**Candidate Genes in Sulfur Related Pathways Identified in *Brassica Oleracea* Near QTLs**

Gracie Gordon, Michael Lorentsen, and James Helzberg

**Abstract:**

In a nation plagued by fast food induced obesity, nutrition is a critical area of research. Many people in the United States, and around the world, are not getting enough nutrients in their diets. This project focused on mineral uptake, assimilation, and storage in *Brassica oleracea*. Our team focused on sulfur and its role in amino acid synthesis, glucosinolate production, and the glutathione pathway. We used the literature to locate genes in *Arabidopsis thaliana* that were associated with sulfur pathways and near QTLs, provided by NC State. We used BLAST and IGB (interactive genome database) to locate the genes close to the QTLs in *B. oleracea*. We were able to locate fourteen candidate genes, near the QTLs, that should be pursued by broccoli breeders to produce plants with higher levels of sulfur related nutrients that will help improve human health and nutrition.

**Introduction:**

*Brassica oleracea* (broccoli) is a nutrient-rich food that provides many essential vitamins and minerals for humans, one of which is sulfur.  Nimni *et al.* (2007) explains how sulfur is necessary for the synthesis of the essential amino acids cysteine and methionine and provides detoxification functions such as neutralizing toxic compounds and free radicals. Broccoli also delivers a source glutathione that helps store sulfur in the liver, and contains glucosinolates (GSH), which may help prevent cancer. Plants also need Sulfur to synthesize cysteine and methionine, as well as to rid themselves of toxic compounds. These essential functions make sulfur a vital mineral.

Leustek *et al.* (2000) and Buchner *et al.* (2004) explain how sulfur travels through *B. oleracea.* Sulfur is first taken up from the soil by the roots as sulfate using hydrogen gradients.  The SULTR genes encode sulfur transferase molecules and help move sulfate through the plant.  SULTR1 helps transfer the sulfate through the roots.  SULTR2 transfers the sulfate through the phloem to the leaves and SULTR4 uses a hydrogen gradient or ATPase to move it into the plastids. Here the sulfate can be assimilated via reduction or is stored in the vacuole. The assimilation pathway involves O-acetylserine and APS reductase as well as glucosinolates to form sulfur-containing molecules. Sulfate uptake is down regulated when the plant includes many sulfur containing compounds or has high levels of O-acetylserine and uptake is down regulated when these levels are low.  Any sulfur sequestered in the vacuole does not play a part in this feedback system, (Koprivova *et al*., 2013; Maruyama-Nakashita *et al*., 2004; Hawkesford, 2000).

The model organism *Arabidopsis thaliana* was used as a starting place to find orthologous genes that play a role in sulfur uptake and storage in *B. oleracea*. Many genes of interest were found in the literature and pursued in *B. oleracea,* (listed in Table 2). This paper attempts to locate the orthologous genes from *A. thaliana* in *B. oleracea* close to quantitative trait loci (QTLs) that are associated with sulfur uptake, assimilation, and storage. The QTLs were located by researchers at NC State and placed on one of the nine *B. oleracea* chromosomes that they assembled. In finding genes associated with the QTLs, we can breed broccoli for sulfur enrichment. Growing broccoli with more sulfur content could help combat sulfur deficiencies in our diets, helping to produce amino acids, detoxify compounds in our bodies, and potentially help prevent cancer.

**Methods:**

BLAST: We searched the literature to find genes in *A. thaliana* that were related to sulfur uptake and storage. These genes are listed in Table 2. We were also provided three QTLs by NC State, shown in Table 1, on the *B. oleracea* genome and two on the *B. rapa* genome.  We located the amino acid sequences for each gene of interest in *A. thaliana* from NCBI and used these sequences to run a tBLASTn against the *B. oleracea* and *B. rapa* genomes.  We recorded hits we found that were within 2Mbp of the given QTLs with an E-value of 10-6 or smaller.

