**Identifying SSR Markers for Cold Tolerance Gene Orthologs in the Blueberry Genome to Improve Selective Breeding**

Spencer Chadinha

Laboratory Methods in Genomcis

Spring, 2013

**Abstract**

Current methods of artificial selection are limited by the time it takes for crops to mature and the desired traits to develop. We are collaborating on a project to create a genetic screening method to reduce the time consumed by the selective breeding process in *Vaccinium corymbosum* (blueberry). We are generating SSR markers for genes associated with stress tolerance that breeders can use to identify plants with particular combinations of traits. I found genes associated with cold tolerance, which is a major factor in blueberry crop yield – especially in variable weather conditions. I found genes involved in the CBF (Cis-repeat binding factor) mediated cold tolerance pathway and COR (cold response) genes that produced dehydrin proteins. After searching for amino acid sequences and orthologs in the blueberry genome, I found 3 different SSR markers for 15 different genes from an original list of 28 genes involved in cold tolerance.

**Introduction**

Blueberry cultivation could benefit greatly form a decrease in selective breeding times. The US is the leading producer of blueberries in the world and blueberries are the second most profitable fruit crop in the US (Agricultural Marketing Resource Center, 2013). In addition to their economic impact, blueberries have many health benefits including high amounts fiber and vitamin C, low amounts of fat, and polyphenols that have putative antioxidants and anti-inflammatory abilities (US Highbush Blueberry Council 2013). Revenue created by the blueberry is driving selective breeding for more stress tolerant crops and crops with larger fruit sizes.

Selective breeding of agricultural products has led to plants that yield more produce, larger produce, and enhanced natural traits that society deems desirable. For example, seedless watermelons are a product of intense selective breeding (McCuistion and Wehner, 2010). In today’s process of selective breeding, each trait is analyzed qualitatively after the plant has produced its fruit. While this process has generated bigger and better crop yields, this method of selective breeding has a major limitation. Whatever trait we are selecting for, regardless of what it is, can only be detected after the crop has finished growing. This is a serious time investment, especially if the crop lacks the desired trait.

We are collaborating with North Carolina State University, Washington State University, and the U.S. Department of Agriculture to create a more efficient method of selective breeding. We hope to find genetic markers we can use to create a genetic screen that will speed up the process of selecting desirable traits.

 For my particular part of the project, I chose to focus on cold tolerance. The ability of blueberry plants, specifically the flower buds, to survive the cold is one of the most important issues for blueberry cultivar success (Rowland *et. al., 2012)*. Increasing cold tolerance in blueberries will help expand the current range in which blueberries can be grown. Creating a genetic screen will alleviate problems with identifying cold tolerance between cultivars that are grown in different sites and temperature conditions.

 Cold tolerance is mediated by many genes through several complex pathways, therefore, I chose to focus on the pathway most unique to cold tolerance and on proteins most up-regulated during cold acclimation. I chose the CBF signaling pathway because it has the least overlap with other stress response pathways while still regulating a significant portion of the COR genes. I chose Dehydrin proteins based on expert recommendation – in past studies blueberry experts found that dehydrin production increased significantly when the blueberry plant was acclimating to cold.

 The CBF signaling pathway detects low temperatures through an unknown mechanism. Some experts believe that the rigidity of the cell membrane is the signal (Chinnusamy, 2007). The calcium second messenger system relays the external detection of low temperatures to ICE transcription factors, which in turn activate the CBF transcription factors. The CBF transcription factors are ultimately responsible for activating the cold response genes (COR genes) that code for the proteins that carry out the process of cold acclimation (Chinnusamy *et. al, 2007*). Because CBF transcription factors are near the beginning of the transcription cascade, the levels of these proteins have a significant effect on the cold tolerance of blueberries and other plants.



Image from Chinnusamy 2007

**Figure 1: The CBF Mediated Cold Response Pathway.** This figure is a graphical representation of the currently know interactions between all identified proteins involved in the CBF signaling pathway. This figure does not include the dehydrin proteins that were part of the study, however, they would be included among the line labeled COR genes near the bottom of the figure.

