2014

Sequence Analysis of Maize Yellow Stripe3 Candidate Genes

Dennis B. DePaolo
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/masters_theses_2

Part of the Botany Commons, Genetics Commons, Molecular Biology Commons, Molecular Genetics Commons, and the Plant Biology Commons

Recommended Citation
DePaolo, Dennis B., "Sequence Analysis of Maize Yellow Stripe3 Candidate Genes" (2014). Masters Theses. 81.
https://scholarworks.umass.edu/masters_theses_2/81

This Open Access Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
SEQUENCE ANALYSIS OF MAIZE YELLOW STRIPE3 CANDIDATE GENES

A Thesis Presented

by

DENNIS BERNARD DEPAOLO

Submitted to the Graduate School of the University of Massachusetts Amherst in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

September 2014

Plant Biology
SEQUENCE ANALYSIS OF MAIZE YELLOW STRIPE3 CANDIDATE GENES

A Thesis Presented

By

DENNIS BERNARD DEPAOLO

Approved as to style and content by:

_________________________________________
Samuel Hazen, Chair

_________________________________________
Elizabeth Vierling, Member

_________________________________________
Dong Wang, Member

_________________________________________
Elsbeth Walker, Director
Plant Biology Graduate Program
ACKNOWLEDGEMENTS

I would like to thank Associate Professor Samuel Hazen for giving me the opportunity to join his laboratory group to complete my Masters thesis. The support and guidance was paramount and encouraged me to continue pushing forward and maximized the quality of my graduate research experience.

I would like to thank Professor Elsbeth Walker for her guidance which greatly influenced my decisions as a young scientist. I also want to thank Professors Dong Wang and Elizabeth Vierling for serving on my committee and providing advice and assistance when needed.

Susan Capistran your relentless work keeps the Plant Biology program at UMASS running smoothly with your well spirited and gracious personality that outshines the darkest moments of the graduate career.

I want to give a special thanks to the Plant Biology Graduate students who have been more like family than friends. This group really had something special in the support and dedication to both science and each other. These are memories and people I will never forget.

To my family and friends, thank you for always believing in my abilities.
The work presented here focuses on the molecular mechanism of phytosiderophore secretion in graminaceous plants. In maize, yellow stripe3 (ys3) is a mutant that is deficient in its ability to secrete iron-chelating compounds of the mugineic acid family known as phytosiderophores. Phytosiderophores are specific to grasses and are used for the acquisition of iron. Genetic linkage mapping of the ys3 locus lead to a region of interest on chromosome 3 defined by marker UMC1773. The sequence of eleven candidate genes (GRMZM2G390345, GRMZM2G390374, GRMZM2G342821, GRMZM5G800764, GRMZM2G502560, GRMZM5G849435, GRMZM2G105766, GRMZM5G876835, GRMZM2G036976, GRMZM2G502563, miR167g) revealed several small deletions and point mutations within the coding regions of four candidate genes (GRMZM2G390345, GRMZM2G342821, GRMZM2G105766 and GRMZM2G502560). To further investigate these candidate genes, knockout (Mutator) lines were used in an effort to recapitulate the ys3 phenotype. I characterized a GRMZM2G390345 mutant and determined that the mutation in GRMZM2G390345 did not cause the ys3 phenotype. For the three remaining candidate genes, GRMZM2G342821, GRMZM2G105766 and GRMZM2G502560 further analysis is needed. The closest homolog of OsTOM1 (an effluxer of mugineic acid) GRMZM2G603306 which, lies ~7 Mbp outside the linked marker, has also been suggested to be yellow stripe3. Preliminary sequence analysis uncovered a 4 bp insertion within the 9th exon of this gene that shifts the reading frame and leads to a premature stop codon. Three of the genes analyzed remain plausible candidates for yellow stripe3.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td></td>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>Importance of iron in plants and humans</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Strategies of iron acquisition</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Strategy I iron acquisition</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Strategy II iron acquisition</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Phytosiderophores and maize</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Iron deficiency response in maize</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Microbial degradation and the lack of a diurnal rhythm in maize</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Mechanism of phytosiderophore release: exocytosis vs transport</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Maize mutant yellow stripe3</td>
<td>8</td>
</tr>
<tr>
<td>II. MATERIALS AND METHODS</td>
<td>Plant growth conditions</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Growth on soil</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Foliar iron application</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Analysis of Z. mays Illumina Mutator lines</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Analysis of gene expression in Zea mays</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RNA isolation...........................................................................................................10
First strand cDNA synthesis..............................................................................10
Quantitative RT-PCR........................................................................................11
Analysis of genomic DNA sequencing.............................................................13
Plant genomic DNA extraction.........................................................................13
Plant genomic DNA extraction with Qiagen Dneasy Kit.................................13
Direct PCR.........................................................................................................14
Gel Purification..................................................................................................14
DNA sequencing...............................................................................................14
Sequence alignment..........................................................................................15

III. RESULTS.........................................................................................................16
    Part I: Characterization of GRMZM2G390345.............................................24
    Part II: Characterization of GRMZM2G342821...........................................32
    Part III: Characterization of GRMZM2G105766........................................37
    Part IV: Characterization of GRMZM2G502560........................................39
    Part V: Characterization of GRMZM2G390374...........................................39

IV. DISCUSSION....................................................................................................43

APPENDIX: SUPPLEMENTAL FIGURES: GENE SEQUENCING DATA.............53

BIBLIOGRAPHY....................................................................................................93
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Primer sequences used for qRTPCR analysis</td>
<td>12</td>
</tr>
<tr>
<td>2. Gene candidates surrounding marker UMC1773 targeted for sequence analysis</td>
<td>18</td>
</tr>
<tr>
<td>3. Primers used to sequence genomic DNA of candidate genes</td>
<td>21</td>
</tr>
<tr>
<td>4. Primers used to amplify GRMZM2G390345 insertional mutants</td>
<td>29</td>
</tr>
<tr>
<td>5. Genotype and phenotype data for GRMZM2G390345 insertional mutant population</td>
<td>30</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Strategies for iron uptake from the soil</td>
<td>3</td>
</tr>
<tr>
<td>2. Protein model and Gene model of GRMZM2G390345</td>
<td>26</td>
</tr>
<tr>
<td>3. Multiple protein alignment of GRMZM2G390345 (Protein Kinase)</td>
<td>27</td>
</tr>
<tr>
<td>4. qPCR of GRMZM2G390345 (Kinase) normalized to GAPDH</td>
<td>28</td>
</tr>
<tr>
<td>5. Maize GRMZM2G390345 Mu insertion line (Mu1007255)</td>
<td>31</td>
</tr>
<tr>
<td>6. Protein model and gene model of GRMZM2G342821</td>
<td>34</td>
</tr>
<tr>
<td>7. Multiple protein alignment of GRMZM2G342821 (Calmodulin Binding Protein)</td>
<td>35</td>
</tr>
<tr>
<td>8. Multiple protein alignment of DUF4005 domain within GRMZM2G342821</td>
<td>36</td>
</tr>
<tr>
<td>9. Protein model and gene model of GRMZM2G105766</td>
<td>38</td>
</tr>
<tr>
<td>10. Gel electrophoresis of GRMZM2G390374 cDNA</td>
<td>41</td>
</tr>
<tr>
<td>11. qPCR of maize GRMZM2G390374 (RHD3-like) normalized to GAPDH</td>
<td>42</td>
</tr>
<tr>
<td>12. Map of maize chromosome 3</td>
<td>48</td>
</tr>
<tr>
<td>13. Gene model of ZmTOM1 showing SNPs and indels</td>
<td>49</td>
</tr>
<tr>
<td>S1. Genomic alignment of GRMZM2G390345 (B73 vs ys3)</td>
<td>53</td>
</tr>
<tr>
<td>S2. Genomic alignment of GRMZM2G342821 (B73 vs ys3)</td>
<td>59</td>
</tr>
<tr>
<td>S3. Genomic alignment of GRMZM2G390374 (B73 vs ys3)</td>
<td>63</td>
</tr>
<tr>
<td>S4. Genomic alignment of GRMZM2G502560 (B73 vs ys3)</td>
<td>88</td>
</tr>
<tr>
<td>S5. Genomic alignment of GRMZM2G105766 (B73 vs ys3)</td>
<td>90</td>
</tr>
<tr>
<td>S6. Genes present on chromosome 3 between (118,288,00-119,675,000) that were targeted for sequence analysis</td>
<td>92</td>
</tr>
</tbody>
</table>
Importance of iron in plants and humans

Iron (Fe) is an essential micronutrient throughout all kingdoms including plants, animals, bacteria, fungi and archea. Plants require iron for many aspects of life. Iron is often a limiting factor to plant growth and is a cofactor in many plant redox reactions (Rouhier, 2010). Levels of iron within the plant are tightly regulated. Too much iron can lead to oxidative stress, while lack of iron leads to inhibition of photosynthesis. Iron is used in the production of chlorophyll, which also demonstrates the importance of iron in energy production. On earth today iron is the fourth most abundant element and the second most common heavy metal making up ~5% of the weight of the earth’s crust. Although abundant in soil, at neutral pH iron is usually bound in iron oxides or other ferric forms, which renders the iron insoluble and thus biologically unavailable for plant acquisition (Guerinot and Yi, 1994). To deal with the problem of iron availability, plants have evolved two distinct methods for obtaining iron from the environment known as Strategy I and Strategy II (Marschner et al., 1986).

