Increasing Iron Uptake in *Brassica oleracea*

By Phoebe Parrish and Emily Keator

**Abstract**

 Iron is an important micronutrient for both humans and plants. Therefore increasing iron uptake in foods such as broccoli (*Brassica oleracea*) is an important step towards improving human nutrition worldwide. To work towards this goal, we compared iron uptake, storage and transport genes that have been characterized in the scientific literature to orthologs in the sequenced *B. oleracea* genome. Using tools such as NCBI’s BLAST and the Integrated Genome Browser (IGB), we found a number of genes of interest near quantitative trait loci (QTL) that had been experimentally identified previously. Many of these candidate genes could have contributed to the observed phenotype of increased iron uptake in *B. oleracea*. Using these data, our collaborators can characterize the alleles that may lead to increased iron uptake, transport and storage. In the future, this knowledge will contribute to the improvement of human health globally.

**Introduction**

 For this project, we characterized genes involved in iron uptake, transport and storage in the broccoli species *Brassica oleracea*. Our goal was to help our collaborators at North Carolina State University (NCSU) with their project to improve the nutritional value of broccoli. Based on their previous research, they were able to provide us with quantitative trait loci (QTL) in the broccoli genome that were involved in an increase in concentration of certain nutrients in the florets of broccoli plants. We were assigned iron as our nutrient of interest.

 Increasing iron uptake in plants is an especially important goal, as iron deficiency is the leading nutritional disorder in the world today (Grotz and Guerinot, 2006). Iron is also important for the plants themselves, as it is the micronutrient that they require in the highest levels (Kobayashi and Nishizawa, 2012). There are many reasons for this high iron requirement, including the fact that that iron is a common cofactor for enzymes. It is also highly reactive, and is therefore involved in many redox reactions. For example, iron forms part of the electron transport chain (ETC) in photosynthesis and cellular respiration. This ETC is made up of cytochromes, a type of hemeprotein containing iron in the center of a porphyrin ring. However, iron’s reactivity can also have negative effects on plants. When it is unchelated and present in large amounts, iron can encourage the formation of reactive oxygen species, which have many toxic effects (Kobayashi and Nishizawa, 2012). Therefore iron homeostasis must be tightly controlled.

 Iron uptake, transport and storage have been researched extensively in the model plant *Arabidopsis thaliana*. However, not much is known about these processes in *B. oleracea*, and genes that may contribute to excess iron accumulation in broccoli have yet to be definitively identified and located. Therefore we decided to compare *A. thaliana* genes of interest obtained from the literature to the *B. oleracea* genome. This comparison is valid because *B. oleracea* and *A. thaliana* are quite closely related, with both species forming part of the Brassicaceae family.

*Uptake:* Both *A. thaliana* and *B. oleracea* are non-graminaceous, meaning they are not part of the grass family. Non-graminaceous plants use strategy I iron uptake. A summary of this method can be seen in Figure 1 below.



**Figure 1:** A brief comparison of the genes and mechanisms involved in iron uptake in strategies I and II (from Kobayashi and Nishizawa, 2012). Many of these processes will be discussed in more detail later.

 Strategy I iron uptake involves three major gene families. The H+-ATPase family contains the AHA genes, with *AHA1*, *AHA2*, and *AHA9* being the members most closely involved in iron uptake (Harper *et al.*,1990; Sussman, 1994; Schmid *et al.*, 2014). These genes code for proteins that extrude hydrogen ions into the soil. Fe3+ ions (the form in which iron is most commonly found in the soil) are more soluble in an acidic environment; therefore the AHA genes “facilitate the dissolution of Fe3+” by lowering the pH of the environment around the plant’s roots (Schmid *et al.*,2014).

 The next step in this process involves the ferric-chelate reductase oxidase family, specifically *FRO2*. This gene codes for a protein that reduces chelated Fe3+ ions to Fe2+, separating these ions from the chelate (Schmid *et al.*, 2014). This reduction step is necessary for iron to be taken up by the proteins of the ZIP family, especially *IRT1* and *IRT2*.

 *IRT1* is the major cation uptake gene. It encodes for a transmembrane protein that takes up many heavy metal cations, including iron, zinc and copper. It is responsible for the majority of iron uptake in soils that lack iron. It is expressed in external cell layers of the root to uptake iron from the soil, as well as in the flowers, to provide iron to maturing pollen grains (Vert *et al.*, 2002). These three uptake genes work together as shown in Figure 2 below.