 Dehydrins are hydrophilic molecules that are part of a class of proteins known as late embryogenesis abundant proteins (LEA proteins). All dehydrins are associated with dehydration stress in plants, including cold tolerance. There are currently four putative dehydrin conformations, each with a specific function: lipid and protein binding dehydrins, radical scavenging dehydrins, metal binding dehydrins, and cryoprotective dehydrins. All dehydrins lack a well-defined secondary structure, however, it is hypothesized that the lack of secondary structure contributes to their activity (Rorat, 2006).

 To create a time saving method of selective breeding we used short sequence repeats (SSRs), which are highly polymorphic stretches of DNA, to identify particular alleles for genes of interest (Van Belkum *et. al, 1998*). Ideally, with a library of SSRs linked to different alleles, blueberry cultivators can breed particular plants to create a blueberry bush that has all of the desired genotypes.

 Using the knowledge we gain from researching blueberries, we hope to collect a list of genes for our individual topics that we can use to predict increased stress tolerance. From that list of genes, we will generate primers for SSRs that are linked to individual genes. Using these primers, blueberry breeders can quickly screen each blueberry bush and its offspring for SSRs linked to desirable alleles.

**Methods**

 **Choosing Genes of Interest**

I found all of my genes of interest from cold tolerance literature found on PubMed. KEGG pathways is another useful resource for identifying genes involved in specific pathways, unfortunately, the CBF mediated cold response pathway was not annotated for Arabidopsis in KEGG.

 I used the protein database search on NCBI to obtain amino acid sequences for each gene of interest. I used amino acid sequences from close relatives of blueberry when available – such as *Vitis vinifera* (grape) and  *Fragaria ananassa* (strawberry). When close relatives of blueberry did not have an available amino acid sequence for the gene of interest, I used *Arabidopsis thaliana*.

 I used the TextEdit program on Mac to save each individual amino acid sequence. Each text file was converted to a plain text format from the format menu. The method we used to find orthologs in the blueberry genome required the use of plain text files.

 **GenSAS**

We attempted to use the annotation tool GenSAS to locate SSRs for the genes of interest. Unfortunately, GenSAS was still under development and its data output was difficult to decipher and use.

 In addition to attempting to use GenSAS as an SSR tool, we each tested another function of GenSAS. I tested the getorf function, which was supposed to find open reading frames. The getorf tool did produce data, unfortunately it was not presented in a way that I could understand just by reading the getorf guide. Images of the product are provided in the appendix.

 **Running the Local tBLASTn**

To identify potential orthologs of genes of interest in the blueberry genome, we ran a local tBLASTn. tBLASTn takes the amino acid sequences we found for the genes of interest and compares them to the six reading frames of a nucleotide blueberry genome input. We used the program Terminal version 2.3 on a Mac desktop to run the tBLASTn. The first step to running the tBLASTn was to create the database through which the tBLASTn would run. We created the database on Mac desktops using a copy of the blueberry genome and the following code:

cd Desktop/ncbi-blast-2.2.27+

./bin/makeblastdb -in Sorted\_454Scaffolds.txt -input\_type fasta -dbtype nucl -title blueberry\_Genome

bin/tblastn -query query.txt -db Sorted\_454Scaffolds.txt

 The first and second lines of code establish the database on the desktop. The third line of code indicates which amino acid sequence to BLAST against the blueberry genome. To choose a particular text file to BLAST against the genome, replace ‘query.txt’ with the name of your text file. The file must end in .txt for the BLAST to work.

 The terminal window will display the results of your BLAST by decreasing bit score, beginning with the highest and moving down. Following the list of genes is an in depth analysis of each hit. We took the first hit, with the highest bit score, and ignored all other hits for simplicities sake.



**Figure 1. An example of the overview of a tBLASTn output from the local blast**. This list of hits, which scaffold each hit is on, the bit score of each hit, and the e-value of each hit is the first output from the tBLASTn.

**Figure 2. An example of the secondary output from a tBLASTn using the local blast**. The secondary output is an in-depth analysis of each hit. In addition to the bit score and e-value, the secondary output also shows percent identity, number of gaps, location of the hit in the scaffold, and part of the hit sequence.