IGB: We used the Integrated Genome Browser (IGB, 2014) program to search for genes within 1Mbp of each QTL.  We explored each gene that was annotated in this program and followed up these findings with research in the literature to verify any connection to sulfur uptake and storage, glucosinolates, or glutathione (GSH).

Table 1 . QTLs for *B. oleracea* and *B. rapa*

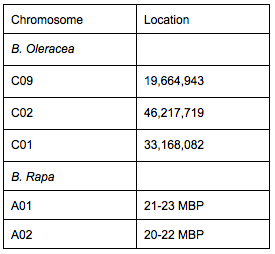


Table 1. This table displays the QTLs that the researchers at NC State provided.

Combined: We also used the BLAST function in IGB to verify that the descriptions in IGB matched the BLAST results.  In genes that lacked an overlap, we took the amino acid sequence from a tBLASTn and a nucleotide translated sequence from IGB at the location of the gene. We then performed a BLAST2 between the amino acid sequences to verify that the nucleotides were the same. Using information from BLAST and IGB, we compiled a list of candidate genes, researching their function to provide more support that they are good candidate genes. From these data, we were able to compile a list of potential primers for each candidate gene within roughly 200,000bp from the gene’s location, based on a list of primers provided by researchers at NC State (Appendix).

**Results:**

Using the BLAST and IGB methods we were able to find fourteen candidate genes close to the QTLs, which are summarized in Figure 2 and Table 3. The proteins these genes encode can be grouped into major categories: transport proteins, proteins associated with pathways involving sulfur storage or detoxification and proteins that help synthesize sulfur-containing molecules such as glucosinolates, and glutathione. Each of these groups of proteins plays important roles in the uptake, assimilation, and or storage of sulfur in *B. oleracea.* It was interesting that these molecules were all expressed in the roots and leaves since the root is the site of uptake and the leaves the site of assimilation, but most were expressed in the flower and stem as well, (Phoenix Bioinformatics Corporation, 2014).

The gene of interest, phosphate transporter like protein may seem out of place at first, but Figure 1 shows that phosphate transporter and sulfur transporters are extremely similar in terms of transmembrane domains, therefore this gene should be examined further, (Smith *et al*., 2000).

Table 2. Genes of interest BLASTed against *B. oleracea*

|  |  |
| --- | --- |
| Gene ID | Function |
| AT1G75270 | GSH-dependent dehydroascorbate reductase |
| AT1G19570 | GSH-dependent dehydroascorbate reductase |
| AT2G25080 | Glutathione peroxidase |
| AT3G24170 | Glutathione reductase, cytosolic |
| AT1G02920 | Glutathione transferase |
| AT3G43800 | Glutathione transferase-like protein |
| AT3G09270 | Putative glutathione transferase |
| AT3G13110 | Serine acetyltransferase (Sat-1) |
| AT3G59760 | Cysteine synthase oasC |
| AT3G03630 | O-acetylserine (thiol) lyase; cysteine synthase |
| AT1G36370 | Putative serine hydroxymethyltransferase |
| AT3G17390 | S-adenosylmethionine synthetase like |
| AT4G01850 | S-adenosylmethionine synthetase 2 |
| AT3G02470 | S-adenosylmethionine decarboxylase |
| AT4G13940 | Adenosylhomocysteinase |
| AT4G23100 | Gamma-glutamylcysteine synthetase |
| AT5G13550 | Sulfate transporter AST68 (Sultr2;1) |

Table 2. This table gives the gene ID for *Arabidopsis thaliana* and the function or each gene that is associated with sulfur uptake, assimilation, and storage, (Nikifora et al., 2003; Smith *et al*., 2000).