In humans, iron-deficiency anemia is the most common and widespread nutritional disorder in the world. Around 2 billion individuals are affected by iron deficiency anemia world wide, the majority of whom are women and children (Balarajan et al., 2011). Around the world, cereals, specifically wheat, rice and corn, make up over half of the daily human caloric intake. Humans depend solely on their diet to acquire adequate iron, which is needed to make oxygen-carrying proteins such as hemoglobin and myoglobin. The human body typically stores iron for when needed, however, low
iron levels over a prolonged period of time can lead to iron-deficiency anemia. Because grains are the staple crops for many underdeveloped countries, the problem with anemia in these countries can be compounded due to the low iron content of these grains. Understanding iron acquisition by cereal plants and storage within the grains could facilitate bio-fortification as a remedy for iron deficiency anemia (Hirschi, 2009).

**Strategies of iron acquisition**

Iron exists in two oxidation states $\text{Fe}^{+3}$, ferric and $\text{Fe}^{+2}$, ferrous. In the presence of oxygen $\text{Fe}^{+2}$ is rapidly oxidized to $\text{Fe}^{+3}$, which is poorly soluble in water and precipitates as iron oxides. Ions of the heavy metals (such as Fe, Zn, Mn and Cu) do not freely cross the cell membrane and rely on transporters that allow for metal uptake. Plants have evolved two distinct mechanisms to accomplish iron uptake known as Strategy I and Strategy II.

**Strategy I iron acquisition**

Strategy I iron acquisition (Figure 1A) involves acidification of the rhizosphere and is used by the majority of plants except for grasses. In Strategy I, ATP-dependent H$^+$ pumps located on the plasma membrane of root epidermal cells pump protons out into the rhizosphere, lowering the pH of the soil and increasing iron solubility. After soil acidification, an integral membrane ferric reductase/oxygenase enzyme, FRO2, located on the plasma membrane, reduces $\text{Fe}^{+3}$ to the more soluble $\text{Fe}^{+2}$ form. The $\text{Fe}^{+2}$ is then transported through the iron-regulated transporter (IRT1) into the roots where it can be used by the plant (Vert et al., 2002).
Figure 1: Strategies for iron uptake from the soil. (A.) Strategy I. Acidification of the rhizosphere solubilizes Fe(III) and provides the substrate for the enzyme ferric chelate reductase (FRO2), which in turn provides Fe(II) for transport into the cell by IRT1. (B.) Strategy II. Phytosiderophores (PS) are produced by the root cells, and are released into the rhizosphere by an unknown mechanism. PS bind to Fe(III) in the soil, and the resulting Fe(III)-PS complexes are taken up into the root cells by the transporter YS1. C. Rice uses the typical grass Strategy II mechanism, but can also take up Fe(II) through IRT.
Strategy II iron acquisition

In grasses, a unique method of iron acquisition, referred to as Strategy II, has evolved (Figure 1B). Strategy II iron acquisition uses secondary metabolites known as phytosiderophores that act as iron chelators. Phytosiderophores describe a handful of compounds from the mugineic acid family including mugineic acid (MA), 3-hydroxymugineic acid (HMA), 2'-deoxymugineic acid (DMA), avenic acid, distichonic acid, 3-epihydroxymugineic acid (epiHMA) and 3-epihydroxy-2'-deoxymugineic acid (epiHDMA). The biosynthetic pathway of phytosiderophores has been well studied. They are synthesized from nicotianamine, which is originally derived from the well-characterized methionine biosynthesis pathway (Ohata et al., 1993, Higuchi et al., 1999, Bashir et al., 2006). In Strategy II, grasses secrete phytosiderophores into the rhizosphere through a currently unknown process. Once in the rhizosphere the siderophores chelate and solubilize Fe$^{3+}$, which can then be taken back up into the root by the YELLOW STRIPE1 (YS1) transporter and used by the plant (Curie et al., 2001). While the iron uptake mechanism in grasses has been known for some time, the initial mechanism of phytosiderophore secretion by the roots remains the final uncharacterized step toward the complete understanding of iron acquisition in grasses.

Phytosiderophores and maize

Phytosiderophores were first discovered in the mid-1970s and described as iron chelating compounds found in oat and rice root washings (Takagi 1976). Phytosiderophore structure and properties were further described physiologically five years later (Sugiura et al., 1981). As Strategy I and II became recognized as separate categories for iron acquisition, maize, like other graminaceous plants was classified as a
Strategy II plant that acquires iron from the release and uptake of phytosiderophores (Romheld and Marschner, 1986). The single type of phytosiderophore released by maize was identified as DMA (Mori and Nishizawa, 1987). There are differences in phytosiderophore secretion between maize and other grass species regarding the type of phytosiderophore synthesized, amount synthesized, amount secreted and timing of secretion (Bernards et al., 2013).

**Iron deficiency response in maize**

Synthesis and secretion of mugineic acids are dramatically enhanced under Fe-deficient conditions. The upsurge of secretion is correlated with increased expression of phytosiderophore biosynthetic genes and also increased expression of Fe-phytosiderophore transporters (Kobayashi, et al., 2005). This process is referred to as the iron deficiency response. In maize there is variation between genotypes relating to the iron deficiency response. Genotypes B73 and Mo17 respond differentially to limiting iron in hydroponic growth with contrasting iron chlorosis phenotypes. These two genotypes were compared by deep RNA-Sequencing (Urbany et al., 2013). Over 400 significantly regulated transcripts (FDR < 0.05) within both inbred lines were identified at limiting iron growth conditions (Urbany et al., 2013). This study took a broad view at understanding differential iron deficiency associated chlorosis between B73 and Mo17 and represents a valuable resource for differentially expressed genes upon iron limitation and chlorosis. It is proposed that the methionine salvage pathway and sterol metabolism, among others pathways underlie the contrasting iron deficiency related chlorosis phenotypes of the two inbred lines.
Microbial degradation and the lack of a diurnal rhythm in maize

Phytosiderophore secretion in most cereals follows a distinct diurnal rhythm with a pronounced secretion peak following sunrise (Marschner et al., 1986). This copious secretion is thought to aid the plant in overcoming bacterial acquisition of phytosiderophores. Bacteria have been isolated from barley roots that can utilize phytosiderophores as their only carbon source for growth, and microbial decomposition of phytosiderophores is a well-documented phenomenon (Shi et al., 1988; Takagi et al., 1988; Watanabe and Wada, 1989). Bacterial degradation can lead to severe Fe-deficiency chlorosis in nutrient solution cultures (von Wirén et al., 1993). Although disadvantages exist for continuous release of phytosiderophores, continuous phytosiderophore release from 15-d old maize (Fe-chlorosis tolerant cultivar ‘Alice’) during 14-h of daylight has been observed (Yehuda et al., 1996). More recently, it was determined that DMA is released from multiple maize hybrids during both dark and light periods (Bernards et al., 2000; Ueno et al., 2009). This study also compared mRNA expression of two different ferric-phytosiderophore transporters (ZmYS1 in maize and HvYS1 in barley). The relative mRNA level of the maize transporter, ZmYS1, did not show a diurnal rhythm. In contrast, the relative mRNA level of HvYS1 in barley exhibited a distinct diurnal rhythm, reaching a maximum level immediately before the onset of phytosiderophore secretion.

Phytosiderophore release from maize cultivars tested did not follow a distinct peak in secretion, which is in contrast to typical secretion peaks observed in other cereal species. Maize is known to be an iron sensitive crop. Release of phytosiderophore did not correspond with the severity of Fe deficiency chlorosis when different Fe-efficient and Fe-inefficient maize cultivars had been tested (Lytle and Jolley, 1991).
Mechanism of phytosiderophore release: exocytosis vs transport

Studies of barley roots also suggest a correlation between secretory vesicles and the release of phytosiderophores. Secretory vesicles are typically membrane bounded vesicles derived from the Golgi apparatus and contain material that is to be released from the cell. Secretory vesicles laden with ribosomes were detected at the plasma membrane facing the rhizosphere suggesting that these specific vesicles bud off of the rough endoplasmic reticulum. Vesicles were observed using transmission electron microscopy. Swollen vesicles were observed just prior to sunrise and the release of phytosiderophores. After phytosiderophore secretion the vesicles were no longer observed (Nishizawa and Mori, 1987; Negishi et al., 2002). This implies a possible role for exocytosis in phytosiderophore secretion. However these vesicles could be secreting a variety of compounds and would likely not secrete phytosiderophore alone.

