scaffold02058	4.00E-86	527	CUFF.51711.1	-	gi|256574658	*Rhododendron x*
scaffold00173	1.00E-89	626				
scaffold01081	2.00E-92	894				
scaffold00197	7.00E-78	630				
ANS (anthcyanidin synthase)						
LDOX (leucoanthocyanidin dioxygenase)						
Grape Gene ID	100233142					
scaffold01231	3.00E-67	542	CUFF.43605.1	Bra019350	EC1.14.11.19	*Brassica rapa*
scaffold00795	5.00E-27	440	CUFF.41155.1	GSVIVT01008907001	-	*Vitis vinifera*

Table 11 continued

scaffold02001	9.00E-23	731	CUFF.48069.1	Potri.001G080600.1	EC-1.14.11.9	*Populus trichocarpa*
scaffold00099	3.00E-21	446	13607_t	-	-	-
			model.g5123.t1	-	-	-
scaffold00236	1.00E-20	446	26739_g.1	-	gi|508724321	*Theobroma cacao*
scaffold00085	4.00E-20	386	model.g7342.t1	-	-	-
AOMT (anthocyanin O-methyltransferase)						
Grape Gene ID	100233134	100233087				
scaffold01424	7.00E-27	294	CUFF.43261.1	GSVIVT01015245001	EC-2.1.1.104	*Vitis vinifera*
scaffold00820	9.00E-25	294	CUFF.29739.1	Potri.009G099800.1	EC-2.1.1.104	*Populus trichocarpa*
	1.00E-52	294	-	-	-	-
UFGT (UDP glucose: flavonoid-3-O-glucosyltransferase)						
Grape Gene ID	100233099					
scaffold00184	1.00E-79	1931	CUFF.35538.1	Potri.013G118700.1	EC-2.4.1.115	*Populus trichocarpa*
scaffold00427	2.00E-39	2990	-	-	-	-
scaffold00024	1.00E-38	899	CUFF.22249.1	Potri.006G023100.2	EC-2.4.1.203	*Populus trichocarpa*
			CUFF.22248.2	Potri.016G021000.1	EC-2.4.1.215	*Populus trichocarpa*
			CUFF.22248.1	Glyma19g03451.1	EC-2.4.1.215	*Glycine max*
scaffold00191	9.00E-37	866	CUFF.29439.1	-	gi|460369398	*Solanum lycopersicum*
ANP (anthocyanin permease)						
Grape Gene ID	225544292	100268149				
scaffold00396	1.00E-86	16847				
scaffold011287	5.00E-59	1313	CUFF.60596.1	-	gi|255537876	*Ricinus communis*
scaffold00026	4.00E-52	718	CUFF.42152.2	-	gi|255556131	*Ricinus communis*

Table 11 continued

scaffold00031	1.00E-51	1115	49220_t	-	-	-
			model.g17874.t1	-	-	-
scaffold01562	2.00E-49	584	CUFF.44535.1	-	gi|225470571	*Vitis vinifera*
scaffold00225	3.00E-49	1310	CUFF.38877.1	-	gi|470105037	*Fragaria vesca*
scaffold00111	3.00E-48	569	CUFF.5494.1	-	gi|508708719	*Theobroma cacao*
Myb transcription factors (VvmybA1-3)						
DDBJ Accession	AB427165					
scaffold03342	4.00E-27	251	CUFF.51789.1	-	gi|359300580	*Vaccinium corymbosum*
scaffold00046	4.00E-27	275	CUFF.14288.1	-	gi|269784590	*Diospyros kaki*
scaffold01323	6.00E-24	239	CUFF.43191.1	-	gi|297746469	*Vitis vinifera*
			CUFF.43192.1	-	gi|356551232	*Glycine max*
			CUFF.43193.1	-	gi|356551232	*Glycine max*
scaffold04355	3.00E-23	233	CUFF.52976.1	-	gi|462413412	*Prunus persica*
scaffold03562	4.00E-23	239	CUFF.52597.1	-	gi|13346178	*Gossypium hirsutum*
			CUFF.52595.1	-	gi|225435749	*Vitis vinifera*
			CUFF.52596.1	-	gi|13346178	*Gossypium hirsutum*

[a] the hit was aligned to scaffold00900 in Newbler assembly
[b] the hit was aligned to scaffold00001 in Newbler assembly

Table 12. Putative anthocyanin-related genes of blueberry with scaffold location and association with each linkage map.

Gene	Scaffold No.	Linkage map and linkage group
PAL (phenylalanine ammonia lyase)	scaffold01851	--
	scaffold00017	Draper 11, Interspecific hybrid 08, Cranberry 01
	scaffold03037	--
4CL (4-coumarate:CoA ligase)	scaffold00005	Draper 02, Cranberry 04, Diploid 04
	scaffold00480	--
	scaffold00322	--
	scaffold00005	Draper 02, Cranberry 04, Diploid 04
	scaffold01505	--
CHI (chalcone isomerase)	scaffold02594	Diploid 11
	scaffold04077	--
	scaffold01181	Cranberry 04
CHS (chalcone synthase /stilbene synthase)	scaffold00032	Jewel 01
	scaffold00104	Jewel 03
	scaffold00594	--
	scaffold00318	--
DFR (dihydroflavonol 4-reductas)	scaffold02629	--
	scaffold05151	--
	scaffold00768	Interspecific hybrid 04
	scaffold03229	--
	scaffold00902	Draper 07
	scaffold00556	Draper 02
F3H (flavanone 3-hydroxylase)	scaffold00293	Jewel 06
	scaffold00009	--
	scaffold00102	--
	scaffold00191	Draper 05, Jewel 05, Diploid 02
	scaffold00079	Draper 02
F3'H (flavonoid 3' hydroxylase)	scaffold00197	Interspecific hybrid 06
	scaffold00677	--
	scaffold00681	Jewel 10
	scaffold00681	--
	scaffold01245	--
	scaffold00387	--
	scaffold01081	Jewel 10, Cranberry 08
	scaffold00173	--
	scaffold00412	--
F3'5'H	scaffold00387	--

Table 12 continued

(flavonoid-3,5'-hydroxylase)	scaffold02058	--
	scaffold00173	Florida 05
	scaffold01081	Jewel 10, Cranberry 08
	scaffold00197	Jewel 03, Florida 06
ANS (anthcyanidin synthase)	scaffold01231	Draper 09
	scaffold00795	--
	scaffold02001	--
LDOX (leucoanthocyanidin dioxygenase)	scaffold00099	--
	scaffold00236	Draper 03, Jewel 04, Interspecific hybrid 04
	scaffold00085	Draper 04
AMOT (anthocyanin O-methyltransferase)	scaffold01424	--
	scaffold00820	Draper 12
UFGT (UDP glucose: flavonoid-3-O-glucosyltransferase)	scaffold00184	Draper 02, Interspecific hybrid 05, Diploid 01
	scaffold00427	--
	scaffold00024	--
	scaffold00191	Draper 05, Jewel 05, Diploid 02
ANP (anthocyanin permease)	scaffold00396	Draper 04, Interspecific hybrid 03
	scaffold011287	--
	scaffold00026	Draper 05
	scaffold00031	Draper 10, Interspecific hybrid 06
	scaffold01562	--
	scaffold00225	--
	scaffold00111	Jewel 01
Myb transcription factor	scaffold03342	--
	scaffold00046	Jewel 10
	scaffold01323	--
	scaffold04355	--
	scaffold03562	--

REFERENCES

Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., . . . Moreno, R. F. (1991). Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science (New York, N.Y.), 252*(5013), 1651-1656.

Ahn, S. (2011). Introduction to bioinformatics: Sequencing technology. *Asia Pacific Allergy, 1*(2), 93-97. doi:10.5415/apallergy.2011.1.2.93 [doi]

Anjos, e. S., Antunes, L. E. C., Anthonisen, D. G., Lemões, J. S., & Gonçalves, E. D. (2008). Characterization of blueberry genotypes using molecular markers. / caracterização de genótipos de mirtilo utilizando marcadores moleculares. *Revista Brasileira De Fruticultura, 30*(1), 180-184.

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature, 408*(6814), 796-815. doi:10.1038/35048692.

Armour, J., Alegre, S., Miles, S., Williams, L., & Badge, R. (1999). Minisatellites and mutation processes in tandemly repetitive DNA. *Microsatellites: Evolution and Applications.Oxford University Press, Oxford*, 24-33.

Aruna, M., Austin, M. E., & Ozias-Akins, P. (1995). Randomly amplified polymorphic DNA fingerprinting for identifying rabbiteye blueberry (*Vaccinium ashei* reade) cultivars. *Journal of the American Society for Horticultural Science, 120*(5), 710-713.

Aruna, M., Ozias-Akins, P., Austin, M. E., & Kochert, G. (1993). Genetic relatedness among rabbiteye blueberry (*Vaccinium ashei*) cultivars determined by DNA amplification using single primers of arbitrary sequence. *Genome, 36*(5), 971-977.

Ayeh, K. O. (2008). Expressed sequence tags (ESTs) and single nucleotide polymorphisms (SNPs): Emerging molecular marker tools for improving agronomic traits in plant biotechnology. *African Journal of Biotechnology, 7*(4)

Aza-González, C., Herrera-Isidrón, L., Núñez-Palenius, H., De La Vega, O Martínez, & Ochoa-Alejo, N. (2013). Anthocyanin accumulation and expression analysis of biosynthesis-related genes during chili pepper fruit development. *Biologia Plantarum, 57*(1), 49-55.

Azuma, A., Kobayashi, S., Yakushui, H., Yamada, M., Mitani, N., & Sato, A. (2007). VvmybA1 genotype determines grape skin color. *Vitis-Geilweilerhof-, 46*(3), 154.

Azuma, A., Kobayashi, S., Mitani, N., Shiraishi, M., Yamada, M., Ueno, T., . . . Koshita, Y. (2008). Genomic and genetic analysis of myb-related genes that regulate anthocyanin biosynthesis in grape berry skin. *Theoretical and Applied Genetics, 117*(6), 1009-1019.

Ball, A. D., Stapley, J., Dawson, D. A., Birkhead, T. R., Burke, T., & Slate, J. (2010). A comparison of SNPs and microsatellites as linkage mapping markers: Lessons from the zebra finch (*Taeniopygia guttata*). *BMC Genomics, 11*, 218-2164-11-218. doi:10.1186/1471-2164-11-218 [doi]

Bassil, N., Oda, A., & Hummer, K. (2008). Blueberry microsatellite markers identify cranberry cultivars. Paper presented at the *IX International Vaccinium Symposium 810,* 181-187.

Bassil, N., Oda, A., & Hummer, K. E. (2009). Blueberry microsatellite markers identify cranberry cultivars. *Acta Horticulturae,* (810), 181-186.

Bassil, N. V. (2012). Microsatellite markers: Valuable in *Vaccinium* L. *International Journal of Fruit Science, 12*(1), 288-293.

Bell, D. J., Drummond, F. A., Polashock, J. J., & Rowland, L. J. (2008). Suitability of EST-PCR markers developed in highbush blueberry for genetic fingerprinting and relationship studies in lowbush blueberry and related species. *Journal of the American Society for Horticultural Science, 133*(5), 701-707. doi:http://hdl.handle.net/10113/22526

Bhat, P. R., Krishnakumar, V., Hendre, P. S., Rajendrakumar, P., Varshney, R. K., & Aggarwal, R. K. (2005). Identification and characterization of expressed sequence tags−derived simple sequence repeats markers from robusta coffee variety 'CxR'(an interspecific hybrid of *Coffea canephora*× *Coffea congensis*). *Molecular Ecology Notes, 5*(1), 80-83.

Bian, Y., Ballington, J., Raja, A., Brouwer, C., Reid, R., Burke, M., . . . Brown, A. (2014). Patterns of simple sequence repeats in cultivated blueberries (*Vaccinium* section Cyanococcus spp.) and their use in revealing genetic diversity and population structure. *Molecular Breeding, 34*(2), 675-689.

Boches, P., Bassil, N., & Rowland, L. (2005). Microsatellite markers for *Vaccinium* from EST and genomic libraries. *Molecular Ecology Notes, 5*(3), 657-660.

Boches, P., Rowland, L., & Bassil, N. V. (2006a). Genetic diversity in the highbush blueberry evaluated with microsatellite markers. *Journal of the American Society for Horticultural Science, 131*(5), 674-686. doi:http://hdl.handle.net/10113/21627

Boches, P. S., Bassil, N. V., Hummer, K., & Rowland, J. (2006b). Cross-species amplification of SSRs in the genus *Vaccinium*. *Acta Horticulturae,* (715), 119-127.

Brazelton, C. (2011). World blueberry acreage & production. *US Highbush Blueberry Council,*

Brevis, P. A., Hancock, J. F., Ballington, J. R., & Bassil, N. V. (2008). Impact of wide hybridization on highbush blueberry breeding. *Journal of the American Society for Horticultural Science, 133*(3), 427-437. doi:http://hdl.handle.net/10113/19419

Brooks, S., & Lyrene, P. (1995). Characteristics of sparkleberry x blueberry hybrids. Paper presented at the *Proceedings-Florida State Horticultural Society. 108* 337-338.

Brooks, S. J., & Lyrene, P. M. (1998a). Derivatives of *Vaccinium arboreum*× *Vaccinium* section Cyanococcus: I. morphological characteristics. *Journal of the American Society for Horticultural Science, 123*(2), 273-277.

Brooks, S. J., & Lyrene, P. M. (1998b). Derivatives of *Vaccinium arboreum*× *Vaccinium* section Cyanococcus. II. fertility and fertility parameters. *Journal of the American Society for Horticultural Science, 123*(6), 997-1003.

Brown, A. F., Yousef, G. G., Chebrolu, K. K., Byrd, R. W., Everhart, K. W., Thomas, A., . . . Oliver, R. (2014). High-density single nucleotide polymorphism (SNP) array mapping in *Brassica oleracea*: Identification of QTL associated with carotenoid variation in broccoli florets. *Theoretical and Applied Genetics, 127*(9), 2051-2064.

Brown, T. (2002). Genomes. 2nd edition, Chapter 6, sequencing genomes. (Oxford: Wiley-Liss)

Burgher, K. L., Jamieson, A. R., & Lu, X. (2002). Genetic relationships among lowbush blueberry genotypes as determined by randomly amplified polymorphic DNA analysis. *Journal of the American Society for Horticultural Science, 127*(1), 98-103.

Camp, W. (1945). The north American blueberries with notes on other groups of *Vacciniaceae*. *Brittonia, 5*(3), 203-275.

Chabane, K., Ablett, G. A., Cordeiro, G. M., Valkoun, J., & Henry, R. J. (2005). EST versus genomic derived microsatellite markers for genotyping wild and cultivated barley. *Genetic Resources and Crop Evolution, 52*(7), 903-909.

Chambers, G. K., & MacAvoy, E. S. (2000). Microsatellites: Consensus and controversy. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 126*(4), 455-476.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution, 24*(3), 621-631. doi:msl191 [pii]

Cho, E., Seddon, J. M., Rosner, B., Willett, W. C., & Hankinson, S. E. (2004). Prospective study of intake of fruits, vegetables, vitamins, and carotenoidsand risk of age-related maculopathy. *Archives of Ophthalmology, 122*(6), 883-892.

Cho, Y. G., Ishii, T., Temnykh, S., Chen, X., Lipovich, L., McCOUCH, S. R., . . . Cartinhour, S. (2000). Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics, 100*(5), 713-722.

Collard, B., Jahufer, M., Brouwer, J., & Pang, E. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica, 142*(1-2), 169-196.

Collard, B. C., & Mackill, D. J. (2008). Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society of London.Series B, Biological Sciences, 363*(1491), 557-572. doi:8731828765572240 [pii]

Connor, A. M., Luby, J. J., Tong, C. B., Finn, C. E., & Hancock, J. F. (2002). Genotypic and environmental variation in antioxidant activity, total phenolic content, and anthocyanin content among blueberry cultivars. *Journal of the American Society for Horticultural Science, 127*(1), 89-97.

Costich, D., Ortiz, R., Meagher, T., Bruederle, L. P., & Vorsa, N. (1993). Determination of ploidy level and nuclear DNA content in blueberry by flow cytometry. *Theoretical and Applied Genetics, 86*(8), 1001-1006.

Coulibaly, I., Gharbi, K., Danzmann, R., Yao, J., & Rexroad, C. (2005). Characterization and comparison of microsatellites derived from repeat−enriched libraries and expressed sequence tags. *Animal Genetics, 36*(4), 309-315.

Coville, F. V. (1921). *Directions for blueberry culture, 1921* US Department of Agriculture.

Dakin, E., & Avise, J. (2004). Microsatellite null alleles in parentage analysis. *Heredity, 93*(5), 504-509.

Darrow, G. M., Camp, W., Fischer, H., & Dermen, H. (1944). Chromosome numbers in *Vaccinium* and related groups. *Bulletin of the Torrey Botanical Club,* 498-506.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics, 12*(7), 499-510.

Debnath, S. C. (2009). Development of ISSR markers for genetic diversity studies in *Vaccinium angustifolium*. *Nordic Journal of Botany, 27*(2), 141-148.

DeFaveri, J., Viitaniemi, H., Leder, E., & Merilä, J. (2013). Characterizing genic and nongenic molecular markers: Comparison of microsatellites and SNPs. *Molecular Ecology Resources, 13*(3), 377-392.

Dhanaraj, A. L., Alkharouf, N. W., Beard, H. S., Chouikha, I. B., Matthews, B. F., Wei, H., . . . Rowland, L. J. (2007). Major differences observed in transcript profiles of blueberry during cold acclimation under field and cold room conditions. *Planta, 225*(3), 735-751.