**Figure 2:** Shows the three major genes involved in the uptake of iron from the soil (from Hell and Stephan, 2003).

*Intercellular Transport:* Once iron has entered the roots, is moved through the plant in two major ways: efflux transport through the xylem, and influx transport through the phloem (Kobayashi and Nishizawa 2012). In both cases, it is often complexed with chelates such as citrate and nicotianamine to increase solubility and reduce reactivity (Grotz and Guerinot, 2006).

 An important gene involved in efflux transport is *FRD3*, which is part of the MATE (Multidrug and Toxin Efflux) family of proteins. *FRD3* codes for a protein that facilitates the transport of citrate-chelated iron molecules into the xylem (Kobayashi and Nishizawa 2012). This provides the shoots with a usable form of iron (Grotz and Guerinot, 2006). A gene known as both *FPN1* and *IREG1* is also involved in efflux transport. It is a ortholog of a mammal efflux transporter; however not much else is known about it (Kobayashi and Nishizawa, 2012).

 The yellow stripe-like (YSL) family of proteins is a major player in influx transport. *YSL1* and *YSL2* are two of the important genes that have been identified in the literature. These genes transport iron-nicotianamine complexes through the phloem (Grotz and Guerinot, 2006; Vert *et al.*, 2002).

*Intracellular Transport/Storage:* These two categories have been grouped together because iron storage is often tightly controlled by various intracellular transport proteins. Although we were unable to find many of these genes close to our QTLs, the ones we did find are significant. Localization of iron within cells must be tightly controlled due to its high reactivity and involvement in the formation of reactive oxygen species, which can have toxic effects (Galaris and Pantopoulos 2008). Therefore storage proteins play an important role in iron homeostasis.

 *PIC1*, or permease in chloroplasts, is involved in transporting iron into and out of the chloroplasts. This regulates compartmentalization of iron within the cells (Kobayashi and Nishizawa, 2012). As mentioned earlier, this is important because iron forms part of the electron transport chain for cellular respiration.

 Ferritin is a stored form of iron. The *FER* genes, especially *FER1*, store iron in a readily usable form by complexing up to 4,500 iron ions together at once (Briat *et al.*, 1999). Ferritin is typically located in the amyloplasts and vacuoles (Roschzttardtz *et al.*, 2013), both of which are typical storage areas within the cell.

*Regulation:* There are two major regulatory gene families that control iron uptake. The first is the FER-like iron deficiency induced transcription factors, especially *FIT1*. *FIT1* is root-specific (Grotz and Guerinot, 2006). It forms heterodimers with members of the basic helix-loop-helix transcription factor family such as *bHLH38*, *bHLH39*, *bHLH100* and *bHLH101* to regulate iron uptake in case of deficiency (Kobayashi and Nishizawa, 2012). These genes code for transcription factors that activate the expression of iron uptake genes such as *IRT1* and *FRO2*. The functions of all of the genes described above are summarized in Table 1 on the next page.

**Table 1:** This table provides the category, name, and function of genes involved in iron uptake, transport, and storage in *A. thaliana* that were later also found in the *B. oleracea* genome through BLAST and IGB methods.

|  |  |  |  |
| --- | --- | --- | --- |
| **Category** | **Family Name** | **Gene Names** | **Function** |
| Uptake | H+-ATPase | *AHA1,2,9* | Release H+ ions into the soil, making iron more soluble |
| Ferric-chelate reductase oxidase  | *FRO2* | Reduces chelated Fe3+ → Fe2+ + chelate  |
| Iron-regulated transporter  | *IRT1* | Transports Fe2+ (and other divalent cations) |
| Intercellular Transport | Yellow stripe-like family | *YSL1,2* | Faciliates iron-NA influx into phloem |
| Ferric reductase defective  | *FRD3* | Member of the Multidrug and Toxin Efflux (MATE) family; facilitates iron-citrate efflux into xylem  |
| Iron regulated protein  | *FPN1/IREG1* | Involved in iron efflux into xylem  |
| Intracellular Transport/Storage | Ferritin  | *FER1* | Complexes iron in a readily usable form, up to 4500 molecules at once!  |
| Permease in chloroplasts  | *PIC1* | Regulates Fe storage in chloroplasts  |
| Regulation | FER-like iron deficiency induced transcription factor | *FIT1* | Positively regulates Fe uptake genes such as IRT1 and FRO2; induced when Fe deficiency is sensed in the roots. Forms heterodimers with bHLH-family proteins to perform its functions  |
| Basic helix-loop-helix transcription factors  | *bHLH38,39,100,101* | Positively regulate Fe uptake genes |

**Methods**

 Our research took place in 4 main steps: (1) we did background research on our pathway and possible genes of interest in the scientific literature. (2) We then used BLAST to compare our genes of interest to the broccoli genome, and (3) we later identified our BLAST results on the Integrated Genome Browser (IGB, 2014) and also searched +/- 1 Mbp on either side of the QTL to find any other potential genes of interest. After we had identified our best candidate genes, (4) we identified SSR primers that would allow researchers at NCSU to mark genes of interest and track the effects of different alleles on a plant’s phenotype.