**Finding SSR Markers**

 To find SSR markers, we took a copy of the scaffold the blueberry ortholog was located on and saved it in a plain text document. We produced SSR markers by submitting the text file of the scaffold sequence to the SSR finder tool on the Vaccinium website (Genome Database for *Vaccinium*, 2011). We changed the settings on the SSR finder to ignore pentanucleotide and tetranucleotide repeats and to only focus on dinucleotide and trinucleotide repeats.

**Figure 6. Sample SSR readout from Vaccinium.org**. This is the readout that vaccinium.org’s SSR tool produces. The readout shows the number of base pairs in the repeat, the repeating motif, the number of repeats, start and stop location, forward and reverse primers, and primer length for each SSR from left to right.

**Choosing the Best SSR Marker and Primer Sequence**

 The SSR tool on vaccinium.org produces a list of all SSRs on the scaffold and information such as where the SSR is located, how many repeats it has, and shows available primer sequences. Given this information, we chose three primer pairs that we felt were the best options to use to screen for the genes of interest.

 My criteria for determining the best SSR marker and primer pair was, in order of importance, length of the SSR repeat, proximity to target gene, and dinucleotide repeats were more desirable than trinucleotide repeats. Unfortunately, some of the SSRs that matched these criteria had primer pairs that were not specific enough to be useful. .

**Results**

 I found fifteen genes that had three different primer pairs of the appropriate quality to use as a genetic screen. Unfortunately, not every gene had an ortholog in the current assembly of the blueberry genome. Any hit that had an e-value above 0.001 was ignored because the percent identity with the input sequence was too low to be considered an ortholog. Several genes of interest did not have hits in the blueberry genome with low enough e-values to be considered orthologs Additionally, some genes with orthologs in the blueberry genome did not produce usable primer pairs or produced very poor primer pairs. Thus, of the 28 initial genes of interest, only 15 produced usable SSRs with primers. A complete list of SSR markers and primer pairs for each gene is include in Appendix B.

|  |  |
| --- | --- |
| Regulation Protein | Dehydrin Protein |
| CBF2/DREB1C | COR47 |
| ICE1 | ERD10 |
| MYB15 | COR6.6 |
| ZAT12 | COR15A |
| LOS2 | COR78 |
| RAP2.1 | WCS19 |
| HOS9 | CAS15A |
| HOS10 | HVA1 |
| ZAT10 |  |
| RAP2.6 |  |
| VrCBF4 |  |
| ESK1 |  |
| RAV1 |  |

**Table 1: List of genes target for SSR markers.** Proteins under the regulation proteins column are all involved in the CBF signaling pathway, acting as either transcription factors or repressors. Proteins under the dehydrin proteins column are all dehydrin proteins. Genes highlighted green have viable SSR markers and primers, genes in red did not have viable SSR markers or primers. Genes highlighted in green produced usable primer pairs, genes in red had orthologs without primer pairs, and genes highlighted purple all hit the same spot on the blueberry genome, suggesting the sequences used to blast against the blueberry genome were too similar to be distinguished from each other. Genes that did not have orthologs are: ABF1, ABF2, CBF1, CBF3, CORA, ABF3, and ABF4.

**Discussion**

**Limitations of Artificial Selection Through Genomics**

While the SSR markers for the genes in this paper will help blueberry breeders select more cold tolerant blueberry bushes, we are only correlating genetic markers to phenotypes. This method does not help us understand how the alleles of each gene promote cold tolerance, this method only helps us identify cold tolerant alleles.

 A second limitation to this method is the amount of resources that would be required to create a complete SSR library that blueberry breeders can use to screen their crops. Identifying genes related to cold tolerance, finding their orthologs in blueberry, and identifying the best primer pairs for SSR markers is a time consuming process and paying researchers to complete the library would be extremely expensive. However, once the initial time and costs have been paid, the genomic method of selective breeding will be thorough and streamlined.

**Conclusion**

 The SSR markers and their primer pairs in this paper will help build a more complete library for the genetic screening process. Although only fifteen cold tolerance genes with useful primer pairs were found, blueberry cultivators can still use them to select for more cold resistant offspring in future plant crosses. Despite the limitations of the genomic approach to selective breeding today, the knowledge and technology are heading towards a point where powerful genetic screens can be created.