Dhanaraj, A. L., Slovin, J. P., & Rowland, L. J. (2004). Analysis of gene expression associated with cold acclimation in blueberry floral buds using expressed sequence tags. *Plant Science, 166*(4), 863-872.

Die, J. V., & Rowland, L. J. (2013). Advent of genomics in blueberry. *Molecular Breeding, 32*(3), 493-504.

Draper, A., & Hancock, J. (2003). Florida 4B: Native blueberry with exceptional breeding value. *Journal of American Pomological Society,*

Finn, C. E., Luby, J. J., Rosen, C. J., & Ascher, P. D. (1993). Blueberry germplasm screening at several soil pH regimes. I. plant survival and growth. *Journal of the American Society for Horticultural Science, 118*(3), 377-382.

Gabaldón, T., & Koonin, E. V. (2013). Functional and evolutionary implications of gene orthology. *Nature Reviews Genetics, 14*(5), 360-366.

Garvin, M., Saitoh, K., & Gharrett, A. (2010). Application of single nucleotide polymorphisms to non−model species: A technical review. *Molecular Ecology Resources, 10*(6), 915-934.

Georgi, L., Herai, R. H., Vidal, R., Carazzolle, M. F., Pereira, G. G., Polashock, J., & Vorsa, N. (2012). Cranberry microsatellite marker development from assembled next-generation genomic sequence. *Molecular Breeding, 30*(1), 227-237.

Georgi, L., Johnson-Cicalese, J., Honig, J., Das, S. P., Rajah, V. D., Bhattacharya, D., . . . Vorsa, N. (2013). The first genetic map of the American cranberry: Exploration of

synteny conservation and quantitative trait loci. *Theoretical and Applied Genetics, 126*(3), 673-692.

Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R., Dunn, M., . . . Briggs, S. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science (New York, N.Y.), 296*(5565), 92-100. doi:10.1126/science.1068275.

Gonzali, S., Mazzucato, A., & Perata, P. (2009). Purple as a tomato: Towards high anthocyanin tomatoes. *Trends in Plant Science, 14*(5), 237-241.

Gould, K., Davies, K. M., & Winefield, C. (2008). *Anthocyanins: Biosynthesis, functions, and applications* Springer Science & Business Media.

Green, E. D. (2001). Strategies for the systematic sequencing of complex genomes. *Nature Reviews Genetics, 2*(8), 573-583.

Green, P. (1997). Against a whole-genome shotgun. *Genome Research, 7*(5), 410-417.

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

Gupta, P., Rustgi, S., Sharma, S., Singh, R., Kumar, N., & Balyan, H. (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Molecular Genetics and Genomics, 270*(4), 315-323.

Gupta, V., Estrada, A., Blakley, I., Reid, R., Patel, K., Meyer, M., . . . Loraine, A. (2015). RNA-seq analysis and annotation of a draft blueberry genome assembly identifies candidate genes involved in fruit ripening, biosynthesis of bioactive compounds, and stage-specific alternative splicing. *Gigascience, 4*(1), 1-22. doi:10.1186/s13742-015-0046-9

Hatier, J. B., & Gould, K. S. (2009). Anthocyanin function in vegetative organs. *Anthocyanins* (pp. 1-19) Springer.

He, F., Mu, L., Yan, G., Liang, N., Pan, Q., Wang, J., . . . Duan, C. (2010). Biosynthesis of anthocyanins and their regulation in colored grapes. *Molecules, 15*(12), 9057-9091.

He, J., & Giusti, M. M. (2010). Anthocyanins: Natural colorants with health-promoting properties. *Annual Review of Food Science and Technology, 1*, 163-187.

Hiirsalmi, H. (1977). Inheritance of characters in hybrids of *Vaccinium uliginosum* and highbush blueberries. *Annales Agriculturae Fenniae (Finland),*

Hokanson, K., & Hancock, J. (1993). The common lowbush blueberry, *Vaccinium angustifolium* Aiton, may be an allopolyploid. *Canadian Journal of Plant Science, 73*(3), 889-891.

Holton, T. A., & Cornish, E. C. (1995). Genetics and biochemistry of anthocyanin biosynthesis. *The Plant Cell, 7*(7), 1071.

Huang, S., Ding, J., Deng, D., Tang, W., Sun, H., Liu, D., . . . Meng, M. (2013). Draft genome of the kiwifruit *Actinidia chinensis*. *Nature Communications, 4*

Jaakola, L., Maatta, K., Pirttila, A. M., Torronen, R., Karenlampi, S., & Hohtola, A. (2002). Expression of genes involved in anthocyanin biosynthesis in relation to anthocyanin, proanthocyanidin, and flavonol levels during bilberry fruit development. *Plant Physiology, 130*(2), 729-739. doi:10.1104/pp.006957.

Jarne, P., & Lagoda, P. J. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution, 11*(10), 424-429.

Jones, C., Edwards, K., Castaglione, S., Winfield, M., Sala, F., Van de Wiel, C., . . . Daly, A. (1997). Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding, 3*(5), 381-390.

Kalt, W., Joseph, J. A., & Shukitt-Hale, B. (2007). Blueberries and human health: A review of current research. *Journal-American Pomological Society, 61*(3), 151.

Kalt, W., MacKinnon, S., McDonald, J., Vinqvist, M., Craft, C., & Howell, A. (2008). Phenolics of *Vaccinium* berries and other fruit crops. *Journal of the Science of Food and Agriculture, 88*(1), 68-76.

Kobayashi, S., Goto Yamamoto, N., & Hirochika, H. (2005). Association of VvmybA1 gene expression with anthocyanin production in grape (*Vitis vinifera*) skin-color mutants. *Journal of the Japanese Society for Horticultural Science (Japan),*

Kobayashi, S., Ishimaru, M., Hiraoka, K., & Honda, C. (2002). Myb-related genes of the kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis. *Planta, 215*(6), 924-933.

Kobayashi, S., Goto-Yamamoto, N., & Hirochika, H. (2004). Retrotransposon-induced mutations in grape skin color. *Science (New York, N.Y.), 304*(5673), 982. doi:10.1126/science.1095011.

Koes, R., Verweij, W., & Quattrocchio, F. (2005). Flavonoids: A colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science, 10*(5), 236-242.

Koonin, E. V. (2005). Orthologs, paralogs, and evolutionary genomics 1. *Annu.Rev.Genet., 39*, 309-338.

Kumar, P., Gupta, V., Misra, A., Modi, D., & Pandey, B. (2009). Potential of molecular markers in plant biotechnology. *Plant Omics J, 2*(4), 141-162.

Kumar, S., Banks, T. W., & Cloutier, S. (2012). SNP discovery through next-generation sequencing and its applications. *International Journal of Plant Genomics, 2012*

Lagercrantz, U., Ellegren, H., & Andersson, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research, 21*(5), 1111-1115.

Lee, J., Finn, C. E., & Wrolstad, R. E. (2004). Anthocyanin pigment and total phenolic content of three *Vaccinium* species native to the pacific northwest of north America. *Hortscience, 39*(5), 959-964.

Leigh, F., Lea, V., Law, J., Wolters, P., Powell, W., & Donini, P. (2003). Assessment of EST-and genomic microsatellite markers for variety discrimination and genetic diversity studies in wheat. *Euphytica, 133*(3), 359-366.

Lepiniec, L., Debeaujon, I., Routaboul, J., Baudry, A., Pourcel, L., Nesi, N., & Caboche, M. (2006). Genetics and biochemistry of seed flavonoids. *Annu.Rev.Plant Biol., 57*, 405-430.

Levi, A., & Rowland, L. (1997). Identifying blueberry cultivars and evaluating their genetic relationships using randomly amplified polymorphic DNA (RAPD) and simple sequence repeat-(SSR-) anchored primers. *Journal of the American Society for Horticultural Science, 122*(1), 74-78.

Litt, M., & Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics, 44*(3), 397-401.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., . . . Law, M. (2012). Comparison of next-generation sequencing systems. *BioMed Research International, 2012*

Lyrene, P. M., & Brooks, S. J. (1996). Use of sparkleberry in breeding highbush blueberry cultivars. *Journal of Small Fruit & Viticulture, 3*(2-3), 29-38.

Lyrene, P. M. (1997). Value of various taxa in breeding tetraploid blueberries in Florida. *Euphytica, 94*(1), 15-22.

Lyrene, P. M. (2011). First report of *Vaccinium arboreum* hybrids with cultivated highbush blueberry. *Hortscience, 46*(4), 563-566.

Lyrene, P. (1991). Fertile derivatives from sparkleberry× blueberry crosses. *Journal of the American Society for Horticultural Science, 116*(5), 899-902.

Lyrene, P., & Olmstead, J. (2012). The use of inter-sectional hybrids in blueberry breeding. *International Journal of Fruit Science, 12*(1-3), 269-275.

Lyrene, P., & Perry, J. (1982). Production and selection of blueberry polyploids in vitro. *Journal of Heredity, 73*(5), 377-378.

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in Genetics, 24*(3), 133-141.

McCallum, S., Woodhead, M., Jorgensen, L., Gordon, S., Brennan, R., Graham, J., . . . Olmstead, J. W. (2012). Developing tools for long-term breeding of blueberry germplasm for UK production. *International Journal of Fruit Science, 12*(1-3), 294-303.

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature Reviews Genetics, 11*(1), 31-46.

Metzker, M. L. (2005). Emerging technologies in DNA sequencing. *Genome Research, 15*(12), 1767-1776. doi:15/12/1767.

Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J. H., . . . Lewis, K. L. (2008). The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* linnaeus). *Nature, 452*(7190), 991-996.

Moerman, D. E. (1998). *Native American ethnobotany* Timber Press Portland.

Moyer, R. A., Hummer, K. E., Finn, C. E., Frei, B., & Wrolstad, R. E. (2002). Anthocyanins, phenolics, and antioxidant capacity in diverse small fruits: *Vaccinium*, *Rubus*, and *Ribes*. *Journal of Agricultural and Food Chemistry, 50*(3), 519-525.

Nagaraj, S. H., Gasser, R. B., & Ranganathan, S. (2007). A hitchhiker's guide to expressed sequence tag (EST) analysis. *Briefings in Bioinformatics, 8*(1), 6-21. doi:bbl015.

Oetting, W. S., Lee, H., Flanders, D., Wiesner, G., Sellers, T., & King, R. (1995). Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics, 30*(3), 450-458.

Olmstead, J. W., Armenta, H. P. R., & Lyrene, P. M. (2013). Using sparkleberry as a genetic source for machine harvest traits for southern highbush blueberry. *Horttechnology, 23*(4), 419-424.

Pashley, C. H., Ellis, J. R., McCauley, D. E., & Burke, J. M. (2006). EST databases as a source for molecular markers: Lessons from helianthus. *The Journal of Heredity, 97*(4), 381-388. doi:esl013.

Penner, G. A., Bush, A., Wise, R., Kim, W., Domier, L., Kasha, K., . . . Fedak, G. (1993). Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. *PCR Methods and Applications, 2*(4), 341-345.

Polashock, J., Zelzion, E., Fajardo, D., Zalapa, J., Georgi, L., Bhattacharya, D., & Vorsa, N. (2014). The American cranberry: First insights into the whole genome of a species adapted to bog habitat. *BMC Plant Biology, 14*, 165-2229-14-165. doi:10.1186/1471-2229-14-165.

Pompanon, F., Bonin, A., Bellemain, E., & Taberlet, P. (2005). Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics, 6*(11), 847-846.

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

Prior, R. L., Cao, G., Martin, A., Sofic, E., McEwen, J., O'Brien, C., . . . Krewer, G. (1998). Antioxidant capacity as influenced by total phenolic and anthocyanin content, maturity, and variety of *Vaccinium* species. *Journal of Agricultural and Food Chemistry, 46*(7), 2686-2693.

Prior, R. L., & Wu, X. (2006). Anthocyanins: Structural characteristics that result in unique metabolic patterns and biological activities. *Free Radical Research, 40*(10), 1014-1028.

Qu, L., Hancock, J., & Whallon, J. (1998). Evolution in an autopolyploid group displaying predominantly bivalent pairing at meiosis: Genomic similarity of diploid *Vaccinium darrowi* and autotetraploid *V. corymbosum* (*Ericaceae*). *American Journal of Botany, 85*(5), 698. doi:85/5/698 [pii]

Rassmann, K., Schlötterer, C., & Tautz, D. (1991). Isolation of simple−sequence loci for use in polymerase chain reaction−based DNA fingerprinting. *Electrophoresis, 12*(2−3), 113-118.

Routray, W., & Orsat, V. (2011). Blueberries and their anthocyanins: Factors affecting biosynthesis and properties. *Comprehensive Reviews in Food Science and Food Safety, 10*(6), 303-320.

Rowland, L.J., Mehra, S., Dhanaraj, A., Ogden, E.L., & Arora, R. (2003a). Identification of markers associated with cold tolerance in blueberry. *Acta Hort*. 625:59-69.

Rowland, L. J., Mehra, S., Dhanaraj, A. L., Ogden, E. L., Slovin, J. P., & Ehlenfeldt, M. K. (2003b). Development of EST-PCR markers for DNA fingerprinting and genetic relationship studies in blueberry (*Vaccinium*, section cyanococcus). *Journal of the American Society for Horticultural Science, 128*(5), 682-690.

Rowland, L. J., Hancock, J. F., & Bassil, N. V. (2011). Blueberry. In K. Folta, & C. Kole (Eds.), *Genetics, genomics, and breeding of berries* (pp. 1-40). Enfield, NH: Science Publishers.

Rowland, L. J., Ogden, E. L., Bassil, N., Buck, E. J., McCallum, S., Graham, J., . . . Haynes, K. G. (2014). Construction of a genetic linkage map of an interspecific diploid blueberry population and identification of QTL for chilling requirement and cold hardiness. *Molecular Breeding, 34*(4), 2033-2048.

Rowland, L., & Levi, A. (1994). RAPD-based genetic linkage map of blueberry derived from a cross between diploid species (*Vaccinium darrowi* and *V. elliottii*). *Theoretical and Applied Genetics, 87*(7), 863-868.

Rowland, L., Mehra, S., & Arora, R. (2002). Identification of molecular markers associated with cold tolerance in blueberry. Paper presented at the *XXVI International Horticultural Congress: Biotechnology in Horticultural Crop Improvement: Achievements, Opportunities and 625,* 59-69.

Rowland, L. J., Alkharouf, N., Darwish, O., Ogden, E. L., Polashock, J. J., Bassil, N. V., & Main, D. (2012). Generation and analysis of blueberry transcriptome sequences from leaves, developing fruit, and flower buds from cold acclimation through deacclimation. *BMC Plant Biology, 12*, 46-2229-12-46. doi:10.1186/1471-2229-12-46.

Rungis, D., Bérubé, Y., Zhang, J., Ralph, S., Ritland, C. E., Ellis, B. E., . . . Ritland, K. (2004). Robust simple sequence repeat markers for spruce (*Picea* spp.) from expressed sequence tags. *Theoretical and Applied Genetics, 109*(6), 1283-1294.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America, 74*(12), 5463-5467.

Schierwater, B., & Ender, A. (1993). Different thermostable DNA polymerases may amplify different RAPD products. *Nucleic Acids Research, 21*(19), 4647-4648.

Schlautman, B., Fajardo, D., Bougie, T., Wiesman, E., Polashock, J., Vorsa, N., . . . Zalapa, J. (2015). Development and validation of 697 novel polymorphic genomic and EST-SSR markers in the American cranberry (*Vaccinium macrocarpon* Ait.). *Molecules, 20*(2), 2001-2013.

Schlötterer, C. (1998). Microsatellites. *Molecular Genetic Analysis of Populations: A Practical Approach, 2*, 237-261.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., . . . Wilson, R. K. (2009). The B73 maize genome: Complexity, diversity, and dynamics. *Science (New York, N.Y.), 326*(5956), 1112-1115. doi:10.1126/science.1178534.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology, 18*(2), 233-234.

Seeb, J., Carvalho, G., Hauser, L., Naish, K., Roberts, S., & Seeb, L. (2011). Single−nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources, 11*(s1), 1-8.

Semagn, K., Bjørnstad, Å., & Ndjiondjop, M. (2006). An overview of molecular marker methods for plants. *African Journal of Biotechnology, 5*(25)

Senan, S., Kizhakayil, D., Sasikumar, B., & SHEEJA, T. E. (2014). Methods for development of microsatellite markers: An overview. *Notulae Scientia Biologicae, 6*(1), 1-13.

Shi, C. Y., Yang, H., Wei, C. L., Yu, O., Zhang, Z. Z., Jiang, C. J., . . . Wan, X. C. (2011). Deep sequencing of the *Camellia sinensis* transcriptome revealed candidate genes for major metabolic pathways of tea-specific compounds. *BMC Genomics, 12*, 131-2164-12-131. doi:10.1186/1471-2164-12-131.

Solovchenko, A., & Schmitz-Eiberger, M. (2003). Significance of skin flavonoids for UV-B-protection in apple fruits. *Journal of Experimental Botany, 54*(389), 1977-1984. doi:10.1093/jxb/erg199.

Sonah, H., Deshmukh, R. K., Sharma, A., Singh, V. P., Gupta, D. K., Gacche, R. N., . . . Sharma, T. R. (2011). Genome-wide distribution and organization of microsatellites in plants: An insight into marker development in *Brachypodium*. *Plos One, 6*(6), e21298.