The most notable genes that were found were the ABC transporter protein, O-acetylserine, cysteine synthase, glutaredoxin, and APS reductase. O-acetylserine, cysteine synthase and APS reductase are vital for production of the essential amino acid cysteine, (Phoenix Bioinformatics Corporation, 2014). Without cysteine, proteins would not be able to form correctly, rendering them non functional. ABC transporters and glutaredoxin are involved in synthesis of glucosinolates. Glucosinolates are of particular interest because they have detoxification function, including prevention of cancer, (Nimni *et al*., 2007). All these genes are associated with pathways that are vital to the plant, either producing essential amino acids or keeping the plant free of toxins.

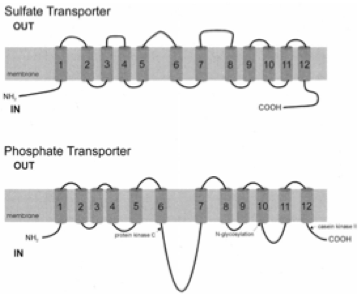


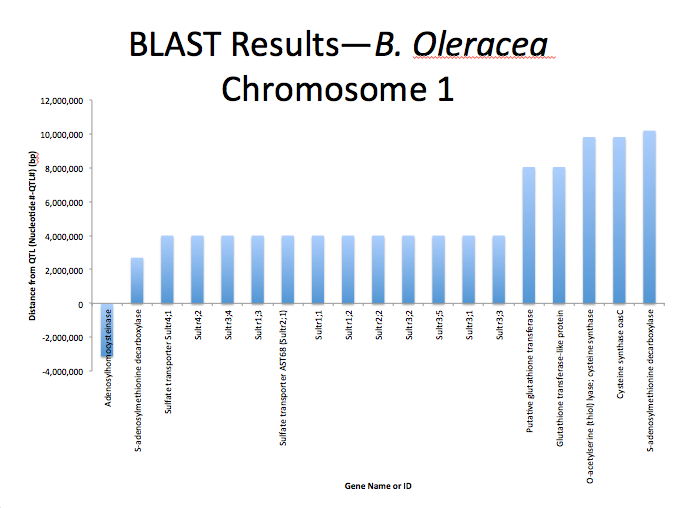
Figure 1. This image shows similarity between sulfur transporters and phosphate transporters, (Smith *et al*., 2000).

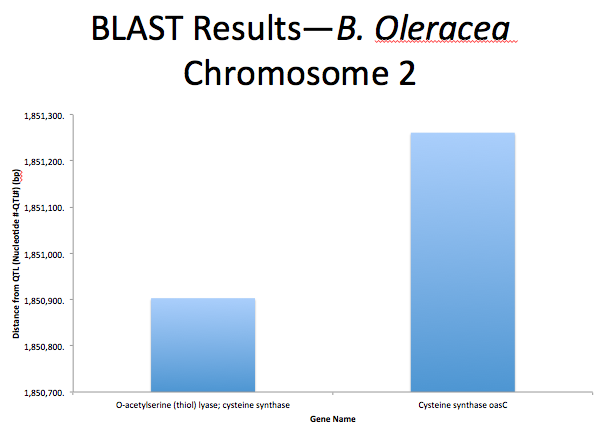
Table 3. Genes of interest located in *B.oleracea*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Chromosome | Gene | Pathway | Expression | Method |
| 1 | SULTR | Sulfur transport | Roots, stems, leaves | BLAST |
| 1 | ABC transporter | Glucosinolate biosynthesis, Transferase via ATPase | Leaves, root, flower, stem | IGB |
| 1 | Glutaredoxin | Glucosinolate biosynthesis, glutathione disulfide pathway | Leaves, flower, root, stem | IGB |
| 1 | Major facilitator superfamily protein | Transmembrane transport | Leaves, flower, root | IGB |
| 2 | O-acetylserine | Cysteine synthesis, glucosinolate pathway | Root, flower, leaves, stem | BLAST |
| 2 | Cysteine synthase | Cysteine synthesis, glucosinolate pathway | Root, flower, leaves, stem | BLAST |
| 2 | Glutathione transferase-like protein | Glutathionee system | Roots, leaves | IGB |
| 2 | Phosphate transporter like protein | Phosphate transport, possibly related to SULTR genes | Roots, stems, leaves | IGB |
| 2 | Glutaredoxn | Glucosinolate biosynthesis, glutathione disulfide pathway | Leaves, flower, root, stem | IGB |
| 9 | O-acetylserine | Cysteine synthesis, glucosinolate pathway | Root, flower, leaves, stem | BLAST |
| 9 | APS reductase | Cysteine biosynthesis, glucosinolate pathway, reduces sulfate | Leaves, flower, root, stem | BLAST |
| 9 | Glutathione S-transferase | Glutathionee system | Roots, leaves | IGB |
| 9 | Protein disulfide isomerase like protein | Catalyzes disulfide bonding | Leaves, flower, root, stem | IGB |
| 9 | S-adenosyl-L-methionine-dependent methyltransferases superfamily protein | Transfers methyl groups | Leaves, flower, root, stem | IGB |