 Each new SSR marker added to the SSR library will increase the effectiveness of the blueberry genomic screen, allowing blueberry breeders to identify crops that have greater cold tolerance. As more cold tolerant blueberry bushes are bred, the yearly yield of blueberries will increase – benefiting the agricultural economy and making the health benefits of blueberries more accessible.

 **Works Cited**

Agricultural Marketing Resource Center. Blueberries [Internet]. Ames (IA): Agricultural Marketing Resource Center, Iowa State University; 2013 [cited 2013 May 9]. Available from <http://www.agmrc.org/commodities__products/fruits/blueberries/>

Chinnusamy V., Zhu J., and Zhu J-K. *Cold stress regulation of gene expression in plants.* TRENDS in Plant Science 12(10): 444-451 (2007).

Davik J., Koehler G., From B., Torp T., Rohloff J., Eidem P., Wilson R. C., Sønsteby A., Randall S. K., and Alsheikh M. *Dehydrin, alcohol dehydrogenase, and central metabolite levels are associated with cold tolerance in diploid strawberry (Fragaria spp.)*. International Journal of Plant Biology, (September 2012).

Genome Database for *Vaccinium*. Kannapolis (NC): Plants for Human Health Institute, North Carolina State University, and Washington State University; 2011 [cited 2013 May 9]. Available from <http://www.vaccinium.org>

Guy C. *Molecular Responses of Plants to Cold Shock and Cold Acclimation.* J. Mol. Microbiol. Biotechnol. 1(2): 231-242 (1999).

Heidarvand L., Amiri R. M. *What happens in plant molecular responses to cold stress?* Acta Physiol Plant. 32:419–431 (2010).

McCuistion F. and Wehner T. C. Seedless Watermelon Breeding [Internet]. Raleigh (NC): Cucurbit Breeding Horticultural Science, North Carolina State University; 2010 June [cited 2013 May 9]. Available from <http://cuke.hort.ncsu.edu/cucurbit/wmelon/seedless.html>

Puhakainen T., Hess M., Mäkeäl P., Svensson J., Heino P., and Palva E. T. *Overexpression of multiple dehydrin genes enhances tolerance to freezing stress in Arabidopsis.* Plant Molecular Biology 54: 743-753 (2004)

Rorat T. *Plant Dehydrins – Tissue Localization, Structure and Function.* Cellular and Molecular Biology Letters 11: 536-556 (2006).

Rowland L. J., Bassil N. V., Brown A. F., Drummond F. A., Hancock J. F., Olmstead J. W., Polashock J. J., Alkharouf N., Finn C., Hanes S., Mockler T., Campbell M., Graham J., McCallum S., Buck E., and Storey R. *Generating Genomic Tools for Blueberry Improvement.* Grant proposal (2012).

Siddiqua M., Nassuth A. Vitis *CBF1 and* Vitis *CBF4 differ in their effect on Arabidopsis abiotic stress tolerance, development and gene expression.*Plant, Cell & Environment. 34(8): 1345–1359 (2011).

Thomashow M. F. *Role of Cold-Responsive Genes in Plant Freezing Tolerance.* Plant Physiol. 118: 1–7 (1998).

U.S. Highbush Blueberry Council. Health Benefits of Blueberries [Internet]. Folsom (CA): U.S. Highbush Blueberry Council; 2013 [cited 2013 May 9]. Available from <http://www.blueberrycouncil.org/health-benefits-of-blueberries/blueberry-nutrition/>

Van Belkum A., Scherer S., Van Alphen L., and Verburgh H. *Short Sequence DNA Repeats in Prokaryotic Genomes.* American Society for Microbiology 62(2): 275-293 (1998).

Zhu J., Dong C-H., and Zhu J-K. *Interplay between cold-responsive gene regulation, metabolism and RNA processing during plant cold acclimation*. Current Opinion in Plant Biology, 10:290–295 (2007).

**Acknowledgements**

I would like to acknowledge Allen Brown at NCSU, Jeannie Rowland at USDA, and Doreen Main at WSU for their work on the blueberry genome, without which this study would not have been posslbe. I would also like to acknowledge the Davidson College Laboratory Methods in Genomics class of 2013 for helping to develop a procedure for using the local blast and analyzing the effectiveness of the GenSAS tool.