(1) *Research:* We started out by searching through the scientific literature (mostly by way of NCBI, Google Scholar, and the Davidson College Library website) to find genes involved in uptake, transport and storage of iron in *A. thaliana*, since this species is a model plant for scientific research and a relative of *B. oleracea*. We sorted our genes into four categories, based on their functions in *A. thaliana*: uptake, intercellular transport, intracellular transport/storage, and regulation.

(2) *BLAST:* We then took the genes of interest we found in step 2 and searched for their amino acid sequences on NCBI. We used tBLASTn on the Vaccinium website (Vaccinium, 2014) to find similar sequences in the *B. oleracea* genome. We recorded hits with e-values of less than 1x10-6 that were located within 10 Mbp on either side of our QTLs. We recorded these results as our original genes of interest.

(3) *IGB:* We used the locations of our genes of interest from BLAST to identify them in the *B. oleracea* genome on IGB. We also used the browser function on IGB to identify new genes of interest that were located within an approximate range of 1 Mbp on either side of each of our QTLs. Occasionally we expanded this range to include genes that looked like good candidates but were slightly farther away. We took the +1 reading frame amino acid sequence for these genes from IGB and used the BLASTp program on NCBI (2014) to compare these sequences to the *A. thaliana* genome. We were then able to research each specific gene on TAIR (2014) to confirm its location, function, and expression in the plant.

(4) *SSRs:* We received a file containing all of the SSRs for the *B. oleracea* genome. We chose primers that would allow the NCSU researchers to use PCR to isolate areas of the genome containing our genes of interest. We selected these primers based on a +/- 100,000 base pair range on either side of the QTL. Some subjective criteria were also applied, as we were also attempting to increase the number of repeats per primer. However mainly we tried to find the SSR closest to our gene of interest.

**Results**

 Our results from BLAST revealed that many of the genes we found in the literature are very likely present in the *B. oleracea* genome. Many of them were within 10 Mbp of our QTLs as well. We confirmed many of these BLAST results on IGB, but limited our IGB search to approximately 1 Mbp on either side of each QTL. Only a few of our previously identified BLAST genes of interest were within this narrower range. However, as seen in Figure 4, we were able to find a number of other promising candidate genes based on their IGB descriptions.

 Figure 4a shows two genes of interest on chromosome 1. The heavy metal transport/detoxification superfamily protein functions in metal ion transport and binding. *ZIP9* is a homolog of *IRT2*, which has a similar function to *IRT1*. Our genes of interest from Figure 5b were also involved in ion transport. They included at MATE efflux family protein and a multidrug resistance protein. The MATE efflux family protein was identified as FRD3 through the BLAST/TAIR method, and the multidrug resistance protein was identified as *MdtK* through the same method. The protein encoded by this gene functions in ion transport.

 We found the most candidate genes on chromosome 5, shown in Figure 4c. The zinc transporter and ferritin were identified as *IRT1* and *FER1*, respectively, through the BLAST/TAIR method on IGB. Both MATE efflux family proteins were also identified as *FRD3* through the same method. The heavy metal transport/detoxification superfamily protein is a homolog of *ATX1*, which functions in metal ion transport and binding.

 The AHA gene of interest shown in Figure 4d was identified through the BLAST method was confirmed through the BLAST/TAIR method on IGB. The ferrochelatase gene (also seen in Figure 4d) encodes for the terminal enzyme in heme biosynthesis. As heme is an important stored form of iron, we thought this would be important to record. This gene was also expressed in the petals. The ferrochelatase in Figure 4e had the same information on TAIR as the one in Figure 4d.

 Neither the MATE efflux family protein nor the heavy metal transport/detoxification superfamily protein in Figure 5f had much information associated with them on TAIR; however both function in ion transport.

 All of the genes we identified through the BLAST/TAIR method on IGB were expressed in the petals, and also during the petal differentiation and expansion stage. This suggests that they could very likely be involved in the increased iron levels in the florets of certain *B. oleracea* plants observed by our colleagues at NCSU.