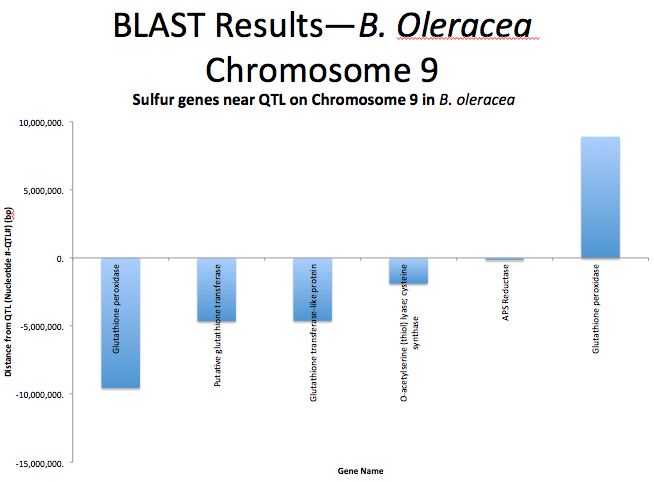
This table summarizes information about the genes of interet including method, location, expression data, and associated pathways, (Phoenix Bioinformatics Corporation, 2014).

A

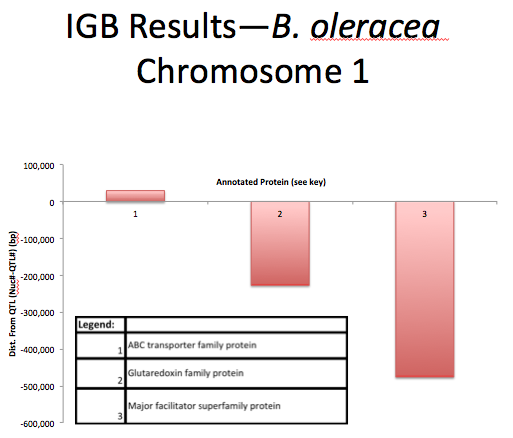


B

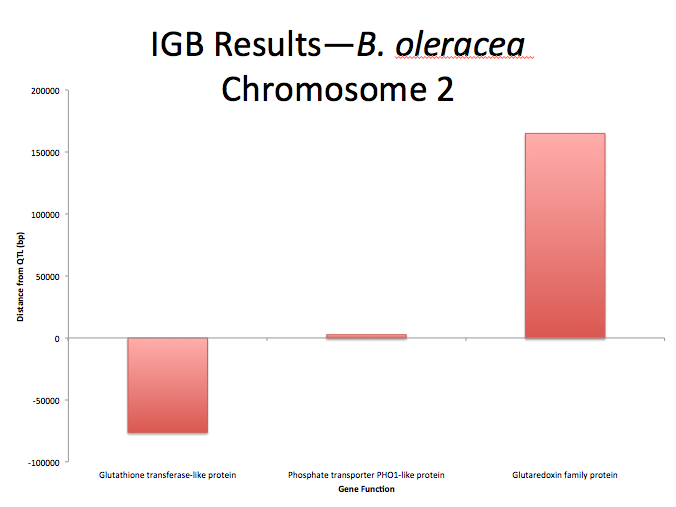
C



D



E



F

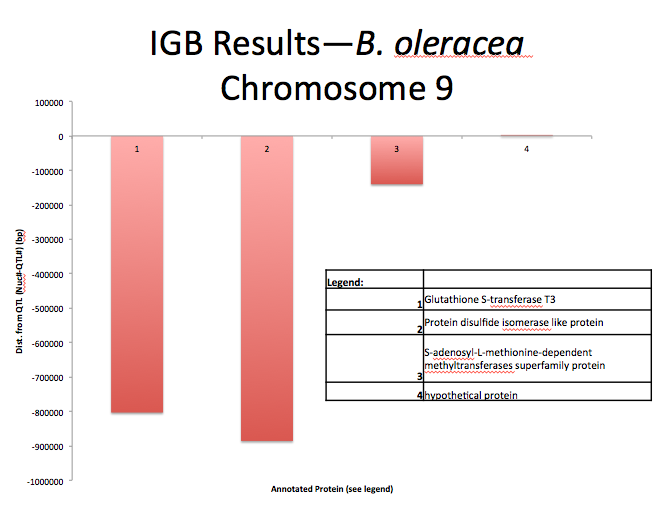


Figure 2. Graph visualizes BLAST (blue, A-C) and IGB (red, D-F) results for each chomosome containing a QTL for sulfur. The distance each gene is from the QTL is shown with the bars. The y-axis shows the QTL is at position zero. A negative number represents a smaller bp number where a positive number is a bigger bp number than the QTL.

**Discussion:**

This study located many genes vital to sulfur uptake, assimilation and storage that were near our QTLs using both BLAST and IGB, but we did not see much correlation of genes between our two methods. Although there were genes that belonged to the same family they were located in different positions on different chromosomes. These genes included cysteine synthase, glutaredoxin, glutathione transferase and sulfur related transferase.

Since there was not much overlap between BLAST results and IGB results we used the BLAST result closest to a QTL, APS reductase, to see if the sequence of the gene was in IGB, but not annotated. When the amino acid sequences from the same location of the BLAST hit was compared with IGB there was 100% identity, meaning the APS Reductase gene was in IGB, but was not annotated.  