Song, G., & Hancock, J. F. (2011). *Vaccinium*. *Wild crop relatives: Genomic and breeding resources* (pp. 197-221) Springer.

Song, G., & Sink, K. (2004). Agrobacterium tumefaciens-mediated transformation of blueberry (*Vaccinium corymbosum* L.). *Plant Cell Reports, 23*(7), 475-484.

Springob, K., Nakajima, J., Yamazaki, M., & Saito, K. (2003). Recent advances in the biosynthesis and accumulation of anthocyanins. *Natural Product Reports, 20*(3), 288-303.

Strik, B. (2005). Blueberry: An expanding world berry crop. *Chronica Horticulturae, 45*(1), 7-12.

Strik, B. C., & Yarborough, D. (2005). Blueberry production trends in north America, 1992 to 2003, and predictions for growth. *Horttechnology, 15*(2), 391-398.

Tammi, M. T. (2003). The principles of shotgun sequencing and automated fragment assembly. *Center for Genomics and Bioinformatics, Karolinska Institute, Stockholm, Sweden,*

Tan, L., Wang, L., Wei, K., Zhang, C., Wu, L., Qi, G., . . . Liang, J. (2013). Floral transcriptome sequencing for SSR marker development and linkage map construction in the tea plant (*Camellia sinensis*). *PloS One, 8*(11), e81611.

Tanaka, Y., & Ohmiya, A. (2008). Seeing is believing: Engineering anthocyanin and carotenoid biosynthetic pathways. *Current Opinion in Biotechnology, 19*(2), 190-197.

Tautz, D., & Renz, M. (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research, 12*(10), 4127-4138.

Taylor, D. R., & Ingvarsson, P. K. (2003). Common features of segregation distortion in plants and animals. *Genetica, 117*(1), 27-35.

Thiel, T., Michalek, W., Varshney, R., & Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics, 106*(3), 411-422.

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., . . . Rokhsar, D. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science (New York, N.Y.), 313*(5793), 1596-1604. doi:313/5793/1596 [pii]

van Dijk, E. L., Auger, H., Jaszczyszyn, Y., & Thermes, C. (2014). Ten years of next-generation sequencing technology. *Trends in Genetics, 30*(9), 418-426.

Vander Kloet, S. P. (1988). *The genus Vaccinium in north America.* Agriculture Canada.

Vander Kloet, S., & Dickinson, T. (2009). A subgeneric classification of the genus *Vaccinium* and the metamorphosis of *V.* section Bracteata nakai: More terrestrial and less epiphytic in habit, more continental and less insular in distribution. *Journal of Plant Research, 122*(3), 253-268.

2, R. K., Graner, A., & Sorrells, M. E. (2005). Genic microsatellite markers in plants: Features and applications. *Trends in Biotechnology, 23*(1), 48-55.

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

Wang, M. L., Barkley, N. A., & Jenkins, T. M. (2009). Microsatellite markers in plants and insects. part I: Applications of biotechnology. *Genes, Genomes and Genomics, 3*(1), 54-67.

Wiering, H., & De Vlaming, P. (1984). Genetics of flower and pollen colors. *Monographs on Theoretical and Applied Genetics, 9*, 49-67.

Willson, M. F., & Whelan, C. J. (1990). The evolution of fruit color in fleshy-fruited plants. *American Naturalist,* , 790-809.

Winter, P., & Kahl, G. (1995). Molecular marker technologies for plant improvement. *World Journal of Microbiology and Biotechnology, 11*(4), 438-448.

Wu, X., Beecher, G. R., Holden, J. M., Haytowitz, D. B., Gebhardt, S. E., & Prior, R. L. (2006). Concentrations of anthocyanins in common foods in the united states and estimation of normal consumption. *Journal of Agricultural and Food Chemistry, 54*(11), 4069-4075.

Yousef, G. G., Lila, M. A., Guzman, I., Ballington, J. R., & Brown, A. F. (2014). Impact of interspecific introgression on anthocyanin profiles of southern highbush blueberry. *Journal of the American Society for Horticultural Science, 139*(2), 99-112.

Yu, J., Hu, S., Wang, J., Wong, G. K., Li, S., Liu, B., . . . Yang, H. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science (New York, N.Y.), 296*(5565), 79-92. doi:10.1126/science.1068037.

Zafra−Stone, S., Yasmin, T., Bagchi, M., Chatterjee, A., Vinson, J. A., & Bagchi, D. (2007). Berry anthocyanins as novel antioxidants in human health and disease prevention. *Molecular Nutrition & Food Research, 51*(6), 675-683.

Zane, L., Bargelloni, L., & Patarnello, T. (2002). Strategies for microsatellite isolation: A review. *Molecular Ecology, 11*(1), 1-16.

Zhang, H., Gao, S., Lercher, M. J., Hu, S., & Chen, W. H. (2012). EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. *Nucleic Acids Research, 40*(Web Server issue), W569-72. doi:10.1093/nar/gks576.

Zifkin, M., Jin, A., Ozga, J. A., Zaharia, L. I., Schernthaner, J. P., Gesell, A., ... & Constabel, C. P. (2012). Gene expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. Plant physiology, 158(1), 200-224.

Żmieńko, A., Samelak, A., Kozłowski, P., & Figlerowicz, M. (2014). Copy number polymorphism in plant genomes. *Theoretical and Applied Genetics, 127*(1), 1-18.

APPENDICES

Appendix A. Markers with associated scaffold information in diploid blueberry linkage map.
DI01a

Locus	Position	Scaffold No.	Scaffold Size (bp)
KAN11381_236	0	scaffold00504	379205
CRA_scf10688_365	24.214	-	
KAN16539_297	24.823	scaffold01612	48003
scaf00155_116	36.268	scaffold00215	354536
KAN11381_242	49.175	scaffold00504	
KAN41661_184	53.119	scaffold00018	779024
KAN-16471_282	55.03	-	
KAN41365_189	60.959	scaffold00009	855368
GVC-C102_226	68.005	scaffold24087	623
berc203	71.755	-	
KAN16539_316	73.583	scaffold01612	
berc128a	78.159	-	
berc345	81.212	scaffold54887	418
KAN1853_209	89.40	scaffold00478	222982
berc781a	93.67	-	
berc781a	99.796	-	
berc263c	101.249	scaffold03900	6992
berc104	107.284	scaffold00493	341725
berc229	123.821	scaffold00056	1035298
KAN-11440_291	131.408	scaffold00187	406777
SL192	134.039	scaffold00378	263052
KAN-11057	137.325	scaffold00257	327594
VCB-C09527_226	139.312	scaffold01712	129556
VCB-C09527_228	139.312	scaffold01712	
berc51b	141.303	scaffold01279	73229
berc396	144.118	scaffold00490	333366
all03191	147.595	scaffold00184	384554
berc647	151.144	scaffold00460	230906
Total			6,173,208

Appendix A continued

DI01b

Locus	Position	Scaffold No.	Scaffold Size (bp)
pink-00254	0	-	
CA169	18.562	scaffold00101	551401
berc136	20.712	scaffold01327	69185
SIZ1-2_87	22.964	scaffold00789	467809
GVC-V61g11	25.897	scaffold00421	246565
KAN-11281	26.509	-	
KAN-2328	26.547	scaffold00466	228325
OPC-09a	27.551	-	
CA243a	30.313	scaffold01961	30957
CA243b	31.577	scaffold01961	
Pr031818815	37.748	scaffold00112	463758
ripe-00787	39.032	scaffold16491	808
scaf00062_123	42.306	scaffold00092	719514
KAN12244_225	46.999	scaffold00110	688201
VCB-BH-1DV1YK_163	50.377	scaffold00160	482281
VCB-BH-1DV1YK_172	51.769	scaffold00160	
berc351	67.888	scaffold00250	333465
OPC-02a	67.99	-	
berc102	80.036	scaffold01353	66800
OPAO-02b	87.665	-	
berc60	91.534	scaffold00716	159544
KAN23741_286	94.021	scaffold01428	60224
KAN-11260	98.244	scaffold00041	654636
Pr031818814	99.348	-	
Pr031818813	99.751	-	
OPC-09b	108.611	-	
scaf00007-2_518	111.333	scaffold00021	740937
KAN11205_265	114.727	scaffold00142	425733
KAN11205_341	116.316	scaffold00142	
Pr031818818	117.17	-	
SL161	122.23	scaffold00678	170302
berc301d	134.822	scaffold00393	312446
KAN2260_320	140.552	scaffold00003	1081109
KAN2260_291	144.015	scaffold00003	
Total			7,954,000

Appendix A continued

DI02

Locus	Position	Scaffold No.	Scaffold Size (bp)
berc301c	0	scaffold00393	312446
DFR00528A_325	15.274	-	
DFR00528A_262	16.385	-	
CHS00491B_261	30.317	-	
CHS00491A_281	31.067	-	
CHS00491A_314	31.818	-	
scaf00113_218_234	33.741	scaffold00210	433997
CA221a	45.694	-	
KAN-11200_133	49.253	scaffold00097	492098
GVC-V23g01_213	52.557	scaffold00664	172771
GVC-V23g01_210	52.557	scaffold00664	
OPB-05b	56.608	-	
VCB-C12195_318	59.016	-	
VCB-C12195_324	60.107	-	
VCB-C08295_247	60.459	-	
KAN24885_236	62.725	scaffold00191	761721
NA292Sb	64.758	scaffold00384	260086
berc369	66.076	-	
scaf00100_267	71.616	scaffold00227	455788
CRA_scf11l_210	71.723	-	
CRA_scf11l_216	73.263	-	
berc542a	77.744	scaffold00384	
berc279	82.037	-	
KAN12346_210	91.455	scaffold01061	103192
KAN1875_206	98.571	scaffold03812	7225
leaf-00186	103.738	-	
CHS00519A_341	107.623	-	
CHS00519A_306	109.495	-	
berc536	118.383	-	
berc542b	124.922	scaffold00384	
KAN15306_147	131.634	scaffold00455	289009
OPAG-14a	138.012	-	
berc341	142.599	scaffold00789	467809
OPAA-15a	153.162	-	
SL309	159.977	-	
IP5PII-2_286	176.242	scaffold00315	440179

Appendix A continued

IP5PII-2_234	198.706	scaffold00315	
Total			4,196,321

Appendix A continued

DI03

Locus	Position	Scaffold No.	Scaffold Size (bp)
KAN11106_210	0	scaffold00053	691932
OPC-02b	2.424	-	
berc401	12.186	scaffold00230	346202
CRA_scf12916_311	13.224	-	
berc98a	15.793	scaffold00175	390882
berc98b	15.793	scaffold00175	
green-00012	23.539	scaffold00053	
OPV-08c	27.05	-	
KAN11106_216	30.315	scaffold00053	
GVC-C66a_306	36.176	-	
scaf00090_204	37.31	scaffold00183	384964
Pr031818821_185	40.063	scaffold01748	147235
Pr031818821_193	40.508	-	
leaf-00158a	43.639	scaffold00137	569231
berc218	47.482	scaffold00880	128286
OPU-01b	49.045	-	
berc54b	50.918	scaffold00274	540986
OPB-04a	52.101	-	
UBC211b	54.873	-	
ripe-01939	58.096	scaffold01011	188051
OPAR-19	59.713	-	
VCB-C06669_298	62.697	scaffold00156	828406
CA1785S	62.766	-	
berc97	65.666	scaffold00150	418438
berc222	70.288	scaffold00458	232048
berc56b	75.441	-	
green-00058	78.032	-	
berc539	81.605	scaffold00006	975988
scaf00216_302	92.418	scaffold00143	425556
bud789-00617b	103.595	-	
GVC-V62a08	115.234	scaffold00817	139766
CA636S	145.774	-	
CRA_scf15903c_297	147.088	-	
Total			6,407,971

Appendix A continued

DI04

Locus	Position	Scaffold No.	Scaffold Size (bp)
berc85	0	scaffold00224	349623
berc56a	2.954	-	
CA1105	6.178	scaffold00037	675259
berc209a	18.135	scaffold00706	162195
berc209b	18.135	scaffold00706	
scaf00281_147_156	35.541	scaffold00416	355028
OPU-03b	39.057	-	
CHI01251A_236	41.179	-	
4CL00900A_206	41.867	-	
GVC-C703a_232	46.985	scaffold01261	75288
scaf00001-4_268	47.989	scaffold00005	1202632
scaf00001-4_280	48.177	scaffold00005	
CHI01251A_240	51.073	-	
scaf00001-3_288	51.907	scaffold00005	
leaf-00158b	54.744	scaffold000137	569231
OPU-01c	56.387	-	
leaf-00248	63.938	-	
berc230	65.671	scaffold00767	148352
ANPER00666-2_278	68.478	scaffold00396	256767
Pr031818828_184	72.217	scaffold00267	319274
Pr031818828_179	72.217	scaffold00267	
OPAR-12b	73.907	-	
scaf00001-4_292	77.13	scaffold00005	
SL256	84.953	-	
Total			4,113,649

Appendix A continued

DI05

Locus	Position	Scaffold No.	Scaffold Size (bp)
GVC-V23g01_202	0	scaffold00664	172771
GVC-V24d10b_202	18.322	-	
GVC-NA721	27.555	scaffold00051	1004216
OPAA-17b	35.503	-	
berc2	37.526	scaffold03900	6992
berc24	38.671	scaffold00305	298661
CHS00014B_223	41.415	-	
CHS00014B_186	42.596	-	
berc363	53.946	scaffold01086	98903
GVC-NA113	64.99	-	
GVC-V41c07	70.391	-	
UBC34c	77.303	-	
VCB-C09467_292	86.887	scaffold01585	49359
CA344_157	96.475	-	
CA344_153	96.475	-	
KAN-11118	101.17	scaffold01218	206135
scaf00004-3_294	104.244	scaffold00020	816324
GVC-C71_104	105.687	-	
CA54	107.133	-	
SL151	119.298	-	
scaf00013_245	142.412	scaffold00045	719855
scaf00295_170	161.062	scaffold00469	354999
Pr031818820_157	170.653	-	
Pr031818820_173	171.434	-	
CRA_scf26r_273	175.841	-	
CRA_scf26r_286	175.841	-	
berc62	178.325	scaffold00046	230778
SL40	192.62	-	
Total			3,958,993

Appendix A continued

DI06

Locus	Position	Scaffold No.	Scaffold Size (bp)
berc54a	0	scaffold00274	540986
GVC-C608	11.824	scaffold00497	217934
GVC-C625	14.148	scaffold00343	275381
KAN-11482_260	22.122	scaffold00893	126450
GVC-C322_147	39.459	scaffold00487	231649
berc133	40.216	-	
KAN41355_257	48.171	-	
KAN41355_265	51.803	-	
CA1343	62.203	scaffold00391	257699
CA1049	64.191	scaffold00391	
SL42	68.673	-	
berc123	70.378	-	
berc187b	74.664	scaffold10236	2490
GVC-V64f07_239	80.609	scaffold00217	353892
berc523	87.246	scaffold00883	127734
OPP-16a	91.237	-	
berc138	93.646	scaffold00024	734356
berc485	100.574	scaffold00022	865675
scaf00169_272	102.512	-	
scaf00169_265	103.753	-	
berc366	105.93	scaffold00407	250157
berc51a	113.739	scaffold01279	73229
OPP-09b	132.366	-	
OPB-05a	140.281	-	
Total			4,057,632

Appendix A continued

DI07

Locus	Position	Scaffold No.	Scaffold Size (bp)
berc268	0	-	
KAN79_198	4.711	scaffold00013	1396472
scaf00069_308	17.17	scaffold00013	
KAN29551_268	20.732	scaffold00015	808311
Pr031818823	32.178	scaffold00102	579877
KAN-11199_263	49.642	scaffold00561	195343
KAN-11199_254	49.642	scaffold00561	
CA191	59.746	scaffold00704	162966
berc263a	66.014	scaffold03900	6992
berc472a	69.775	scaffold00892	126485
berc618	77.301	scaffold00840	134655
VCB-C00694_307	79.434	scaffold01109	161903
Pr031818819	81.765	-	
CRA_scf112c_144	83.686	-	
CRA_scf112c_157	85.28	-	
berc488b	88.884	-	
Pr031818811	90.251	scaffold00759	150188
OPAO-19a	91.129	-	
CA66	93.39	-	
berc895	103.216	scaffold00158	405259
VCB-C04624_95	110.707	scaffold00528	243098
OPAQ-01c	118.09	-	
GVC-V22a02_171	122.298	-	
GVC-V22a02_173	122.298	-	
KAN42567_202	130.303	scaffold04452	5904
GVC-V31e03_247	135.524	-	
KAN24806_169	135.835	scaffold00095	957275
GVC-V31e03_240	149.824	-	
KAN27020_261	153.53	scaffold00576	191504
Total			5,526,232

Appendix A continued

DI08

Locus	Position	Scaffold No.	Scaffold Size (bp)
OPV-08a	0	-	
OPP-12c	26.046	-	
NA292Sa	32.402	scaffold00384	260086
berc5	37.664	scaffold01279	73229
ripe-00122	41.519	-	
scaf00009-4_310	45.475	scaffold00023	778774
berc210	46.393	-	
DFR00528A_325	49.024	-	
DFR00528A_262	50.057	-	
berc149a	55.317	-	
OPV-08d	67.118	-	
Pr031818817	77.049	-	
OPP-14d	84.396	-	
berc893	90.11	-	
GVC-C136_300	92.698	scaffold01642	46057
KAN-11109	94.615	scaffold00011	840785
GVC-C089	95.477	-	
berc496	96.973	scaffold00018	779024
OPV-08b	99.676	-	
berc824	101.855	-	
SL188	117.188	-	
Total			2,777,955

