SSR Markers for Genes that Regulate Fruit Firmness in Blueberries

Stewart Dalton

Laboratory Methods in Genomics

Spring, 2013

**Abstract**

 As the popularity of Blueberries (*Vaccinium corymbosum*) has significantly increased over the past few decades, so has their status as an important commercial crop in the United States. A major issue facing the blueberry industry is the short shelf life of blueberries as blueberries lose their firmness quickly after they are harvested. I investigated which genes play an important role in making blueberries soft. In the course of this investigation, I identified 23 SSR primer pairs that can be used by other investigators to breed and screen for blueberries that will stay firmer for longer. These blueberries will not only satisfy consumers’ desire for firmer blueberries but will also improve the blueberry industry.

**Introduction**

Blueberry (*Vaccinium corymbosum*) demand and production has more than doubled since the late 1970’s. A major reason for this demand is the increasing awareness of food nutritional value. Blueberries are known to have excellent antioxidant levels among fresh fruits. In addition, the United States saw more than a 60% increase in commercial blueberry acreage between 1990 and 2005 (Vicente *et al*., 2007). Blueberries have become an important commercial crop and any improvements that can be made to the fruit would improve their market value and the consumer’s experience. A significant problem that faces the blueberry industry is the short post-harvest life of the fruit (Jiménez-Bermúdez *et al.,* 2002). Blueberries lose their firm texture quickly once they have been harvested. Most consumers prefer firm blueberries as opposed to soft, mushy ones. In my research, I identified genes, and their associated SSR primer pairs, responsible for the fruit softening process. Other researchers may use my research to breed blueberry plants with the desired alleles and produce blueberry fruits with optimal firmness.

 The softening process begins with alterations to the primary cell wall and middle lamella. The primary cell wall is made of a network of polysaccharides such as pectic substances, hemicellulose and cellulose that surrounds the cells of a plant and its fruit. As a fruit ripens, these polysaccharides undergo depolymerization and are solubilized. The middle lamella, which is made of pectin, acts as the glue that holds the primary cell walls of the fruit together (Proctor and Peng, 2001). The process that leads to the disassembly of the primary cell wall and middle lamella is what causes a fruit to soften during ripening (Dotto *et al*., 2006). However, the biochemical pathway by which fruits undergo the softening process is not fully understood (Dotto *et al*., 2006). In addition, the classification of the blueberry as a climacteric or non-climacteric fruit is unclear as the blueberry exhibits characteristics of both. Climacteric fruits are characterized as having a respiratory burst and increased ethylene production at the onset of ripening. Non-climacteric fruit do not display such a dramatic increase in respiration at the onset of ripening, although an upsurge in ethylene production is sometimes seen (Zifkin *et al*., 2012). As more literature was available on the topic, I investigated the climacteric pathway of fruit softening.

Fig. 1- the primary cell wall and its constituent polysaccharides are diagramed. The middle lamella, which acts as the glue that holds cell walls together, is also pictured (PolySac3DB).

 The key regulator in fruit ripening in climacteric fruits is the plant hormone ethylene. An increase in ethylene biosynthesis initiates the ripening and softening process in plants (White, 2002). The pathway begins when S-adenosyl-methionine (SAM) is converted into 1-aminocyclopropane-1-carboxylic acid (ACC). This reaction is catalyzed by ACC synthase (*ACS2*). ACC oxidase (*ACO1)* catalyzes the oxidization of ACC into ethylene.

Ethylene

ACC

SAM

Ethylene directly affects the transcription of many ripening-related genes, which leads to the softening of fruits (Fonseca *et al.*, 2005). However, I do not fully understand the process by which ethylene interacts with these genes, as the literature was unclear in this matter. Six genes in the tomato (*LeETR1-6)* encode for ethylene receptors. In the absence of ethylene, these receptors inhibit the expression of ethylene-responsive genes, which causes the fruit to be more resistant to softening. A mutation of the *AtETR1* gene (ortholog of *LeETR3)* leads to the *Never ripe (Nr)* phenotype. Fruits displaying the *Nr* phenotype do not ripen and, thus, stay firmer for longer. The *AtETR1* gene is the only ETR gene whose expression changes significantly during ripening. The production of *AtETR1* during ripening reduces the ethylene perception of ethylene-responsive genes. Thus, the fruit does not ripen but stays firm, as it cannot respond to ethylene production (White, 2002).

Fig. 2- this diagram depicts the conversion of SAM into ethylene, which is an important plant hormone that initiates the ripening process in climacteric fruits. (NSF, year).