This shows that IGB does not contain a complete set of annotations. Ideally the annotations in IGB would correspond to information produced by BLAST, because using these methods together would help form a strong case for the genes actually being there. IGB is also a very easy way to search around a QTL for certain genes, so adding more annotations would make the process of searching for genes much easier. Our colleagues who performed searches for other minerals with these tools had varying levels of success. Some had results with very little overlap, but other groups had almost every gene appear in both methods. Hopefully in the future more annotations will be added so the information in both methods will correspond in all cases.

With the information about the location of the genes of interest, we were able to compile a list of SSR primers that researchers at NC State can use to look for markers so that they can quickly breed broccoli with these traits. We found tri-nucleotide repeats within 100,000bp of the location of the gene of interest with a high number of repeated units (Appendix A). When tri-nucleotides were rare, we provided options of di-nucleotides with high repeats, accompanied by the best tri-nucleotide candidates available.

Ideally, more research will be performed on these genes, using the primers provided, to verify their roles in *B. oleracea* and to breed sulfur enriched broccoli*.* With this information breeders at NC State will be able to implement breeding to optimize these genes of interest, which will increase levels of sulfur, glutathione, and glucosinolates in the plant. Raising levels of these nutrients can help prevent cancer and ensure the production of methionine and cysteine in the plant. These improvements would make broccoli even more nutritious and could help improve human health and nutrition, as well as aid in the fight against cancer.