Alternatively, it has been shown that phytosiderophores are secreted along with potassium (K⁺) in equimolar amounts, suggesting a possible symport mechanism. Anion channel blockers were used to restrict the amount of phytosiderophore being secreted from barley roots suggesting phytosiderophores are secreted as monovalent anions and also an involvement of a possible transport protein in phytosiderophore secretion (Sakaguchi et al., 1999).

Recently a Major Facilitator Superfamily (MFS) transporter with phytosiderophore efflux activity, named Transporter of Mugineic Acid (TOM) has been identified. This gene was discovered in rice and barley (Nozoye et al., 2011). TOM1 in rice is expressed on a diurnal rhythm that coincides with phytosiderophore secretion and is also up-regulated when plants are starved of iron. Over-
expression of TOM1 lead to increased tolerance to iron deficiency, while repression of TOM1 causes a decrease in the secretion of phytosiderophores (Nozoye et al., 2011). Heterologous expression of TOM1 in oocytes revealed that TOM1 is able to efflux phytosiderophores but not nicotianamine, the precursor to phytosiderophore synthesis. Potentially this transport protein is solely responsible for the efflux of phytosiderophores from the roots into the rhizosphere.

**Maize mutant yellow stripe3**

*Yellow stripe3* is a recessive mutation in maize that results in defective phytosiderophore secretion. The *ys3* plant is able to produce normal levels of phytosiderophore, but these are retained within the root system and are not secreted (Neumann, 1999). The Walker lab fine mapped *ys3* to an interval on chromosome 3 defined by marker *umc1773*. Sequencing of the predicted genes within the fine mapped interval did not uncover large insertions or deletions. However it did uncover point mutations that change conserved amino acid residues as well as a small deletion (Table 2). Maize *ys3* is a natural mutation that was isolated from a field crop in the late 1960’s. The genetic background is not known, however upon sequencing *ys3*, B73 and W22 genotypes I found *ys3* to be more similar to W22 than B73. W22 is not nessiserily the genetic background of *ys3*. The reference genome published online and used for comparison was B73. Next, I present the materials and methods used in this study.
CHAPTER II

MATERIALS AND METHODS

Plant growth conditions

Growth on soil

_{Zea mays} accessions B73, W22, insertional mutants and mutant yellow stripe3_ seeds were sown into a 3:1 (Soilless mix: coarse sand) mixture of potting mix (Fafard Canadian Growing Mix 2) pretreated with Gnatrol (Valent Bioscience Corporation, Libertyville, IL). All maize was grown in a greenhouse. Greenhouse conditions for soil grown plants were 16 hours light, 8 hours dark at 21-23°C.

Foliar iron application

A 2 mg/mL solution of ferric ammonium citrate ((Fisher cat# I72-500, Waltham, MA) was sprayed directly onto the aboveground parts of the plants. The exact formulation of ferric ammonium citrate cannot be determined, owing to variability of the salts formed. Spraying was started when the leaves of the mutant plants began to show interveinal chlorosis, and continued every 3 to 4 days throughout the rest of the plant’s life cycle.

Analysis of _Z. mays_ Illumina Mutator lines

Maize Mutator lines are available from the Mu-illumina collection (Williams et al., 2010). Maize locus _mu-illumina_218271.3 was identified correlating to a Mutator insertion within the 3rd exon of GRMZM2G390345. Seeds were ordered from the Institute of Molecular Biology (University of Oregon) and 15 seeds were sown along side both wildtype W22 and ys3 into a 3:1 (Soilless mix: coarse sand) mixture of potting mix (Fafard Canadian Growing Mix 2) pretreated with Gnatrol (Valent Bioscience
Corporation, Libertyville, IL). Greenhouse conditions for soil grown plants were 16 hours light, 8 hours dark at 21-23°C.

**Analysis of gene expression in *Zea mays***

**RNA isolation**

*Z. mays* tissue up to 100 mg was ground using a mixer mill (Retsch, Newtown, PA). Chrome Steel Beads 3.2 mm in diameter (454 grams) (BioSpec Products, Bartlesville, OK) were placed in a 1.5mL eppendorf tube and frozen with liquid N\textsubscript{2}. Total RNA was isolated using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA), followed by DNase I treatment (Ambion, Austin, Texas). Total RNA concentration was quantified using a NanoDrop and confirmation of RNA quality was performed by visualizing on a 1X TAE gel stained with ethidium bromide.

**First strand cDNA synthesis**

First strand cDNA was synthesized from 600 ng total RNA using SuperScriptIII reverse transcriptase (Invitrogen, Carlsbad, CA) and oligo-dT primers. Reactions without reverse transcriptase were also used for each sample in order to detect any genomic DNA contamination. Equal amounts of each RT reaction were used as templates for PCR amplifications. Integrity of RNA samples and equality of cDNA amounts were examined by amplifying ZmGAPDH genes using primers shown in Table 1. PCR reactions were performed using Ex- Taq polymerase (Takara, Madison, WI) under the following conditions: an initial denaturation step of 94°C for 2 min, followed by 25 cycles of 94°C for 30 sec, 57°C for 30 sec, 72°C for 45 sec, with a final elongation step of 72°C for 10 min. In the cases where the PCR products were sequenced, a proofreading DNA polymerase Phusion (Thermo Scientific, Rockford, IL) was used in PCR reactions under
the following conditions: an initial denaturation step of 98°C for 2 min, followed by 35 cycles of 98°C for 30 sec, an annealing temperature that is 3° higher than the lower Tm primer for 30 sec, 72°C for 1 min, with a final elongation step of 72°C for 7 min.

**Quantitative RT-PCR**

GRMZM2G390345 and GRMZM2G390374 gene expression was normalized to *ZmGAPDH*. The sequences of these primers are shown in Table 1. The primer efficiency of GRMZM2G390345, GRMZM2G390374 and the housekeeping gene were calculated empirically by performing serial dilutions of the appropriate cDNA template with each primer set and using REST software (Qiagen, Valencia, CA).

The Brilliant II SYBR Green QPCR Master Mix (Stratagene, La Jolla, CA) was used to perform the qRT-PCR reactions. A two-step thermocycler program was used with the following conditions: 5 min at 95 °C followed by 40 cycles of 10 sec at 95 °C and 30 sec at 60 °C. The quality of the reactions was then verified through inspecting the melting curves. All the qRT-PCR reactions were performed using three technical replicates. The mean of the three technical replicates for each reaction was calculated to determine the final threshold cycle (Ct) value for each condition. The comparative ΔCt method was then used to produce a quantitative representation of gene expression by normalizing the Ct values of both GRMZM2G390345 and GRMZM2G390374 to *ZmGAPDH* values. The primer efficiencies were accounted for in the calculation of gene expression using the geNorm algorithm (Vandesompele, De Preter et al. 2002), and the addition of water instead of cDNA was included as the negative control for each condition.
<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRMZM2G390345</td>
<td>Forward: GGAGGAGACATGGCTGAGGTTATC</td>
</tr>
<tr>
<td></td>
<td>Reverse: GCCACTTGCAAAGTTTCTGGTCAG</td>
</tr>
<tr>
<td>GRMZM2G390374</td>
<td>Forward: CACGTTTGTGGTGTTGAACC</td>
</tr>
<tr>
<td></td>
<td>Reverse: GTCATCCTCTCTCTCTCTGTTCC</td>
</tr>
<tr>
<td>ZmGAPDH</td>
<td>Forward: CCTGCTTCTCATGGATGGTT</td>
</tr>
<tr>
<td></td>
<td>Reverse: TGGTAGCAGGAAGGAAACA</td>
</tr>
</tbody>
</table>
Analysis of genomic DNA-sequencing

Plant genomic DNA extraction

A 0.2g sample of Z. mays leaf was ground using a mixer mill (Retsch, Newtown, PA) and 3.2mm Chrome Steel Beads (454 grams) (BioSpec Products, Bartlesville, OK) in liquid nitrogen in a 1.5mL eppendorf tube. 600 µl of DNA extraction buffer (DEB) (100mM NaCl, 50mM Tris-Cl pH 8.0, 25mM EDTA, 1% sodium dodecyl sulfate (SDS) and 10mM b-mercaptoethanol (BME) was added. The suspension was mixed by vortexing. Samples were placed at 65°C for 10 minutes, then allowed to cool on ice. 0.25µl of 5M KOAc was added as the tubes were incubated on ice for an additional 30 minutes. Samples were spun at 17000xg for 10 minutes at 4°C. Supernatants were removed and filtered through Miracloth to new 1.5 mL eppendorf tubes containing 600µl of isopropanol, inverted to mix. Samples were spun at 17000 xg for 1 minute. Supernatants were poured off, and 100µl of 70% ethanol was added to rinse the pellet. Samples were dried and the pellet was dissolved in 90µl of Tris-EDTA buffer T_{10E_{5}}. Once completely dissolved 10µl of 3M NaOAc pH 5.2 was added and then mixed followed by adding 200µl of 70% ethanol. Samples were spun for 10s at 17000 xg and supernatants were poured off, and pellets were rinsed with 70% ethanol and allowed to dry at room temperature. Pellets were resuspended in 50µL TE.