**Appendix**

**Figure 2. Information input screen for GenSAS**. This image is the user interface for GenSAS. The input sequence – in this case scaffold – that we want annotated goes in the top left panel then one or multiple annotation jobs can be chosen from the bottom left panel. The bottom right panel shows the tasks that are being processed by GenSAS. All results can be accessed from the task queue and will only be stored for a limited amount of time.



**Figure 3. Output from getorf**. This image shows the output from the getorf annotation function. This particular output was difficult for me to understand, additionally, it only showed the first 200 results without any indication of how many results it found. When I told the program designer I could not understand the output I was a way to download the full results page, unfortunately, I was unable to find the results page.

**Appendix B**

**SSR markers and their primer pairs**. Each gene in this list has three different primer pairs that can be used for a genetic screen. The location of each ortholog in the blueberry genome is listed above the three primer pairs. For each primer pair, the number of repeats, repeat motif, PCR product length, and location on the scaffold are listed.

COR47 (dehydrin) found in scaffold 00093 (query sequence starts at 55523 in scaffold)

1)

 - for primer: GAAAAATCAGCCCCAAAAGAG

 - rev primer: ATAATCGTCCCCACTTACAACG

 - repeats (ga) x 19, PCR product: 165 bp, and starts at 210233

2)

 -for primer: ATCCAGTGGTCTAGCAGAGGAG

 -rev primer: CCCACTTAACCCAGCAAAAATA

 -repeats (ag) x 13, PCR product : 281 bp, and starts at 162503

3)

 -for primer: GATTGACTGTTTTTCTCCGCTT

 -rev primer: CTCATCAGAAACCAGTGATCCA

 -repeats (ag) x 11, PCR product : 179 bp, and starts at 76616

DREB1C/CBF2 found in scaffold 00009 (query sequence starts at 488186 in scaffold)

1)

 - for primer: CACACAAGACGTGTATGCAGAA

 - rev primer: CCCTCATTATGAAATGGCTATG

 - repeats (ag) x 24, PCR product: 298 bp, and starts at 625376

2)

 -for primer: AGTTCTAAACCGATTGTGCGTT

 -rev primer: AATTCCAACCTAACTGCCAGAA

 -repeats (ag) x 10 PCR, product : 291 bp, and starts at 479995

3)

 -for primer: ATGTTAATGTCACTTGGGCCTC

 -rev primer: AGCTTCTCTGTGACTACGTGGC

 -repeats (ag) x 15 PCR, product : 250 bp, and starts at 359942

ICE1 found in scaffold 00051 (query sequence starts at 55523 in scaffold)

1)

 - for primer: GATTTATGCACAACAAGGCTCA

 - rev primer: GACAGAAGAGTTAGCCAAGGGA

 - repeats (aat) x 13, PCR product: 237 bp, and starts at 27931

2)

 -for primer: GTTTGTGAACTCAGAGGCAGTG

 -rev primer: TATGTGACCCCAGTACACAAGC

 -repeats (ag) x 11, PCR, product : 281 bp, and starts at 162503

3)

 -for primer: GATTGACTGTTTTTCTCCGCTT

 -rev primer: CTCATCAGAAACCAGTGATCCA

 -repeats (ag) x 11, PCR product : 268 bp, and starts at 24413

ZAT12 found in scaffold 000815 (query sequence starts at 76639 in scaffold)

1)

 - for primer: AGGATCCATTTCCCTCTTGATT

 - rev primer: CATGCAGCAGATAGGTTAAATCTG

 - repeats (ga) x 15, PCR product: 266 bp, and starts at 3883

2)

 -for primer: GAAGAAATCGTACCTTTGGACG

 -rev primer: AACACATATTGAATAACGGCCC

 -repeats (ag) x 6, PCR product : 153 bp, and starts at 106056

3)

 -for primer: GAATTTTTGCCACAAAGGTCTC

 -rev primer: CCTATATTTCTTGAACCCAGCG

 -repeats (aat) x 6, PCR product : 248 bp, and starts at 70960

HOS10 found in scaffold 000067 (query sequence starts at 237891 in scaffold)