Appendix A continued

DI09

Locus	Position	Scaffold No.	Scaffold Size (bp)
berc836	0	-	
ripe-00162a	20.728	-	
OPP-14b	24.169	-	
berc798	27.102	scaffold03548	8075
scaf00069_238	28.819	scaffold00013	1396472
KAN79_190	32.622	scaffold00013	
pink-00018a	32.853	scaffold00734	153931
berc196	33.131	-	
berc143	40.93	scaffold00333	359928
KAN40732_232	45.963	scaffold00208	358366
berc555	48.748	scaffold00208	
GVC-C179_213	58.755	scaffold28956	564
GVC-V32g09_292	65.962	scaffold01091	97635
OPAR-13c	70.218	-	
GVC-C725_144	71.596	-	
KAN24598_271	73.666	scaffold00408	340872
KAN24598_350	74.608	scaffold00408	
AOMT00197A_257	74.792	-	
KAN43117_202	81.035	scaffold00078	521410
berc680	86.066	scaffold00354	271555
bud789-02577c	94.462	scaffold00074	602287
Total			4,111,095

Appendix A continued

DI10

Locus	Position	Scaffold No.	Scaffold Size (bp)
KAN27020_261	0	scaffold00576	191504
Pr031818816_224	3.801	scaffold01914	32875
Pr031818816_189	5.349	scaffold01914	
NA186	15.27	scaffold00248	334508
GVC-C613	20.213	scaffold00665	288184
KAN24806_169	27.121	scaffold00095	957275
GVC-NA146	41.376	scaffold01114	93297
KAN17147_172	52.863	scaffold00098	491118
GVC-C455	56.091	scaffold00602	398509
KAN27179_195	60.827	-	
KAN27179_191	61.594	-	
KAN24307_192	80.339	scaffold00129	437798
berc130	90.139	scaffold01640	46217
KAN3462_172	94.732	scaffold00147	421988
Pr031818812_198	100.537	scaffold00095	
Pr031818812_195	100.537	scaffold00095	
berc292	110.724	-	
KAN79_198	117.522	scaffold00013	1396472
GVC-C634_272	122.264	scaffold01325	69520
scaf00069_308	125.08	scaffold00013	
berc198	126.901	-	
KAN27356_207	130.632	scaffold01466	57687
KAN24973_194	135.204	-	
KAN27743_233	135.387	scaffold00049	620373
KAN40732_228	142.968	scaffold00208	358366
KAN40732_251	145.9	scaffold00208	
berc48	146.275	scaffold00211	645650
KAN13486_205	160.364	-	
KAN13486_195	160.364	-	
KAN43117_235	171.613	scaffold00078	521410
KAN24598_287	191.593	-	
Total			7,362,751

Appendix A continued

LG11

Locus	Position	Scaffold No.	Scaffold Size (bp)
CHI03186-1_128	0	scaffold02594	14782
CHI03186-1_124	8.472	scaffold02594	
CRA_scf12916_311	14.797	-	
VCB-C03938_274	20.882	-	
OPAA-11a	28.626	-	
CA325_227	37.182	scaffold00889	126730
CA325_205	37.182	scaffold00889	
SL413	43.596	-	
Total			141,512

Appendix B. Markers with associated scaffold information in tetraploid blueberry "Jewel" linkage map.

J01

Locus	Scaffold No.	Scaffold size (bp)
36828a	scaffold00879	128298
36828b	scaffold00879	
Kan-11067	-	
TP28237	scaffold00628	260262
Vac288573	-	
TP110243	scaffold00628	
TP38086	scaffold0045	719855
TP109245	scaffold0045	
TP14818	scaffold0045	
TP109706	scaffold00538	202917
TP60254	scaffold00810	140768
Con600b	scaffold00810	
TP67359	scaffold00111	463851
TP87508	scaffold00006	975988
50936a	scaffold01057	103844
94595a	-	
TP109470	scaffold00268	360762
TP73709	scaffold00111	
TP24636	-	
TP4195	scaffold01057	
Con278	scaffold01362	65893
TP25456	scaffold00363	269267
TP55569	scaffold01449	58643
TP59531	scaffold28777	566
TP33921	scaffold00164	398974
TP104128	scaffold00164	
TP82022	scaffold00164	
TP96396	-	
TP99353	scaffold04607	5639
TP20020	scaffold01476	57064
TP105869	scaffold00032	703146
TP90799	-	
TP75195	scaffold00032	
TP40460	-	
TP95218	scaffold00764	149169

TP67352	scaffold00764	
TP58518	scaffold01602	48351
TP13136	scaffold00051	1004216
TP65047	-	
TP108503	scaffold00388	258737
TP112282	scaffold00305	298661
TP72590	scaffold00072	531288
TP2208	scaffold00048	620420
TP96725	scaffold00048	
TP68938	scaffold00305	298661
TP12388	scaffold00048	
TP77044	scaffold00048	
Kan-11151a	-	
GVC-V24d10b	-	
TP76523	scaffold00072	531288
82973	scaffold00195	555132
Total		9,211,660

J01U

Locus	Position	Scaffold No.	Scaffold size (bp)
430a	unmapped	scaffold16964	788
VCB-C09467a	unmapped	scaffold01585	49359
Con78c	unmapped	-	
1164	unmapped	-	
Total			50147

Appendix B continued

J02

Locus	Scaffold No.	Scaffold size (bp)
43a	scaffold00336	279056
TP16789	scaffold00306	297789
TP104336	scaffold00794	143788
TP81982	-	
TP721	scaffold00794	
TP35239	scaffold00684	266926
TP112748	-	
TP106924	scaffold00515	210655
TP2599	scaffold00336	279056
TP53035	scaffold00046	629757
Vac97025	-	
TP62208	scaffold01097	97143
TP110196	scaffold00788	144407
TP19675	scaffold00149	420419
TP56599	scaffold72745	
TP68096	scaffold00234	343835
TP99585	scaffold00498	217686
TP113473	-	
TP81471	scaffold00234	
TP3055	scaffold00186	383081
TP41833	scaffold01191	83470
TP100758	scaffold00295	305088
TP38854	scaffold00358	278998
TP37254	-	
TP4905	scaffold00358	
TP11191	scaffold01191	
TP39993	scaffold00358	
SL247b	scaffold01205	81508
TP98336	-	
CER62a	-	
TP62542	scaffold00218	353305
TP67652	scaffold00416	355028
TP32755	scaffold01052	104520
TP112121	scaffold01261	75288
TP68010	-	

TP55896	-	
TP49265	scaffold00499	244404
TP45278	scaffold00854	249440
Total		5,844,647

J02U

Locus	Position	Scaffold No.	Scaffold size (bp)
TP22212	unmapped		
TP90749	unmapped	scaffold00163	775497
TP117429	unmapped	scaffold00732	193511
TP84489	unmapped	-	
TP73157	unmapped	scaffold01181	123103
TP55711	unmapped	scaffold01052	104520
TP107810	unmapped	scaffold00741	152631
Kan-11199a	unmapped		
Vac124324b	unmapped	-	
43b	unmapped		
Total			1,349,262

Appendix B continued

J03a

Locus	Scaffold No.	Scaffold size (bp)
TP26949	scaffold00060	565451
TP3224	scaffold00060	
TP74612	scaffold00040	657903
TP95384	scaffold00294	404674
TP27248	scaffold00294	
TP17735	-	
Vac127278a	-	
TP35242	scaffold00040	
TP26544	scaffold01618	47616
TP110323	scaffold00040	
TP81898	scaffold00040	
TP82477	scaffold11399	
TP116340	scaffold00235	668704
TP10775	scaffold00235	
TP15419	scaffold01618	
TP3066	scaffold00040	
TP18995	scaffold00191	761721
TP18352	scaffold00191	
TP20232	scaffold00191	
TP75070	scaffold00191	
TP48286	scaffold00138	505320
TP14086	scaffold00138	
TP108998	scaffold00848	132412
VCCk4b	-	
TP3853	scaffold02872	
TP98897	scaffold00024	734356
TP68171	scaffold00024	
TP100237	scaffold01308	70654
TP117392	scaffold00104	749716
TP48250	scaffold00104	
TP11035	scaffold00104	
118a	scaffold01199	82117
TP34440	scaffold00181	387676
718b	scaffold00378	263052
TP40276	scaffold00520	207452
TP7594	scaffold00520	

TP41772	scaffold00270	
TP18182	scaffold00270	646955
TP38894	scaffold00581	
TP12701	scaffold00637	270162
TP77295	scaffold01399	62802
TP6452	scaffold00633	180169
TP33471	scaffold00633	
NA961a	scaffold00019	766197
Total		8,165,109

Appendix B continued

J03aU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP41079	unmapped	scaffold00522	295195
TP105796	unmapped	-	
TP35686	unmapped	scaffold00419	348991
TP56506	unmapped	scaffold00662	173125
TP70939	unmapped	scaffold03018	
TP107696	unmapped	scaffold00206	362265
TP21961	unmapped	-	
TP1649	unmapped	-	
TP47496	unmapped	scaffold03845	
TP67064	unmapped	scaffold00574	370960
TP103089	unmapped	scaffold12018	
TP91382	unmapped	scaffold00235	
TP99758	unmapped	scaffold00053	691932
TP31252	unmapped	-	
TP39820	unmapped	scaffold00522	
TP23921	unmapped	scaffold00868	129985
TP51636	unmapped	scaffold00040	
TP90891	unmapped	scaffold03845	
TP16944	unmapped	scaffold00670	171494
NA1778	unmapped	-	
503b	unmapped	scaffold01153	89301
IP5PII1a	unmapped	scaffold00315	440179
TP93350	unmapped	scaffold00531	204293
Con274c	unmapped		
VCCs10b	unmapped	-	
Total			3,277,720

Appendix B continued

J03b

Locus	Scaffold No.	Scaffold size (bp)
TP96030	scaffold00541	402017
TP46432	scaffold00401	421380
TP3026	scaffold00582	189502
TP101278	scaffold01280	73057
TP20219	-	
TP33728	scaffold00615	183994
TP103281	scaffold00324	335631
TP17943	scaffold00096	545778
TP112041	scaffold01309	70422
TP16313	scaffold00011	840785
TP15791	scaffold00262	14345
Kan-11051	-	
TP47513	scaffold00567	194978
TP104540	scaffold00422	246326
TP60394	-	
TP43567	scaffold00030	709418
TP56252	scaffold00023	778774
TP26099	scaffold00210	433997
TP6116	-	
TP19095	scaffold00256	327961
TP27989	scaffold00153	628812
TP39867	scaffold05449	
TP77321	scaffold00153	
TP25471	scaffold00153	
TP65159	scaffold00153	
Kan-11049	scaffold00153	
TP98153	scaffold00172	1123822
Total		7,520,999

Appendix B continued

J03bU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP58006	unmapped	scaffold00301	302349
TP104243	unmapped	scaffold04942	5206
TP31453	unmapped	scaffold09737	2746
TP64027	unmapped	scaffold00201	499705
TP40647	unmapped	scaffold00215	354536
TP83674	unmapped	scaffold02251	22419
TP91335	unmapped	scaffold00153	
TP59089	unmapped	scaffold03946	
TP86452	unmapped	scaffold00215	
TP91014	unmapped	scaffold00215	
Total			1,186,961

Appendix B continued

J03c

Locus	Scaffold No.	Scaffold size (bp)
789a	-	
TP66456	scaffold00007	907789
TP77496	scaffold03106	
TP5103	scaffold01776	38845
NA741	scaffold00033	692296
TP14011	scaffold00196	376196
Vac123588	-	
TP15771	scaffold24452	
TP41015	scaffold00007	907789
Total		2,922,915

J03cU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP76603	unmapped	scaffold00660	173306
TP50286	unmapped	scaffold00033	692296
TP54168	unmapped	scaffold00197	375923
TP101105	unmapped	scaffold00171	393957
TP76130	unmapped	scaffold00352	405778
TP37276	unmapped	scaffold01640	46217
TP28631	unmapped	scaffold80404	
52320a	unmapped	scaffold00784	145357
580	unmapped	scaffold00046	629757
Total			2,862,591

Appendix B continued

J03d

Locus	Scaffold No.	Scaffold size (bp)
TP43662	scaffold00934	119966
TP31591	scaffold00934	
TP96708	scaffold00934	
TP94639	scaffold00420	487829
TP72189	scaffold00420	
TP47338	scaffold82822	
TP58920	scaffold01026	109342
TP60773	scaffold01026	
TP37999	scaffold00119	452897
TP54405	scaffold00308	296153
TP107305	scaffold00395	420445
TP92762	scaffold00053	691932
TP50133	scaffold00053	691932
TP95994	scaffold00053	691932
TP88665	-	
TP93446	scaffold00053	691932
TP54570	-	
TP54867	-	
Kan-2584	-	
TP74880	-	
TP50442	-	
TP109277	scaffold01146	90049
TP7556	scaffold00130	437528
Total		5,181,937

J03dU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP93379	unmapped	scaffold00314	422472
TP44990	unmapped	scaffold00048	620420
TP89492	unmapped	scaffold00308	
TP6984	unmapped	scaffold00308	
TP34419	unmapped	-	
Total			1,042,892

Appendix B

continued

J04a

Locus	Scaffold No.	Scaffold size (bp)
TP14469	scaffold00638	405928
503c		89301
51576a	scaffold01632	46826
TP11546	scaffold00220	352288
TP44932	scaffold00220	
TP22373	scaffold04252	115752
286	-	
TP60388	scaffold00691	257645

TP67880	scaffold00236	341406
TP116959	scaffold00691	
Total		1,609,146

J04U

Locus	Position	Scaffold No.	Scaffold size (bp)
TP80956	unmapped	scaffold00344	275181
TP88342	unmapped	scaffold02902	11378
TP81239	unmapped	scaffold00409	274813
TP48521	unmapped	scaffold02902	
812123a	unmapped		
MAH13a	unmapped	-	
53769a	unmapped		
53769b	unmapped		
Total			561,372

Appendix B continued

J04b

Locus	Scaffold No.	Scaffold size (bp)
50936b		103844
TP72468	scaffold00111	463851
TP29122	scaffold01218	206135
TP53410	scaffold01218	
TP16839	scaffold01378	64870
Con78a		384554
TP51164	scaffold00020	816324
TP97420	scaffold00006	975988
CA855a	-	
TP878	scaffold00651	175031
TP96642	-	
TP69426	scaffold00073	529575
TP7416	scaffold00501	217406
TP69527	scaffold01691	43510
TP91901	-	
TP102509	-	
TP65540	-	
TP108314	scaffold00472	225529
Total		4,206,617

J04bU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP1122	unmapped	scaffold00022	865675
TP21956	unmapped	scaffold00022	
50936c	unmapped		
Total			865,675

Appendix B continued

J05

Locus	Scaffold No.	Scaffold size (bp)
TP3551	scaffold00412	632986
TP106905	scaffold00218	353305
AP12	-	
TP43058	-	
TP80292	-	
TP106208	-	
TP1394	scaffold02984	10847
TP99603	scaffold01036	185344
TP91905	scaffold00068	545623
TP33771	-	
TP20867	scaffold00590	273175
TP52998	scaffold00068	
TP13245	scaffold00068	
TP93505	scaffold00068	
TP24195	scaffold06563	3938
Vac287779c		
TP13525	scaffold00423	245963
Vac123022	scaffold00489	219673
TP102609	scaffold00245	573520
TP19883	scaffold00423	
TP52862	-	
TP100289	scaffold00423	
TP12124	scaffold00105	471899
TP31009	scaffold01012	111356
1107b	scaffold00015	808311
1107a	scaffold00015	
TP45101	scaffold00009	855368
NA800		308232
TP3552	scaffold00504	379205
TP29148	scaffold00504	
NA172	scaffold00437	240902
TP65407	scaffold01017	205468
TP34270	scaffold01017	
TP67319	scaffold01017	
TP110038	scaffold01612	48003
TP23277	scaffold00165	755211

73891c

Total	7,228,329

J05U

Locus	Position	Scaffold No.	Scaffold size (bp)
TP18495	unmapped	scaffold00123	673109
SL282b	unmapped	scaffold76127	355
Kan-1829b	unmapped	-	
Total			673,464

Appendix B continued

J06a

Locus	Scaffold No.	Scaffold size (bp)
TP110946	scaffold00307	296234
TP47056	-	
TP95617	scaffold00122	640349
TP72538	-	
TP51655	scaffold00310	391849
TP114253	scaffold00580	189768
TP94271	-	
TP45214	scaffold00121	453778
VCCs10d	-	
TP84056	scaffold01252	76262
TP19936	scaffold00364	268968
TP14064	scaffold02000	29413
TP18926	scaffold00364	268968
TP37985	scaffold00771	147254
TP7778	scaffold01252	76262
TP15543	scaffold01252	76262
TP35016	scaffold00771	147254
TP28465	scaffold00113	563029
TP64690	scaffold00154	414012
TP27258	-	
TP15633	scaffold00154	414012
TP17580	scaffold01822	36850
TP56115	scaffold01637	46562
TP74231	scaffold01108	2172
TP104871	scaffold00217	353892
TP21619	scaffold00293	305308
TP56156	scaffold00525	422446
TP45093	-	
TP100463	-	
TP44464	scaffold00403	251634
TP28913	scaffold00497	217934
TP18773	scaffold00596	307834
337b	scaffold00054	606533
TP36232	scaffold00391	257699
TP11483	scaffold00391	257699
Total		7,520,237