 A family of genes that is stimulated by ethylene is the expansin gene family. Expansins (*EXPA1-3*)are enzymes that loosen the cell walls of fruits and promote the expansion of the primary cell wall (Scherban *et al*., 1995). Polysaccharides that compose primary cell walls are long polymers that remain tightly bound to one another through noncovalent interactions such as hydrogen bonds. Expansins weaken the hydrogen bonds between the polysaccharides and make the polysaccharides more susceptible to attack by a variety of cell-wall-degrading enzymes. As the polysaccharides in the cell wall are depolymerized, the fruit softens. The overexpression of the *LeEXP1* gene leads to softer fruits while silencing of the gene leads to firmer fruits. In *Arabidopsis*, at least 24 expansin genes have been found (Cosgroove, 2000).

 As expansin proteins loosen the bonds of the polymers that compose the cell walls of fruits, polysaccharides such as cellulose and hemicellulose become more exposed and vulnerable to attack by cell-wall-degrading enzymes. These cell-wall-degrading enzymes are directly activated by the production of ethylene. However, the way by which ethylene activates the cell-wall-degrading is not well understood. One such group of enzymes is the polygalacturonases *(PG1-2)*. These enzymes depolymerize pectins by cleaving the galacturonic acid backbone of the pectin polymers. An enzyme that works in cooperation? with the *PG* proteins is pectinmethylesterase (*PMEU1*). Pectinmethylesterase de-esterifies pectin, which makes them vulnerable to attack by PG’s (Proctor and Peng, 2001). Ethylene induces an accumulation of *PG* mRNA in pears and tomatoes. The beta-galactosidases (*TBG4*) comprise another group of cell-wall-degrading enzymes that is activated by ethylene. An additional group of enzymes important in the fruit softening process is the Xyloglucan endotransglucylase/hydrolases (*XTH*) enzyme group, which is also induced by ethylene. These enzymes catalyze the cleavage of beta-D-glucan backbones and catalyze the re-ligation xyloglucan molecules. A final group of enzymes that are positively regulated by ethylene are the beta-D-xylosidases (*XYL*). These enzymes catalyze the break down of xylan of the matrix polysaccharide (Fonseca *et al.*, 2005). All of these genes are induced by ethylene which leads to the softening of fruit. If blueberries can be bred that have less active cell-wall-degrading enzymes or produce less ethylene, these blueberries may be able to remain firmer for a longer period of time.

Fig. 3- this schematic outlines how ethylene is related to and activates the important enzymes catalyze the depolymerization of the polysaccharides of the cell wall and middle lamella.

 A final group of enzymes that play an important role in the degradation of pectin are the pectate lyases (PL). Four such genes were identified in *Arabidopsis*: *AT59, AT1G11920, AT5G55720* and *AT2G23900.* PL enzymes work by depolymerizing the pectin in the middle lamella and primary cell wall. PL enzymes carry out a beta-elimination reaction, which cleaves the pectin polymers into oligosaccharides. PL enzymes play a more significant role in the softening of fruits than other enzymes such as pectinmethylesterase and the polygalacturonases. In addition, the reduction in the amount of PL mRNA leads to firmer fruit without having significant affects on other important characteristics of a fruit such as it’s color or weight (Jiménez-Bermúdez *et al.,* 2002). Thus, breeding for blueberries that produce lower amounts of PL mRNA may lead to blueberries that will remain firmer for longer and can be used commercially.

 Another collection of genes that are important in fruit softening is a group of single gene mutations found in tomatoes known as *RIN, CNR, NOR* and *NR*. *RIN, CNR, NOR* and *NR* are pleiotropic and code for regulatory proteins that control the ripening process, including the production of ethylene. The *NR* gene encodes for a protein similar to that of the *Arabidopsis* ethylene receptor *ETR1*. The *NOR* and *RIN* mutations are already used commercially in extending the shelf lives of tomatoes, although how the *NOR* and *RIN* genes are used to do so was not clear in the literature (Seymour *et al.,* 2002). The *CNR* gene is important as it acts as an activator of the *RIN* gene (Martel *et al.,* 2011). In both *RIN* and *NOR* mutants the expression of *TDR4*, which plays an important role in the ripening process, is inhibited. A similar MADS-box gene to *TDR4* is *DEFH28. DEFH28* is also inhibited by the *RIN* and *NOR* mutants (Seymour *et al.,* 2002).

 The *RIN* gene, which is a MADS-box gene (*MADS-RIN)*, acts as a master regulator in the ripening pathway. It does so by interacting with the promoters of many genes involved in the ripening process including the genes responsible for ethylene production, ethylene perception and cell-wall-degrading enzymes. Decreased ethylene production is found in *RIN* mutant fruit. Thus, *RIN* plays an important role in the regulation of ethylene production. The promoter of the *ACS2* gene, which codes for ACC synthase, is targeted by *RIN* while the promoter of the *ETR1* gene is enriched *RIN*. *RIN* also indirectly inhibits the *ACO1* gene by binding to the promoter of the *HB1* gene. *HB1* binds to the promoter of *ACO1* and is necessary for its expression. The *RIN* protein also associates with the promoters of the *PG2* and *EXPA1* genes, which causes them to be silenced in the *RIN* mutant fruits resulting in more firm fruits. *RIN* also binds to the *NOR* promoter causing strong ripening inhibition. The *RIN* gene plays a variety of roles in making a fruit remain firmer as it inhibits other genes that are vital to the softening process (Figure 4). For this reason, breeding for blueberries that express the *RIN* mutation may be commercially important to extend the shelf lives of blueberries.