****

**(a)**

**(b))**

**(c))**

****

**(d))**

****

**(f))**

**(e))**

**Figure 4:** Maps of candidate genes identified through BLAST and IGB. Red letters indicate QTLs, blue letters indicate genes that were only found on IGB, black letters indicate genes that were also found through BLAST, and **^** indicates genes that were identified through the BLAST/TAIR method. **(a)** Results for our QTL on chromosome 1. **(b)** Results for our QTL on chromosome 3. **(c)** Results for our QTL on chromosome 5. **(d)** Results for our QTL on chromosome 8. **(e)** Results for our first QTL on chromosome 9. **(f)** Results for our second QTL on chromosome 9.

 Table 2 shows our SSR results. We were able to find three primer pairs within 100,000 base pairs of each of our genes of interest.

**Table 2:** The results for our SSR primers. The information is based on an Excel file sent to us by our collaborator Dr. Brown at NCSU.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chromosome & Gene** | **ID #** | **Forward sequence** | **Reverse sequence** | **Size** | **Location (start of SSR)** |
| 1: Heavy metal transport/detox | 366 | TCTTGATCATTACTCCTTTT  | TCTTGATCATTACTCCTTTT  | 349 | 2113720 |
| 1: ZIP9 | 530 | CATTCATTCTTACACAAGGT  | ACAATCAATACGTTTTAGGA  | 391 | 3006991 |
| 3: FRD3 | 1906 | TCTTGTGTCATAAACTTCTGT  | AGGTCTCTAGAACCCTAAAA  | 400 | 12414633 |
| 3: Multidrug resistance | 1906 | TCTTGTGTCATAAACTTCTGT  | AGGTCTCTAGAACCCTAAAA  | 400 | 12414633 |
| 5: IRT1 | 4918 | ATAAATAATGGACAATGCAC  | GCTTTCTCTCTCTTTCTCTC  | 271 | 41531310 |
| 5: FER1 | 5086 | TTAAAGCAGTGGAGTTTTAC  | CTCTTGATGTGTTCTGATTT  | 243 | 42530670 |
| 5: Heavy metal transport/detox | 5123 | CAAAAACGACTCTACTCAAC  | CAAAAAGGTAATCATCTCTG  | 228 | 42725498 |
| 5: FRD3(1) | 5230 | AATTTTAGATGCAACTTTTG  | AGTGCAGACTCGAAATATAA  | 353 | 43547983 |
| 5: FRD3(2) | 5245 | TTTTCTCAGTGAAACAAAAT  | CCCACCTATCTCTCTCTAAT  | 310 | 43620523 |
| 8: AHA | 3918 | GGGATAGAATTAGAGTTGGT  | TAATTTGGTAAAAAGACAGC  | 397 | 32391138 |
| 8: Ferrochelatase | 4138 | GAGTTACCATTAGAGCATGA  | TAACGGAAACTATCCTAAAA  | 119 | 33824048 |
| 9-1: Ferrochelatase | 581 | CTTTTGAGTTCTGACTTCTG  | TTCTTTTGTTCATCCTCTAA  | 312 | 3507903 |
| 9-2: MATE efflux | 5544 | AATTGTGGATGGTAATAGTG  | ATCTTTTATCTTCGTCCTTT  | 307 | 45564249 |
| 9-2: Heavy metal transport/ detox | 5603 | CCTCCTCCTTTATTATTACC  | ACATGATGATGCCTAATAAC  | 337 | 45926388 |

 These primers can be used to remove select sections of the genome and analyze their sequences. In this way, researchers from NCSU can identify which alleles of a gene code for specific phenotypes.

**Discussion**

 We were able to find a number of promising candidate genes. Many of the genes of interest present in Figure 4 are especially strong candidates, since they were identified as important in the literature and were identified through both methods 2 and 3. However, we were also able to find multiple genes on IGB (method 3) that had not been identified through method 2, such as the MATE efflux family genes that after further investigation turned out to be *FRD3*. This suggests that it is best to use both approaches in order to maximize the number of candidate genes found.

 Another suggestion I have for future researchers is that they should broaden their window to greater than 1 Mbp on either side of the QTL when searching on IGB. When Dr. Brown visited and listened to our presentation, he noted that many promising candidate genes were actually about 1.2 Mbp base pairs away from the QTL. This suggests that our window for IGB searches could perhaps be widened to 1.5 Mbp.

 For our first two QTLs, it is relatively easy to select the best candidate gene. For chromosome 1, I think the *ZIP9* gene that we identified through both BLAST and IGB methods is the most likely candidate gene. Although this gene is slightly farther than 1 Mbp away from the QTL, it is known to be involved specifically in iron uptake. The same can’t necessarily be said for the heavy metal transport and detoxification superfamily protein. For chromosome 3, I think the *FRD3* gene is our most likely candidate, as it was identified through our research, BLAST and IGB methods (1-3).