Plant genomic DNA extraction with Qiagen DNeasy Kit

1 g of plant shoot tissue was ground using a ceramic mortar and pestle in liquid nitrogen. The powdered tissue was transferred to a 15mL falcon tube, and Qiagen DNeasy Plant Maxi kit protocol (Qiagen, Valencia, CA) was followed. Total DNA
concentration was quantified using a NanoDrop spectrophotometer (Thermo Scientific, Rockford, IL), and confirmed on 1X TAE gel stained by ethidium bromide.

**Direct PCR**

PCR was run under the following conditions: an initial denaturation step of 95ºC for 5 min, followed by 35 cycles of 95ºC for 15 sec, an annealing temperature that is 5º higher than the lower Tm primer for 30 sec, 72ºC for 60 sec, with a final elongation step of 72ºC for 1 min.

**Gel purification**

PCR products were analyzed using gel electrophoresis before being purified and sent for sequencing. 1% agarose gels were made using 0.6g agarose (BullsEye, St. Lois, MO) in 60ml of deionized water. 1.5µL of ethidium bromide 1% solution (Fisher BioReagents) was incorporated into the gel to visualize the DNA. 20µL of PCR product was combined with 1.0µL 6X loading dye and loaded into the appropriate lane. Gels were run for 60 minutes at 100V and 400mA. Once complete the gel was placed on a UV lamp to visualize the DNA bands for extraction. The appropriate band was removed from the gel and further purified using Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine,CA). Once purified the DNA fragments were ready for sequencing.

**DNA sequencing**

All purified PCR products were sequenced using GENEWIZ (Cambridge, MA). PCR products ranging from 500-1000 bp were purified and submitted for sequencing at a concentration of ~2ng/L in a 10µL sequencing reaction. Primers were diluted to 5µM and 5µL were used in each reaction. Sequencing reads were tested and given both a QS
(Quality Score) and CRL (Contiguous Read Length) score to assess the quality of the reads sequenced.

**Sequence alignment**

All genomic sequencing was carried out by GENEWIZ (Cambridge, MA). Reads which passed the quality control were trimmed to remove low quality sequence from the terminal ends of each read. Sequencher software (Gene Codes Corporation, Arbor, MI) was used for sequence alignment. Assembly parameters were adjusted accordingly to use the optimal algorithm, including minimum match percentage and minimum overlap between reads. Reads were scored by base pair on quality and chromatographs were analysed to determine if each mismatch between the sample sequence and the reference sequence contained a high quality read.
CHAPTER III

RESULTS

I have evaluated and sequenced 11 candidate coding genes out of 47 genes (Supplemental Fig. 6) located between GRMZM2G105766 and genetic marker UMC1773. Marker UMC1773 has been shown to cosegregate with the ys3 mutation (Table 2). When selecting candidate genes for analysis I considered positioning relative to UMC1773, as well as available gene annotations, giving specific attention to genes involved in signaling, vesicle trafficking, exocytosis or transport. I also considered expression data. Priority was given to genes that showed high expression in roots where phytosiderophore secretion occurs. Of the 11 genes sequenced, four of the candidates GRMZM2G390345, GRMZM2G342821, GRMZM2G105766 and GRMZM2G502560, were found to contain polymorphisms within the coding regions that may cause the ys3 phenotype. A fifth candidate, GRMZM2G390374 contains a large (397 bp) deletion within the 21st intron that could potentially affect transcript splicing. This gene was also investigated further.

Of the four candidate genes that contain polymorphisms within their coding regions, an insertional mutant from the Mu-illumina collection was available for GRMZM2G390345. Analysis of a segregating population with an insertion in GRMZM2G390345 revealed that the population was in fact segregating for the insertion, as verified through genotyping (Table 5; Figure 5) and qPCR analysis (Figure 4). One half of the insertional mutant population showed phenotypes of stunted growth and discoloration at the three leaf stage after three weeks of growth. Foliar application of ferric ammonium citrate was not able to rescue this phenotype, and within 2 weeks of
foliar iron application one half of the insertional mutant population became necrotic and died, while all of the \( y_{33} \) individuals planted along side survived and became iron repleat. Of the three remaining candidates with coding region polymorphisms, GRMZM2G105766 encodes a calmodulin dependent protein kinase with high root expression. Within the coding region of this gene two single nuceotide polymorphisms (SNPs) caused nonsynonymous substitutions and a 4 bp deletion (\( \Delta \text{TCCG} \)) shifts the open reading frame on the messanger RNA (mRNA) (Fig. 9). GRMZM2G342821 encodes a protein of unknown function with weak similarity to calmodulin binding protein. It is expressed highly in the endosperm, but also shows expression within the root tissue. GRMZM2G342821 contains four SNPs within the coding region that lead to nonsynonymous substitutions (Fig. 6). GRMZM2G502560 is specific to maize and a likely pseudogene which encodes a small protein with no functional annotation. A single base pair (\( \Delta \text{T} \)) deletion within the coding region disrupts the reading frame. Further analysis of these three genes will be needed to determine if they are involved with phytosiderophore secretion.
Table 2: Gene candidates surrounding marker UMC1773 targeted for sequence analysis. All polymorphisms within coding regions (exons) of yS3 compared to the B73 reference genome are reported. Base pairs are numbered from the translational start site (ATG).

<table>
<thead>
<tr>
<th>Gene Identifier</th>
<th>Annotation</th>
<th>Expression</th>
<th>Polymorphism in exons</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRMZM2G390345*</td>
<td>Serine/threonine protein kinase</td>
<td>High in embryo, immature tassel and immature cob</td>
<td>Non-synonymous substitutions: bp44 (C→T) (Ser→Phe) bp3645 (A→G) (Ser→Gly) Bp3672 (G→A)(Asp→Asn) bp3790 (G→T) (Gly→Val) bp 4102 (A→G) (Glu→Gly) Synonymous substitutions: bp1604 (T→C) (Ser→Ser) bp1779 (C→T) (Asp→Asp) bp3659 (C→T)(Ser→Ser) Bp3671 (T→C) (Leu→Leu) bp3677 (A→G) (Gln→Gln) bp3875 (C→T) (Gly→Gly) bp4097 (G→C) (Pro→Pro) Deletions: bp3762 to 3792 27bp deletion→GGCC (MKKNLSSAPGHD) (GHDV)</td>
</tr>
<tr>
<td>GRMZM2G390374</td>
<td>RAS-like GTPase Root Hair Defective3</td>
<td>High in anthers</td>
<td>Synonymous substitutions: bp104 (C→T) (Val→Val)</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
<td>Expression</td>
<td>SNP Changes</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
</tr>
</tbody>
</table>
| GRMZM2G390374 | RAS-like GTPase Root Hair Defective3 | High in anthers | bp1122 (A→G)  
Leu→Leu  
bp2305 (A→T)  
Ala→Ala  
Non-synonymous substitutions:  
bp35 (G→C)  
(Leu→Phe)  
bp39 (T→C)  
(Ser→Pro)  
bp1152 (G→C)  
(Leu→Phe)  
bp1304 (C→T)  
(Pro→Leu)  
Synonymous substitutions:  
bp104 (G→A)  
(Ala→Ala)  
bp1245 (C→T)  
(Ala→Ala)  
bp1506 (C→T)  
(Thr→Thr) |
| GRMZM2G342821* | Protein of unknown function weak similarity to calmodulin binding protein | High in endosperm and immature tassel | |
| GRMZM5G800764 | No functional annotation | High in embryos Very low in roots. Maize specific gene. | None  
Deletion:  
bp148 (T→deleted)  
Shifts open reading frame |
| GRMZM2G502560* | No functional annotation | No data available | None |
| GRMZM5G849435 | TBC1 domain (vesicular transport) Rab-GTPase | No data available | None |
| GRMZM2G105766* | Calmodulin dependent protein kinase | High in roots | Non-synonymous substitutions:  
bp765 (C→G)  
(Ser→Trp)  
bp1030 (C→T)  
(Pro→Leu)  
in new reading frame  
(Pro→Ser)  
Deletion:  
bp861→864  
(4bp→deletion)  
Shifts ORF |
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Functional Annotation</th>
<th>Expression Level</th>
<th>Annotation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRMZM5G876835</td>
<td>No functional</td>
<td>Low in bundle sheath</td>
<td>Synonymous</td>
<td>Substitution: bp724 (A→G) (Lys→Lys)</td>
</tr>
<tr>
<td></td>
<td>annotation</td>
<td>cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRMZM2G036976</td>
<td>No functional</td>
<td>Embryo</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td></td>
<td>annotation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRMZM2G502563</td>
<td>No functional</td>
<td>No data available</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>miR167g</td>
<td>annotation</td>
<td>miRNA</td>
<td>14bp deletion</td>
<td>within the loop of miR167g within ys3. Deletion also present in W22.</td>
</tr>
</tbody>
</table>
Table 3: Primers used to sequence genomic DNA of 11 candidate genes. Gene name (left), primer name (center) and primer sequence 5’ – 3’ (right).

