Erich Baker

Dr. Malcolm Campbell

BIO 343

9 May 2013

Identification of Drought Stress Genes in *Vaccinium corymbosum* and Selection of PCR Primers for Amplification of Nearby Simple Sequence Repeats

**Abstract**

The blueberry, *Vaccinium corymbosum*, is an important agricultural product in many areas of the United States, and like most crops it is susceptible to moderate and severe drought. In an effort to help researchers selectively breed for drought-resistant blueberries, I performed a literature search to identify genes in model organisms that may play a role in drought-resistance pathways. I identified orthologs of known drought-resistance genes. I utilized a web-based program to locate simple sequence repeats (SSRs) near the genes of interest, and I selected PCR primers that could be used to amplify these SSRs. The SSRs can be used to help researchers genotype individual plants and identify alleles, promoting selective breeding efforts.

**Introduction**

*Vaccinium corymbosum*, the blueberry, is the second-most harvested berry crop in the United States, with over 560 million pounds of blueberries harvested annually and worth over $850 million in 2012 (Blueberries, 2012). In North Carolina alone, over 28 million pounds are harvested, with a value of approximately (U.S., 2012). Blueberries contain many purportedly-healthful compounds, including phytonutrients, vitamins C and K, manganese, fiber, and more (Health, 2012). Compounds such as these may play significant roles in cardiovascular and brain health.

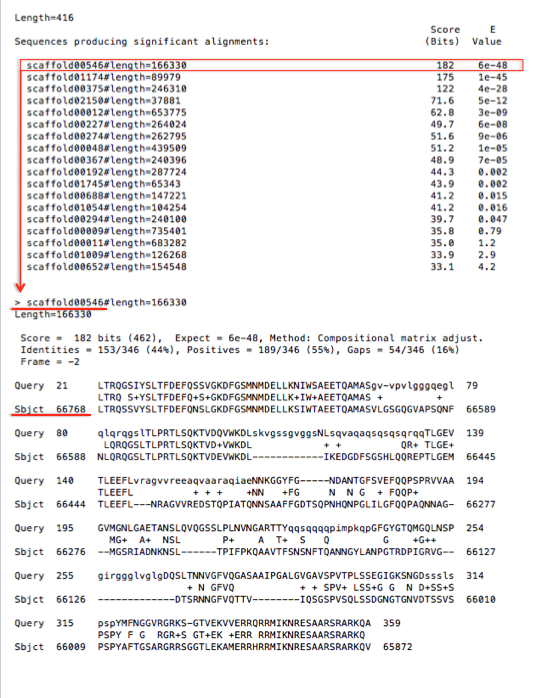
Because of the nutritive and commercial value of blueberries, there is research interest in the biology of blueberries. We have partnered with researchers at North Carolina State University and the North Carolina Research Center in order to study the blueberry genome, with the goal of helping breeders produce better blueberries. I examined genetic pathways involved in resistance to drought stress. Soil water content of less than 50% can lead to blueberry plants producing fewer and/or smaller berries (Longstroth, 2012). The generation of a drought resistant crop could improve annual blueberry harvests, increasing profits for growers and making blueberries more available and less expensive to the public. I first studied universal drought resistance mechanisms in the model plant *Arabidopsis thaliana*, and then I searched the blueberry genome for orthologs to *Arabidopsis* drought resistance genes.

Once I identified an ortholog of a drought resistance gene in the blueberry genome, I searched the genome for nearby simple sequence repeats, or SSRs. SSRs are frequently mutated by DNA polymerase, which has a tendency to make errors when synthesizing repeat units of DNA. These SSRs serve as molecular markers for the nearby gene. I produced PCR primers that could be used to clone and amplify these SSRs, so that blueberry growers and researchers could identify blueberry plants that contain an allele associated with a particular SSR. Hopefully, cloning these SSRs will give blueberry growers the ability to identify certain alleles of drought-resistance genes and ultimately selectively breed for a more drought-resistant crop.

As a part of this project, I also tested the recently developed Genome Sequence Annotation Server, or GenSAS. The goal of this program was to automatically annotate features of the blueberry genome using different computer programs. I tested a program that has been integrated into GenSAS called BLAT. BLAT is similar to BLAST, and it uses an index derived from assemblies of entire genomes to predict genes. The GenSAS program has been recently developed, and as one of its first users, I attempted to see if it could be used to produce viable annotations.  
  **Methods**

I focused on the drought stress genes that I found to be most well-characterized in the literature, which were abscisic acid- (ABA) dependent genes in the model plant *Arabidopsis thaliana*. I took genes that I had identified as involved in drought-resistance in *Arabidopsis* and searched for them in NCBI’s protein database (REF). When I found a protein on NCBI, I saved its amino-acid FASTA sequence to a plain-text ile so that I could BLAST it against the blueberry genome to search for alignments.