**Appendix:**

Table 3. SSR primers for genes of interest in *B. oleracea*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Gene | Chr# | Gene Start | Primer | Location | Motif | Size | Forward | Reverse |
| ABC transporter family protein | 1 | 33,197,530 | 1 | 33,309,207 | (TTC)12 | 268 | AATATTCAAGAATGCAAAAG | TAACCATATAACAATCCCAC |
|  |  |  | 2 | 33,188,799 | (AGA)6 | 327 | TGTTGAGAAAGGACAAGTAG | AGGACTTGGGTAATAGAGAC |
|  |  |  | 3 | 33,188,485 | (CCG)5 | 396 | TATATTAGGCACAAAAGAGG | CTACTTGTCCTTTCTCAACA |
| Glutaredoxin family protein | 1 | 32,941,169 | 1 | 32,779,679 | (ACTA)9 | 234 | AAATAGTACAAGAGTTTGCG | ATCATTTGAAAGTTTGACAC |
|  |  |  | 2 | 33,066,223 | (ACC)5 | 356 | TTTACCAAAATACGCTCTAC | TACTTCTCCTTCCAACAAC |
|  |  |  | 3 | 32,931,524 | (ATT)5 | 353 | GTCAAAAACAAAAACTGAAC | TTGTTAAGGATTTTCTGAAG |
| Major facilitator superfamily protein | 1 | 32,710,746 | 1 | 32,779,679 | (ACTA)9 | 234 | AAATAGTACAAGAGTTTGCG | ATCATTTGAAAGTTTGACAC |
|  |  |  | 2 | 32,775,112 | (CTT)6 | 260 | GATTCAACATCATTATCCAC | AAGAGTATCCAAAGGTCTTC |
|  |  |  | 3 | 32,777,644 | (GAG)6 | 185 | CCAAGTTAATATGCTACGTT | ACTAATCCTCATTCTTCTCC |
| Glutathione transferase-like protein | 2 | 46,141,482 | 1 | 46,140,715 | (TC)13 | 341 | GTGTCAAGATATGAGAGGTG | ACCAACAAATAACAAAGAAA |
|  |  |  | 2 | 46,124,066 | (AT)12 | 270 | TCAGCCAGAATTATTGTATT | TCAGCCAGAATTATTGTATT |
|  |  |  | 3 | 46,080,404 | (TTA)5 | 370 | AATCCCCATGAGTAAATTAG | CTAATTTTCTGCAATCACTT |
| Phosphate transporter PHO1-like protein | 2 | 46,220,368 | 1 | 46,187,707 | (TCT)5-(TCT)7 | 326 | TCTTTTACCAGGATCAATAA | CACTTACCGTAGTTGTTTTC |
|  |  |  | 2 | 46,217,127 | (CAAA)5 | 328 | CTTTGTTTCAAAGATTGTTC | TAGGTGTTTCATAAGGGTTA |
|  |  |  | 3 | 46,201,795 | (TC)11 | 124 | AATGTTTGTTTAACAGAGGA | ACATATTTAAATCAAAAGCG |
| Glutaredoxin family protein | 2 | 46,052,574 | 1 | 45,979,906 | (TTA)11 | 357 | TATCCATGTCATATTTTTCC | CTGATATTCTCTCTGTCCAA |
|  |  |  | 2 | 45,930,506 | (AT)15 | 266 | GCAAAAACTGTTATCAAATC | TGATCAATGGTTTAATTTTT |
|  |  |  | 3 | 45,860,034 | (CT)13 | 132 | TTCTTCCTCCTCTTTCTAAT | AATAGGGAGAGAGATTTGAG |