Fig. 4- this schematic outlines how the *RIN* gene inhibits the expression of *ACS2, HB1, EPA1* and *PG2*, which leads to firmer fruits. I have also pictured that *CNR* activates *RIN* and that *RIN* activates *ETR1*.

**Methods**

 I identified the genes and proteins that I used for investigation through a literature search. Almost all of the genes that I found came from *Solanum lycopersicum* (tomato) or *Arabidopsis thaliana*. I needed to identify the blueberry orthologs to these genes. Once I identified a gene or protein of interest, the first step was to find the amino acid sequence the codes for the protein. I accomplished this amino acid sequence search at the National Center for Biotechnology Information (NCBI, year) website as seen in Figure 5 (I have selected the gene *EXPA1* as an example to illustrate the procedures of my investigation).



Fig. 5- I executed a search for the gene *EXPA1* on the home page of the NCBI website.

Once I executed the search, I selected the gene from a model organism or an organism similar to the blueberry (such as *Solanum lycopersicum* or *Arabidopsis thaliana*) as seen in Figure 6.



Fig. 6- the gene of interest was selected for investigation. Genes of closely related species to the blueberry, such as grapes or strawberries were acceptable to use. In addition, other plants such as the tomato and *Arabidopsis* were frequently used to find orthologs as these plants are well-studied species.

I found the amino acid sequence of the protein of interest by choosing a protein from the list of “Related Sequences”. Once I selected a protein, I found the amino acid sequence on the bottom of the web page under the heading, “Origin”. I copied and pasted the amino acid sequence into a “TextEdit” file and saved. It is important to save the TextEdit files in FASTA format. I saved the file as a “Plain Text File” and I put the extension “.txt” at the end of the file name. Once I saved the amino acid sequence in the proper format, I performed a local amino acid BLAST against the blueberry genome. First, I produced a local database into the Terminal using this command line code: “cd Desktop/ncbi-blast-2.2.27+”. When the database was generated, the line “w11304:ncbi-blast-2.2.27+ stdalton$” appeared indicating that the Terminal was ready for the BLAST commands. An amino acid BLAST was being performed, thus, I used the command “bin/tblastn -query EXPA1.txt -db Sorted\_454Scaffolds.txt”. An example of these procedures can be seen in Figure 7.



Fig. 7- these are the Terminal commands for running an amino acid BLAST. The second line in the picture shows where I entered the database in which I was working (cd Desktop/ncbi-blast-2.2.27+). The third line in the picture indicates the specific TextFile I was BLASTing in the Sorted\_454Scaffolds (bin/tblastn -query EXPA1.txt -db Sorted\_454Scaffolds.txt).

Once the Terminal processed the query, a list of possible scaffolds on which the gene of interest might be located appeared on the Terminal (Figure 8). In most cases, I used the first result with the smallest E-value. However, I did not use any scaffold with an E-value greater than 1e-04 or 1x10-4.



Fig. 8- these are the BLAST results indicating the possible scaffolds on which the gene of interest is located. The possible scaffolds appear on the far left of the figure with the base pair length of the scaffold next to it. On the far right of the figure, the E-values of the corresponding scaffolds are displayed.

In this example, the scaffold “00393” was used. Once I identified the desired scaffold, I searched for the scaffold using the “command F” function, in local blueberry genome in the file “Sorted\_454Scaffolds.txt”. Once I found the scaffold, I copied and pasted the whole sequence a new TextEdit file. I saved this file in FASTA format (Figure 9). At the top of the TextEdit I used a “greater than” sign (>) followed, with no spaces, by “scaffold#”. The next line starts the scaffold sequence.



Fig. 9- an example of FASTA format in which the TextEdit must be saved. Note that the full scaffold is not displayed in this figure.

Once the scaffold is saved as a TextEdit file in FASTA format, it is ready to be uploaded onto Vaccinium (REF), which is a genome database for the blueberry. On the home page of the website, I selected the “Tools” tab. Under the “Tools” tab I selected the “SSR” option. During my investigation, only dinucleotide and trinucleotide motifs were used. All other options were “zeroed” (Figure 10). Once the necessary information is filled out and the scaffold is uploaded, I hit “Submit” and waited for the SSR results to be emailed to me. The time it took for the SSR reports to arrive at my inbox varied depending on how large the scaffold was that I submitted.



