**Candidate Genes Involved in Iron Uptake and Storage in *B. oleracea***

Emily Keator**\*** and Phoebe Parrish**\***

**Introduction**

The element iron is essential for life. In humans, one of its most important functions is in hemoglobin, which brings oxygen to our cells; in plants, it functions in the electron transport chain of photosynthesis, through which plants produce energy. However, as Grotz *et al.* (2006) noted, plants have difficulty obtaining iron due to its poor solubility under aerobic conditions. Though the earth’s crust is rich in iron, one-third of the world’s soil is iron-deficient, meaning that iron is a limiting nutrient. Because iron is limited in supply for primary producers, moving up the food chain, we see that iron deficiency is the most common nutritional disorder in humans. NSCU researchers hope to increase the concentration of iron in *Brassica oleracea*, broccoli. As part of the broccoli breeding project, the researchers are mapping the *B. oleracea* genome. Previous research at NCSU determined that certain randomly screened plants showed an increased iron concentration in the broccoli florets. We researched proteins in a model organism that might also be relevant to iron uptake and storage in broccoli. We then obtained their nucleotide sequences, and searched for highly similar sequences in the *B. oleracea* genome, as well as recording existing annotations in the *B. oleracea* genome for genes related to iron uptake and storage. Only those close to quantitative trait loci (QTLs) that had been previously determined by NCSU researchers were considered strong candidates for causing the iron increase in some plants. We also selected SSR (single sequence repeat) primers near each final candidate gene to allow for identification.

Generally, non-graminaceous (non-grass) plants such as *Arabidopsis thaliana* and broccoli use what is known as “Strategy I” iron uptake (Fig. 1); grass plants use “Stratgey II.” Strategy I iron uptake involves the reduction of ferric chelates in the soil, so that iron can be taken up into the cells as Fe2+­­ (Kobayashi and Nishizawa, 2012). Schmid *et al* (2014)*.* described how AHA, an H+-ATPase, pumps H+ ions across the membranes of root cells into the soil. With the pH of the soil dropping, iron becomes more soluble, and ferric-chelate reductases (FRO1/FRO2) transport electrons into the soil so that Fe3+-chelate complexes can be reduced to Fe2+ and chelate, thus freeing iron in a form that can be absorbed by the plant. IRT1, a transport protein expressed in the outer cells of roots and localized to the plasma membrane, brings Fe2+ ions into the cell.



**Figure 1:** Strategy I iron uptake from the soil. Note the presence of AHA (green), FRO2 (red), and IRT1 (blue). Image courtesy Hell et. al, 2003.

Once iron has been absorbed into the cells, a protein called FRD3 transports citrate in. Citrate chelates iron inside of the plant, which in turn allows iron to be transferred through the xylem (Kobayashi and Nishizawa, 2012). In the phloem, YSL1 and YSL2 proteins are responsible for transporting complexes of iron and nicotianimine (a chelate) into other plant cells (Grotz *et al.*, 2006). Once iron is in the cell, it may be used for a variety of purposes such as building a variety of proteins, forming the iron-sulfur complex used in photosynthesis, and making heme, which is destined for the electron transport chains of photosynthesis and cellular respiration (Fig. 2). Iron is a highly reactive element that can form reactive oxygen species and become highly toxic to the plant cells. Thus, excess iron is sequestered in vacuoles, plastids (especially chloroplasts and amyloplasts), and mitochondria, or stored in ferritin (Hell *et al*., 2003).



**Figure 2:** Potential pathways for iron once it is inside the cell. Note that it is complexed with nicotianimine (NA) before moving along these pathways, which include storage in ferritin and other organelles. Image courtesy Hell et. al., 2003.

Ferritin (FER1) is a protein that binds up to 4500 iron ions in a readily-usable form. FER1 is usually expressed only in conditions of excess iron, and tends to accumulate in plasmids and vacuoles (Briat *et al*., 1999). Additional proteins move iron into the vacuoles and organelles. PIC1 transports iron into the chloroplast, MIT into the mitochondria, and IREG1 into vacuoles (Kobayashi and Nishizawa, 2012). VIT1 is another vacuolar membrane iron transporter, which localized to vacuolar membranes and brings Fe2+ ions into the vacuoles. NRAMP3 and 4 bring iron back out of the vacuole and resist the actions of VIT1 (Roschzttardtz *et al*., 2013)

 In addition to proteins that directly interact with iron ions for uptake and storage, there are several transcription factors that regulate iron uptake and storage proteins to prevent iron starvation or excess iron accumulation. These transcription factors (FIT1, bHLH38,39,100,101) upregulate IRT1 and FRO2 when iron is low in the soil. FIT1 forms heterodimers with bHLH (basic helix-loop-helix) 38 and 39 to control the transcription of IRT1 and FRO2 (Kobayashi and Nishizawa, 2012; Wang *et al*., 2013). Our research focused on finding iron-related genes close to the QTL’s provided to us (Table 1).