Appendix B continued

J06b

Locus	Scaffold No.	Scaffold size (bp)
TP116908	scaffold01396	63013
66b	-	
TP80345	scaffold00703	163099
66a	-	
GVC-V64f07c	-	
TP40006	scaffold00391	257699
GVC-V64f07b	-	
Total		483,811

Appendix B continued

J07

Locus	Scaffold No.	Scaffold size (bp)
TP10274	-	
TP34540	scaffold00172	1123822
TP6869	scaffold00030	709418
GVC-V52d01	-	
TP20267	-	
TP965	-	
TP8085	scaffold00380	262222
TP10107	scaffold00259	443186
TP91420	scaffold02780	12477
TP102061	scaffold02780	
TP35268	scaffold00571	193725
TP107607	-	
Kan-2133	-	
TP52753		
NA398b	-	
Total		911,610

J07U

Locus	Position	Scaffold No.	Scaffold size (bp)
TP66	unmapped	scaffold00012	831255
TP19252	unmapped	scaffold00030	709418
TP44308	unmapped	scaffold03266	9165
TP35038	unmapped	scaffold00210	433997
TP109237	unmapped	scaffold00211	645650
TP109184	unmapped	scaffold00390	257714
TP110644	unmapped	scaffold00508	352154
TP99837	unmapped	scaffold03266	9165
TP46237	unmapped	-	
TP16675	unmapped	scaffold00211	
1095b	unmapped	scaffold00256	327961
VCB-C14758b	unmapped	-	
Total			3,576,479

Appendix B continued

J08a

Locus	Scaffold No.	Scaffold size (bp)
TP13388	scaffold01511	54197
TP89358	scaffold10680	
COPII	-	
TP44704	scaffold00740	153174
Kan-11199c		
TP81095	scaffold00334	379997
Con652a	-	
TP33044	-	
TP7787	scaffold06893	
TP16606	scaffold00289	306473
TP27557	-	
TP32396	scaffold00289	
TP33294	scaffold00180	388487
TP3071	scaffold00180	
TP80027	scaffold00434	241217
CA23	scaffold00128	438941
TP69790	scaffold00434	
TP1063	-	
TP24949	scaffold00729	154815
453		328835
TP2436	scaffold00719	158925
TP38166	-	
TP100411	-	
GVC-V31e03f		
GVC-V31e03c		
TP107925	scaffold01563	50870
GVC-V22a02c	-	
TP113036	scaffold00629	181304
TP66351	scaffold00074	602287
TP22681	scaffold00280	310193
81215a	scaffold00673	170915
Total		3,920,630

Appendix B continued

J08aU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP108466	unmapped	scaffold00136	432078
TP3977	unmapped	scaffold00136	
TP20811	unmapped	scaffold01185	84285
TP67482	unmapped	scaffold00506	216598
TP32271	unmapped	scaffold00347	660065
RL15	unmapped	scaffold00455	238658
GVC-V22a02b	unmapped	-	
Total			1,631,684

Appendix B continued

J08b

Locus	Scaffold No.	Scaffold size (bp)
CA169	-	
Vac288135	-	
TP97411	scaffold00340	306374
TP90876	scaffold00161	401621
TP97198	scaffold03650	
656d	scaffold01169	86900
TP8643	-	
TP16244	scaffold01015	181697
TP58524	scaffold04169	6367
TP96121	scaffold02284	21543
TP76686	scaffold02284	
TP58155	scaffold01453	58561
TP35861	scaffold00102	579877
TP43484	scaffold00120	451729
TP69817	scaffold05975	4282
TP5320	scaffold00855	131834
TP39216	scaffold00500	217606
TP36560	scaffold00500	
Total		2,448,391

J08bU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP39845	unmapped	-	
TP29534	unmapped	scaffold0074	602287
TP62140	unmapped	-	
GVC-V31e03a	unmapped		
Kan-11199b	unmapped		
Total			602,287

Appendix B continued

J09

Locus	Scaffold No.	Scaffold size (bp)
TP99530	scaffold00361	309809
TP22368	scaffold00379	262243
Vac96445a	-	
TP93808	scaffold00335	280041
CA1553b	scaffold00383	267642
GVC-V61b10	-	
GVC-V32g09	scaffold01091	97635
TP66666	scaffold01636	46601
TP74107	scaffold00193	378270
TP23160	scaffold01254	75869
TP53833	-	1718110
Total		1,718,110

J09U

Locus	Position	Scaffold No.	Scaffold size (bp)
TP14392	unmapped	scaffold00274	540986
TP93548	unmapped	scaffold02705	13193
TP59681	unmapped	scaffold01261	75288
TP104276	unmapped	scaffold00688	166476
TP94848	unmapped	scaffold00485	264905
Total			1,060,848

Appendix B continued

J10a

Locus	Scaffold No.	Scaffold size (bp)
TP223	scaffold00028	817853
TP20664	scaffold00014	887145
TP78815	scaffold00028	
TP14851	scaffold00028	
TP108095	-	
Vac123749	-	
TP39244	scaffold00014	
TP38254	scaffold00058	579313
Vac110398b	-	
TP92779	scaffold00004	1245574
TP94297	scaffold00632	180458
TP92780	scaffold00399	448249
TP39388	scaffold00399	
TP41374	scaffold00632	
33.6	scaffold01254	75869
TP76010	scaffold00712	160346
TP39404	scaffold00712	
TP24823	scaffold00247	335996
991a	scaffold00247	
TP90349	scaffold01019	110748
TP34524	scaffold00252	330555
CA94b	scaffold00990	114518
TP59352	scaffold01772	38962
TP61010	scaffold00004	
TP92083	scaffold00681	236166
TP19800	scaffold00681	
TP60515	scaffold00410	421262
628a	scaffold00608	184985
TP67524	-	
Total		6,167,999

Appendix B continued

J10aU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP76669	unmapped	scaffold00247	
TP49002	unmapped	-	
TP110201	unmapped	scaffold00064	552209
TP98754	unmapped		302818
TP50763	unmapped	scaffold01181	123103
93226a	unmapped		
GVC-V52C09b	unmapped	-	
Total			978,130

Appendix B continued

J10b

Locus	Scaffold No.	Scaffold size (bp)
CHI2c	-	
Vac124930	-	
TP35878	scaffold00239	338672
TP65878	scaffold00209	358017
TP7726	scaffold00209	
TP24678	scaffold01223	79503
TP6813	scaffold00213	555703
TP28408	scaffold01223	
TP60410	scaffold00410	421262
TP40172	scaffold00018	779024
TP5283	scaffold00617	355452
TP17427	scaffold01469	167251
VCCb3	scaffold01081	100076
TP113527	scaffold00284	308249
TP2343	scaffold00146	710646
TP27265	scaffold00146	
Total		4,173,855

Appendix B continued

J10c

Locus	Scaffold No.	Scaffold size (bp)
524a	scaffold00046	629757
MCYAII	-	
Kan-4548a	-	
1233a	-	
Vac287779a	scaffold00001	1275046
MAH13b	-	
CA1785s	-	
CA112	scaffold00866	130337
TP13713	scaffold00001	
TP81889	scaffold00001	
6ms4e04		
TP21630	scaffold00981	116055
Vac287779b	scaffold00001	1275046
TP54379	-	
524b	scaffold00046	3426241
Total		342,6241

J10cU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP27309	unmapped	scaffold00219	353182
TP14985	unmapped	-	
TP71480	unmapped	scaffold00131	436091
TP113414	unmapped	scaffold00219	
TP5208	unmapped	scaffold00282	315548
TP16688	unmapped	-	
Kan-11082c	unmapped	-	
Kan-4548b	unmapped	-	
52236a	unmapped	-	
Total			1,104,821

Appendix B continued

J11a

Locus	Scaffold No.	Scaffold size (bp)
TP104400	scaffold01138	193554
TP111410	scaffold00092	719514
TP110784	scaffold00337	366165
TP10482	-	
TP101494	scaffold00421	246565
TP42218	scaffold00348	274192
TP25721	scaffold00735	153736
TP90901	scaffold00421	
TP43631	-	
TP38943	scaffold00421	
TP76233	scaffold00512	305430
TP95021	scaffold00123	673109
TP94478	scaffold00876	266877
TP59149	scaffold01938	31831
Vac126383a	-	
TP117940	466	228325
VCB-BH-1DV1YKb	scaffold00160	482281
TP80736	-	
VCCj9c	scaffold00092	
VCCj1a	scaffold03147	9739
TP28540	-	
TP65644	-	
TP34187	scaffold00693	165513
NA619b	scaffold00160	
TP45143	-	
TP87818	scaffold00693	
TP44099	scaffold00110	688201
TP71582	scaffold00693	
TP72871	-	
TP43100	scaffold00067	775372
TP77008	scaffold00693	
TP76704	-	
TP14041	scaffold00118	551576
NA619a		
TP42453	scaffold00118	

TP8228	scaffold01122	92468
TP10805	-	
TP19706	scaffold01122	
TP77409	-	
TP15376	scaffold01107	94789
TP29352	scaffold00103	489561
TP37687	scaffold00513	211737
TP10321	scaffold00148	421443
TP61028	scaffold00041	654636
Total		8,096,614

J11aU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP43015	unmapped	scaffold00876	
TP110824	unmapped	scaffold00421	
TP55298	unmapped	-	
TP13531	unmapped	scaffold00118	
799c	unmapped	scaffold00421	
VCB-BH-1DV1YKa	unmapped	scaffold00160	

Appendix B continued

J11b

Locus	Scaffold No.	Scaffold size (bp)
47089	-	
TP96133	scaffold00010	843235
TP103028	scaffold00329	281605
TP109639	scaffold00329	281605
TP30539	scaffold00329	281605
TP55997	scaffold00329	281605
TP67409	-	
TP42928	scaffold00471	225797
TP71604	scaffold00471	
TP90658	scaffold00002	1526805
TP15600	scaffold00002	
TP43657	scaffold00471	225797
TP3322	scaffold00002	
GVC-V62h04	-	
TP21283	scaffold00002	
TP79660	scaffold00002	
TP3686	scaffold00003	1081109
TP22481	scaffold00021	740937
TP74315	scaffold05485	
TP62324	scaffold00118	551576
TP69019	scaffold00752	261255
157a	scaffold00881	128233
Total		4,741,509

J11bU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP7285	unmapped	scaffold00002	
TP20971	unmapped	scaffold00002	
TP102483	unmapped	-	
1792b	unmapped		

Appendix B continued

J12a

Locus	Scaffold No.	Scaffold size (bp)
TP97478	scaffold00012	831255
TP62475	scaffold01672	44704
TP42800	scaffold00095	957275
TP76995	scaffold00147	421988
TP115165	-	
TP27068	scaffold00406	368770
TP8905	scaffold00057	581046
TP102542	scaffold00406	
TP118091	scaffold00057	
TP11939	scaffold01640	46217
TP10687	scaffold01330	68928
TP26100	scaffold00783	145460
TP74776	-	
TP60976	scaffold01330	
TP19805	scaffold01074	100730
TP77512	scaffold01074	
TP83375	scaffold01092	97505
TP106340	scaffold00107	467516
TP102138	scaffold02668	13669
TP64920	scaffold00162	780715
TP106898	scaffold00774	146858
TP112887	scaffold00167	606697
TP11906	scaffold00920	246533
TP88091	scaffold00898	125923
TP32377	-	
TP18089	scaffold00303	300255
VCCj3a		
TP37696	scaffold00303	
VCCj3b		
TP3327	scaffold00602	398509
TP44931	scaffold00303	
TP33689	scaffold01114	93297
TP39044	scaffold00727	157017
TP60038		
TP52206	scaffold00125	1018538
TP64381	scaffold00303	

TP101375	scaffold00125	
TP110247	scaffold00248	334508
TP95286	scaffold00018	779024
TP104469	scaffold00665	288184
TP101929	-	
TP102229	scaffold00350	272822
TP24005	scaffold00018	
TP18610	scaffold00248	
TP42461	scaffold00018	
TP86995	scaffold00151	641592
TP32781	scaffold00151	
TP112776	scaffold00223	351036
TP32528	scaffold00134	434152
NA41	-	
TP9058	scaffold00134	
TP104769	scaffold00151	
TP55260	scaffold00929	121508
Con668	-	
Vac126607b	scaffold01914	32875
TP55259	scaffold03191	9532
TP64332	scaffold00090	501067
TP43663	-	
TP6922	scaffold00090	
TP71982	scaffold00061	563063
TP105168	scaffold00090	
CA794a	-	
Total		12,348,768

Appendix B continued

J12aU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP1159	unmapped	scaffold00095	
TP80903	unmapped	scaffold00656	174111
TP26350	unmapped	scaffold00125	
TP20103	unmapped	scaffold00151	
TP92859	unmapped	scaffold00223	
52320b	unmapped	scaffold00784	145357
Vac126607a	unmapped	scaffold01914	
Kan-11138	unmapped	-	
Total			319,468

Appendix B continued

J12b

Locus	Scaffold No.	Scaffold size (bp)
TP71362	scaffold00191	761721
TP51767	scaffold00155	946270
GVC-V22g08a	scaffold00155	
TP47009	scaffold00155	
VCBC12195	scaffold00155	
TP43291	scaffold00155	
TP82443	scaffold00227	455788
VCCi2b	scaffold00008	1199225
VCCk4a	-	
TP33976	-	
TP14075	scaffold00086	506086
TP85295	scaffold00083	513550
1099	scaffold00008	
662b	scaffold00863	195450
Total		4,578,090

J12bU

Locus	Position	Scaffold No.	Scaffold size (bp)
TP69094	unmapped	scaffold00369	328062
TP391	unmapped	scaffold00191	
TP75437	unmapped	scaffold92387	
131	unmapped	scaffold00138	505320
GVC-V21e04	unmapped	-	
Total			833,382

Appendix C. Markers with associated scaffold information in tetraploid blueberry "Draper" linkage map.

D01

Locus	Scaffold No.	Scaffold Size (bp)
ScKan11151b	scaffold00349	278175
ScVCBC09467a	-	
TP21913	scaffold579	190855
TP66406	scaffold00039	659609
FL430	scaffold16964	788
TP16488	scaffold00039	
TP111210	scaffold00039	
TP73519	scaffold00185	461508
MICA855	-	
TP74601	scaffold00144	536255
MICA344	-	
TP77102	scaffold00039	
TP65523	-	
TP118334	-	
TP140	scaffold00164	398974
TP83186	-	
TP48136	scaffold01035	106970
TP115798	-	
TP115023	scaffold00070	1171683
TP59202	scaffold00111	463851
FLCon78b	-	
TP80084	scaffold00020	816324
TP75621	scaffold00020	
TP49993	scaffold00651	175031
TP107547	scaffold00020	
TP45023	scaffold00020	
TP58157	scaffold00651	
TP33974	scaffold00006	975988
TP43077	scaffold00006	
FL94595b	-	
TP70593	scaffold03115	9896
TP79503	scaffold00073	529575
TP100369	scaffold00073	
TP55121	scaffold00395	420445
ScATDGK13	-	

TP99730	scaffold00073	
TP54290	-	
TP51992	scaffold00332	280382
TP21341	scaffold04170	6366
TP62897	scaffold04170	
ScVc127155a	-	
Total		7,482,675

Appendix C continued

D02

Locus	Scaffold No.	Scaffold size (bp)
TP65030	-	
TP46312	scaffold00404	646915
ScCER62b	-	
TP105994	scaffold00404	
TP71935	scaffold00079	519913
TP105172	scaffold00653	174560
ScVc96355a	-	
TP109905	scaffold00416	355028
TP17740	scaffold00005	1202632
ORSL247b	-	
TP11435	scaffold00323	285260
TP48964	scaffold00323	
ORSL247a	-	
ORSL247c	-	
TP103341	scaffold00218	353305
TP36095	scaffold00358	278998
TP71729	scaffold01441	59340
TP90423	scaffold00988	114920
TP27827	scaffold00732	193511
TP44158	scaffold00732	
TP11039	scaffold00732	
ScVc97025	scaffold00267	319274
TP88746	scaffold00467	312321
ScVc124324c	-	
TP48802	scaffold00360	467034
TP62382	scaffold00556	197458
TP108957	scaffold00886	127204
TP21522	-	
TP22699	scaffold05094	5041
ScVc124324a	-	
TP40828	scaffold00184	384554
TP66713	scaffold00035	791315
ScVc124324b	-	
FL1007	scaffold01753	39875
TP34184	scaffold00037	675259
TP26258	scaffold00037	

TP74187	scaffold00037	
TP33648	-	
TP53919	-	
Total		7,503,717

Appendix C continued

D03

Locus	Scaffold No.	Scaffold size (bp)
TP60753	scaffold00304	300232
TP72836	scaffold00255	329260
TP60401	scaffold00007	907789
TP892	-	
TP31938	-	
ScVc326725a	-	
TP90607	scaffold00029	785184
TP94869	scaffold01224	79435
TP18323	-	
TP111807	-	
TP99861	scaffold00657	296015
TP30737	scaffold00713	172217
ScVCBC06669	scaffold00156	828406
TP99238	scaffold00614	187375
TP89330	scaffold01345	67636
TP80421	scaffold01222	79690
TP68945	scaffold00254	475302
TP112016	scaffold00124	719687
TP27951	scaffold03526	8152
TP37109	scaffold02461	17207
TP2213	scaffold00152	417541
TP66585	scaffold00124	
TP56611	scaffold00124	
TP37969	scaffold00152	
TP71680	-	
ScFLS3	scaffold00236	341406
TP7244	scaffold01078	100407
TP12763	scaffold01078	
TP1978	scaffold00124	
TP114677	scaffold00124	
TP74341	-	
TP43461	-	
TP104831	-	
TP77275	scaffold00124	
TP23903	scaffold00124	
TP91332	scaffold00156	