Fig. 10- the webpage where the scaffolds were submitted. The “Hexanucleotides” and “Pentanucleotides” options were zeroed.

The results arrived as Excel files. I selected three SSR primer pairs (if at least three were generated) for each gene. I selected the SSR primer pairs based primarily off of two criteria. The first was the number of motif repeats. A larger number of repeats was preferred over a shorter number. In addition, I considered how close the repeats were to the scaffold in the genome. I selected closer repeat motifs as opposed to ones farther away. However, sometimes I had to sacrifice one criterion for another. For instance, many times I found a set of SSR primer pairs that contained a large number of motif repeats but was also far away from the gene of interest. Thus, I introduced subjectivity into the process by picking that SSR primer pair even though it was not ideal. Also, I did not use SSR primer pairs that were without forward or reverse primers nor did I use SSR primer pairs that contained a large number of “CTCTCTCTCT…” repeats.

**Results**

 In Appendix A, I have presented a list of the SSRs that I identified. I was able to obtain three SSR primer pairs for 22 of the genes of interest, one for one of the genes of interest and zero for three genes of interest. Table 1 lists all of the genes of interest and how many primer pairs I was able to identify.

|  |  |
| --- | --- |
| Gene of Interest | Number of SSR Primer Pairs  |
| *AT59* | Three |
| *PME3* | Three |
| *PGA4* | Three |
| *AT2G23900* | Three |
| *TCTR2* | Three |
| *AT1G11920* | Three |
| *AT5G55720* | Three |
| *TBG4* | Three |
| *PG1* | Three |
| *PG2* | Three |
| *XTH8* | Three |
| *XYL4* | Three |
| *EXPA1* | Three |
| *EXPA2* | Three |
| *EXPA3* | Three |
| *MADS-RIN* | Three |
| *ACS2* | Three |
| *ACO1* | Three |
| *ETR1* | Three |
| *ETR6* | Three |
| *TDR4* | Three |
| *DEFH28* | Three |
| *PMEU1* | One |
| *CNR* | Zero |
| *NOR* | Zero |
| *HB1* | Zero |

Table 1- the genes of interest that I investigated and the number of SSR primer pairs that I was able to identify.

Another tool used during this investigation was GenSAS. GenSAS stands for “Genome Sequence Annotation Server” and can be used for genome structural annotation. The website offers a variety of tools for its users. I sampled the tool named “FGENESH” with a scaffold that I know to contain a gene. This tool is an intrinsic gene predictor. However, the tool was not effective and output data yielded zero results. A picture of the yielded output results is shown in Figure 11.



Fig. 11- the GenSAS tool “FGENESH” yielded no output results.

**Discussion**

 With the identification of the genes involved in the fruit softening process, researchers can use these genes to establish ways to extend the shelf lives of blueberries. Other investigators can use the SSR primer pairs that I have found to generate PCR products of different sizes linked to particular alleles. PCRs for different alleles will be of different weights and produce different banning patterns when tested on gels. Thus, researchers with large cultivers of blueberry plants can breed blueberry plants to produce offspring with the desired genotype. The researchers can screen each germinating blueberry with the set of PCRs that was generated from the SSR primer pairs. Using this screening method, investigators will be able to quikly tell if the young blueberry plant possesses the desired alleles and will express the desired phenotype or not. Thus, the process by which firmer blueberries can be obtained will be expediated as researchers will no longer have to wait until blueberries are fully grown to observe if they posses the desired trait or not.

The blueberry is an important commercial crop in the United States. Any improvement that can be made to the blueberry’s quality will not only benefit the consumer but also all who contribute to the production of the blueberry. It is of the upmost importance to make all strides necessary to improve the quality, production and profits made by the blueberry industry as such gains will also benefit the U.S. economy. With the use of my research, other investigators will be able to cross breed and produce blueberries that will stay firmer for longer. These blueberries will increase consumer satisfaction and the sales of the blueberry fruit.

**Appendix**

**1. PGA4 found in Scaffold 00119 (query sequence starts at base 268649 on scaffold)**

1)

For Primer CCCCCATTTTACAATTCACACT

Rev Primer GGACTTCAATTGATGCTAAGGG

repeats (ct) x34 PCR product= 283 bp & start at base 216423

2)

For Primer TATAAGGGGAAATGTTGGATCG

Rev Primer AAAACTAGGAGGGTGGCTTCTC

repeats (ct) x32 PCR product= 292 bp & start at base 193059

3)

For Primer CCCACAAGAACACACCTCTTAG

Rev Primer CACCTTCAGTCCCAGAAACATT

repeats (ct) x30 PCR product= 284 bp & start at base 282049

**2. PME3 found in Scaffold 00668 (query sequence starts at base 93678 on scaffold)**

1)