**Methods**

 We searched the existing literature on *Arabidopsis thaliana*, to identify genes involved in iron uptake and storage. We developed a list of candidate genes, obtained their FASTA protein sequences at NCBI (National Library of Medicine, 2014; Table 2). Candidate proteins for which a sequence could not be found were omitted at this stage. We ran a tBLASTn on each sequence against the *Brassica oleracea* genome (“B. oleracea genome v1”) using the Genome Database for Vaccinium (NCSU, 2014). We calculated the distance of the candidate gene from the QTL, and moved forward with those that were within 10 million base pairs (10 Mbp).

**Table 1:** QTL’s linked to iron accumulation in broccoli florets

|  |
| --- |
| ***B. oleracea* Iron QTLs****Name Position** |
| C01 | 1,923,307 |
| C03 | 13,636,922 |
| C05 | 43,302,017 |
| C08 | 33,379,515 |
| C09-1 | 3,558,943 |
| C09-2 | 45,002,150 |

|  |
| --- |
| **Table 2: Candidate Genes and Accession Numbers** |
| **Process** | **Gene** | **Accession #** |
| *Uptake* | *AHA1* | NP\_179486.1 |
|  | *AHA2* | NP\_001190870.1 |
|  | *AHA9* | NP\_001185450.1 |
|  | *FRO1* | AAE27309 |
|  | *FRO2* | AEE27308 |
|  | *IRT1* | NP\_567590.3 |
|  |  |  |
| *Intercellular Transport* | *FRD3* | NP\_001154595.1 |
|  | *FPN1/IREG1* | AEC09539.1 |
|  | *YSL1* | NP\_567694.2 |
|  | *YSL2* | NP\_197826.2 |
|  |  |  |
| *Storage* | *FER1* | NP\_195780 |
|  | *VIT1* | AEC05496 |
|  | *IREG1* | AEC09539 |
|  | *PIC1* | AEC06383 |
|  |  |  |
| *Regulation* | *FIT1* | AEC08086 |
|  | *bHLH38* | Q9M1K1.1 |
|  | *bHLH39* | Q9M1K0.1 |
|  | *bHLH100* | NP\_850349.1 |
|  | *bHLH101* | NP\_196035.1 |
|  |  |  |

**Table 2:** Candidate gene and accession number for FASTA sequences (National Library of Medicine, 2014)

**Table 2:** Candidate gene list developed through research and accession numbers used to obtain FASTA sequences, courtesy of National Library of Medicine, 2014.

We used the genome browser IGB (UNC Charlotte, 2014) to search for the genes we had found on BLAST. We also scanned 1 Mbp to either side for additional candidate genes that were not identified in BLAST. We BLASTed all sequences of interest against the *A. thaliana* genome, and followed links for the best hits to TAIR (2014), where we obtained further information on their expression and function.

 We identified SSR primers for all candidate genes. We manually searched the file of SSR primers. To be considered a valid SSR primer pair for the gene, the primers had to be within 100 kb of the gene’s start. For each SSR, we recorded the forward and reverse sequence, the size, the location, and the SSR Primer ID assigned by NCSU.

**Results**

 BLASTing our candidate genes, we found many hits within 10 million bp of our QTL’s. All hits recorded had E-values of less than 1 x 10-6. On chromosome 1, YSL1,2, FIT1, AHA1,2,9, and IRT1 were within 10 Mbp. On chromosome 3, FRD3, FIT1, and AHA1,2,9 were within 10Mbp. FPN1/IREG1, FIT1, IRT1, FER1, FRD3, and AHA1,2,9 were within 10 Mbp on chromosome 5. On chromosome 8, FER1, bHLH38&39, bHLH100&101, AHA1,2,9, IRT1, and FRO1,2 were all within 10 Mbp. Chromosome 9 had 2 QTLs. Within 10 Mbp of C09-1 were FRO1, AHA1,2,9, YSL1,2, and PIC1, while YSL1,2, AHA1,2,9, and IRT1 were within 10 Mbp of C09-2.