TP30043	-	
TP70435	scaffold00156	
TP22608	scaffold00156	
TP58885	scaffold00130	437528
TP88895	scaffold00156	
TP97463	scaffold43033	468
TP96943	scaffold00206	362265
TP5996	scaffold01410	120589
TP34118	scaffold00156	
TP22008	scaffold00330	568207
TP24735	scaffold00662	173125
TP45061	scaffold00443	239184
TP14149	scaffold00353	278153
TP80372	scaffold00196	376196
FL812123a	-	
TP91783	scaffold08216	3339
TP33302	scaffold08216	
FL789b	-	
TP75653	scaffold00088	663201
TP86092	scaffold00088	
Total		9,335,196

Appendix C continued

D04

Locus	Scaffold No.	Scaffold size (bp)
FL254	-	
FL458	scaffold00638	405928
MINA961b	-	
TP20145	scaffold00531	204293
TP24935	scaffold00882	127954
TP16647	scaffold00531	
TP88172	scaffold00045	719855
FL51576b	scaffold01632	46826
TP117626	scaffold00045	
TP69846	scaffold00045	
TP72721	scaffold00315	440179
TP95955	scaffold00045	
ScIP5PII1c	scaffold00315	
TP50486	-	
TP105001	-	
TP14197	scaffold00698	245510
TP76652	scaffold00192	379244
TP50176	scaffold11516	2059
TP13286	scaffold00572	360637
TP112173	scaffold00055	606438
TP103091	scaffold11516	
TP88111	scaffold00572	
TP68304	-	
TP72553	scaffold00085	506979
ORVc288573	-	
TP22385	scaffold00085	
TP5637	scaffold00085	
TP94729	-	
TP74609	-	
TP43937	scaffold00879	128298
TP39794	scaffold00192	
TP101410	scaffold01037	106789
TP64266	scaffold00758	150192
ScKan11067	-	
TP3426	scaffold01037	
TP52815	scaffold00520	207452

TP77746	scaffold00085	
TP91921	scaffold00055	
TP81081	scaffold01270	73985
TP11103	scaffold00055	
TP113917	scaffold00572	
TP96627	scaffold00781	145712
TP29160	scaffold01700	42976
TP49297	scaffold00181	387676
TP80904	scaffold00085	
TP8239	scaffold01700	
TP42927	scaffold00181	
MIVCCj1c	scaffold3147	9739
TP33148	scaffold00181	
TP31473	scaffold00235	668704
FL36828c	-	
TP42138	scaffold00394	257179
FL118b	scaffold01199	82117
FL286	-	
Total		6,306,721

Appendix C continued

D05

Locus	Scaffold No.	Scaffold size (bp)
TP49598	scaffold01089	98179
TP61784	scaffold00187	406777
TP58115	scaffold01209	160640
TP73842	scaffold00284	308249
TP89124	scaffold00381	261236
TP18151	-	
TP38269	scaffold00076	526042
TP74136	-	
TP94844	scaffold00187	
TP91269	-	
ScCDF5	scaffold01052	104520
FL987b	scaffold01007	112348
TP104076	scaffold00077	1002871
TP5482	scaffold00077	
TP9449	scaffold00294	404674
TP50917	scaffold01320	69706
TP32774	scaffold00026	1090302
TP39228	-	
TP82863	-	
FL987a	scaffold01007	
TP99293	scaffold00294	
TP44475	-	
ScVc127278b	-	
TP106394	-	
TP83149	scaffold00026	
TP9374	scaffold00731	154426
TP39778	scaffold00400	253230
TP51162	scaffold01225	79363
TP101021	scaffold00250	333465
TP43789	-	
TP47832	scaffold00097	492098
TP110849	scaffold00155	946270
TP52081	scaffold00155	
TP59881	scaffold00191	761721
TP91781	scaffold01225	

TP30031	scaffold00227	455788
FL93226b	-	
TP36941	-	
TP3472	-	
TP68383	scaffold00096	545778
TP31718	scaffold00096	
TP38138	scaffold00096	
TP27546	scaffold00158	405259
TP32569	scaffold00057	581046
TP26069	scaffold01361	65995
Total		9,619,983

Appendix C continued

D06

Locus	Scaffold No.	Scaffold size (bp)
TP36946	scaffold00580	189768
TP48003	scaffold00798	143184
TP25800	-	
TP2903	scaffold00433	341952
TP16404	-	
TP68115	-	
TP17472	scaffold00983	115780
TP50920	scaffold00122	640349
TP95065	scaffold00122	
TP82166	scaffold00580	
TP69624	scaffold01341	67870
TP97642	scaffold01341	
TP47585	-	
TP56718	scaffold00008	1199225
TP86240	-	
TP40711	-	
TP75554	scaffold00113	563029
TP98773	scaffold00069	542167
TP94775	scaffold00113	
TP91003	scaffold00069	
TP78948	scaffold00217	353892
TP30229	scaffold00113	
TP80935	scaffold00084	774246
TP30229	scaffold00464	436785
TP89077	scaffold00113	
TP60125	scaffold00728	156495
TP18528	scaffold00145	630605
TP50454	scaffold02011	29174
TP98824	-	
TP111587	scaffold00679	320651
TP22889	-	
TP85406	-	
TP116301	scaffold40794	480
TP36725	-	
TP5913	scaffold40794	
TP112470	scaffold00887	222547

TP60675	scaffold00887	
TP2621	scaffold00069	
TP105555	scaffold00887	
TP10142	scaffold00494	447887
TP7663	scaffold00679	
TP109691	scaffold00145	
TP15586	scaffold02011	
TP41516	-	
TP24322	-	
TP3136	scaffold00049	620373
TP72207	scaffold01167	87313
TP97148	scaffold00800	142858
TP92184	scaffold85227	334
TP75418	scaffold01085	98977
TP28909	scaffold00971	116985
TP3278	scaffold00971	
TP48047	scaffold01220	79808
TP54730	scaffold02000	29413
TP2584	scaffold00106	471133
TP27511	-	
TP115154	-	
TP71621	-	
FL1555a	scaffold00366	267589
TP1534	-	
TP14942	scaffold00121	453778
TP18935	scaffold02807	12262
TP741	scaffold00088	663201
TP54474	scaffold00088	
TP76630	scaffold02475	16714
Total		10,236,824

Appendix C continued

D07

Locus	Scaffold No.	Scaffold size (bp)
TP84990	-	
TP95276	scaffold00172	1123822
TP24098	-	
ScKan11049	-	
TP42086	scaffold00172	
TP33184	scaffold00172	
TP21609	scaffold00172	
TP111278	scaffold00023	778774
TP69989	scaffold00023	
TP113682	scaffold00474	224701
TP32909	scaffold00637	270162
TP90136	scaffold01524	53394
TP97379	scaffold00865	130357
TP40535	scaffold01524	
TP84979	scaffold00637	
TP108334	scaffold00262	323160
TP5956	scaffold00262	
TP65890	scaffold01779	38784
TP41100	scaffold00547	200344
TP47922	scaffold01709	42473
TP24505	scaffold00902	125008
TP85206	scaffold01240	77593
OR5ms2d01	-	
TP65223	scaffold00541	402017
TP64338	scaffold01221	79738
TP42195	scaffold01221	
TP24719	scaffold00066	664038
TP53859	scaffold00080	517888
TP90029	scaffold00385	259933
TP66812	-	
TP67917	-	
TP10786	scaffold00319	412323
TP58606	-	
TP5635	scaffold00847	226972
TP108533	-	
TP90328	-	

TP2827	scaffold00589	187984
Total		6,139,465

Appendix C continued

D08

Locus	Scaffold No.	Scaffold size (bp)
TP27013	scaffold00411	303062
MINA824a	-	
TP108387	-	
TP40617	scaffold01978	30087
TP13255	scaffold01978	30087
ScVCBC00694	scaffold01109	161903
TP882	scaffold00557	429699
TP63835	scaffold00300	328835
TP66989	scaffold00434	241217
TP104332	-	
TP104242	scaffold00161	401621
TP26964	scaffold01301	71272
TP60587	scaffold00300	328835
TP4146	scaffold00622	182296
TP51789	scaffold00273	438216
TP7864	scaffold00180	388487
TP106957	scaffold00273	438216
FL656a	scaffold01169	86900
TP64012	scaffold01561	51003
TP23870	scaffold00974	116784
TP66483	scaffold00974	116784
TP77630	scaffold01015	181697
TP6890	scaffold00418	247728
TP90448	-	
TP27733	scaffold70050	187156
TP96838	scaffold00334	379997
TP9029	scaffold02026	28679
TP107118	scaffold09128	3019
TP40861	scaffold02026	
ScKan11199c	-	
TP71690	-	
TP71592	scaffold00296	304664
TP96010	scaffold00120	451729
TP38778	scaffold00120	
TP64038	scaffold00102	579877
TP91842	scaffold00102	

TP44060	scaffold01596	48780
TP57776	scaffold04900	5261
TP85344	scaffold00015	808311
TP24094	scaffold00015	
Total		7,372,202

Appendix C continued

D09

Locus	Scaffold No.	Scaffold size (bp)
TP3244	scaffold00127	612582
TP31131	-	
TP5575	scaffold00078	521410
TP112269	scaffold01432	188231
TP67666	scaffold00127	
TP61744	scaffold01091	97635
TP109214	scaffold00659	173440
TP4286	scaffold00298	450682
TP4251	scaffold08270	3321
TP18160	-	
TP77515	scaffold01510	54303
MICA642	-	5508
TP22445	-	
TP5485	scaffold00237	340609
TP74524	scaffold00192	379244
TP110674	scaffold00852	132134
TP63855	scaffold00127	
TP38839	-	
TP51466	scaffold00659	173440
TP66674	-	
TP82202	scaffold00908	124240
TP69922	-	
TP40780	scaffold01231	78910
TP69661	-	
TP64171	-	
Sc3ms2g09	-	
TP42271	-	
TP19290	-	
TP7228	-	
TP27332	scaffold00016	1283226
TP21500	scaffold00078	521410
TP106745	scaffold00129	437798
TP42631	scaffold00129	
TP19988	scaffold00189	536467
TP106791	scaffold00212	357161
TP28315	scaffold00232	345946

TP78400	scaffold00298	
TP85992	scaffold00298	
TP87306	scaffold00354	271555
TP65700	scaffold00379	262243
TP55422	scaffold00379	
TP111506	scaffold00379	
TP114389	scaffold00413	354085
TP79268	scaffold00506	216598
TP12923	scaffold00554	197898
TP59986	scaffold00641	177986
TP21463	scaffold00908	
TP54522	scaffold00908	
TP59871	scaffold01139	90874
TP69231	scaffold01184	84627
TP45497	scaffold01231	
TP9828	scaffold01443	59280
TP65700	scaffold01980	30041
TP10477	scaffold02118	26143
TP22575	scaffold02118	
TP68641	scaffold09837	2692
Total		8,591,719

Appendix C continued

D10

Locus	Scaffold No.	Scaffold size (bp)
TP69119	scaffold00828	242836
TP66816	scaffold00283	308694
TP45782	scaffold00283	
TP91338	-	
TP44089	-	
TP79910	scaffold00441	240129
TP117839	scaffold00018	779024
TP109340	scaffold02230	23129
TP111874	scaffold02230	
TP67061	scaffold00174	502756
TP33401	scaffold00457	358574
TP59534	scaffold00457	
TP104406	scaffold00532	203592
TP112113	scaffold00532	
TP106097	scaffold00532	
TP86367	scaffold00457	
TP11930	scaffold00960	117435
TP80632	scaffold01660	45369
TP80984	scaffold00447	235364
TP94489	-	
TP10691	scaffold00456	500317
TP93226	scaffold00292	366913
TP47143	scaffold00404	646915
TP116759	scaffold00209	358017
TP22485	scaffold00408	340872
TP15274	-	
TP92183	-	
TP74877	scaffold04671	5544
ScCHI2a	-	
TP11137	scaffold00642	177616
TP68592	scaffold00209	
TP90779	scaffold00510	212258
TP110896	scaffold00352	405778
TP94763	scaffold00456	
Sc5ms2C09a	-	
FL319a	scaffold00239	338672

MICA94a	scaffold00990	114518
TP18373	scaffold02163	24830
TP52654	scaffold00331	280640
TP62284	scaffold00141	554342
TP23128	scaffold00141	
MICA787	scaffold00836	135665
TP9274	scaffold00141	
TP5681	scaffold00331	
TP5325	scaffold00168	455615
TP6054	scaffold01053	104420
TP104344	scaffold00168	
FL1030a	scaffold00031	1794986
ScKan4737b	-	
TP1354	scaffold00168	
TP76203	scaffold00577	191076
ScKan4737a	-	
TP43328	scaffold00064	552209
TP10302	scaffold00141	
TP31240	scaffold00141	
TP4416	scaffold00712	160346
TP19217	scaffold00299	302818
TP76236	scaffold00064	
TP5971	-	
TP74501	scaffold00470	449939
TP54534	scaffold00064	
TP22450	scaffold00031	1794986
TP99554	scaffold00031	
TP25793	scaffold00031	
TP62194	scaffold00470	
TP52215	scaffold00031	
TP84065	scaffold00678	170302
TP76532	scaffold00678	
TP21993	scaffold01005	112604
TP75149	scaffold00299	
FL1030b	scaffold00031	
TP6410	scaffold00004	1245574
TP33589	scaffold00031	
TP12951	scaffold00115	459576
TP69465	scaffold00064	
TP92002	scaffold00632	180458

TP29575	scaffold00014	887145
TP67598	scaffold00265	758944
TP29804	scaffold01096	123286
ORVc118588b	-	
TP10287	scaffold00014	
TP58394	scaffold00115	
TP111330	scaffold00399	448249
TP75016	scaffold00014	
TP25988	scaffold00058	579313
TP91571	scaffold01799	37894
TP35518	scaffold00058	
TP47780	scaffold04666	5550
TP53388	scaffold01096	
TP74710	scaffold00028	817853
TP93329	scaffold00028	
TP58599	scaffold00028	
TP40028	scaffold00014	
TP49816	scaffold00004	
Total		18,601,412

Appendix C continued

D11

Locus	Scaffold No.	Scaffold size (bp)
TP58236	-	
TP7397	scaffold00043	873961
TP93962	scaffold01938	31831
TP50470	scaffold01938	
TP76829	scaffold00789	467809
TP60817	scaffold01145	90261
TP112772	scaffold01531	53046
TP61746	scaffold01451	58604
TP102642	scaffold00488	220061
TP15023	scaffold00421	246565
TP67259	scaffold01145	
TP50952	scaffold00789	
TP8806	scaffold00285	308232
TP113812	scaffold00348	274192
TP39560	-	
TP28606	scaffold00488	
TP57253	-	
TP21675	scaffold00421	
TP61870	scaffold00216	354129
TP115753	scaffold00043	
TP85982	-	
MIVCCj9a	-	
TP112120	scaffold00216	
TP45836	scaffold00273	438216
TP5131	scaffold01138	193554
ScVc126383b	scaffold00112	463758
MIVCCj1b	scaffold03147	9739
MIVCCj9b	-	
TP111627	scaffold00362	444635
TP69388	scaffold01451	
TP5121	scaffold01115	93208
TP48068	scaffold00110	688201
TP25021	scaffold00110	
ScBHIDVIYKa	-	
FLNA619a	-	
TP105618	scaffold00140	906998

TP16520	scaffold00529	288865
TP113018	scaffold00118	551576
TP90902	scaffold00118	
TP59988	scaffold00140	
TP103977	scaffold01107	94789
TP118034	scaffold00183	384964
TP23807	scaffold01269	74048
TP4248	scaffold00148	421443
TP6038	scaffold00249	333694
TP78423	scaffold00183	
TP9986	scaffold00705	186066
TP9985	scaffold00017	1144947
TP104690	scaffold00689	262212
TP96041	scaffold00021	740937
TP4241	scaffold00689	
TP48553	scaffold00021	
TP116700	scaffold00689	
TP116232	scaffold00021	
TP22854	scaffold00228	347222
TP56191	scaffold00249	
TP94960	scaffold00228	
TP27999	scaffold00021	
TP21385	scaffold00249	
TP6393	scaffold00228	
TP19309	scaffold00021	
TP96422	scaffold00598	186687
TP112763	scaffold00121	453778
TP65565	scaffold00017	
TP115074	scaffold00021	
TP81792	scaffold00202	368835
TP21386	scaffold00003	1081109
TP97606	scaffold00505	216617
TP14166	scaffold00002	1526805
TP41775	scaffold00002	
Sc6ms2h04	-	
ScKan1593	-	
TP59962	scaffold00002	
TP74568	scaffold00002	
TP24843	scaffold00505	
TP86290	scaffold00002	

TP45201	scaffold00329	281605
TP56198	scaffold00010	843235
TP24775	-	
TP17291	scaffold00010	
MINA247a	-	
Total		16,006,434

Appendix C continued

D12

Locus	Scaffold No.	Scaffold size (bp)
TP2378	scaffold01067	225044
TP26302	scaffold01067	
TP57154	scaffold01061	103192
TP71448	scaffold00707	161888
TP28786	scaffold00779	65995
TP109482	scaffold00008	1199225
TP56551	scaffold00779	
TP64641	scaffold00008	
TP18785	-	
TP75932	scaffold00008	
MIVCCk4b	-	
ScCDF7	scaffold00450	234850
TP76296	scaffold00083	513550
TP42528	-	
MIVCCi2b	scaffold00008	
TP15069	scaffold00573	193153
TP78724	scaffold00648	176192
TP84010	scaffold00573	
TP109463	scaffold00138	505320
TP7373	scaffold00172	1123822
TP51212	scaffold00090	501067
TP44488	scaffold00172	
TP89192	scaffold00863	195450
ScCA794b	-	
MICA794a	-	
MICA483	-	
TP108244	-	
TP87097	scaffold00158	405259
TP81077	scaffold00138	
TP3588	scaffold00863	
TP88208	scaffold00158	
Sc2ms2g08b	-	
TP56621	scaffold00929	121508
TP91414	scaffold00158	
TP53879	scaffold00155	946270
TP52274	scaffold00155	

TP90837	scaffold00191	761721
TP7512	scaffold00155	
TP103841	scaffold00155	
TP23850	scaffold00191	
TP9358	scaffold00191	
TP50688	scaffold31318	543
TP21295	scaffold00848	132412
TP3281	scaffold02624	14345
TP65086	scaffold00191	
TP106877	scaffold00018	779024
TP116710	scaffold00367	267548
TP39024	scaffold00665	288184
TP22365	-	
TP92239	scaffold00873	170595
TP37238	scaffold00040	657903
TP113146	-	
TP67977	scaffold00820	138784
TP39659	scaffold00125	1018538
TP99108	-	
TP115710	scaffold00098	491118
TP77554	-	
TP11603	scaffold00125	
TP32416	scaffold00414	253871
TP48691	scaffold00418	247728
TP113948	scaffold00098	
TP77696	-	
TP95103	scaffold00133	435715
TP20165	scaffold01725	41214
TP71152	scaffold00133	
TP86380	scaffold00162	780715
TP37201	scaffold01725	
TP4621	scaffold00800	142858
TP115419	scaffold00774	146858
TP3764	scaffold00167	606697
TP87963	scaffold00783	145460
TP18150	-	
TP26721	scaffold00098	
TP11473	scaffold00774	
TP71436	scaffold01289	72221
TP10833	scaffold00820	

TP11768	-	
TP42038	scaffold00783	
TP72741	scaffold01334	68633
TP8345	-	
TP51127	-	
Total		14,334,470

Appendix D. Markers with associated scaffold information in the Florida interspecific hybrid linkage map.