For Primer TACGTTGGTCGTACCTCTCCTT

Rev Primer CGGTATGGTAGCAGTTAGGAGG

repeats (tc) x22 PCR product= 119 bp & start at base 110733

2)

For Primer AGGAGGAAGAGGAGAGAGAGGA

Rev Primer GGCACTCAAATTTTCTCAAACC

repeats (ga) x14 PCR product= 211 bp & start at base 151233

3)

For Primer ATCGAGTCACATTCTCCCTTGT

Rev Primer GGGAATTGTGATGGAGAGAGAG

repeats (ct) x14 PCR product= 290 bp & start at base 103518

**3. AT59 found in Scaffold 00222 (query sequence starts at base 219924 on scaffold)**

1)

For Primer TGGATAGTGTTCTAATGCACGC

Rev Primer CCTTTTGCTATTGTGATGCGTA

repeats (ga) x20 PCR product= 273 bp & start at base 11521

2)

For Primer AAGGGGTGTTCAACAGAACATA

Rev Primer GATTTCTGTTTTGTAGCGAGGG

repeats (ga) x20 PCR product= 264 bp & start at base 224648

3)

For Primer CACAAAGCGACGGAAATATACA

Rev Primer CTGTGAGCTATGAAACACCGAC

repeats (ag) x16 PCR product= 271 bp & start at base 185214

**4. PMEU1 found in Scaffold 3067 (query sequence starts at base 14197 on scaffold)**

1)

For Primer AAAGAGATTGGCCGTAATCAAG

Rev Primer TGGACAACCGACAGTGATACTC

repeats (ta) x8 PCR product= 298 bp & start at base 12476

**5. AT1G11920 found in Scaffold 13 (query sequence starts at base 142962 on scaffold)**

1)

For Primer TGGGAAAGAGATGTCTCCGTTA

Rev Primer GGTTGATGGACGAACTGATGT

repeats (ag) x20 PCR product= 300 bp & start at base 87393

2)

For Primer ACAAATGGCACACAGAAGACAT

Rev Primer GCAACTGGAGTAATCCTGGAAC

repeats (ca) x16 PCR product= 170 bp & start at base 51659

3)

For Primer CTCAGATCCCATCTCTGATTCC

Rev Primer CGAAAAGCTAATTCACCAAAGG

repeats (ag) x16 PCR product= 220 bp & start at base 267173

**6. AT5G55720 found in Scaffold 1047 (query sequence start at base 84801 on scaffold)**

1)

For Primer AAACGGTTTAAAGGTACGTTCG

Rev Primer GAAGTGCAGATGTTTCAGCAAG

repeats (ct) x14 PCR product= 295 bp & start at base 1811

2)

For Primer GCCTCTCTCGCTCTTTGATTT

Rev Primer GAAGTGCAGATGTTTCAGCAAG

repeats (tc) x11 PCR product= 210 bp & start at base 8462

3)

For Primer TGTGATGTTGTGAGCTGAGATG

Rev Primer CTTCCCTACCTCGTCTTCCTTT

repeats (ag) x10 PCR product= 178 bp & start at base 73113

**7. TBG4 found in Scaffold 780 (query sequence start at base 24863 on scaffold)**

1)

For Primer TCTTTTCTGCCCTTCTTTTGAG

Rev Primer AACCAATCCCCTCTCAAGATTT

repeats (ag) x13 PCR product= 184 bp & start at base 29907

2)

For Primer TCTTTTCTGCCCTTCTTTTGAG

Rev Primer AACCAATCCCCTCTCAAGATTT

repeats (ga) x10 PCR product= 184 bp & start at base 29828

3)

For Primer ACAGCATGCAGAAACATTATGG

Rev Primer CAGGGGACAAGGCAAAAATA

repeats (tg) x8 PCR product= 287 bp & start at base 95163

**8. PG1 found in Scaffold 1540 (query sequence start at base 60890 on scaffold)**

1)

For Primer CGATCTTCTTCGTCTGGTTCTT

Rev Primer ATTGATTGATCGATTCGTAGGG

repeats (ag) x17 PCR product= 136 bp & start at base 54357

2)

For Primer ATTGGAGTTGCAAGTGAGGAAT

Rev Primer ACGGTACGGTAGGTAGGTAGGT

repeats (ct) x11 PCR product= 268 bp & start at base 45817

3)

For Primer GGGACTTTGCCCCTACTTTAAT

Rev Primer GGGCCTAAAAGTCAAAGTTCCT

repeats (ta) x8 PCR product= 264 bp & start at base 72078

**9. PG2 found in Scaffold 695 (query sequence start at base 67133 on scaffold)**

1)

For Primer TCTTCTCAGGCTCTTAACCAGG

Rev Primer AACATCTCCATTCTCCCGACT

repeats (ga) x16 PCR product= 159 bp & start at base 54518

2)