 When we looked for our BLAST results using IGB, the annotation software confirmed the location of every BLAST result, with a few exceptions. The BLAST hit for YSL1 on chromosome 1 was not annotated on IGB. However, the BLAST hit for YSL 1 and YSL2 were overlapping, so it is likely that IGB called the gene YSL2 instead. On chromosome 5, IGB did not annotate FPN1/IREG1 or FIT1, though “hypothetical” or “conserved hypothetical” proteins were listed at these points. For QTL C09-2, we had a BLAST hit for YSL1 nearby. On IGB, a gene labeled in this location was labeled “Gb|AAF18539.1.”

 In addition to confirming most of our BLAST results, we scanned 1 million bp to the left and right of the QTLs on IGB in search of annotated genes related to iron that may have been missed in our research and BLAST. Several genes, though they lacked specific names, coded for a protein class that could be related to iron uptake and storage. For example, we found three heavy metal/detoxification superfamily proteins, near QTL’s C01, C05, and C09-2. These three proteins function in metal ion binding and transport. Each of the three hits localized to different membranes within the cell (Phoenix Bioinformatics Corp., 2014). On chromosome 3, we found MdTk, a multidrug resistance protein, that functions in ion binding and transport. On chromosome 5, we had two hits for MATE efflux family proteins, homologs of FRD3. Additionally, near QTL’s C08 and C09-1 we found the gene for ferrocheletase, the final enzyme in the heme biosynthesis pathway (Phoenix Bionformatics Corp., 2014). Figure 3 depicts the confirmed BLAST results and any new genes found on IGB within the given distances of the six QTLs.



**Figure 3:** *B. oleracea* chromosomes 1, 3, 5, 8, and 9 depicted with their QTLs. Genes listed in blue were discovered on IGB. Those in black were researched, BLASTed, and confirmed on IGB. Note the scale of the chromosomes changes for each QTL, depending on whether a valid hit existed outside of 1 million bp.

According to TAIR (Phoenix Bioinformatics Corp., 2014), our final candidate genes are expressed mainly during flower differentiation and expansion stages, or in pollen, anthers, or petals. Most of the proteins localize to various cell membranes. The TAIR results for the proteins expressed by our final candidate genes, along with the SSR primer IDs, are summarized in Table 3. Detailed primer information is given in Table 4.

**Table 3:** Summary of final candidate proteins. Data on protein function, localization, and expression courtesy of Phoenix Bioinformatics Corp., 2014. SSR Primer ID numbers are those assigned by NCSU researchers.

|  |
| --- |
| **Table 3: Final Candidate Genes** |
|  | **Protein** | **Function** | **Localized** | **Expression** | **SSR Primer ID** |
| **C01** | Heavy metal transport/detoxification superfamily protein | Metal ion transport and binding | Nucleus, plasmodesma | Petal differentiation and expansion | 366 (C01) |
| **C01** | ZIP9 [IRT1 on BLAST] | IRT3 homolog; transports Zn, Fe, other cations | Nucleus, plasmodesma | Petal differentiation and expansion | 530 (C01) |
| **C03** | FRD3  | Iron-citrate transport; MATE efflux family protein | Plasma membrane | Petal differentiation and expansion | 1906 (C03) |
| **C03** | MdTk | Ion transporter/ antiporter | Chloroplast, plasma membrane | Petal differentiation and expansion | 1906 (C03) |
| **C05** | IRT1 | Fe ion transport, responds to Fe starvation | Membrane | Petal differentiation and expansion | 4918 (C05) |
| **C05** | FER1 | Fe ion homeostasis and ferric iron binding | Plasma membrane | Petal differentiation and expansion | 5086 (C05) |
| **C05** | Heavy metal transport/detoxification superfamily protein | Metal ion transport and binding | Plasma membrane | Pollen, petal, petiole | 5123 (C05) |
| **C05** | FRD3 | Iron-citrate transport; MATE efflux family protein | Plasma membrane | Roots, flower, petiole | 5230 (C05) |
| **C05** | FRD3 | Iron-citrate transport; MATE efflux family protein | Plasma membrane | Roots, flower, petiole | 5245 (C05) |
| **C08** | AHA | H+-ATPase; ion transport | Plasma membrane, Golgi, vacuolar membrane | Petal differentiation and expansion | 3918 (C08) |
| **C08** | Ferrochelatase | Terminal enzyme of heme biosynthesis | Chloroplast | Petal differentiation and expansion | 4138 (C08) |
| **C09-1** | Ferrochelatase | Terminal enzyme of heme biosynthesis | Chloroplast | Petal differentiation and expansion | 581 (C09) |
| **C09-2** | MATE efflux family protein | Ion transport | Unknown | Unknown | 5544 (C09) |
| **C09-2** | Heavy metal transport/detoxification superfamily protein | Exact function unknown | Extracellular space (unconfirmed) | Petal differentiation and expansion | 5603 (C09) |