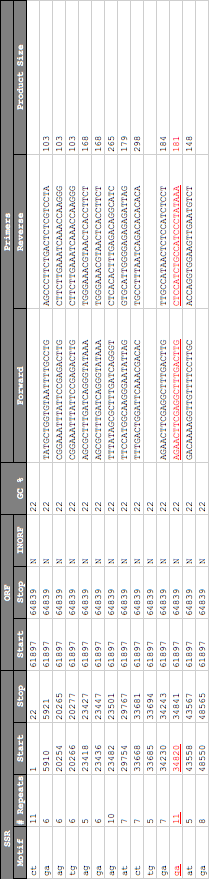
In order to BLAST amino acid sequences of genes of interest against the blueberry genome, I first set up a local BLAST engine on my computer. I did this using a Mac OS X machine. I opened up the Terminal application and entered the following command – “cd Desktop/ncbi-blast-2.2.27+” - where ncbi-blast-2.2.27 was the name of the folder that contained the blueberry genome. The folder was located on the desktop, as indicated by the command. I entered the command – “./bin/makeblastdb -in Sorted\_454Scaffolds.txt -input\_type fasta -dbtype nucl -title blueberry\_Genome” – where Sorted\_454Scaffolds.txt was the text file containing the nucleotide sequences that composed the blueberry genome. This completed the setup of local BLAST.

In order to BLAST a particular amino acid sequence against the blueberry genome, I entered the following command – “bin/tblastn -query XXXXXXXX.txt -db Sorted\_454Scaffolds.txt.” XXXXXXXX.txt should be replaced by the name of the .txt file that contains your query amino acid sequence. The results of an example BLAST for are shown in Figure 1. Higher bit scores and lower E-values correspond to better alignments.

**Figure 1.** Example results of a tBLASTn search. Circled in red is the best alignment the computer identified, its raw bit score, and E-value. Underlined below are details about that alignment, including the scaffold that it is on and the number of the nucleotide on that scaffold where the alignment began.

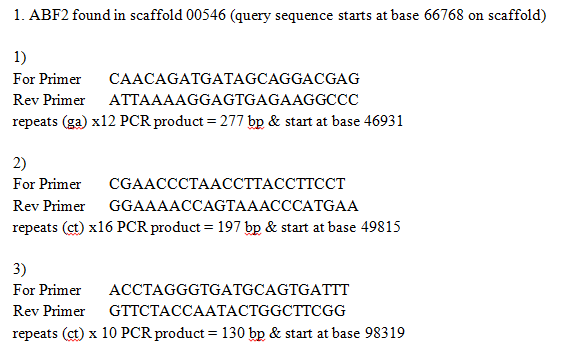
I performed tBLASTn searches, as indicated in the search command, to search the genome (a nucleotide sequence) using an amino acid sequence query. Amino acid queries produced better alignments than nucleotide queries. If the search produced any significant (E-value < 1x10-3) alignment between the query sequence and the genome, I identified the scaffold which contained the lowest E-value hit, and noted the base pair position of the alignment on that scaffold. If I had already identified a gene of interest on the scaffold with the lowest E-value alignment, I selected the alignment of the nest-lowest E-value.

I searched the blueberry genome for the scaffold I had identified. I did this by opening up the file containing the blueberry genome, and using the Cmd+F “find” feature to find the right scaffold within the genome. I copied and pasted the entire scaffold into a separate document, and saved it as a plain-text file. I submitted that file to the SSR identification tool (REF). This tool can be accessed by mousing over the “Tools” header, and selecting “SSR.” Into the SSR server, I entered my email address and attached the .txt file containing the scaffold with the alignment to the gene of interest. When submitting the file, I entered “0” next to tetra- and pentanucleotide repeats so that they would be ignored. SSRs with repeat units ≥4 bases do not produce as much variation as the shorter repeat units (REF). The SSR server returned excel documents to my email that included the position of all di- and trinucleotide SSRs along the query scaffold. The excel file also generated PCR primers that could be used to amplify those SSRs. A sample excel file with relevant data highlighted is included in Figure 2.



**Figure 2.** Example of the Excel file returned by vaccinium.org SSR tool. Important information is highlighted in red, including the repeat motif, the number of times that motif repeats, the number of the scaffold nucleotide where the repeat motif begins, sequences for PCR primers that can amplify the repeats, and the size of the PCR product that would be produced by those primers.

I selected SSRs and their accompanying PCR primers based on the number of repeats a SSR contained and its distance on the scaffold from the gene of interest. Higher-repeat SSRs and close proximity to the gene of interest are more desirable. I generally selected PCR primers for the highest-repeat SSR on the scaffold, for the SSR nearest to the gene of interest, and for an SSR that struck a balance between number of repeats and distance from the gene of interest. I recorded the PCR primers for each SSR, noting the repeat unit, the number of repeats, the length of the PCR product, and the base pair on the scaffold where the primer began. I also recorded the name of the gene of interest and the number of the scaffold the gene was found on. A sample of my finished report on a gene is shown in figure 3.



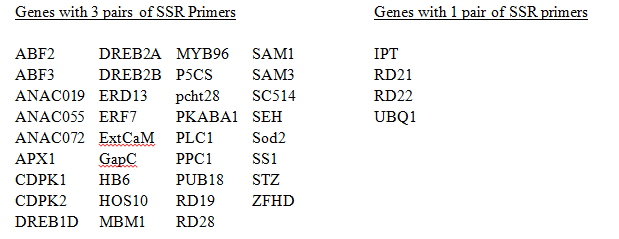
**Figure 3**. Example report for gene of interest ABF2. Includes PCR primers for three different SSRs. Repeat unit, number of repeats, size of PCR product, and location along on the scaffold are included for each.

I also submitted several of my scaffolds to the GenSAS automatic annotation website (REF). I uploaded the scaffold a particular gene of interest was on onto GenSAS, and then selected the BLAT tool as the program I wanted to use to analyze that sequence. I attempted to run the BLAT tool on multiple scaffolds.