For Primer CTCTCCATCATCTGCTGCTCTA

Rev Primer TATGGTATGGGAGCATTGTTGA

repeats (ct) x15 PCR product= 264 bp & start at base 22256

3)

For Primer TGCTCTGAACAAATGGAGAGAA

Rev Primer AGCCAAAATTCACTATAGGCCA

repeats (ct) x8 PCR product= 152 bp & start at base 63425

**10. XTH8 found in scaffold 1511 (query sequence start at base 14570 on scaffold)**

1)

For Primer GATGGACTTGTGTGTGTGTGTG

Rev Primer TGAGCCTTTGAAAACCTCTCTC

repeats (ag) x25 PCR product= 126 bp & start at base 27538

2)

For Primer TTCTGGTCTTGGTACGAATGTG

Rev Primer CGTAATCAACAGTTCACTCCCA

repeats (ga) x12 PCR product= 120 bp & start at base 33443

3)

For Primer GAGGGTGAACAAATTAAATGGC

Rev Primer TGCTTCTCTCCCAAATTTGACT

repeats (tc) x10 PCR product= 283 bp & start at base 16624

**11. XYL4 found in scaffold 73 (query sequence starts at base 148440 on scaffold)**

1)

For Primer GGGCGAAGCCTTTTCTATTATT

Rev Primer TGAACCAGACCACTAAACAGGA

repeats (tc) x14 PCR product= 202 bp & start at base 192915

2)

For Primer TCAGCCAAATCTGAGACTGAAA

Rev Primer AGTGTCCGTGGACTTAGACGAT

repeats (ac) x13 PCR product= 179 bp & start at base 137490

3)

For Primer ATCTTACGGGATAACGGAACCT

Rev Primer CAAAGTGACTAGTCGGTTCAAAG

repeats (ttg) x12 PCR product= 232 bp & start at base 141898

**12. EXPA1 found in Scaffold 00393 (query sequence starts at base 189087 on scaffold)**

1)

For Primer CAGCCCATAGGAAGAGAGAAGA

Rev Primer GAGAGTGATGAGAGAGAGAGAGAGA

repeats (ca) x20 PCR product= 248 bp & start at base 59984

2)

For Primer ATTCCGTCGATTTCTAGGGTTT

Rev Primer TTCTCTCTCTCTCTCTCTCTGCAT

repeats (ag) x10 PCR product= 106 bp & start at base 226011

3)

For Primer ACTCCATCAATCAATGTACCCC

Rev Primer GAAGATGGATGGATGGTTGTTT

repeats (tc) x9 PCR product= 183 bp & start at base 243836

**13. EXPA2 found in Scaffold 00018 (query sequence starts at base 51867 on scaffold)**

1)

For Primer AGACACCTCCTCTCTCTCCCTT

Rev Primer ACGTGCTCTTATCGGACATTTT

repeats (ct) x35 PCR product= 253 bp & start at base 93625

2)

For Primer CCCATGCACAACTCTTGATTTA

Rev Primer CCAGACTGACCCACACAGAGTA

repeats (tc) x20 PCR product= 175 bp & start at base 425378

3)

For Primer TGAGCATGGAACACTTCATAGG

Rev Primer GAAATGGCACCAGTTTCTCTTC

repeats (aag) x19 PCR product= 239 bp & start at base 126251

**14. EXPA3 found in Scaffold 00343 (query sequence starts at base 10714 on scaffold)**

1)

For Primer GTTAACAAGTGTCACGACCGAA

Rev Primer CCCAGAACGGGAGTATCATTAC

repeats (ga) x12 PCR product= 240 bp & start at base 12172

2)

For Primer CGGGAAAATGAAGAGAGAGAGA

Rev Primer ATCATAGTGACTACCCCGTGCT

repeats (ag) x11 PCR product= 249 bp & start at base 11599

3)

For Primer GCACGGAGAGAGAGAGAGAGAG

Rev Primer ATGGCTCAGATTTGTTTGAGGT

repeats (ag) x11 PCR product= 296 bp & start at base 12217

**15. TCTR2 found in Scaffold 00005 (query sequence starts at base 409965 on scaffold)**

1)

For Primer TGAGATGAGCGAGTGAGATACG

Rev Primer AATCTCAGCCGTCTATCCATGT

repeats (ag) x18 PCR product= 293 bp & start at base 456344

2)

For Primer GAGCTCAATGAATTGCTTGTTG

Rev Primer AATGGACTAGCCGTAGCTCTTG

repeats (ct) x17 PCR product= 228 bp & start at base 615700

3)

For Primer AATAGTGTCAGTGGAGAAGCCC

Rev Primer TCATGATGTTGGACTTTTGACC

repeats (ag) x16 PCR product= 272 bp & start at base 533901

**16. MADS-RIN found in Scaffold 000225 (query sequence starts at base 100878 on scaffold)**

1)