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Primer</th>
<th>Sequence (5’ – 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRMZM2G390345</td>
<td>oGRMZM2G390345_73..95</td>
<td>GTCTAGAACCAAATGCCGCGAGA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_335..357</td>
<td>TCTTGGAGTTGGTGACTCTGCTC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_490..512</td>
<td>CCGCCTTTGCAATATTTTATCG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_602..624</td>
<td>CATTTCTTTCTAGCCCGCTTTTG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_719..697</td>
<td>GGAAATCGAAGAAGAACAGAAGT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_2..22</td>
<td>CAACCAAACCCGCCACTTTTG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_622..645</td>
<td>TGATTTTCAGAGCGGTTTTCAGG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_1101..1078</td>
<td>GCAACCTTCAGAGAAATGGTATT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_1250..1227</td>
<td>TGTCAGAGCCGATCTTTGAGGAGC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_1109..1132</td>
<td>GTAGAGGACAAGGGCGAGAAAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_1897..1874</td>
<td>GCACCTTTCAGAGAAATGTGTTT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_1747..1760</td>
<td>TGGGTACTGTGGAGGGGAGAGCAT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_1854..1879</td>
<td>GTATACAAAAGAGCCAAAGGAAAACC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_2000..2023</td>
<td>CTCGTCGCAACTTCATGCTGCTT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_2277..2252</td>
<td>AATGAATATCAGCGGATGGGCTGAAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_3035..3012</td>
<td>QAATATCTGTTGGGCAATGGTGAT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_3271..3296</td>
<td>AAAACATTTCGATAAATCGCCATA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_3479..3457</td>
<td>AAAAGCTGATTCTCTGCACATCT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_2420..2442</td>
<td>AAACAAGAAGTGAGGGACTCACA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_4028..4003</td>
<td>ATATCGTGGTGAGAAAAGGAAAAAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_4407..4428</td>
<td>AAGGACAGTGACGGTGAGGAGCT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_4714..4691</td>
<td>ACCTGACCCGTTGACTTCAGT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_4428..4407</td>
<td>CTATGAGCGCCCTCAGGGGAGC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_4158..4181</td>
<td>CTTGAGCAGCAGAAGAGAGCATCGC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_4869..4846</td>
<td>CCTCCTTCCAGAGCCTACACT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390345_4277..4300</td>
<td>GACAGGATCCGGAATCTCCCTCA</td>
</tr>
<tr>
<td>GRMZM2G390374</td>
<td>oGRMZM2G390374_361..335</td>
<td>TCAATGTTTCTTTCATTTTCATTTCA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_77..100</td>
<td>QAATATGTGGGCTGCCTCAGTAAA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_187..211</td>
<td>TGCCGAGCGAATTGAGAAATAAGT</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_240..264</td>
<td>ATAACCTCTAGTTTCCACATGTGA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_818..840</td>
<td>GCCACAAATGTGAAACACAGGAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_1137..1118</td>
<td>QAAGACAGCGGGGTAGAGAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_1172..1149</td>
<td>CGAAGCAGATACCGGCTTGGAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_1236..1213</td>
<td>GATGCTGATGAGATGGGAGGAGC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_1633..1655</td>
<td>AGTGCACTCCATGTCAGAGGAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_768..790</td>
<td>ATGTCGAGAGAGAGAGAGAGAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_1738..1716</td>
<td>TGGTACTTTTCTGTCTACCTCAA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_2050..2027</td>
<td>GCAATGCACACACAGTCAATTAC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_1414..1437</td>
<td>GACAGGAGTCGGTTGAGATAGGA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_2569..2547</td>
<td>CCCAAGAAGACATTACCTGAC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_1604..1629</td>
<td>TGGTAACTTTTCTGCTCTACCTGA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_2949..2935</td>
<td>GACCTTTTACACTCTTCACCTAC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_2285..2303</td>
<td>ATACCCATTACACACTCCTGAA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_2460..2484</td>
<td>ACTGATGAGAGAGAGAGAGAGAG</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_3301..3324</td>
<td>TTTGCTTGGCAATTTCAGAGCATA</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_3456..3432</td>
<td>CATGGCGAACACCTGTCACATAAAC</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G390374_3489..3466</td>
<td>TGTTTATTGGCAGCCTTTTCAGC</td>
</tr>
<tr>
<td>Gene Symbol</td>
<td>Start Position</td>
<td>Stop Position</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>---------------</td>
</tr>
<tr>
<td>GRMZM5G849435</td>
<td>oGRMZM2G849435_2137</td>
<td>...2111</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_2691</td>
<td>...2716</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_2897</td>
<td>...2920</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_3010</td>
<td>...3087</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_3146</td>
<td>...3121</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_3602</td>
<td>...3627</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_3831</td>
<td>...3854</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_3930</td>
<td>...3903</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_4061</td>
<td>...4036</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_4708</td>
<td>...4734</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G849435_4874</td>
<td>...4851</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G105766</td>
<td>56_80</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G105766_139</td>
<td>...164</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G105766_426</td>
<td>...402</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G105766_607</td>
<td>...629</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G105766_1071</td>
<td>...1048</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G876835</td>
<td>67...93</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G876835_143</td>
<td>...165</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G876835_512</td>
<td>...534</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G876835_742</td>
<td>...720</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G876835_1108</td>
<td>...1086</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G876835_1210</td>
<td>...1230</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G876835_1424</td>
<td>...1403</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G876835_1520</td>
<td>...1497</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976</td>
<td>175...198</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_331</td>
<td>...353</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_744</td>
<td>...762</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_707</td>
<td>...726</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_925</td>
<td>...906</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_1072</td>
<td>...1095</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_1434</td>
<td>...1456</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_1368</td>
<td>...1348</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_1669</td>
<td>...1647</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_1912</td>
<td>...1934</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_2206</td>
<td>...2184</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G036976_2509</td>
<td>...2487</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563</td>
<td>76...98</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563_286</td>
<td>...308</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563</td>
<td>652...630</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563_715</td>
<td>...737</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563_1283</td>
<td>...1261</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563_2897</td>
<td>...847</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563_917</td>
<td>...938</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563_1261</td>
<td>...1261</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2G502563</td>
<td>1478...1451</td>
</tr>
<tr>
<td></td>
<td>oGRMZM2GmiR167g</td>
<td>595...618</td>
</tr>
</tbody>
</table>
Part I: Characterization of GRMZM2G390345

One \textit{ys3} candidate, GRMZM2G390345, encodes a serine/threonine protein kinase that could be involved in signaling phytosiderophore secretion. Sequencing of GRMZM2G390345 revealed a 27 bp deletion and 5 point mutations in \textit{ys3} that were not present in either W22 or B73 inbred lines (Figure 2). The 27 bp genomic deletion is in the 3’ region of the protein that is not highly conserved (Figure 3). The kinase domain of this protein is located between amino acids 34-244, and no mutations were found within this domain (Figure 2). Multiple insertional \textit{Mutator (Mu)} lines have been generated for this locus and are available through the maize genetic stock center.

Seeds were available from the \textit{Mu}-illumina collection with an insertion in the third exon of GRMZM2G390345 chr3:119643504–119644141 designated \textit{Mu1007255} (Figure 2). The seeds were collected from a single self-fertilized ear (ear number: S144-6) and were expected to segregate for the insertion. The position of this \textit{Mu}-element should produce a full knockout of the candidate gene if the plant is homozygous for the insertion. To confirm \textit{Mu1007255} produced a full knockout I used qPCR to measure transcript abundance in plant shoots that were either homozygous for the insertion or homozygous wildtype. I also compared these data to expression of GRMZM2G390345 in the \textit{ys3} background (Figure 4). If GRMZM2G390345 is responsible for the \textit{ys3} phenotype, a full knockout would result in plants that show an iron deficiency phenotype.

All seeds pertaining to \textit{Mu1007255} were cultivated in the greenhouse along side both W22 and the \textit{ys3} mutant 311F. The plants were grown for 3 weeks until they reached the three leaf stage before harvesting leaf tissue to genotype each individual from the segregating population. The primers used to genotype these plants are listed in Table
4. After tissue was collected for genotyping and qPCR, I began foliar feeding with ferric ammonium citrate to test if the chlorotic phenotype could be rescued. The ys3 phenotype can be corrected through foliar application of ferric iron chelates.

After the first two weeks of growth, ys3 plants showed definitive signs of iron deficiency chlorosis (Figure 5). The Mu population consisting of 14 individuals segregated for the insertion (Table 5). However, the expected yellowing phenotype was not observed in the segregating insertional mutants, suggesting that a mutation in GRMZM2G390345 does not cause a ys3 phenotype. Therefore, analysis of the Mu line (Figure 5) (Table 5) suggests GRMZM2G390345 does not encode for ys3.