| O-acetylserine (thiol) lyase; cysteine synthase | 2 | 42,994,233 | 1 | 42,972,095 | (AT)13 | 321 | CCAGATCTTATGGTGTAGAA | ACTGATTTCAAGTGAGAGTG |
|  |  |  | 2 | 42,987,955 | (AAT)5 | 207 | AAAATAAACTGCTTTGAATG | CTTTTTAAAATGCAAAAGAC |
|  |  |  | 3 | 42,915,459 | (CAG)8 | 368 | TCGAAGTATATCCCTCTACA | ACATGCTTATGTTTTTGTTT |
| Cysteine synthase oasC | 2 | 48,068,980 | 1 | 47,991,398 | (TCT)10 | 312 | AAATGAAAGCACTAAACTTG | AAAAGAAGGAAAAATCTCAT |
|  |  |  | 2 | 47,953,568 | (AT)19 | 357 | ACATATGCTTAAAAGGTTGA | TGTCTGTTCATAACCCTTAC |
|  |  |  | 3 | 48,044,700 | (TCA)6 | 278 | AGCTAATTAGTGAAGCAAAA | TATCCACTCAGAAGAAACAC |
| Glutathione S-transferase T3 | 9 | 18,861,741 | 1 | 18,806,940 | (TCC)10 | 354 | CGGTCAGTTGATTAAAATAG | CAAGAGTGTAACCAGAAAAC |
|  |  |  | 2 | 18,890,697 | (TAT)5 | 374 | AACGTTAACACCTTGTAAGA | GGAAAGACATGAATTAGTGA |
|  |  |  | 3 | 18,903,356 | (TCT)7 | 300 | GGTCACTTTAGATGTACCAA | CGTAACCATATCCACTTAAC |
| Protein disulfide isomerase like protein | 9 | 18,779,237 | 1 | 18,806,940 | (TCC)10 | 354 | CGGTCAGTTGATTAAAATAG | CAAGAGTGTAACCAGAAAAC |
|  |  |  | 2 | 18,763,461 | (GTT)6 | 365 | AAACAAGATTTAGACGACAA | ACTAAAAATTGTTGAAATCG |
|  |  |  | 3 | 18,728,556 | (TCA)5 | 317 | ACAAATCAAACTTTGGAATA | CTCGTGTTGATAGAGAAAAG |
| S-adenosyl-L-methionine-dependent methyltransferases superfamily protein | 9 | 18,779,237 | 1 | 18,806,940 | (TCC)10 | 354 | CGGTCAGTTGATTAAAATAG | CAAGAGTGTAACCAGAAAAC |
|  |  |  | 2 | 18,763,461 | (GTT)6 | 365 | AAACAAGATTTAGACGACAA | ACTAAAAATTGTTGAAATCG |
|  |  |  | 3 | 18,728,556 | (TCA)5 | 317 | ACAAATCAAACTTTGGAATA | CTCGTGTTGATAGAGAAAAG |
| O-acetylserine (thiol) lyase; cysteine synthase | 9 | 17,786,685 | 1 | 17,630,637 | (TTA)12 | 236 | TCTTTAATCATTGGCTACAT | GATGAATTTATTGTGCTTTC |
|  |  |  | 2 | 17,769,727 | (TTC)8 | 138 | TAAAATCAAAGAAACCAAAG | CTCTGAGAGATTCAAATACG |
|  |  |  | 3 | 17,818,692 | (AAT)5 | 336 | TTATTAAACACTTCGGTTTC | CACTCATCTCACTCTTTTGT |
| APS Reductase | 9 | 19,517,334 | 1 | 19,422,241 | (AAT)12 | 371 | ATATCTGGTGTTGATAGTCG | ATAGTAATCCGATCAATGTG |
|  |  |  | 2 | 19,478,107 | (AAT)5 | 155 | CAAAAGCTAAATATCTGGAA | CCTTATGCTATCGTATTGAC |
|  |  |  | 3 | 19,592,220 | (TTC)7 | 350 | ATGATGATTATGTTTCTTCG | TACATAGTTGCTTTCATCCT |