 On the other hand, since chromosome 5 had so many hits, it is difficult to say which gene is the best candidate. It seems likely here that more than one gene affected the phenotypic change observed by our collaborators at NCSU for this QTL. However the *IRT1* gene of interest seems like an especially good candidate, partly because it is about 1.2 Mbp away from the QTL. In a personal comment, Dr. Brown said his past work has suggested that this 1.2 Mbp distance applies to many good candidate genes. However, there are a number of other good candidates to keep in mind: both of the *FRD3* genes on chromosome 5 are closest to the QTL, and the *FER1* gene is closer than *IRT1* and was identified through both methods 2 and 3.

 The QTL on chromosome 8 had very few genes of interest overall; therefore it is once again relatively simple to select the best candidate gene. Of the two hits on chromosome 8, I think the *AHA* family gene would be the most likely candidate. It is about 1 Mbp away from the QTL, and was also identified through both methods 2 and 3. The *AHA* family is also definitely involved in iron uptake and transport. Ferrochelatase, on the other hand, was only identified through method 3.

 We also found very few genes of interest for our two QTLs on chromosome 9. The ferrochelatase near the first QTL is the only candidate. I am unable to make a clear judgment call between the two genes of interest near the second QTL, since they both had very little information associated with them on TAIR.

 However, these judgment calls about the best candidate gene are not definitive. Thanks to the information provided in Table 2, researchers at NCSU will be able to determine experimentally which gene (or genes) causes a *B. oleracea* plant to take up excess iron and store it in the florets. Using these primers, scientists in the lab will be able to match alleles with phenotypes, and eventually determine the best combination of alleles that will cause the plant to take up, transport, and store more iron than usual while avoiding any adverse effects. This will bring us another step closer to our goal of increasing the nutrient content of broccoli and improving human health worldwide.

**Acknowledgments**

We would like to thank Dr. Alan Brown and Dr. Charles David from North Carolina State University for their assistance and advice, as well as the data that they shared with us. We would also like to thank Dr. A. Malcolm Campbell at Davidson College for his advice and guidance throughout the project.

**Works Cited**

BLAST. Genome Databse for *Vaccinium*. 2014. [online]. Available from: http://dev.vaccinium.org/tools/blast. Accessed 2014 Mar.

Briat JF, Lobréaux S, Grignon N, Vansuyt G. Regulation of plant ferritin synthesis: how and why. Cellular and Molecular Life Sciences, 1999. 56:155-166.

Harper JF, Manney L, DeWitt ND, Yoo MH, Sussman MR. The *Arabidopsis thaliana* plasma membrane H+-ATPase multigene family. The Journal of Biological Chemistry, 1990. 265(23):13601-13608.

Galaris D, Pantopoulos K. Oxidative stress and iron homeostasis: mechanistic and health aspects. Critical Reviews in Clinical Laboratory Sciences, 2008. 45(1):1-23.

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

Hell R, Stephan UW. Iron uptake, trafficking, and homeostasis in plants. Planta, 2003. 216:541-551.

IGB (Integrated Genome Browser). 2014. [online]. www.igbquickload.org/quickload. 2014 Apr-May.

Kobayashi T, Nishizawa NK. Iron uptake, translocation, and regulation in higher plants. Annual Review of Plant Biology, 2012. 63:131-52.

NCBI. BLAST: Standard Protein Blast. 2014. [online]. http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE\_TYPE=BlastSearch&BLAST\_SPEC=&LINK\_LOC=blasttab&LAST\_PAGE=blastn. Accessed 2014 Apr.

Roschzttardtz H, Conéjéro G, Divol F, Alcon C, Verdeil JL, Curie C, Mari S. New insights into Fe localization in plant tissues. Frontiers in Plant Science, 2013. 4:1-11.

Schmid NB, Giehl RF, Döll S, Mock HP, Strehmel N, Scheel D, Kong X, Hider RC, von Wirén N. Feruloyl-CoA 6’-hydroxylase1-dependent coumarins mediate iron acquisition from alkaline substrates in *Arabidopsis*. Plant Physiology, 2014. 164:160-172.

Sussman MR. Molecular analysis of proteins in the plant plasma membrane. Annual Review of Plant Physiology and Plant Molecular Biology, 1994. 45:211-34.

TAIR. 2014. [online]. Available from: http://www.arabidopsis.org. Accessed 2014 Mar-May.

Vert G, Grotz N, Dédaldéchamp F, Gaymard F, Guerinot ML, Briat JF, Curie C. IRT1, an *Arabidopsis* transporter essential for iron uptake from the soil and for plant growth. The Plant Cell, 2002. 14:1223-1233.