FL01

Locus	Scaffold No.	Scaffold size (bp)
H10718A	scaffold00248	334508
PHYB-2	scaffold00730	476079
F11409	-	
F19165B	-	
Pr4657	scaffold00867	130315
Vc123749	scaffold00028	817853
F15451	scaffold00233	344278
409	scaffold00357	271161
656	scaffold01169	86900
F115411	scaffold00121	453778
Pr11658	scaffold00220	352288
F17988	-	
Va5307B	scaffold02480	16660
H840A	-	
Va3830	scaffold02480	
H12016B	scaffold00069	542167
H4858C	scaffold00895	126206
Vc110398	scaffold00004	1245574
F112769B	scaffold01837	35891
49053	scaffold00545	200758
61874	scaffold01182	85421
H2199A	scaffold00452	233796
F112171	scaffold00213	555703
Pr10249	scaffold00146	710646
F17755C	scaffold00048	620420
F110659A	scaffold00735	153736
982	scaffold00058	579313
F16790	scaffold00687	167522
H791A	-	
F111127	-	
Total		8,540,973

Appendix D continued

FL02

Locus	Scaffold No.	Scaffold size (bp)
H12697B	scaffold00051	1004216
94530	-	
H405B	scaffold02718	13076
H1364B	scaffold00290	394755
NA480	-	
F110164	-	
H5407C	scaffold01209	160640
F1507C	scaffold00311	465090
Va10024	scaffold00311	
H3503B	scaffold00121	453778
H12370B	scaffold00561	195343
F15713	-	
F18589	scaffold00222	351369
H6049C	scaffold00208	358366
Total		3,396,633

Appendix D continued

FL03

Locus	Scaffold No.	Scaffold size (bp)
H14689A	scaffold00094	496738
MCYA1-1	scaffold00038	914781
F13440	scaffold01459	58207
F111046C	-	
F111046A	-	
1233	scaffold00157	532231
H2368A	-	
vco01-6ms4-f07	scaffold00217	353892
Va4721	-	
F111046D	-	
H5829A	scaffold00110	688201
H2012A	scaffold00389	478881
H14397B	-	
F14902B	scaffold00354	271555
F14902A	scaffold00354	
F14244C	scaffold00354	
H4011B	scaffold00496	269026
H11448C	-	
H12089B	scaffold01855	35253
F14244B	scaffold00354	
Pr7332D	scaffold00373	266051
CA112	scaffold00866	130337
52483a	scaffold00298	450682
H4618B	scaffold00396	256767
1254a	scaffold00298	
MAH1-3a	scaffold00361	309809
335	scaffold00078	521410
421a	scaffold00465	334516
Total		6,368,337

Appendix D continued

FL04

Locus	Scaffold No.	Scaffold size (bp)
Pr7306B	scaffold00768	222253
Contig646	scaffold00990	114518
H5430A	scaffold00490	333366
H3511B	scaffold09746	2739
FL319	scaffold00239	338672
F112477	-	
F1443	scaffold00444	238784
319	scaffold00239	
857	scaffold09219	2978
FLS-3	scaffold00236	341406
H8203B	scaffold00138	505320
H15418B	-	
MAH1-1b	scaffold00454	404705
H2578A	scaffold75015	358
F115493	scaffold00124	719687
F15898B	scaffold00124	
H6282B	scaffold00614	187375
F112175C	-	
H6377B	scaffold01693	43325
Pr4087	scaffold00550	591194
H5275B	-	
Pr11307	-	
F11129	scaffold00432	428686
247	scaffold00178	390482
FL628	scaffold00608	184985
Vac.123588	scaffold00344	275181
Vac.326725	scaffold01748	147235
F12008	-	
Va5312	scaffold00116	459447
Total		5,932,696

Appendix D continued

FL05

Locus	Scaffold No.	Scaffold size (bp)
F111684	scaffold00184	384554
Va12547	-	
F11079	-	
337	scaffold00054	606533
F15105	scaffold00090	501067
F16499	scaffold00054	
Va5347	-	
H12121A	scaffold01219	79859
503	scaffold01219	
F110305	scaffold00064	552209
1315B	scaffold00173	392570
H15247B	scaffold00296	304664
H683A	scaffold01169	86900
H7353B	scaffold01015	181697
Contig652	scaffold00334	379997
H5913B	scaffold21048	673
H12539B	scaffold00015	808311
H3680B	-	
H12370A	-	
H7353A	scaffold01015	
H7416B	scaffold01169	
F112888	scaffold01378	64870
Va4950	scaffold01466	57687
Va12555	scaffold00109	553887
CA855Fc	-	
Total		4,955,478

Appendix D continued

FL06

Locus	Scaffold No.	Scaffold size (bp)
Va5429	scaffold00002	1526805
F112408	scaffold00544	200909
F111954	-	
Va1111	scaffold00070	1171683
Va4175	-	
Va12402	scaffold01449	58643
H24A	-	
F17113	scaffold00470	449939
F112618A	scaffold00470	
H15362B	scaffold00609	184955
93226	scaffold06536	3956
F1401B	-	
F112074B	-	
H4536C	-	
1107	scaffold00015	808311
Pr5336	scaffold03833	7166
VCB-C-09527	scaffold01725	41214
CDF-5b	scaffold25726	601
FL989b	scaffold00076	526042
H2662B	scaffold00063	552545
WBCll-5	-	
H6365B	scaffold01168	86909
H1513A	-	
Pr4717	scaffold00063	
Pr7327	scaffold01209	160640
H10753A	scaffold00031	1794986
987	scaffold01007	112348
CDF-5a	scaffold25726	
H1349A	-	
H11038B	-	
Pr734	-	
Pr11538	-	
Pr15293	scaffold00009	855368
Va4551	scaffold01639	46326
Va10822	scaffold00197	375923
F1158B	scaffold00381	261236

Va4595 Va5708	- scaffold11687	2021
Total		9,228,526

Appendix D continued

FL07

Locus	Scaffold No.	Scaffold size (bp)
Va12903	scaffold00195	555132
KAN11151	scaffold00349	278175
Va537	scaffold00559	611963
MAH1-2	-	
F17050	scaffold01362	65893
Va5285C	scaffold02167	24787
1164c	-	
Va12947	scaffold00718	256838
H10394A	scaffold00072	531288
F111219	-	
Pr12789	-	
H15215A	-	
H6260A	scaffold03404	8593
1285	scaffold00246	336437
H5857A	scaffold00046	629757
H4457B	-	
H5568A	scaffold00598	186687
VCB-C-09467	scaffold01585	49359
H4011A	scaffold00496	269026
CA344	scaffold01362	
FL430	scaffold16964	788
CON78	scaffold00184	384554
H12687A	-	
430B	scaffold16964	
430A	scaffold16964	
FLSL282	scaffold76127	355
H10542B	scaffold00268	360762
SL282	scaffold76127	
H2336B	scaffold00043	873961
H10542A	scaffold00268	
H11516A	-	
ATDGK1-3	-	
SL42b	-	
H11985A	scaffold00059	571268
H394A	scaffold00694	165416
H7858B	-	

H15347A	scaffold00688	166476
1164d	-	
F110189C	-	
Total		6,327,515

Appendix D continued

LFL08

Locus	Scaffold No.	Scaffold size (bp)
H5112B	-	
505d	scaffold00016	1283226
NA1040	scaffold00261	560908
COP1-3	scaffold00160	482281
799	-	
WBll-1	scaffold00845	133277
93105B	scaffold00103	489561
51628	-	
H4668A	-	
H2200B	scaffold00017	1144947
H5387B	-	
93105A	scaffold00103	
Vco011ms1g12	-	
H7161A	scaffold40090	484
Va10141	scaffold00735	153736
H12213A	-	
H1659A	-	
H13129B	-	
785a	scaffold00202	368835
H5138B	-	
KAN1593	scaffold00003	1081109
Va4115	scaffold00499	244404
785b	scaffold00202	
Pr12571	scaffold00091	500066
CON302b	-	
Vac.124324	-	
Total		6,442,834

Appendix D continued

FL09

Locus	Scaffold No.	Scaffold size (bp)
458	scaffold00638	405928
F16608	scaffold00493	341725
Va11833	scaffold00746	151644
Pr482	scaffold00141	554342
H1400A	-	
Pr13596	scaffold00277	311847
H7310A	scaffold00379	262243
H11162B	scaffold00084	774246
F12887B	scaffold01333	68770
H3783A	scaffold00519	279991
Pr9820	scaffold10459	2386
43a	scaffold00336	279056
F12887A	scaffold01333	
F12581	-	
1007	scaffold01753	39875
H10079C	-	
F115002A	scaffold00129	437798
F13159	scaffold02837	12055
F17898	-	
Pr4574	scaffold00193	378270
64	scaffold00035	791315
CER6-1	scaffold00404	646915
Vac.96355	-	
505b	scaffold00016	1283226
CA236	scaffold00016	
Total		7,021,632

Appendix D continued

FL10

Locus	Scaffold No.	Scaffold size (bp)
52320	scaffold00784	145357
H246B	scaffold00367	267548
H5138A	-	
IPK2a-2	scaffold00095	957275
39	scaffold00095	
H13851B	scaffold00819	156914
H6242A	scaffold00002	1526805
H15005A	scaffold01213	80479
H6913A	-	
H13851A	scaffold00819	
H15067B	scaffold00029	785184
H1312B	scaffold01764	39249
H1312A	scaffold01764	
254	-	
CA794	-	
H15102A	-	
CA421	scaffold00029	
H3603A	-	
H167C	scaffold00317	289313
Va10223	-	
Total		4,248,124

Appendix D continued

FL11

Locus	Scaffold No.	Scaffold size (bp)
Pr5693	-	
197	-	
H2711B	-	
H4955B	-	
1397	scaffold00153	628812
CBF-1B	scaffold00023	778774
CBF-1A	scaffold00023	
CBF-2	scaffold00023	
F114073	scaffold00329	281605
KAN11082	scaffold00153	
1095	scaffold00256	327961
Pr9751	-	
Pr12468	scaffold00404	646915
F19969B	scaffold00210	433997
Va6917	-	
66	scaffold00650	332226
1555a	scaffold00366	267589
H6416B	scaffold00044	636084
H5876C	-	
F17536	scaffold01283	72767
Va12392	scaffold00397	256557
Va7174	scaffold00047	626560
Va13433	-	
Pr15172	scaffold01859	35154
Total		5,325,001

Appendix D continued

FL12

Locus	Scaffold No.	Scaffold size (bp)
FL118	scaffold01199	82117
213	-	
H9281A	-	
CDF-7a	scaffold00450	234850
H7277B	scaffold00054	606533
MAH1-4	-	
H5877A	scaffold00307	296234
H4744B	-	
H6943B	-	
H2656A	scaffold00008	1199225
38168	scaffold00155	946270
H2034B	scaffold00008	
H1257A	-	
CDF-7c	scaffold00450	
H4449A	scaffold00459	231054
H9772B	scaffold00459	
CON274	-	
VCCK4	-	
H15102B	-	
CDF-7d	scaffold00450	
VCCI2a	scaffold00008	
VCCI2b	scaffold00008	
H9818A	-	
VCB-C-12195	scaffold00155	
CDF-7b	scaffold00450	
H1691B	-	
H2198C	scaffold00307	
H2627B	scaffold00004	1245574
H9877A	scaffold00172	1123822
CON302a	-	
1568	scaffold00155	
H10700B	-	
Va662	scaffold00271	315724
NA377	-	
Total		6,281,403

Appendix E. Markers with associated scaffold information in cranberry linkage map.

C01

Locus	Position	Scaffold No.	Scaffold size (bp)
scf45d	0	scaffold00010	843235
scf191a	0.9	-	
vccj1c	1.4	-	
scf55c	5.5	-	
scf34s	8.6	-	
scf32j	13.1	-	
scf306f	17.1	scaffold00202	368835
scf210s	33.9	scaffold00017	1144947
scf210w	34.2	scaffold00017	
scf21n	35.8	scaffold00017	
lg1296a	40.4	-	
scf6i	40.7	-	
scf275d	52.9	-	
scf13a	61.1	-	
scf35k	66.3	-	
NA619	68.3	-	
scf439	76.8	-	
scf2001	79.4	-	
vccj9	81.6	-	
NA1040	88.5	scaffold00261	560908
scf1655c	102.8	-	
ctg259	107.9	-	
Total			29,17,925

Appendix E continued

C02

Locus	Position	Scaffold No.	Scaffold size (bp)
scf3298	0	scaffold00072	531288
scf105g	4.4	scaffold00195	555132
2ms4d10b	4.9	-	
5ms2b12	17.1	-	
scf1172	22.7	scaffold00132	545651
scf2177	30.7	-	
CA855F	37.4	-	
scf6955c	49.5	-	
scf144d	49.6	scaffold77273	352
SCAR5697	51.4	-	
SCAR5859	69.2	-	
NA1713	77.1	scaffold00006	975988
scf24k	82.8	scaffold00395	420445
ctg600	89.1	-	
scf43g	107.2	scaffold00469	354999
scf222a	108.1	-	
scf26r	111.2	-	
Total			3,383,855

Appendix E continued

C03

Locus	Position	Scaffold No.	Scaffold size (bp)
scf203h	0	-	
lg51a	9.3	-	
scf303c	23.5	-	
scf5k	24.2	-	
NA1792	29.6	scaffold00354	271555
3ms2g09	33.2	scaffold01091	97635
scf36l	60.9	-	
scf31h	69.7	-	
scf137c	77.2	-	
6ms4e4b	84.2	-	
scf2s	99.3	-	
scf1594	100.1	-	
Total			369,190

Appendix E continued

C04

Locus	Position	Scaffold No.	Scaffold size (bp)
lg6523b	0	-	
2ms2g09	0.8	-	
CA1413b	14.6	scaffold01034	107857
CA1413a	27.2	scaffold01034	
scf4860	31.5	scaffold01181	123103
NA800F	38.3	-	
scf22m	57.9	-	
scf19e	68.3	-	
scf3514	78.5	-	
scf2505a	80.6	-	
lg729b	86.8	-	
scf1p	96.3	scaffold00005	1202632
Total			1,433,592

Appendix E continued

C05

Locus	Position	Scaffold No.	Scaffold size (bp)
scf239d	0	scaffold00074	602287
2ms2a02	10.4	-	
scf283b	28.4	scaffold00604	186206
scf207d	32	-	
lg13662a	33.4	-	
scf27L	48.4	-	
vccj5	55.8	-	
NA824	58.1	-	
scf112c	61.7	-	
ctg652	87.1	-	
lg9279a	92	-	
scf17d	97.1	-	
ctg428	102.1	-	
scf44a	105	scaffold01472	57179
Total			845,672

Appendix E continued

C06

Locus	Position	Scaffold No.	Scaffold size (bp)
scf37h	0	-	
scf15b	22.7	-	
scf79c	27.4	-	
scf6213	36.6	-	
scf248	44.1	-	
CA39R	48.8	scaffold00122	640349
Total			640,349

Appendix E continued

C07a

Locus	Position	Scaffold No.	Scaffold size (bp)
NA172	0	scaffold00437	240902
NA800c	5.1	-	
ctg480	6.1	-	
scf10688	22.7	-	
scf142e	23.9	-	
scf5304	28.9	-	
SCAR2122	30.5	-	
scf9025	55.1	-	
scf2253d	67.7	-	
Total			240,902

C07b

Locus	Position	Scaffold No.	Scaffold size (bp)
ctg130	0	-	
scf511	1.2	scaffold00082	514963
scf20g	2.8	-	
Total			514,963

Appendix E continued

C08

Locus	Position	Scaffold No.	Scaffold size (bp)
scf280l	0	-	
scf41c	12	-	
CA325	13.8	scaffold00889	126730
scf2000b	27.4	-	
scf12916	36.5	-	
scf300f	48.4	scaffold70683	369
scf25m	50.2	-	
scf14j	55.5	scaffold00174	502756
CA187F	69.2	scaffold00283	308694
ctg704	70.7	-	
scf23c	71.6	-	
vccb3	76.4	scaffold01081	100076
lg15420a	82.1	-	
Total			1,038,625

Appendix E continued

C09

Locus	Position	Scaffold No.	Scaffold size (bp)
scf262a	0	-	
scf111	6.7	-	
scf72c	13.8	scaffold00250	333465
scf8l	21.8	-	
scf258d	36.4	-	
Total			333,465

Appendix E continued

C10

Locus	Position	Scaffold No.	Scaffold size (bp)
CA933	0	scaffold00167	606697
vccj3	8.6	-	
scf9e	18.7	scaffold00500	217606
lg21768b	19.8	-	
CA794F	41.5	-	
CA421	42.4	-	
Total			824,303

Appendix E continued

C11a

Locus	Position	Scaffold No.	Scaffold size (bp)
scf15903c	0	scaffold00047	626560
CA1413d	3.8	scaffold01034	107857
scf94a	8.2	scaffold00352	405778
SCAR5051	11.8	-	
Total			1,140,195

C11b

Locus	Position	Scaffold No.	Scaffold size (bp)
vccj1a	0	scaffold03147	9739
vccj1b	0.1	scaffold03147	
scf6341b	6.3	-	
scf3072b	20.6	scaffold01038	420509
scf171f	44	-	
scf46g	50.5	scaffold00116	459447
scf108b	52.2	-	
scf30g	57	-	
scf12i	57.4	scaffold00034	688467
Total			1,578,162

Appendix E continued

C12

Locus	Position	Scaffold No.	Scaffold size (bp)
scf4b	0	-	
scf2882	3.5	-	
scf6355	5.7	scaffold00066	664038
lg28559a	7.6	-	
vccj1d	9.4	scaffold03147	9739
scf39e	10	-	
Total			673,777

Appendix F. Information of 96 polymorphic SSR markers that were used in diploid blueberry linkage group analysis, include marker names, repeat motif, forward and reverse primer sequences, estimated product size, and scaffold alignment.