For Primer TGGGGCTATGATACAGGTCTCT

Rev Primer AGTGCCCAGGAAAGAATATCTG

repeats (tc) x26 PCR product= 250 bp & start at base 235971

2)

For Primer GGTGTGAAGCCAAGAGAGAGAG

Rev Primer AAACAGCAGGTGTCTGTGCTAA

repeats (ct) x19 PCR product= 227 bp & start at base 76235

3)

For Primer AGTAGGAAAGGATCAACCGTCA

Rev Primer CGAGCAAACTGGTCATTTGTAA

repeats (ag) x12 PCR product= 163 bp & start at base 124829

**17. ACS2 found in Scaffold 00661 (query sequence starts at base 7341 on scaffold)**

1)

For Primer GGGAAGATTTACTCATTGCTGG

Rev Primer AAGGATACTGATGCCCAATCAT

repeats (gt) x11 PCR product= 263 bp & start at base 65085

2)

For Primer ACGTTCCGGTACATGTATTCAA

Rev Primer CCTGTCAAAAATCAAAGGGAAC

repeats (ga) x11 PCR product= 260 bp & start at base 92206

3)

For Primer GTTGAAGTTTCAAAGGAATGGC

Rev Primer GGCGCAGGTTTTACTGAATTAC

repeats (ag) x10 PCR product= 142 bp & start at base 62088

**18. ACO1 found in Scaffold 02113 (query sequence starts at base 38007 on scaffold)**

1)

For Primer TCAAACACATTGGATAGTTCGG

Rev Primer CCTTCAAAAATCGTAAGCCACT

repeats (ta) x8 PCR product= 228 bp & start at base 40187

2)

For Primer AGTTAAAGCCCACTTTTCTCCC

Rev Primer AGGTGTTTGGTGGTTTTTATGG

repeats (ct) x5 PCR product= 216 bp & start at base 18179

3)

For Primer CCAAAGGATCAAGAAACTGGAC

Rev Primer TAGGAATCTGCATAGGGGAAGA

repeats (tca) x4 PCR product= 283 bp & start at base 32445

**19. ETR1 found in Scaffold 00228 (query sequence starts at base 111269 on scaffold)**

1)

For Primer AGATAAAGTCAAGGCTTTTGCG

Rev Primer AAGGTTTTCTGGGGGTAGGTAG

repeats (ga) x17 PCR product= 155 bp & start at base 294830

2)

For Primer TGGAAAGAGGAAACCCTAAACA

Rev Primer CATAAAGCAAGATGGAGTGCAA

repeats (ga) x14 PCR product= 135 bp & start at base 233889

3)

For Primer ACTAGACTGGGACTTTCCTCCC

Rev Primer GGAGTTTGGGAGAGGGTAGAGT

repeats (ct) x13 PCR product= 242 bp & start at base 171814

**20. ETR6 found in Scaffold 01249 (query sequence starts at base 40815 on scaffold)**

1)

For Primer TAGTACTTGATGCGCTCTACGC

Rev Primer GTTAGGGAGACTCAAGGGAGGT

repeats (ca) x18 PCR product= 275 bp & start at base 45590

2)

For Primer CCCTAACCCTACCTTTCGTTCT

Rev Primer TGTTTTTGGTCTCTCCCTCTGT

repeats (ct) x14 PCR product= 193 bp & start at base 39265

3)

For Primer ATAGCCCTAGGTCCACATGAAA

Rev Primer ACATCTAACGGTTTGGATCGTC

repeats (tg) x11 PCR product= 176 bp & start at base 9339

**21. TDR4 found in Scaffold 00235 (query sequence starts at base 196799 on scaffold)**

1)

For Primer TGGGAAAATCTGAGAGAGAGAGA

Rev Primer TTGGGGGAATCTAGTACCCTTT

repeats (ag) x27 PCR product= 180 bp & start at base 188891

2)

For Primer GAATTTGAGAGGCTTGTCTTGG

Rev Primer ACTCGTAGCAACCCATCTTTGT

repeats (ag) x13 PCR product= 212 bp & start at base 162898

3)

For Primer TGTACTGTGGGGAGATGACTTG

Rev Primer TGTTCGAAAGAACACAATGACC

repeats (tc) x12 PCR product= 292 bp & start at base 166162

**22. AT2G23900 found in Scaffold 00418 (query sequence starts at base 67136 on scaffold)**

1)

For Primer GAGCTTTGAAGTTGCAGTTCCT

Rev Primer AAGTTTAACCGTTCGTAAGACGAC

repeats (ga) x22 PCR product= 142 bp & start at base 50090

2)

For Primer GCACAAACTCCTATCTCCAACC

Rev Primer CTGCGATGAATTCTGTGAAAAG

repeats (ag) x15 PCR product= 205 bp & start at base 102033

3)