**Table 4:** SSR primer information, including the start nucleotide of the SSR sequence. SSR Primer ID numbers are those assigned by NCSU researchers.

|  |
| --- |
| **Table 4: SSR Primers** |
|  | **Protein** | **Forward Seq** | **Reverse Seq** | **Size** | **Location** **(Start of repeat)** | **SSR Primer ID** |
| **C01** | Heavy metal transport/detox.  | TCTTGATCATTACTCCTTTT  | TCTTGATCATTACTCCTTTT  | 349 | 2113720 | 366 (C01) |
| **C01** | ZIP9 [IRT1 on BLAST] | CATTCATTCTTACACAAGGT  | ACAATCAATACGTTTTAGGA  | 391 | 3006991 | 530 (C01) |
| **C03** | FRD3  | TCTTGTGTCATAAACTTCTGT  | AGGTCTCTAGAACCCTAAAA  | 400 | 12414633 | 1906 (C03) |
| **C03** | MdTk | TCTTGTGTCATAAACTTCTGT  | AGGTCTCTAGAACCCTAAAA  | 400 | 12414633 | 1906 (C03) |
| **C05** | IRT1 | ATAAATAATGGACAATGCAC  | GCTTTCTCTCTCTTTCTCTC  | 271 | 41531310 | 4918 (C05) |
| **C05** | FER1 | TTAAAGCAGTGGAGTTTTAC  | CTCTTGATGTGTTCTGATTT  | 243 | 42530670 | 5086 (C05) |
| **C05** | Heavy metal transport/detox.  | CAAAAACGACTCTACTCAAC  | CAAAAAGGTAATCATCTCTG  | 228 | 42725498 | 5123 (C05) |
| **C05** | FRD3 | AATTTTAGATGCAACTTTTG  | AGTGCAGACTCGAAATATAA  | 353 | 43547983 | 5230 (C05) |
| **C05** | FRD3 | TTTTCTCAGTGAAACAAAAT  | CCCACCTATCTCTCTCTAAT  | 310 | 43620523 | 5245 (C05) |
| **C08** | AHA | GGGATAGAATTAGAGTTGGT  | TAATTTGGTAAAAAGACAGC  | 397 | 32391138 | 3918 (C08) |
| **C08** | Ferrochelatase | GAGTTACCATTAGAGCATGA  | TAACGGAAACTATCCTAAAA  | 119 | 33824048 | 4138 (C08) |
| **C09-1** | Ferrochelatase | CTTTTGAGTTCTGACTTCTG  | TTCTTTTGTTCATCCTCTAA  | 312 | 3507903 | 581 (C09) |
| **C09-2** | MATE efflux  | AATTGTGGATGGTAATAGTG  | ATCTTTTATCTTCGTCCTTT  | 307 | 45564249 | 5544 (C09) |
| **C09-2** | Heavy metal transport/detox.  | CCTCCTCCTTTATTATTACC  | ACATGATGATGCCTAATAAC  | 337 | 45926388 | 5603 (C09) |

**Discussion**

 In our research, we began with a broad list of genes involved in uptake and storage of iron that could be responsible for the floret’s increase in iron. Careful BLASTing, followed by IGB scanning, led us to eliminate some of our initial genes, and to include some new ones. Many of the genes that we researched and BLASTed were also on IGB, a secondary confirmation of their location near the QTLs.

 Though we researched many genes, only a few proved to be within +/- 1 million bp of the QTLs. Therefore, we included several researched and BLASTed genes that were over 1 Mbp away, but which we felt were valid candidates, such as IRT1 (1,764,219 bp away on C05; see Fig. 3C).