No.	Marker[a]	SSR	Forward Primer	Reverse Primer	Product Size[b]	Scaffold No.
1	4CL00900A	(TTCT)3	GGTTTCAGCAGAGGTCACTTTT	TCCCTCTCGATCTCCATACCTA	199	scaffold00005
2	ANPER01018C	(TTC)5	GAGCTACTGTGAGGACTTCCGT	TCATTTGCTCCTTAAGTTGGGT	278	scaffold00031
3	AOMT00011B	(TTC)9	CAAGAAACCTGGGATTTGAGTC	GAAACTACACCCACATCTGCAA	223	-
4	AOMT00197A	(TCT)6	GAACCAGAACCAGATCCAGAAG	CAAGGAAAATAGCAAAGTTGGG	232	scaffold00046
5	CHI01251A	(AAT)7	GGTTGCCAACTAAGGACGTATC	CATCACGAAGTTGTTCCTTGAA	223	scaffold01181
6	CHI01251B	(AAT)6	ATATACGGATTGCCAAAGAGGA	TCACGTTATAGTGCATGTCGAA	265	scaffold01181
7	CHS00014B	(GTA)16	CACCATTCTTATTTAATCGCCC	GTAAGCGAGAGAGGGAAGTTCA	208	scaffold00032
8	CHS00491A	(CCT)5	CTGGAAACATAAGAAACCTGGG	GTAGCAGAAACTGCTCCCACAT	293	-
9	CHS00491B	(GAG)8	AATGAAGCAAAGGCATAAGACC	GAAGTTCCTCTCCGAACTTTGA	255	scaffold00414
10	CHS00519A	(TGG)4	AGCCTAGAAGCAGCAACACTCT	GGTACTTCGATCAACACCCTCT	294	scaffold00104
11	DFR00528A	(ATA)9	CACGAAAGTAGACGCTGTGAAA	ATTTGAGCTGTCCAAAACACG	248	-
12	F3'H00479C	(TTC)5	AACAAATCAACTTCGGTGGAGT	TCTGACAATGAGCGACTGTCTT	248	scaffold00681
13	F3'H01438B	(TCC)7	ATCCGATCGTCACATAGTTCCT	TTGAGCTTCTCTCTTTTGCCTT	222	scaffold00412
14	KAN11106	(TCT)6	CTCTGCACTCATCAATCCCA	GCCGAGAAGATTTTCGATCA	195	-
15	KAN11205	(GA)16	GCCAATTGGTACGGAAGCTA	TAGTTGGGGGACACCTCATT	249	scaffold00142
16	KAN11381	(TGA)10	AGCGTATGGAAGAGGAAGCA	CCCAGTATGCTAGCAAAGCAA	222	scaffold00504
17	KAN11408	(GT)7	CCGCCTCTTGTGTCTGAAAT	ATGGACGGTACCTATTCCCC	192	scaffold00256
18	KAN11799	(TG)10	AGCCTTAACTGGCTTCCGTT	AGACTCACCGAATCCCATTG	171	scaffold00467
19	KAN12244	(TC)9	GCGTTTGAATAATTTTGGGC	CATTCTTGCATTGCTGGCTA	179	scaffold00110
20	KAN12346	(TC)6	GAAATGGGCGTTATTCG	GCTTGACCTTTCCCACCATA	204	-

Appendix F continued

21	KAN129	(TATG)5	GCAACTTGTGGAGTGGTAACAA	CAATGAAGGCACAAGCACAT	272	scaffold00228
22	KAN13486	(TC)6	CTCCTTGAAGCCAAAGCAAC	AGTACCAAAGCGCTGCAAAT	207	-
23	KAN15306	(TC)10	AGCAGCTGCCTAAACCGTAA	CCCATTTTGAACAAACAGCA	139	scaffold00455
24	KAN16539	(TG)8	TGAACAGTTCGGATCGTCAA	GTCTTCCTTAACGCGTGCAT	279	scaffold01612
25	KAN17147	(GGA)7	AGGAGCAATTTGTTGGTTCG	GGGCTATGCAAGCTAACAGC	160	scaffold00098
26	KAN1853	(GA)8	CTCAATCCCAGGTCAATGGT	ATTCGTTGGCATCGAGAATC	171	scaffold00478
27	KAN1875	(CT)12	GTTGCTTAATGGTGGTGGCT	GGAAGCGAGAAGAAGAGGGT	188	-
28	KAN20177	(TC)7	ACGTCATTTATCTCGGACGC	AGGAGCAAAGGAGATGGGTT	125	scaffold00075
29	KAN2237	(TC)7	TTGGGTACGGGATTCTTCCT	AGCTGATGGGCTTTGGTCTA	169	scaffold00075
30	KAN2260	(AG)7	ACTCAAACTGGACCAAACCG	AGAGAGGAGTTGGATCGGGT	237	-
31	KAN23334	(GA)12	TCCAAACCTCCAATTTGTCC	GGCTATCGATCCAATCCAAA	220	scaffold01836
32	KAN23741	(AG)8	CATGGATCTTGGGCTAGAAAA	CGGTTATGGGATTGGCATAC	252	-
33	KAN24307	(TC)11	GGGGTCAAAGGGTTTCATTT	AAACGGCTGAGAAATGGATG	179	scaffold00129
34	KAN24523	(AATA)8	TTTATTCTCCACACGCTCCC	AAAAGGTGCTGCCTTTTTCA	211	scaffold00018
35	KAN24598	(CT)6	CAACAGCTGCCCCTATTTGT	TCTGATCTGAGGGAGGATGG	258	-
36	KAN24806	(GT)11	CTTTCGGGTGTGTGTTGGTT	GAGATTGGGCAAATGCAAGT	191	scaffold00095
37	KAN24885	(AG)13	AACCGTGATACTACGTCGGC	GGCTCTGTTTGACTTTTCCG	243	scaffold00191
38	KAN24973	(TTTA)5	CACTCACGGGACAGTAGCAA	GGCCGAGAAAGGTATCAGGT	182	-
39	KAN25281	(AG)17	GCACTCATACTCCCCACACA	ACCCCAAATGGAAATCAACA	205	scaffold00550
40	KAN27020	(GA)8	CAAAGCCCAAACAATCAACA	GCCCAGGTGATCTCTCACTC	231	scaffold00576
41	KAN27179	(AG)11	TATGTCTGTGGTGGTGGTGG	CGGAGTGCAGAGTCAAGTCA	175	-
42	KAN27356	(TG)13	TGGAAGCCGTTCAACTTCTC	CATTGGCGTAGGGTTTGATT	165	scaffold01466

Appendix F continued

43	KAN27743	(AC)14	ATTTACCCAAGGCCCAAAAA	TGTTCCTCGAGTTAATGGGG	229	scaffold00049
44	KAN29551	(AG)12	TGAACGGATTGTACCCTTCC	TGCTGTGTTTGAGTCTTCCG	239	scaffold00015
45	KAN3462	(CT)6	TCTCTTCACGACTCCTCCGT	TCACAACATCTGTGCCAACA	143	scaffold00147
46	KAN40732	(TAT)16	AAAACGCGATGAATGGAGAG	TGCGCCAAATTATGAATGAA	216	scaffold00208
47	KAN41355	(TC)21	AGGTTGGGAAGGCACTTTTT	TGAATACATCGAGCACACGC	248	scaffold00436
48	KAN41365	(AG)18	GTCGCCTTCTTCCTCTTCCT	CCCACCTCAAGACTTACCCA	175	-
49	KAN41661	(GAG)5	TCCAAGGGTTCCAAACAAAG	CAATTCTGCAGGTTCGTTCA	193	-
50	KAN42567	(GA)13	CCAACCACAAGTGAGCAGAA	TGGTAAGCCTCCAACGGTTA	185	scaffold04452
51	KAN43117	(CT)12	AGGCTGTTGCTGGAACATTT	AAACATTGGCAAACAGAGGC	191	scaffold00078
52	KAN43405	(CT)14	GAACGAAAGTGGTCCGTTTG	GTTTTGCGGGACCTATTTGA	223	scaffold00581
53	KAN711	(ATT)10	CCAGTGGCAACTCCAAGACT	GCATTGAGACTACCTAACAACGC	225	-
54	KAN79	(GA)12	TCCCGTGTTATGGTCCTTGT	GCCAGGTTTCTCTTTCCCTC	180	-
55	scaf00001-3	(TGG)4	AATCTTTCAAACCCTGAAGCAC	GAAGGGAAAGATATTCCCAAGG	264	scaffold00005
56	scaf00001-4	(TAT)10	CTGTCAACATGATGTGCAACTT	TGAATACAAAAGGACTTCACCG	285	scaffold00005
57	scaf00004-3	(CGA)4	CTTCATAGAAGGAAATGTCCCG	TTCATCTTCATCTTCCTCGTCA	279	scaffold00020
58	scaf00004-4	(CT)5	CTCGGCTCGTTGACTAAAACTT	AGCCCTAATAATTCCCAGAAGC	268	scaffold00020
59	scaf00004-6	(ATC)8	TAATAGGTAGCTGGGCCATGAG	CGTTCCATCTCCTTAATCCTTG	276	-
60	scaf00007-2	(AGA)8	TGAACATGGGGAAGACAAAAG	CCAACCTCGTTTAGGTCAAGAA	300	scaffold00021
61	scaf00009-4	(AAG)4	CGCTGTTGTCCAAATCTTCATA	TCCTTAGCTGAGCTCGGAGTTA	297	scaffold00023
62	scaf00013	(TTC)4	CACAAAGTGAGGTAATCGTCCA	GCGAAAACTACAGGATGCTAGG	238	scaffold00045
63	scaf00021	(TCC)7	GCCCACTTCATAAGATGGCTAC	CTGCTACTTGTAGGGCTAGGGA	167	scaffold00014
64	scaf00033	(TAG)6	TATGATCTCGTTTCGTGCTGAC	ACTCTAGCCAGAGAACGCAAAA	256	-

Appendix F continued

65	scaf00062	(CCA)6	CAACATGAGCTCCGAAAACTCT	AGAGAGAAGCAAATATCGACGG	125	-
66	scaf00069	(AAT)8	TTTCACTTGAATGGTTGGTCAG	GATCAGCCAGTCCTACCTTTTG	292	scaffold00013
67	scaf00074	(AAT)4	GGGTTGGAATCTGATAACCTCA	ATCTTTGTCGTGGGTAGTTGGT	132	scaffold00172
68	scaf00090	(TAT)4	GTTTGAACCTTGGTTGAGCTTC	AATGACCAAACCAGGTGTCTTT	255	scaffold00183
69	scaf00100	(AGA)15	GGACTGCGTTTTAGGGTACTTG	GCGACTATGATTTCGGTACCTC	253	scaffold00227
70	scaf00113	(CCA)4	ACCATTGTCACCACCCTTACTC	ACCACGACGTTTAAGGCTATGT	205	scaffold00210
71	scaf00125	(GAA)4	CTAAAAAGCCAGCTTTAGCAGC	ACGATTGAGCCTTTTTGCTAAG	265	scaffold00209
72	scaf00137	(GTT)5	ATTGTCCGAAACGGACTAGAAA	TTCGACAGATGATATGAAACGC	193	scaffold00105
73	scaf00155	(TAG)6	GAAAACCGTCAAAACCAGTAGC	GGCAGTGGATCAAAAAGAAAGA	100	scaffold00215
74	scaf00156	(CAC)4	CCATCATCACTACCACCATCAC	AGGCAGGGAGAGAGAAATACTG	283	scaffold00283
75	scaf00169	(AAG)9	TGAGTAGGGCCAATATGATGTG	TGGTCATACGACTTTCAAGGTG	262	-
76	scaf00206	(CAT)5	TAACGCCCCTAATTTTGAGAAC	AGAGAAACCGTAGCCTACCTGA	172	scaffold00244
77	scaf00216	(TAC)4	CCGTTCATTTTGTTAGCCTCAT	AGGAGATTGTTAGGGTTCCGTT	275	scaffold00143
78	scaf00223	(TGG)5	GCAATCGATGATTCACAAGAAG	CTTTCCCTTTCATTTCTCTCCA	280	scaffold00364
79	scaf00246	(TG)14	TAAGAAAAGGAGAAGCACAGGC	CATCACACTAGCAGAAGGCATC	292	scaffold00319
80	scaf00267	(CGA)4	TCTCCCCTGTCAACTACACCTT	AAGCTCGTGGAGGTGATTAGAG	173	scaffold00406
81	scaf00281	(GTG)5	TGCTCCCGAATCAAATAAGACT	TCATTGTTTTTGCAGCTACACC	143	scaffold00416
82	scaf00293	(ACG)4	AAGTGCTCTCCAACTCATCTCC	CAATTGTGCCATCATAGCTGTT	101	scaffold00134
83	scaf00295	(GAT)4	TCCGTGTCCGTGCTACATATAC	TCTCTCTCATTACCATGACCCA	160	scaffold00469
84	scaf00306	(GAA)6	GCAATATTTTGAGGAGGAGGTG	CATGTACGTTTGAGTCGATGGT	161	scaffold00459
85	scaf00346	(TC)10	AGGGGCAGACTCATACAGTCAT	CGAAGTGAGTTCCCTTGAGAGT	247	scaffold00493
86	scaf00369	(GTG)5	AATAGTGCCTCTTCCCCAGATT	GGGACTCCACCACTACTACTGC	297	-

Appendix F continued

87	scaf00392	(AAG)4	TCATCGTTAATACATGGTGGGA	CAGACCATTCACATCGGAACTA	299	scaffold00038
88	CRA_lg6523b	-	CCATCTACCACGGCAGAGAT	GCATATTTTGGTTGGATCGG	-	-
89	CRA_scf10688	-	TCACTTTTCTTTCATGCCCC	GTGCTCCCACTAGCCCATAA	-	-
80	CRA_scf112c	-	ATGTGATTCGCGAAGGATTC	GAAATCCGGGGGTGTAAACT	-	-
91	CRA_scf11l	-	TAATGAGTGCTGGTTCTGCG	TTCAAATCCACGTCAGCAAA	-	-
92	CRA_scf12916	-	GGAGATGGATTTGGCAAGAA	ATCCATGTGGCAGCAGTGTA	-	-
93	CRA_scf15903c	-	ACTTACCCACGAGCCTACCA	GAAGGAGAAAGTGACGTCGG	-	-
94	CRA_scf203h	-	AAGTTACAACGGTTCGTGGC	TGCAACATTGTGATGGTCCT	-	-
95	CRA_scf26r	-	ATGATGTTGGATGTGCCTCA	TTCCTCAACAAACCCTCCAC	-	-
96	CRA_scf79c	-	GGTTCTTCGTGGCATGATAGT	CCAAATAACCCAGGAGAGCA	-	-

[a] markers with name initiate with "scaf" were derived from Newbler assembly; "KAN" were obtained from previous SSR mining research (Bian et al., 2014); "CRA" refers to cranberry SSR markers (Georgi et al., 2013)
[b] Expected PCR product size without M-13 (-21) sequence