**Results**

I generated useful SSR primers for 39 genes. I generated three pairs of primers for 35 of the 39 genes, but 4 of the genes presented only one nearby SSR that was suitable for amplification (Appendix A). The genes I generated primers for are listed in table 1. Some genes, which I have not included in this paper, did [NOT?] produce any significant alignments to the blueberry genome.

**Table 1**. Genes for which SSR primers were generated.



In addition, I found literature regarding the potential of select genes to generate drought resistant phenotypes. I found literature examining the biology of these “high-importance” genes, which included MYB96, IPT, and CKX. I identified these genes to examine in detail because the literature showed that they confer significant and well-characterized drought resistance.

The GenSAS BLAT tool did not produce any results. After I submitted my scaffolds for annotation, I received an email informing me that there had been an error and that the annotation could not be completed. I tried this with several different scaffolds, but the GenSAS BLAT tool did not work properly on any of them.

**Discussion**

The PCR primers I have generated could be used to amplify an SSR, which could assist blueberry growers in identifying plants of a particular allele of the gene in question. Particular alleles of a drought-resistance gene are likely associated with a certain form of the SSR because those regions are so prone to mutation. By using the PCR primers I generated to amplify and sequence an SSR near a gene of interest, breeders will be able to determine what allele an individual plant has. Accurate genotyping of drought-resistance genes in blueberries could accelerate the selective breeding efforts of growers by allowing them the ability to breed only homozygous individuals. This would allow growers to produce cultivars that are true-breeding for a drought resistance gene. Enhanced selective breeding could lead to wide-spread growing of drought resistance blueberries and thus increased crop yields.

To further the usefulness of amplifying SSRs for a gene, it would be ideal if the blueberry genome could be more thoroughly sequenced. If there were fewer scaffolds and fewer ambiguous nucleotides in the genome, better SSRs could be identified. In addition, it would be beneficial if more drought-resistance genes could be identified in the blueberry genome. A careful review of the current literature on drought-resistance pathways could help identify key drought-resistance genes that are of greatest importance to identify in blueberry. I focused only on abscisic acid- (ABA) mediated pathways, and so there are likely many ABA-independent pathways that are as yet un-annotated in the blueberry genome.

I also examined the phenotypic impacts of several drought resistance genes that I identified in the blueberry genome. These genes - MYB96, and IPT/CKXs - were shown to confer significant drought resistance. Increased expression of IPT and CKX leads to higher levels of intracellular cytokinins (Nishiyama *et al*., 2011). Higher cytokinin levels confer significantly increased drought resistance, yet also stunt the growth of plants (Nishiyama *et al*., 2011). Plants that express high levels of IPT and CKX would thus be desirable under drought conditions when survival of the plant is more important than plant size. Under non-drought conditions, however, such plants would likely be smaller and undesirable. Perhaps there is potential for breeders to selectively breed for a number of cultivars with different expression levels of these genes, and they can then select for the one that strike the optimal balance between size and drought resistance based on the weather predictions for a growing season.

Increased expression of MYB96 leads to increased cuticular wax deposition and reduced water loss via transpiration (Seo *et al*., 2011). Unlike IPT/CKXs, increased MYB96 expression does not seem to stunt the growth of plants. Therefore, if a MYB96 allele that produces increased expression could be identified, it could be a good candidate gene to selectively breed for. Increased wax deposition could potentially alter the taste and texture of blueberries, however, which might make them undesirable.

The BLAT tool I ran through GenSAS did not produce any results, despite several trials on different scaffolds. If GenSAS were fixed so that it produced automatic annotations, it could potentially be a useful tool for analyzing genomes.

**Acknowledgements**

I would like to thank Dr. Allan Brown of North Carolina State University, Dr. Jeannie Rowland of the U.S. Department of Agriculture, and Dr. Doreen Main of Washington State University for helping me gather this data. I would also like to thank Dr. Malcolm Campbell for his guidance and support throughout this project.

**References**

"Blueberries." *Agricultural Marketing Resource Center*. Agricultural Marketing Resource Center, Jan. 2012. Web. 01 May 2013. <http://www.agmrc.org/commodities\_\_products/fruits/blueberries/>.

"Health Benefits of Blueberries." *Blueberry Council*. U.S. Blueberry Council, n.d. Web. 01 May 2013. <http://www.blueberrycouncil.org/health-benefits-of-blueberries/blueberry-nutrition/>.

Longstroth, Mark. "Drought Stress Symptoms in Blueberries." *MSU Blueberries*. Michigan State University, July 2012. Web. 1 May 2013.

Nishiyama, R., Watanabe, Y., Fujita, Y., Le, D.T., Kojima, M., Werner, T., Vankova, R., Yamaguchi-Shinozaki, K., Shinozaki, K., Kakimoto, T.*, et al.* (2011). Analysis of Cytokinin Mutants and Regulation of Cytokinin Metabolic Genes Reveals Important Regulatory Roles of Cytokinins in Drought, Salt and Abscisic Acid Responses, and Abscisic Acid Biosynthesis. The Plant Cell Online *23,*2169-2183.

Seo, P.J., Lee, S.B., Suh, M.C., Park, M., Go, Y.S., and Park, C. (2011). The MYB96 Transcription Factor Regulates Cuticular Wax Biosynthesis under Drought Conditions in Arabidopsis. The Plant Cell Online *23,*1138-1152.

"U.S. Blueberry Industry." *United States Department of Agriculture Economics, Statistics, and Market Information System*. USDA, May 2012. Web. 01 May 2013. <http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do? documentID=1765>.