1)

 - for primer: ATTCGCGTATCTACACCCAAAC

 - rev primer: ATTTGCCTCCCTTATTTTCCTC

 - repeats (ag) x 19, PCR product: 221 bp, and starts at 208660

2)

 -for primer: ACATGGGCCCTACTCTCCTTAT

 -rev primer: GTCCGGTCATCTCAATTTGTCT

 -repeats (ag) x 11, PCR product : 195 bp, and starts at 275656

3)

 -for primer: GCTTTGTTACAATACCGCATCA

 -rev primer: AGGAGTTGTGCAACCCATACTT

 -repeats (aat) x 5, PCR product : 187 bp, and starts at 243279

HOS9 found in scaffold 01128 (query sequence starts at 105528 in scaffold)

1)

 - for primer: CATTACTGTTTAAGAGCGTGCG

 - rev primer: GGAAGTTGACAAAGACCAGACC

 - repeats (tc) x 12, PCR product: 184 bp, and starts at 93786

2)

 -for primer: GCGCGTATGTGCTTGTAAAATA

 -rev primer: CGTGTAGCGATGAACTCGTTAG

 -repeats (ct) x 11, PCR product : 242 bp, and starts at 94750

3)

 -for primer: GGACGGCTGGAGTAACTGATAG

 -rev primer: ATACAACGATCATCCGGAAAAC

 -repeats (ta) x 5, PCR product : 287 bp, and starts at 107728

MYB15 found in scaffold 00171 (query sequence starts at 43657 in scaffold)

1)

 - for primer: TGGCAATCAACCCTAAGAGATT

 - rev primer: CCTCATCTCTTCTCCCCTCTCT

 - repeats (tc) x 23, PCR product: 152 bp, and starts at 72875

2)

 -for primer: CTACCCATTTTCCCTCTCCTCT

 -rev primer: TCACCTGTGTGTGGGTACTTTC

 -repeats (ct) x 8, PCR product : 287 bp, and starts at 39227

3)

 -for primer: CGGATATTCTATAGCGGGTCAG

 -rev primer: TAGATGTGGTGGGCCAGTAGTA

 -repeats (ta) x 5, PCR product : 253 bp, and starts at 41283

ABF1 found in scaffold 00277 (query sequence starts at 131353 in scaffold)

1)

 - for primer: CCAATATGCTGATGGAATCTGA

 - rev primer: ATCATATTCGCGAGAAGTTTCC

 - repeats (tc) x 19, PCR product: 273 bp, and starts at 116819

2)

 -for primer: TAATGGGAGGGTTGAAAATGTC

 -rev primer: TTGGTGTGTTGTGTAGGTGTGA

 -repeats (tc) x 9, PCR product : 130 bp, and starts at 146125

3)

 -for primer: ATAACAACACCACCACCAACAA

 -rev primer: ACAACCTATTCATGACCCCAAC

 -repeats (aac) x 7, PCR product : 239 bp, and starts at 126581

LOS2 found in scaffold 01453 (query sequence starts at 47942 in scaffold)

1)

 - for primer: CGACGGAAATACTGAAATCCTC

 - rev primer: ACTTGGCGTATTGAGTGTCCTT

 - repeats (ct) x 10, PCR product: 292 bp, and starts at 21100

2)

 -for primer: ACATAAACCAAAAACCCTTCCC

 -rev primer: ACGGTAGTTCACTGATGAGGGT

 -repeats (ct) x 8, PCR product : 186 bp, and starts at 45159

3)

 -for primer: CTTCTGAATATTCCGCTCTTCG

 -rev primer: ACGGTAGTTCACTGATGAGGGT

 -repeats (ct) x 8, PCR product : 260 bp, and starts at 45114

RAP2-1 found in scaffold 00667 (query sequence starts at 78548 in scaffold)

1)

 - for primer: GCAAGACGCCAAATATGTGATA

 - rev primer: CTCTATCTTTTCCTTTCCACCG

 - repeats (ag) x 10, PCR product: 274 bp, and starts at 81226

2)