Due to the possibility that wildtype maize can show a yellowing phenotype when grown in greenhouse conditions, genotypic analysis is required to confirm visual observations. For example, S144-6 #7 showed a yellowing phenotype which was comparable to the ys3 mutant (Fig. 5: panel G). After foliar application of ferric ammonium citrate the yellowing phenotype was corrected, indicating that the plant may be a defective in iron uptake from its roots. However, genotyping revealed that S144-6 #7 was heterozygous for the insertion. Because ys3 is known to be a recessive mutation, this eliminates the S144-6 #7 phenotype from being associated with the insertion within GRMZM2G390345. When analyzing Mutator lines it is important to note that up to 10 other loci within the genome may also contain the Mu-element, which can result in phenotypes that are not directly related to the locus in question.
Figure 2: Protein model and Gene model of GRMZM2G390345. (A) Protein model showing GRMZM2G390345 kinase domain A.A> 34-244. (B) Gene model of GRMZM2G390345 (Kinase) showing the location of the Mu-element (Mu1007255) within the 3rd exon of the gene. Polymorphisms between y53 (Blue) and B73 (Black) are depicted within the 1st and 14th exons. A 27bp deletion was also found within the 14th exon.
Figure 3: Multiple protein alignment of GRMZM2G390345 (protein kinase). Nine A.A. deletion in yellow stripe3 (blue box). Region of protein is not highly conserved.
Figure 4: qPCR analysis of maize GRMZM2G390345 (Kinase) normalized to GAPDH. Shoot expression was measured at the three-leaf stage 21 days past germination. Results reflect three biological replicates with a p-value of 0.002.
Table 4: Primers used to genotype GRMZM2G390345 insertional mutants

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Primer Position</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRMZM2G390345-GSP</td>
<td>Forward</td>
<td>ACCGTGCTTCTTTCAATCCAG</td>
</tr>
<tr>
<td></td>
<td>Reverse</td>
<td>ACTCATGCCAGGTCAGTTCA</td>
</tr>
<tr>
<td>EoMU1</td>
<td>Left Boarder Primer</td>
<td>GCCTCCATTTTCGTCGAATCCC</td>
</tr>
<tr>
<td>EoMU2</td>
<td>Right Boarder Primer</td>
<td>GCCTCTATTTTCGTCGAATCCG</td>
</tr>
</tbody>
</table>
Table 5: Genotype and phenotypic data for the GRMZM2G390345 insertional mutant population. 14 individuals segregating for mu1007255 with a Chi-squared value of 0.092 using a probability of 0.05 and 2 degrees of freedom Phenotypes are recorded at three weeks/three leaf stage prior to supplemental foliar iron application and again 2 weeks after the application of foliar iron.

<table>
<thead>
<tr>
<th>Plant ID</th>
<th>Phenotype</th>
<th>Genotype of mutants</th>
</tr>
</thead>
<tbody>
<tr>
<td>S144-6 #1</td>
<td>Wild-type</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #2</td>
<td>Wild-type</td>
<td>(+/-), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #3</td>
<td>Wild-type</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #4</td>
<td>Wild-type</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #5</td>
<td>Wild-type</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #6</td>
<td>Wild-type</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #7</td>
<td>Yellow</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #8</td>
<td>Yellow</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #9</td>
<td>Wild-type</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #10</td>
<td>Albino</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #11</td>
<td>Pale green</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #12</td>
<td>Wild-type</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #13</td>
<td>Albino</td>
<td>(+/+), (-/-) X</td>
</tr>
<tr>
<td>S144-6 #14</td>
<td>Pale green</td>
<td>(+/+), (-/-) X</td>
</tr>
</tbody>
</table>
Figure 5: Maize GRMZM2G390345 Mu insertion line (Mu1007255). In panels A-N a representative leaf was chosen from W22 (left) and ys3 (center). A representative leaf from each individual Mu plant was taken from a segregating population (right). Panels A-N represent individuals 1-14 respectively.
Part II: Characterization of GRMZM2G342821

GRMZM2G342821 encodes a protein of unknown function with weak similarity to some calmodulin binding proteins. This protein could potentially sense changes in calcium through its interaction with calmodulin that would lead to signaling of PS secretion. External Ca$^{2+}$ ions at the appropriate concentrations and a Ca$^{2+}$ influx are required to elicit exocytosis in *Arabidopsis* root hairs (Schiefelbein et al., 1992). The inability to sense changes in calcium may prevent phytosiderophore secretion if an exocytosis mechanism is involved. Sequencing GRMZM2G342821 in *ys3* uncovered four missense mutations. None of the mutations exist within the calmodulin-binding domain located at position 91-109 and 113-130 of the amino acid sequence (Figure 6).

Two of the four non-synonymous mutations occur within the N-terminus of the protein. The first mutation (bp 35/ G→C)) changes a conserved leucine, a hydrophobic aliphatic amino acid into a phenylalanine, a hydrophobic aromatic amino acid. The second mutation (bp 39/T→C) changes an unconserved serine, which is small and polar residue to a proline, which is a small non-polar residue (Figure 7).

The other two mutations (bp 1152 G→C, bp 1304 C→T) exist within the Domain of Unknown Function (DUF) (amino acids 245-360; Fig. 6) and create non-synonymous substitutions changing amino acid 263 from a leucine to a phenylalanine and amino acid 315 from a proline to a leucine. However, these substitutions are not within the conserved regions of this domain (Figure 8). Due to the positioning of these polymorphisms, it is unlikely they will have an effect on protein function. No insertional mutants were available for this locus in maize or any other grass species (such as *O. sativa* or *B.*
*distachyon* that show clear synteny of this genomic region, so alternative methods must be used to preform further tests on this gene.
Figure 6: Protein model and gene model of GRMZM2G342821. (A) Protein Model showing calmodulin (CaM) binding domain A.A. 91-109, 113-130 and DUF4005 domain A.A. 245-355. (B) Gene model of GRMZM2G342821 (calmodulin binding protein) showing the location of polymorphisms between B73 (Black) and ys3 (Blue). Depicted within the 1st and 5th exons.
Figure 7: Multiple protein alignment of GRMZM2G342821 (calmodulin binding protein). Conserved leucine at the 12th A.A. position is substituted for a phenylalanine and an unconserved serine at the 14th A.A. position is substituted for a proline in *yellow stripe3*. 
Figure 8: Multiple protein alignment of DUF4005 domain within GRMZM2G342821. Conserved residues within the domain are highlighted with more highly conserved residues in red. In yellow stripe3 a leucine at A.A position 263 is replaced with a phenylalanine and a proline at A.A. position 315 is replaced with a leucine residue. Neither substitution lies within the conserved binding region of the protein.
Part III: Characterization of GRMZM2G105766

GRMZM2G105766 codes for a calmodulin dependent protein kinase, which could be involved in signaling phytosiderophore secretion by phosphorylating its target upon calmodulin sensing of changes in intracellular calcium concentrations. GRMZM2G105766 is highly expressed in the roots of maize as well as endosperm as shown on the maize eFP browser (Sekhon, et al., 2011).

The sequence of GRMZM2G105766 revealed two single nucleotide polymorphisms (SNPs) as well as a 4 bp deletion (Figure 9). The first SNP (bp 765/C\text-right\text-G) results in a non-synonymous substitution at amino acid 255 that replaces a serine residue with tryptophan. This mutation does not lie within a conserved active site or a binding pocket of the kinase domain. The 4 bp deletion \(\Delta\text{TCCG}\) (bp 862-865; amino acid 288) disrupts the reading frame of the unconserved C-terminus of the protein. The protein kinase domain comprises amino acid 22-280 and is not affected by the frame shift. The second SNP (bp 1030/C\text-right\text-T) results in a non-synonymous substitution that changes a proline residue to a leucine, however this SNP is located 3’ to the 4 bp insertion where the coding region has already been shifted out of frame. The disrupted reading frame takes the native stop codon out of frame and extends the protein by 74 amino acids into the 3’ UTR of the native transcript. No insertion lines were available for further analysis of this locus.
Figure 9: Protein model and gene model of GRMZM2G105766. (A) Protein Model showing the kinase domain A.A. 22-280. (B) Gene model of GRMZM2G105766 (1.54kb in length) showing the location of polymorphisms between B73 (Black) and ys3 (Blue). Depicted within the single exon of this gene.
Part IV: Characterization of GRMZM2G502560

GRMZM2G502560 codes for a hypothetical protein that is specific to maize. The gene is 203 bp in length and has no functional annotation. The hypothetical protein produced by this gene is 67 amino acids in length and does not contain any conserved domains. Upon sequence analysis a single bp deletion (ΔT) was found within the single exon of this gene (Figure S4). The deletion shifts the reading frame, which disrupts the native stop codon and extends the protein by 46 amino acids into the 3’UTR. Searching both gene homology and gene ancestry in Phytozome has resulted in no homology and no ancestry. Using Basic Local Alignment Search Tool (BLAST) to compare the protein sequence of this gene against the NCBI database, I was only able to retrieve the entered maize target. Due to the uncharacterized nature of this gene and its specificity to maize, it is probable that this locus contains a pseudogene that would have no affect on phytosiderophore secretion.