**Acknowledgements:** Thank you to Dr. Allan Brown, Dr. Charles David, and NC State for collaborating with me on this project. Also, thank you to my Genomics Lab Class at Davidson College for organizing this research and to Dr. Malcolm Campbell for providing guidance and advice through this whole process.

**References:**

1. Bourgis, F., Roje, S., Nuccio, M.L. Fisher, D.B., Tarczynski, M.C., Li, C., Herschbach, C., Rennenberg, H., Pementa, M.J., Shen, T.L., Gage, D.A., Hanson, A.D., 1999. S-methylmethionine plays a major role in phloem sulfur transport and is synthesized by a novel type of methyltransferase. Plant Cell 11:1485-98.
2. Buchner, P., Stuiver, C., Westerman, S., Wirtz., Hell, R., Hawkesford, M., and De Kok, L. 2004. Regulation of Sulfate Uptake and Expression of Sulfate Transporter Genes in Brassica oleracea as Affected by Atmospheric H2S and Pedospheric Sulfate Nutrition. Plant Physiology136:3396-3408.
3. European Bioinformatics Institute, SIB Swiss Institute of Bioinformatics, and Protein Information Resource. The Universal Protein Resource (UniProt). <http://www.uniprot.org/> <accessed May, 2014.>
4. Grossman, A., and Takahashi, H. 2001. Macronutrient utilization by photosynthetic eukaryotes and the fabric of interactions. Annu. Rev. Plant Physiol. Plant Mol. Biol. 52:163-210.
5. Hawkesford, M. J. 2000. Plant responses to sulphur deficiency and the genetic manipulation of sulphate transporters to improve S-utilization efficiency. Journal of Experimental Botany. 51:131-138.
6. Hayes, J.D., Pulford, D.J. 1995. The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer and chemoprotection and drug resistance. Crit. Rev. Biochem. Mol. Bio. 30:445-600
7. Koprivova, A., Giovannetti, M., Baraniecka, P., Lee, B., Grondin, C., Loudet, O., and Kopriva, S. 2013. Natural Variation in the ATPS1 Isoform of ATP Sulfurylase Contributes to the Control of Sulfate Levels in Arabidopsis. Plant Physiology 163:1133-1141.
8. Leustek, T. 2002. Sulfate Metabolism. The Arabidopsis Book.
9. Leustek, T., Martin, M.N., Bick, J., and Davies, J.P. 2000. Pathways and Regulation of Sulfur Metabolism Revealed through Molecular and Genetic Studies. Annual Review of Plant Physiology and Plant Molecular Biology 51:141-165.
10. Maruyama-Nakashita, A., Nakamura, Y., Yamaya, T., Takahashi, H. 2004. A novel regulatory pathway of sulfate uptake in *Arabidopsis* roots: implication of CRE1/WOL/AHK4-mediated cytokinin-dependent regulation. The Plant Journal 38:779-789
11. Nikiforova, V., Freitag, J., Kempa, S., Adamik, M., Hesse, H., and Hoefgen, R. 2003. Transcriptome analysis of sulfur depletion in *arabidopsis thaliana:*  interlacing of biosyntheic pathways provides response specificity. The Plant Journal 33:633-650.
12. Nimni, M., Han, B., and Cordoba F. 2007. Are we getting enough sulfur in our diet? Nutrition & Metabolism 4:24.
13. Phoenix Bioinformatics Corporation. The Arabidopsis Information Resource (TAIR). <http://www.arabidopsis.org/index.jsp> <accessed May, 2014.>
14. Smith, F.W., Rae, A.L., and Hawkesford, M.J. 2000. Molecular mechanisms of phosphate and sulphate transport in plants. Biochimica et Biophysica Acta 1465:236-245.
15. Takahashi, H., Yamizaki, M., Sasakuru, N., Watanabe, A., Leustek, T., Engler, J., Engler, G., Montagu, M.V., Saito, K. 1997. Regulation of sulfur assimilation in higher plants; a sulfate transporter induced in sulfate-starved roots plays a central role in *Arabidopsis thaliana*. Plant Biology 94:11102-11107.
16. US National Library of Medicine (NLM). National Center for Biotechnology Information (NCBI). <http://www.ncbi.nlm.nih.gov/> <accessed February-May, 2014.>