Yoshida, T., Fujita, Y., Sayama, H., Kidokoro, S., Maruyama, K., Mizoi, J., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2010). AREB1, AREB2, and ABF3 are master transcription factors that cooperatively regulate ABRE-dependent ABA signaling involved in drought stress tolerance and require ABA for full activation. Plant Journal *61,*672-685

1. ABF2 found in scaffold 00546 (query sequence starts at base 66768 on scaffold)

1)

For Primer CAACAGATGATAGCAGGACGAG

Rev Primer ATTAAAAGGAGTGAGAAGGCCC  
repeats (ga) x12 PCR product = 277 bp & start at base 46931  
  
2)  
For Primer CGAACCCTAACCTTACCTTCCT  
Rev Primer GGAAAACCAGTAAACCCATGAA  
repeats (ct) x16 PCR product = 197 bp & start at base 49815  
  
3)  
For Primer ACCTAGGGTGATGCAGTGATTT  
Rev Primer GTTCTACCAATACTGGCTTCGG  
repeats (ct) x 10 PCR product = 130 bp & start at base 98319

2. ABF3 found in scaffold 00375 (query sequence starts at base 185340 on scaffold)

1)

For Primer AAATGGGACCTAATGTGGTACG

Rev Primer TCACTCCATAAATCATTGCAGC  
repeats (ag) x23 PCR product = 285 bp & start at base 172031  
  
2)  
For Primer AGCTGGCCTAGGCTTGACTC  
Rev Primer CTTAACTCGGCTCGATAAGGCT  
repeats (ca) x13 PCR product = 288 bp & start at base 183449  
  
3)  
For Primer CACAGAGAGAGAGAGAGAGAGAGAG  
Rev Primer TGCCAAGAGTAGTCATAAGCCA  
repeats (ga) x 18 PCR product = 192 bp & start at base 192091

3. ABF4 found in scaffold 00054 (query sequence starts at base 66789 on scaffold)

1)

For Primer CGAACCCTAACCTTACCTTCCT

Rev Primer GGAAAACCAGTAAACCCATGAA  
repeats (ct) x16 PCR product = 197 bp & start at base 49815  
  
2)  
For Primer ACCTAGGGTGATGCAGTGATTT  
Rev Primer GTTCTACCAATACTGGCTTCGG  
repeats (ct) x16 PCR product = 130 bp & start at base 98319  
  
3)  
For Primer TAGATTGGTGAGATCAGGGAGG  
Rev Primer TGTCATCTGCAACTCTATTCGC  
repeats (ct) x 10 PCR product = 156 bp & start at base 112538

4. ANAC019 found in scaffold 00116 (query sequence starts at base 390369 on scaffold)

1)

For Primer TAATTTGTCCTAAATGCTCCCG

Rev Primer TCCACATTGTACTGAAACTGCC  
repeats (ct) x10 PCR product = 247 bp & start at base 327537  
  
2)  
For Primer CAAGAGAAGAGGAGAGCCAAAA  
Rev Primer TCTCGTTCATGGTTATTGATGC  
repeats (tc) x11 PCR product = 247 bp & start at base 338965  
  
3)  
For Primer CAAGAGAAGAGGAGAGCCAAAA  
Rev Primer TCTCGTTCATGGTTATTGATGC  
repeats (tc) x 17 PCR product = 247 bp & start at base 338783

5. ANAC055 found in scaffold 01837 (query sequence starts at base 43877 on scaffold)

1)

For Primer TTTATAGCGCTTTGATCAGGGT  
Rev Primer CTCACACTTTGAGACAGGCATC  
repeats (ag) x10 PCR product = 265 bp and start at base 23482

2)  
For Primer AGAACTTCGAGGCTTTGACTTG  
Rev Primer CTCCATCTGCCATCCCTATAAA  
repeats (ga) x11 PCR product = 181 bp and start at base 34820

3)