Part V: Characterization of GRMZM2G390374

GRMZM2G390374 is similar to the RAS-like GTPase Root Hair Defective3 (RHD3), which is involved in transport between the endoplasmic reticulum (ER) and Golgi. This protein may be involved in vesicle trafficking, which could pertain to the y3 phenotype if an exocytosis mechanism is involved with PS secretion. Sequencing of GRMZM2G390374 did not uncover any mutations within the coding region of this gene, although a 397 bp deletion was uncovered inside the 21st intron, which could potentially effect transcript splicing (Figure S3). To evaluate whether there was a defect in intron splicing due to this deletion, I amplified cDNA using primers specific to this transcript that included the intron of interest. If the intron in question were retained, we would
expect to see a 1199 bp difference in size of amplified cDNA between W22 and ys3 lines. No such size difference was observed (Figure 10), and the amplicon was sequenced for verification. Sequencing of the GRMZM2G390374 promoter region (1000 bp upstream) revealed many polymorphisms as well as a few small deletions relative to B73 (Figure, S3). To further evaluate this gene and determine that there were no promoter mutations that were affecting gene expression, I measured GRMZ2G390374 mRNA levels in both the W22 and ys3 backgrounds (Figure 11). Tissue samples were taken at the three-leaf stage of development prior to foliar application of iron ammonium citrate. Differences observed in gene expression between W22 and ys3 could be due to a variety of reasons including iron deficiency, necrosis, general stress response or a possible mutation causing over expression. W22 and ys3 do not have the same genetic background, which could also offer some explanation into the differences observed in expression. The data shows that GRMZM2G390374 was expressed, ruling out any promoter mutations preventing expression.
Figure 10: Gel electrophoresis of GRMZM2G390374 cDNA. Gel electrophoresis of cDNA spanning the 21st intron of GRMZM2G390374 amplified from both W22 and ys3 plants. Expected size 509bp.
Figure 11: qPCR analysis of maize GRMZM2G390374 (RHD3-like) normalized to GAPDH. Shoot expression was measured at the three-leaf stage 21 days past germination. Results reflect 3 biological replicates with a p-value of 0.03.
CHAPTER IV
DISCUSSION

Four of the candidate genes (GRMZM2G390345, GRMZM2G342821, GRMZM2G105766 and GRMZM2G502560) identified in the mapped region contained polymorphisms within their coding regions. Of these four genes only one, GRMZM2G390345, which codes for a protein kinase, was available as an insertional mutant from the Mu-illumina collection. Analysis of this insertional mutant, which contains a Mutator element within the third exon of GRMZM2G390345, showed that a knockout of GRMZM2G390345 does not recapitulate the ys3 phenotype. Further analysis of GRMZM2G390345 is not needed for the purpose of evaluating ys3.

Two of the candidate genes (GRMZM2G105766 and GRMZM2G342821) are predicted to be involved with intracellular calcium (Ca$^{2+}$) concentrations through calmodulin binding. Calcium exists in a gradient across the plasma membrane. In the apoplasm, extracellular calcium is either being locked up in structures such as calcium-pectate gels where the egg-box model predicts that calcium is used for crosslinking in homogalacturonan polysaccharides (Caffall and Mohnen, 2009) or is freely exchangeable at the cell wall (Marschner, 2012). Inside the cell, calcium concentrations can vary between different organelles, the majority being stored within the vacuole; concentrations inside the cytosol are very low. Signaling events often involve an influx of calcium across the plasma membrane or release of calcium from the vacuole, where the increase in cytosolic calcium can initiate or alter cellular processes. Calcium influx is required to elicit exocytosis in Arabidopsis root hairs (Schiefelbein et al., 1992) making these two
genes particularly interesting candidates if an exocytosis mechanism is involved, as opposed to a transport mechanism.

It would be interesting to see if calcium channel blockers have an effect on phytosiderophore secretion. The calcium channel blocker verapamil was used on the roots of rice and buckwheat and was shown to decrease both growth and development in a concentration dependent manner (Budagovskaya, et al., 2001). By measuring the amount of phytosiderophores secreted both before and after the application of a drug such as verapamil, it would be possible to obtain evidence that calcium is involved with phytosiderophore secretion. Measuring cytosolic Ca\textsuperscript{2+} concentrations using fluorescent dyes just prior to PS secretion may provide more insight as to whether calcium is indeed involved with phytosiderophore release.

The fourth gene that contains a polymorphism within its coding region (GRMZM2G502560) is specific to maize. No functional annotation is available for this gene and the predicted protein has no conserved domains. Searching both gene homology and gene ancestry in Phytozome resulted in no homology and no ancestry. Using Basic Local Alignment Search Tool (BLAST) to compare the protein sequence of this gene against the NCBI database, I was only able to retrieve the entered maize target. It is likely that this locus contains a pseudogene. Because this locus is specific to maize and not found in any other phytosiderophore secreting species, it is unlikely that this gene is involved with the retention of phytosiderophores in the *ys3* background. This locus is not a high priority for further analysis.

The fifth gene of interest GRMZM2G390374 has high similarity to Root Hair Defective3 (RHD3), which has been shown to be involved in vesicle transport between
the endoplasmic reticulum (ER) and Golgi apparatus. Although no polymorphisms were found within the coding region of this gene, a 397 bp deletion was found within the 21st intron of the gene. Evaluation by amplification (Primers: oGRMZM2G390374_24464…24486 and oGRMZM2G390374_26252…26231) of the cDNA between the ys3 and W22 did not reveal and size differences that would suggest retention of an intron during transcript splicing. Maize inbred lines W22 and B73 are genetically distinct and contain polymorphisms that are specific to each inbred line. Comparing sequences of W22, B73 and ys3 I have observed that ys3 shares more similarity to the W22 background within this small region of the genome, however ys3 is likely not a mutant within the W22 background. Expression of GRMZM2G390374 was much higher in iron-starved ys3 than it was in the iron replete inbred line W22. The cause of this increase in expression is ambiguous at best; however, it should be noted that over expression can also have a dramatic effect on plant phenotype. The data presented here suggest that GRMZM2G390374 is an unlikely ys3 candidate.

There is a gene of interest that lies ~7 Mbp downstream of marker UMC1773 which shares homology with a Major Facilitator Superfamily (MFS) transporter with phytosiderophore efflux activity, named Transporter of Mugineic Acid (TOM). This gene was identified in rice and barley (Nozoye et al., 2011). TOM1 in rice is expressed on a diurnal rhythm that coincides with phytosiderophore secretion and is also up-regulated when plants are starved of iron. Over expression of TOM1 led to increased tolerance to iron deficiency, while repression of TOM1 caused a decrease in the secretion of phytosiderophores (Nozoye et al., 2011). Heterologous expression of TOM1 in oocytes revealed that TOM1 is able to efflux phytosiderophores, but not nicotianamine, the
precursor to phytosiderophore synthesis. These findings imply that TOM1 plays a role in phytosiderophore secretion.

Recently there has been interest to see if maize yellow stripe 3 (ys3) carries a causative mutation in its TOM1 homolog GRMZM2G063306 (Nozoye et al., 2013). Amplification of GRMZM2G063306 (ZmTOM1) cDNA revealed retained introns in 4 different mRNA variants (Nozoye et al., 2013). Two of the mRNA variants retain a single intron. The third variant retains two introns, and the fourth variant is the predicted isoform for wildtype plants. Sequencing of ZmTOM1 should reveal causative mutations for these observed unspliced introns. ZmTOM1 has been suggested to be a ys3 candidate (Nozoye et al., 2013). This locus is worth evaluating as it is located ~7 Mbp outside of our fine mapped region (Chromosome 3: 112,042,104-112,047,482), but in the same map BIN (3.04; Fig. 12).

Preliminary sequencing of GRMZM2G063306 has uncovered a 4 bp insertion (TAAT) in the 9th exon of the gene, which was also present within the cDNA sequence (Figure 13). ZmTOM1 codes for a Major Facilitator Superfamily (MFS) transport protein. Due to the 4 bp insertion a premature stop codon is formed 12 codons downstream of the insertion site after amino acid 238 truncating the remaining 238 amino acids from the C-terminus of the native protein. The insertion disrupts the reading frame of the mRNA and changes the protein coding sequence at amino acid 227, which is located inside the MFS domain. The MFS domain is located between amino acids 42-241 of the native protein sequence. The functionality of the mutated protein could be lost due to a partial loss of the MFS domain, though this would have to be tested. To verify if this is the mutation that is causative to the ys3 phenotype further analysis is needed. No
insertion lines are available for this locus in maize.