Of our candidate genes, FRD3, which is responsible for iron transport, could cause increased iron floret concentration by moving more iron-citrate complexes. Similarly, AHA (on C08; see Fig. 3D) normally secretes H+ ions into the soil to facilitate iron uptake, and increased presence or activity of the protein could increase iron uptake. Finally, FER1 we confirmed close to QTL C05 (see Fig. 3C). Since ferritin is responsible for storing iron ions, and can store a large number of them, it is a strong candidate for increasing iron concentration.

Candidate genes that we identified uniquely on IGB tended to be more broad-category genes there are not iron-specific. For example, MdTk is an ion transporter, and the heavy metal transport/detoxification superfamily proteins transport metal ions. These may not be exclusively expressed to transport iron, but each could still be responsible for the increase in iron. Ferrochelatase is one of the most speculative candidates on our final list. Discovered on IGB, it is the terminal protein in heme biosynthesis and thus interacts with iron. However, whether it could cause an increase in iron concentration is unclear.

Some of the most promising evidence for our candidate genes comes from the TAIR data on their expression. Every final candidate gene in our list is expressed either in structures or at stages related to the florets, where NCSU researchers noted increased iron. Several of our candidates were expressed during the petal differentiation and expansion stage, including FER1, IRT1, ZIP9, FRD3, MdTk, AHA, Ferrochelatase, and heavy metal transport/detoxification superfamily proteins. Similarly, FRD3 and heavy metal transport/detoxification superfamily proteins are expressed in the petiole, flower, and/or petals. Since the florets are the precursors of the flower, this information further supports the candidacy of these genes.

While this paper presents a list of candidate genes for an increase in broccoli floret iron accumulation, it is important to remember that the increase could be due to more than one gene. A central tenet of genomics is that everything is interconnected, and that effects are best studied in context rather than in isolation. Therefore, it could be that several of these candidate genes are working together, such that excess IRT1 is brings more iron into the plant and increased amounts of FER1 allow that iron to be stored. Consequently, it will be necessary to test the genes in concert with each other. Furthermore, considering that many of our candidate proteins are non-discriminatory ion transporters, the increase in iron could be related to the increase in other minerals that NCSU researchers have also noted. Indeed, in a separate but related study, researchers looking into the uptake and storage of zinc found similar genes and even shared a QTL with this experiment (QTL C09-1). Since IRT1 transports both Fe and Zn, that explains identical QTL mapping. Thus, as research moves forward, it will be important to test the final candidate genes laid out in this study with each other, and with genes related to increases of other minerals.

**Acknowledgements**

We would like to thank Dr. Malcolm Campbell and Dr. Charles David for their guidance and supervision. We would also like to thank Dr. Allan Brown and his team at NCSU for the opportunity to participate in this research and their support along the way.

**References**

Briat, JF *et al*. “Regulation of plant ferritin synthesis: how and why.” *Cellular and Molecular Life Sciences*, 1999. 56:155-166.

Grotz, N *et al*. “Molecular aspects of Cu, Fe, and Zn homeostasis in plants.” *Biochemica et Biophysica Acta*, 2006. 1763:595-608.

Hell, R *et al*. “Iron uptake, trafficking, and homeostasis in plants.” *Planta*, 2003. 216:541-551.

Phoenix Bioinformatics Corporation. *The Arabidopsis Information Resource (TAIR)*. 2014. Accessed April 8-15, 2014. <http://www.arabidopsis.org/index.jsp>.

National Library of Medicine. *NCBI Basic Local Alignment Search Tool (BLAST)*. 2014. Accessed Feb. 4-13, 2014. <http://blast.st-va.ncbi.nlm.nih.gov/Blast.cgi>.

NCSU. *Genome Database for Vaccinium*. 2014. Accessed Feb. 11-25, 2014. <http://dev.vaccinium.org>.

Roschzttardtz, H *et al*. “New insights into Fe localization in plant tissues.” *Frontiers In Plant Science*, 2013. 4:1-11.

Schmid, N *et al*. “Feruloyl-CoA 6’-Hydroxylase1-Dependent Coumarins Mediate Iron Acquisition from Alkaline Substrates in Arabidopsis.” *Plant Physiology*, 2014. 164:160-172.

UNC Charlotte. *Integrated Genome Browser (IGB)*. 2014. Accessed March 18-April 15, 2014. <http://bioviz.org/igb/>.

Wang, N *et al*. “Requirement and functional redundancy of Ib subgroup bHLH proteins for iron deficiency responses and uptake in Arabidopsis thaliana.” *Mol Plant*. 2013 Mar; 6(2):503-13.