|  |
| --- |
| For Primer GACAAAAGGTTGTTTTCCTTGC Rev Primer ACCAGGTGGAAGTTGAATGTCT repeats (at) x5 PCR product = 148 bp and start at base 43558  6. ANAC072 found in scaffold 00651 (query sequence starts at base 120766 on scaffold)  1)  For Primer GTACCGGTGAGTGTGTTGAAGA Rev Primer AATAGTGGATTCTGGGGTGAGA repeats (ag) x14 PCR product = 231 bp and start at base 120447  2) For Primer AGAATAGCATGTCCCGACTCTC Rev Primer CAACTAATAGCTCGCTCCCTTT repeats (ct) x14 PCR product = 300 bp and start at base 133843  3) For Primer CGACACAGGTCCTCTCTCTCTC Rev Primer ACCTGCAACTAATAGCTCGCTC repeats (ct) x24 PCR product = 277 bp and start at base 133889  7. Apx1 found in scaffold 00868 (query sequence starts at base 63665 on scaffold)  1)  For Primer TTTCCTTTTCATCGTCAGAGGT Rev Primer ACCTACCATGTGTGAGAATCCC repeats (tc) x14 PCR product = 230 bp and start at base 15575  2) For Primer CAAGCTGGATGAGACTGTTGAG Rev Primer CTTTTTCCCTTCCATGTTTCAC repeats (ac) x7 PCR product = 294 bp and start at base 40973  3) For Primer GCTGTAGAAAAAGCCAAGAGGA Rev Primer TTGTACTTCATTGTCCCGAATG repeats (tc) x6 PCR product = 299 bp and start at base 63790  8. CDPK1 found in scaffold 00154 (query sequence starts at base 315099 on scaffold)  1)  For Primer TCCTCCACAACTACGATGTCTG Rev Primer CAGCGAAAATACCCTATTCCAC repeats (ag) x25 PCR product = 246 bp and start at base 48404  2) For Primer TGTATCTTTCGAATCGCGTAGA Rev Primer CAGGGAAAGGAACATTTTCAAG repeats (att) x17 PCR product = 219 bp and start at base 194378  3) For Primer GTCTTTCTTTGACAACGGAACC Rev Primer CCACGTATATAGTTGTTCGGCA repeats (ct) x9 PCR product = 283 bp and start at base 305851  9. CDPK2 found in scaffold 01799 (query sequence starts at base 51713 on scaffold)  1)  For Primer ACGTTTAATTACGTACGTCCCA Rev Primer ACCGATAAGGACACGAGATTGT repeats (ct) x23 PCR product = 231 bp and start at base 23115  2) For Primer ATGTTCTTTCAGAATCCGTCGT Rev Primer TGCTGAACTATCTTTCAGGGGT repeats (tc) x10 PCR product = 272 bp and start at base 51911  3) For Primer GTTTAAACCGTTACCTTTCCCC Rev Primer GCTAATTATCCTCTCGTGGTGG repeats (ct) x18 PCR product = 282 bp and start at base 52048  10. CKX found in scaffold 00024 (query sequence starts at base 447697 on scaffold)  1)  For Primer GCCTTTGAAGAACAAAACTGCT Rev Primer CGTAACGTAACGAAACGAAGAA repeats (ct) x16 PCR product = 215 bp and start at base 324064  2) For Primer TTTTGTTTGACAGCTCGAAGAC Rev Primer ACGTAAACCGTAACCTTCGTTG repeats (ga) x16 PCR product = 213 bp and start at base 473509  3) For Primer GTAAAGAGAAAGGGCCAAATCC Rev Primer ACCTACCACTACTCCCTCCCTC repeats (ga) x18 PCR product = 140 bp and start at base 480760  11. DREB1D found in scaffold 00009 (query sequence starts at base 488219 on scaffold)  1)  For Primer CACACAAGACGTGTATGCAGAA  Rev Primer CCCTCATTATGAAATGGCTATG repeats (ac) x24 PCR product = 298 bp & start at base 625376  2) For Primer GTCCTACTCGTCGTCGTACCTC Rev Primer GGAAAACCCAGACAAGAATCAG repeats (tc) x20 PCR product = 170 bp & start at base 309958  3) For Primer TAGAGAATGCGATGGGAAATCT Rev Primer GAACGGTAAATCTGACTCCGTC repeats (ct) x 12 PCR product = 168 bp & start at base 681889  12. DREB2A found in scaffold 00313 (query sequence starts at base 149045 on scaffold)  1)  For Primer CACACACAATCTAACAACGCAA  Rev Primer TGTAATTCAATGTGGGACCGTA repeats (ga) x25 PCR product = 223 bp & start at base 94430  2) For Primer CAGAAGCCATGTCACACAAGTT Rev Primer AAAAGAGCAAAGATGGAAGCAG repeats (ga) x14 PCR product = 199 bp & start at base 173268  3) For Primer ATAATGAAGGTGCAATTCGGTC Rev Primer CCCCTAACACACACTAACCACA repeats (ca) x 11 PCR product = 294 bp & start at base 177850  13. DREB2B found in scaffold 00598 (query sequence starts at base 111169 on scaffold)  1)  For Primer ATGCACCCTACCAAAGCTAAAA  Rev Primer GAAAGCCTGAAAAAGCCAGATT repeats (ct) x10 PCR product = 218 bp & start at base 96462  2) For Primer TAAGGGAGATGGGAGCTTATGA Rev Primer TTCCATTAGGGTTTTGACTGCT repeats (aag) x14 PCR product = 269 bp & start at base 105721  3) For Primer GCCTAGGCAGACGACTTTTAGA Rev Primer GAAACGATCACGTTAAAGGAGG repeats (ct) x 7 PCR product = 294 bp & start at base 132523  14. ERD13 found in scaffold 00140 (query sequence starts at base 106588 on scaffold)  1)  For Primer