For Primer TACAAATATGGGGTTCCACCTT

Rev Primer CGACGATTTCAGAGTGACTGAG

repeats (ct) x13 PCR product= 153 bp & start at base 92278

**23. DEFH28 found in Scaffold 00249 (query sequence starts at base 161333 on scaffold)**

1)

For Primer ATTGAACAACAGTTGGAGCAGA

Rev Primer TCAGTGTCGTGTCTATGTCGTG

repeats (ga) x15 PCR product= 258 bp & start at base 144337

2)

For Primer CATGAAAACCCTAGATGCCTTC

Rev Primer TCTCTCAATCACTAGGACCCGT

repeats (ag) x15 PCR product= 241 bp & start at base 202052

3)

For Primer TTACTTCTTTAGCCAAACCCCA

Rev Primer GGCATTGGATGGAAAACTACTC

repeats (ac) x13 PCR product= 283 bp & start at base 185457

**References**

Cosgrove, D.J. (2000). Loosening of Plant Cell Walls by Expansins. Nature 407: 321-

326.

Dotto, M.C., Martínez, G.A., Civello, P.M. (2006). Expression of Expansin Genes in

Strawberry Varieties with Contrasting Fruit Firmness. Plant Physiology and Biochemistry 44: 301-307.

Fonseca, S., Monteiro, L., Barreiro, M.G., Pais, M.S. (2005). Expression of Genes

Encoding Cell Wall Modifying Enzymes is Induced by Cold Storage and Reflects Changes in Pear Fruit Texture

Jiménez-Bermúdez, S., Redondo-Nevado, J., Muñoz-Blanco, Caballero, J.L, López-

Aranda, J.M., Valpuesta, V., Pliego-Alfaro, F., Quesada, M.A., Mercado, J.A. (2002). Manipulation of Strawberry Fruit Softening by Antisense Expression of a Pectate Lyase Gene. Plant Physiology 128: 751-759.

Marín-Rodríguez, M.C., Orchard, J., Seymour, G.B. (2002). Pectate Lyases, Cell Wall

Degradation and Fruit Softening. Journal of Experimental Botany 53: 2115-2119.

Martel, C., Vrebalov, J., Tafelmeyer, P., Giovannoni, J.J. (2011). The Tomato MADS-

Box Transcription Factor RIPENING INHIBITOR Interacts with Promoters Involved in Numerous Ripening Processes in a COLORLESS NONRIPENING-Dependent Manner. American Society of Plant Biologists 157: 1568-1579.

Proctor, A. and Peng, L.C. (1988). Pectin Transitions During Blueberry Fruit

Development and Ripening. Journal of Food Science 54: 384-385.

Seymour, G.B., Manning, K., Eriksson, E.M., Popovich, A.H., King, G.J. (2002). Genetic

Identification and Genomic Organization of Factors Affecting Fruit Texture. Journal of Experimental Botany 53: 2065-2071.

Vicente, A.R., Ortugno, C., Rosli, H., Powell, A.L.T., Greve, L.C., Labavitch, J.M.

(2007). Temporal Sequence of Cell Wall Disassembly Events in Developing Fruits. 2. Analysis of Blueberry (*Vaccinium* Species). Journal of Agriculture and Food Chemistry 55: 4125-4130.

White, Philip J. (2002). Recent Advances in Fruit Development and Ripening: An

Overview. Journal of Experimental Botany 53: 1995-2000.

Zifkin, M., Jin, A., Ozga, J.A., Zaharia, L.I., Schernthaner, J.P., Gesell, A., Abrams, S.R.,

Kennedy, J.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: 200-224.

Genome Database for *Vaccinium*. 11 March 2013. Web. <http://www.vaccinium.org/>

Genome Annotation and Curation (GenSAS). 13 March 2013. Web.

<http://gensas.bioinfo.wsu.edu/home>

National Center for Biotechnology and Information (NCBI). 5 Feb. 2013. Web.

 <http://www.ncbi.nlm.nih.gov/>

Kieber, J., Hansen, M., Yoon, G.M. National Science Foundation (NSF). 2 May 2013.

Web. <http://labs.bio.unc.edu/Kieber/Ethylene%20page.htm>

Sarkar, A. and Perez, S. PolySac3DB. 2 May 2013. Web.

 <http://polysac3db.cermav.cnrs.fr/discover\_cellulose.html>

**Acknowledgements**

 I would like to acknowledge all of the individuals who helped me through this research project. First, I would like to thank Dr. Malcolm Campbell who led the class and assisted me in many ways throughout this investigation. I would also like to thank my classmates who, through collaborative efforts, have helped me in a variety of ways. I would also like to thank Dr. Allen Brown of North Carolina State University, Dr. Doreen Main of Washington State University and Dr. Jeannie Rowland of the USDA for their contributions to the project. Finally, I would like to thank the Davidson College Biology Department.