 -for primer: GGCACTAATTCATCTAGGCTCC

 -rev primer: GAGGAACAAGGATTTGGAGTTG

 -repeats (ca) x 7, PCR product : 298 bp, and starts at 72768

3)

 -for primer: CTCTCTTATTTCTGCCACCCAC

 -rev primer: TGTAACAATTAGAGGGAGGGGA

 -repeats (tc) x 5, PCR product : 128 bp, and starts at 78578

RAP2-6 found in scaffold 00216 (query sequence starts at 246677 in scaffold)

1)

 - for primer: CAATGATTCACACTAAGCTCCC

 - rev primer: TAGGTGGCTTCAAACAAATTCC

 - repeats (tc) x 31, PCR product: 294 bp, and starts at 240883

2)

 -for primer: GTTGAACCATAGCTACCATCGG

 -rev primer: ATGGGGACACTATTTCATTTGG

 -repeats (ga) x 5, PCR product : 240 bp, and starts at 248540

3)

 -for primer: CCAAATGAAATAGTGTCCCCAT

 -rev primer: GGAGAGAGAAAATGGGTTGTTG

 -repeats (gcg) x 7, PCR product : 288 bp, and starts at 248884

RAV1 found in scaffold 00930 (query sequence starts at 63069 in scaffold)

1)

 - for primer: ATTATGTGGAAAGCAAACCGAC

 - rev primer: TAGTTTGTGGCAAATGAGGATG

 - repeats (ct) x 11, PCR product: 189 bp, and starts at 62386

2)

 -for primer: CCACACGAGACATACGCATTAT

 -rev primer: AGGAAACAGAGACTGGGGTGT

 -repeats (ct) x 5, PCR product : 214 bp, and starts at 62994

3)

 -for primer: AGTTGAGAGCAAGAGTGCACAG

 -rev primer: TAAAAAGGTGAAGTAGCCTGCC

 -repeats (ct) x 7, PCR product : 255 bp, and starts at 66036

ZAT10 found in scaffold 02547 (query sequence starts at 19296 in scaffold)

1)

 -for primer: ATTTTTGTCACCGCTCGTACTT

 -rev primer: GAGCTGAAACAGCAGAGTAGCA

 -repeats (ct) x 5, PCR product : 263 bp, and starts at 19147

2)

 -for primer: GGTCTTCAGCAAAGTCTCCACT

 -rev primer: AGACGTGGAGCACGTTTAATTT

 -repeats (ct) x 5, PCR product : 157 bp, and starts at 18203

3)

- for primer: CCGTAGCGTTACTACCCAAAAC

 - rev primer: TTCGTGAGATGGTAGATGTTGC

 - repeats (ct) x 5, PCR product: 149 bp, and starts at 8979

ESK1 found in scaffold 00145 (query sequence starts at 55646 in scaffold)

1)

 -for primer: AAGAGACAGGGAGTAAGGGTCC

 -rev primer: TGCAGTATCAGTTCCCTCTCTTT

 -repeats (ag) x 21, PCR product : 298 bp, and starts at 268361

2)

 -for primer: ACTACTCTGGCAATGCTCAACA

 -rev primer: CCGTACCCTTCGTTCAAAATAA

 -repeats (tc) x 11, PCR product : 272 bp, and starts at 69356

3)

- for primer: AAAACTGCTCTGATCAAGCCTC

 - rev primer: GCTTGTAATATGGTGTGCGTGT

 - repeats (ca) x 9, PCR product: 235 bp, and starts at 58379

CBF4 found in scaffold 00023 (query sequence starts at 138663 in scaffold)

1)

 -for primer: ACACGATTTGAGAATCACCCTT

 -rev primer: GGGGTTGGCTACATGATACATT

 -repeats (ga) x 23, PCR product : 251 bp, and starts at 475004

2)

 -for primer: GGATCAAAGTGCTTTTCCAAGT

 -rev primer: TGGACAAAGGTATTGCATCTTG

 -repeats (ga) x 11, PCR product : 296 bp, and starts at 128901

3)

- for primer: TAAGCAAAATGGCCTGGTTAGT

 - rev primer: CCACTTTCAATTTCGTAGGCTC

 - repeats (ag) x 11, PCR product: 126 bp, and starts at 121561