AGAGAGAGAGGAGGAGGAGGAG Rev Primer TAGAGCTTGATCTTCACCGACC repeats (agt) x24 PCR product = 297 bp and start at base 99061  2) For Primer AGAGAGAGAGGAGGAGGAGGAG Rev Primer TAGAGCTTGATCTTCACCGACC repeats (ag) x25 PCR product = 297 bp and start at base 99133  3) For Primer ATTCCCACGGCAAATTATACTG Rev Primer TCAAGTCCTTCGATTTGTTCCT repeats (tc) x12 PCR product = 231 bp and start at base 101618  15. ERF7 found in scaffold 00202 (query sequence starts at base 30408 on scaffold)  1)  For Primer TCTCTCTAAAACCGAGCCAGTC Rev Primer GAAACCACCACCCAGTCTACTC repeats (ct) x7 PCR product = 218 bp and start at base 30686  2) For Primer CTTTTCCCCCATCTACTCCTTT Rev Primer ACACGTCTTTGCCTGGACTTAT repeats (ga) x13 PCR product = 241 bp and start at base 119738  3) For Primer CTAACCCACGAAAGCTATGTCC Rev Primer AAGTGCACCCTAAGTGCCTAAA repeats (tc) x15 PCR product = 237 bp and start at base 130390  16. ExtCaM found in scaffold 00080 (query sequence starts at base 264255 on scaffold)  1)  For Primer GATGGTGGAGAGTGAGATCCAT Rev Primer CTTTTTGGAAGGCTCAAGCTAA repeats (ag) x5 PCR product = 246 bp and start at base 257045  2) For Primer CTACCTGCGTATTGCTTGTTCA Rev Primer TCTGTGTGTGTGTGTGTGTGTG repeats (tc) x13 PCR product = 254 bp and start at base 279791  3) For Primer GTTGCTTTGTGTGGGAAATGAT Rev Primer TTGTGTAGCTTTGCAAATAGCC repeats (ac) x14 PCR product = 216 bp and start at base 279817  17. GapC found in scaffold 00013 (query sequence starts at base 320988 on scaffold)  1)  For Primer AGATTGCTCGAGATAGCTTTGG Rev Primer AACGAAGGAACGACCTAACGTA repeats (ag) x13 PCR product = 182 bp and start at base 273468  2) For Primer TTGTTATTGGCAAGTGAGATGG Rev Primer AGATGGGAAAATCTAGCCACAA repeats (tc) x11 PCR product = 258 bp and start at base 323097  3) For Primer TTTGATGATTTCTTGTGGGGN Rev Primer GAGAGGGAGCATTTTTGTGAAG repeats (ag) x11 PCR product = 274 bp and start at base 333585  18. HB6 found in scaffold 00240 (query sequence starts at base 82813 on scaffold)  1)  For Primer AGTATTGCCGCCTACAATAGGA Rev Primer ATTGAGTTTACAAGGCTGAGGC repeats (ag) x7 PCR product = 179 bp and start at base 65710  2) For Primer CCTTTCTGGGTTTTCTGTCTTG Rev Primer TCCTTCTTCCTCTAGCCCTTCT repeats (ct) x10 PCR product = 141 bp and start at base 82777  3) For Primer CACGCACATCCTACAGCTAAGA Rev Primer TCAAGACCACTTTTCCTTCGAT repeats (ct) x8 PCR product = 242 bp and start at base 105040  19. HOS10 found in scaffold 00067 (query sequence starts at base 237891 on scaffold)  1)  For Primer ATTCGCGTATCTACACCCAAAC Rev Primer ATTTGCCTCCCTTATTTTCCTC repeats (ag) x19 PCR product = 221 bp and start at base 208660  2) For Primer AGAAGATAGCGAATGGTGTTTG Rev Primer TACAATTAAAAAGGGACGGAGG repeats (at) x7 PCR product = 265 bp and start at base 246621  3) For Primer GCTTTGTTACAATACCGCATCA Rev Primer AGGAGTTGTGCAACCCATACTT repeats (tc) x11 PCR product = 195 bp and start at base 275656  20. IPT found in scaffold 00817 (query sequence starts at base 2 on scaffold)  1)  For Primer AGCCAATAACCTTTCCTGTTGA Rev Primer GTGCCTATCAGCCACTCTTTTT repeats (tc) x9 PCR product = 238 bp and start at base 314842  21. MBM1 found in scaffold 00028 (query sequence starts at base 413851 on scaffold)  1)  For Primer GCTCTTATCTGGTTGTGGCTCT Rev Primer ATGAACATCTGTGAGCAGCAGT repeats (ag) x13 PCR product = 288 bp and start at base 410153  2) For Primer GCTCTTATCTGGTTGTGGCTCT Rev Primer ATGAACATCTGTGAGCAGCAGT repeats (ag) x13 PCR product = 288 bp and start at base 410199  3) For Primer CTCTACCGTAATGACGCTACCC Rev Primer TACTACCGTCCGTCGTCTTCTT repeats (ga) x17 PCR product = 235 bp and start at base 469943  22. MYB96 found in scaffold 00457 (query sequence starts at base 75842 on scaffold)  1)  For Primer TCTCTCTCTCACAAACACACGC  Rev Primer GAGCAACCCCAACAATAGGTTA repeats (ac) x16 PCR product = 223 bp & start at base 46519  2) For Primer TTGCGGTCATTTCTTCTTCTTC Rev Primer TGACATTGCATAGAAGCACACA repeats (tc) x15 PCR product = 174 bp & start at base 59021  3) For Primer CACCTACTCCCTTCGTTCAAAA Rev Primer CCCTAAAGTGCCTCAAGTCATTA repeats (ct) x 16 PCR product = 300 bp & start at base 79191  23. P5CS found in scaffold 00343 (query sequence starts at base 170426 on scaffold)  1)  For Primer TAGCTTGGTACCGGTTTTCCTA Rev Primer GAAATGAGGTCTCAACTGGGTC repeats (tg) x7 PCR product = 160 bp and start at base 169631  2) For Primer GCAACCACGGTCATATGATAAA Rev Primer TCCGAAAGTCTCCTAAAACCCT repeats (ag) x7 PCR product = 255 bp and start at base 177605  3) For Primer GGTTGGTGTAAGATTCGTCCAT Rev Primer TTCAACTCTCACCAGTCTCGAA repeats (ga) x10 PCR product = 108 bp and start at base 189380  24. Pcht28 found in scaffold 03298 (query sequence starts at base 114 on scaffold)  1)  For Primer TAGTCAAGGCGAAATTTGGAAC Rev Primer CTGAGTCAACTATTGCTGCCAA repeats (gaa) x5 PCR product = 217 bp and start at base 2918  2) For Primer TTTCTTATAGTGCACGGGGAGT Rev Primer CCTTACGCAGTTACACTTTGGC repeats (att) x4 PCR product = 177 bp and start at base 5313  3) For Primer ATTCTCATATCGGCCGTCTAAA Rev Primer CACGTCCATAAAAGCAGTCAAA repeats (ag) x8 PCR product = 225 bp and start at base 10002  25. PKABA1 found in scaffold 00012 (query sequence starts at base 630459 on scaffold)  1)  For Primer AGATGCGAAAGCTGTAAAAGGA Rev Primer AGTATCTCCGAGCCACTTTCTG repeats (ag) x15 PCR product = 291 bp and start at base 544222  2) For Primer TGGTGGTCTACTAGATGCGAAA Rev Primer CATGTTTATGGTCCGAGTGCTA repeats (ag) x15 PCR product = 194 bp and start at base 544372  3) For Primer ATAAAAACCCTGCCTTGACTGA Rev Primer TGATCAGAAGCATCTTCTCCAA repeats (ga) x18 PCR product = 233 bp and start at base 650831  26. PLC1 found in scaffold 00254 (query sequence starts at base 241728 on scaffold)  1)  For Primer NCTCTCCTACCTTACTACACACACA Rev Primer AAGGCACTCCGATGGTTAAGT repeats (tc) x14 PCR product = 110 bp and start at base 192135  2) For Primer TGACCATCCCTTCTGAATTTCT Rev Primer TTTCGACAATGCCAACAGATAC repeats (ct) x40 PCR product = 288 bp and start at base 231757  3) For Primer TACACCTCTTCTCTCTCCTGCC Rev Primer GACTGTTACAACTCCGCTTCCT repeats (ca) x12 PCR product = 216 bp and start at base 233794  27. PPC1 found in scaffold 00939 (query sequence starts at base 17909 on scaffold)  1)  For Primer CTCAACGATTGACCAAATAACG Rev Primer ATTCACACTCATGGCTTACCCT repeats (ag) x12 PCR product = 267 bp and start at base 4052  2) For Primer CATAGTTACCGATGTCCGATCA Rev Primer ACTCTGTACCTCTCCTTCTGCC repeats (aga) x10 PCR product = 225 bp and start at base 46809  3) For Primer GCTGTATCAGTCGTTGCTGTTC Rev Primer GATACCTTTGGCCACGTAAAAG repeats (tc) x10 PCR product = 259 bp and start at base 64570  28. PUB18 found in scaffold 00746 (query sequence starts at base 16200 on scaffold)  1)  For Primer CTCTAAACTTTCCGCCTCCTTT  Rev Primer GCTAATCAACCGAAACACCATT repeats (tc) x14 PCR product = 119 bp & start at base 8759  2) For Primer AAAGAGAGGAACGAAACAAACG Rev Primer TTTGTGCTATTAGGGTTTGGCT repeats (tg) x14 PCR product = 286 bp & start at base 26871  3) For Primer GTGTGTGTGTGTGTGTGTGTGA Rev Primer TTTGTGCTATTAGGGTTTGGCT repeats (ga) x 16 PCR product = 201 bp & start at base 26902  29. RD19 found in scaffold 00009 (query sequence starts at base 735401 on scaffold)  1)  For Primer CACACAAGACGTGTATGCAGAA Rev Primer CCCTCATTATGAAATGGCTATG repeats (ag) x24 PCR product = 298 bp and start at base 625376  2) For Primer TAGAGAATGCGATGGGAAATCT Rev Primer GAACGGTAAATCTGACTCCGTC repeats (ag) x12 PCR product = 168 bp and start at base 681889  3) For Primer GAAACGAACAAGGAATAATCGC Rev Primer TCTTCCTCCTTATTTAGCACGC repeats (ct) x10 PCR product = 256 bp and start at base 699162  30. RD21 found in scaffold 11065 (query sequence starts at base 489 on scaffold)  1)  For Primer TGAAATTGTGACAGCTTGTGTG Rev Primer CATCACAAAGCCTATTGATTGC repeats (at) x5 PCR product = 283 bp and start at base 1720  31. RD22 found in scaffold 10193 (query sequence starts at base 2435 on scaffold)  1)  For Primer TCCGTCAGTTCGTTTTATTGTG  Rev Primer ATAACATCTGTTCTTGCAGCGA repeats (ac) x5 PCR product = 243 bp & start at base 3003  32. RD28 found in scaffold 00872 (query sequence starts at base 72552 on scaffold)  1)  For Primer AGAGGACACGAGAAATTTGGAA Rev Primer GAGGATCACCACTAGATCGGAC repeats (tc) x15 PCR product = 300 bp and start at base 82434    2) For Primer AGTTTTTGTTTGGAGCGAAGAG Rev Primer ATTCGTACCCAAGTTTTTCTCG repeats (ag) x13 PCR product = 266 bp and start at base 91970  3) For Primer TCTCGTTAACGTTTGCTCAATG Rev Primer AAAGAGGGGAAAAGAAAGAAGG repeats (ct) x12 PCR product = 203 bp and start at base 93030  33. SAM1 found in scaffold 00742 (query sequence starts at base 28366 on scaffold)  1)  For Primer TGGGATCTATCGATCTAGGCTG Rev Primer TGAAAACGAACCTGTAGACGAA repeats (tc) x8 PCR product = 238 bp and start at base 26720  2) For Primer AAATATTGCCGGCTTGTGATAC Rev Primer TTTTAAGTTCAACTAGGGTGGC repeats (ct) x8 PCR product = 230 bp and start at base 42479  3) For Primer GCCTGGTAAAAGAAACTGCAAA Rev Primer GCGTAACATGTTTAGGTCGGTA repeats (ag) x25 PCR product = 299 bp and start at base 121397  34. SAM3 found in scaffold 00014 (query sequence starts at base 424166 on scaffold)  1)  For Primer TTTGAGGAGAAGTCCAAAGGAG Rev Primer GTTAAGTTTGGGCGACTGAGAG repeats (tc) x15 PCR product = 297 bp and start at base 427651  2) For Primer TTTTCAGTCTCGGACCCTTTTA Rev Primer TGCTAACCTCCACTCCAAAAAT repeats (taa) x22 PCR product = 259 bp and start at base 520280  3) For Primer TGGGGACTCTCTTAGATGCAAT Rev Primer CCAAATACTTGGCCACCTTAGA repeats (ca) x20 PCR product = 282 bp and start at base 590972  35. SC514found in scaffold 00065 (query sequence starts at base 272119 on scaffold)  1)  For Primer AAGTTATCTCTTCATGAGGGCG Rev Primer ACCATCTTCAGATCCAGAAAGC repeats (tc) x16 PCR product = 245 bp and start at base 260303  2) For Primer TTGGGTAAAGAATGGAGATGCT Rev Primer TTAAAGGTCATTTTCTCCCCCT repeats (att) x18 PCR product = 161 bp and start at base 269365  3) For Primer CCATACCCCATGATGATTTCTT Rev Primer ACCTCTCTCTCCCATTTCATCA repeats (ga) x14 PCR product = 173 bp and start at base 274822  36. SEH found in scaffold 00629 (query sequence starts at base 131051 on scaffold)  1)  For Primer CCAAGTAGTGGAACGAAACACA Rev Primer CGGGTTATGACTTATGAGGGAG repeats (tc) x9 PCR product = 222 bp and start at base 111063  2) For Primer GCGAAATCAACAAACTCAGTCA Rev Primer GCCGTGTAGGAATAGTACAGCC repeats (ag) x7 PCR product = 187 bp and start at base 124426  3) For Primer TCTCTTGTTCTTGGGTCGAACT Rev Primer ATTTGTTGTGTGTTTGTCCTCG repeats (tc) x17 PCR product = 147 bp and start at base 144591  37. SOD2 found in scaffold 01664 (query sequence starts at base 5569 on scaffold)  1)  For Primer GGCAAAAGACAAAATTCGAGAC Rev Primer CTCTCTCTCTGCACCATTCCTT repeats (ga) x9 PCR product = 187 bp and start at base 2653  2) For Primer CACTCAAATGTAAGTCGGTCCA Rev Primer TCTCCTCACTCCCAACTCTCTC repeats (tc) x10 PCR product = 259 bp and start at base 25018  3) For Primer CGCTTTCTTACGATTTTTCTGG Rev Primer CACCGCAACTACACTCATCAAT repeats (ag) x11 PCR product = 103 bp and start at base 29658  38. SS1 found in scaffold 01481 (query sequence starts at base 27541 on scaffold)  1)  For Primer CCCACGAGAAGTCTCTCTCTCT Rev Primer ACGAATCACAAGAAGGAGAAGG repeats (tc) x10 PCR product = 239 bp and start at base 8659  2) For Primer TGCTCTTCCATTCTCTCCTCTC Rev Primer CAACATGCTAAATGCCTAACCA repeats (tct) x14 PCR product = 275 bp and start at base 22423  3) For Primer GCTCCAAGCTTCTTGGTGTATT Rev Primer GGGACTTTCACTCTTTGCATTC repeats (ga) x12 PCR product = 261 bp and start at base 26596  39. STZ found in scaffold 02547 (query sequence starts at base 19296 on scaffold)  1)  For Primer CCGTAGCGTTACTACCCAAAAC Rev Primer TTCGTGAGATGGTAGATGTTGC repeats (ct) x5 PCR product = 149 bp and start at base 8979  2) For Primer GGTCTTCAGCAAAGTCTCCACT Rev Primer AGACGTGGAGCACGTTTAATTT repeats (ct) x5 PCR product = 157 bp and start at base 18203  3) For Primer ATTTTTGTCACCGCTCGTACTT Rev Primer GAGCTGAAACAGCAGAGTAGCA repeats (ct) x5 PCR product = 263 bp and start at base 19147  40. UBQ1 found in scaffold 12687(query sequence starts at base 335 on scaffold)  1)  For Primer TCGTTTGCTCTGGTACTCTCTG Rev Primer CATATCCCCATCACTCCATTCT repeats (taa) x4 PCR product = 173 bp and start at base 705  41. ZFHD1 found in scaffold 00085 (query sequence starts at base 50788 on scaffold)  1)  For Primer TTTTTCGATTCTCCTCTCGTTG Rev Primer TAGCCAAACTAGGCCTTAGCAC repeats (aat) x20 PCR product = 219 bp and start at base 43299  2) For Primer TATTCCTTAGGGGAACATGTGG Rev Primer GCGAACCCATCTTCTACAAAAC repeats (ag) x7 PCR product = 136 bp and start at base 53790  3) For Primer TTCTGTGCACACCTGTTCTTTT Rev Primer AGCGTATGAGAGAAATGCAAGA repeats (ttc) x13 PCR product = 247 bp and start at base 76563 |