It has been shown that phytosiderophores are secreted with equimolar amounts of potassium. Anion channel blockers restrict the amount of phytosiderophore being secreted, which suggests a related cation/anion transport mechanism (Sakaguchi et al., 1999). Major Facilitator Superfamily (MFS) transporters are believed to function mainly as proton motive force–driven secondary transporters, catalyzing uniport, symport, or antiport activities (Pao et al., 1998). An MFS transporter such as TOM1, which has been shown to efflux phytosiderophores, could potentially act as a PS/K⁺ symporter. A similar MFS transporter to TOM1, Zinc Induced Facilitator Like1 (ZIFL1) uses K⁺ as a physiological substrate consistent with ZIFL1.3’s function in the regulation of stomatal movements, since potassium fluxes are well-established regulators of guard cell turgor in response to endogenous and environmental cues (Ward et al., 2009). Along these lines, the nitrate (NO₃⁻)/nitrite (NO₂⁻) transporter NarU, which is an MFS transporter also allows potassium ions as the co-transported cation (Yan et al., 2013).
Figure 12: Map of maize chromosome 3. The locations of genomic markers umc1616 and umc1773, and GRMZM2G063306 (ZmTOM1) are depicted in red.
Figure 13: Gene model of ZmTOM1 including SNPs and indels. Green represents B73 and black represents ys3. The black and gray boxes in the gene model depict the exons of ZmTOM1. Shaded in light gray are the exons and introns that were missing from the maize genome databank (MaizeGDB). Enlarged above the gene model the cDNA nucleotide sequence spanning 9th and 10th exon. A 4bp insertion in ys3 cause a frame shift and leads to a premature stop codon 12 codons down stream of the insertion site. This mutation results in a truncated version of ZmTOM1. Two SNPs were detected between wild-type B73 and mutant ys3.
Because maize is recalcitrant to genetic transformation, I suggest the use of RNA interference (RNAi) to knockdown the syntenic ortholog of this gene in the model system *B. distachyon*. If *BdTOM1* is involved in PS secretion a knockdown might recapitulate the yellowing phenotype observed in *ys3*. Genetic complementation of *ys3* with a functional copy of GRMZM2G063306 followed by measurement of phytosiderophore secretion will provide evidence that the mutation found in *ZmTOM1* is responsible for the retention of phytosiderophores in maize mutant *ys3*.

In summary, there are two genes out of the 11 sequenced within the mapped interval that harbor mutations that could possibly be causally related to the *ys3* phenotype (GRMZM2G105766; GRMZM2G342821). No insertional mutants are available for these two loci in maize or any other grass species. Each of these genes may affect intracellular calcium concentrations. A third gene located ~7 Mbp outside our previously mapped region is also a candidate due to its homology to rice *OsTOM1*, which has been shown to efflux phytosiderophores when expressed in oocytes (Nozoye et al., 2011). The 4 bp insertion central to the *ZmTOM1* coding region shifts the reading frame of the mRNA and potentially renders this protein non-functional.

The sequence data presented here narrows the field down to one highly likely candidate (GRMZM2G063306), two potential candidates (GRMZM2G105766; GRMZM2G342821), and one unlikely candidate (GRMZM2G502560), all of which have been shown to have polymorphisms within their coding regions. The most likely candidate GRMZM2G063306 is found quite far (~7 Mbp) outside of the previously mapped region. This could be due to an error in mapping or the use of genetic markers that have not been mapped or have been incorrectly mapped back to the maize genome,
which would distort mapping efforts. The closest homologs of GRMZM2G063306 in both rice and barley have been previously shown to be involved in phytosiderophore secretion (Nozoye et al., 2011) and future efforts should focus on this locus. Alternatively the mapping data is correct and we can use the model system *B. distachyon* to knockdown the orthologs of these remaining candidate genes in effort to recapitulate the *ys3* phenotype, which would provide evidence that we have located a likely candidate.

I generated artificial microRNAs to target the orthologs of the two maize genes in *B. distachyon*. *B. distachyon* also secretes phytosiderophores to acquire iron from the environment. *B. distachyon* chromosome 4 is syntenic to *Z. mays* chromosome 3. Syntenic orthologs with high sequence similarity to GRMZM2G342821 (Bradi4G02460) 45% identity at the protein level and GRMZM2G390374 (Bradi4G02450) 82% identity at the protein level lie in sequential order on *B. distachyon* chromosome 4. Successful knockdown of these candidate genes would determine if it is possible to establish a chlorotic phenotype that is characteristic of iron deficiency.

Because no insertional mutants were available for these loci in maize or any other grass species, I used RNAi to silence gene orthologs in the model species *B. distachyon*. AmiRNAs form hairpin loops that are processed into 21 bp fragments that can be specifically designed to target a gene of interest (Warthmann et al., 2008). That fragment is then used as a probe to locate and cleave complementary native mRNAs. I constructed amiRNAs specific to Bradi4g02460 and Bradi4G02450 using the WMD3 Web MicroRNA Designer (http://wmd3.weigelworld.org/cgi-bin/webapp.cgi). The amiRNA I designed for Bradi4g02460 and Bradi4G02450 have been named amiRNA:EFhand-1 and amiRNA:RHD3-1 respectively.
Hyper-virulent agrobacterium AGL1 was transformed by electroporation with the binary vector pOL001/GW harboring amiRNA:EF-hand-1 or amiRNA:RHD3-1 driven by the maize ubiquitin promoter. This AGL1 strain was then used to create stable transgenic lines by transforming Bd21-3 embryonic callus. I confirmed that the binary vectors did in fact contain the relevant amiRNAs through PCR and sequencing of the plasmids. T2 seeds were plated on ½MS supplemented with 40 µg/ml hygromycin for selection. Also, wild-type Bd21-3 on ½MS as well as ½MS supplemented with 40 µg/ml hygromycin were used as controls. The plates were wrapped in foil and placed at 4°C for 7 days. The T2 transgenic plants from both amiRNA:EF-hand-1 and amiRNA:RHD3-1 germinated on selection as did wildtype Bd21-3 however, these seedlings quickly became necrotic. The transgenic plants did not show resistance to hygromycin and it is unclear if the hptII resistance gene is being silenced or the transgene was never inserted into the regenerated plants. This part of the project was abandoned.
Supplemental figures: S1-S5
Coding regions are highlighted in grey
Polymorphisms/amino acid substitutions are highlighted between B73 and y3
Stop codons highlighted in red

Supplemental Figure 1: Genomic alignment of GRMZM2G390345 (B73 vs y3)
GRMZM2G390345_T2 Alignment B73 vs y3
CLUSTAL O (1.2.1) multiple sequence alignment

---

**APPENDIX**

**SUPPLEMENTAL FIGURES: GENE SEQUENCING DATA**

---
Supplemental Figure 2: Genomic alignment of GRMZM2G342821 (B73 vs ys3)
GRMZM2G342821 Alignment B73 vs ys3
CLUSTAL O (1.2.1) multiple sequence alignment
Supplemental Figure 3: Genomic alignment of GRMZM2G390374 (B73 vs ys3)

**GRMZM2G390374 Alignment B73 vs ys3**

CLUSTAL O (1.2.1) multiple sequence alignment
GRMZM2G390374_B73
AGGTGGGTGGCGGTGTGCAAGGCCTCTACCCAATAGCTGGCGGGGAGAGACGCCTGGAAG
GRMZM2G390374_ys3
AGGTGGGTGGCGGTGTGCAAGGCCTCTACCCAATAGCTGGCGGGGAGAGACGCCTGGAAG

GRMZM2G390374_B73
AGAAGGCATCATATCATATTGGTGGTGGTGTGAATCATGCGCTCGGCCCGGCCATTCTGG
GRMZM2G390374_ys3
AGAAGGCATCATATCATATTGGTGGTGGTGTGAATCATGCGCTCGGCCCGGCCATTCTGG

GRMZM2G390374_B73
GTAGAGGTGTAGGGGCACGAGAGACGCAACTGAACTCCACTGGTGAGGAAGAATGACCTG
GRMZM2G390374_ys3
GTAGAGGTGTAGGGGCACGAGAGACGCAACTGAACTCCACTGGTGAGGAAGAATGACCTG

GRMZM2G390374_B73
GAGGCGTGGTTGTCGAACTCGCGACCATTGTCACACTGCAGGGCACGGACCAGGCGACAG
GRMZM2G390374_ys3
GAGGCGTGGTTGTCGAACTCGCGACCATTGTCACACTGCAGGGCACGGACCAGGCGACAG

GRMZM2G390374_B73
AATTTGGTGGAGACCCAGGCGAAGAAGTGAGGGTGGTAAAAGTGTCGGACTTCAACCGAA
GRMZM2G390374_ys3
AATTTGGTGGAGACCCAGGCGAAGAAGTGAGGGTGGTAAAAGTGTCGGACTTCAACCGAA

GRMZM2G390374_B73
GAGGAAAAGTCCAGAGAAAATGGGAAAAATCATCCAAAATTACCAAATAGTATTTATATC
GRMZM2G390374_ys3
GAGGAAAAGTCCAGAGAAAATGGGAAAAATCATCCAAAATTACCAAATAGTATTTATATC

GRMZM2G390374_B73
CGGAAAGACTAAGTATAGGCGAGGTCCAGAGGTCACAATGAACCAGGTCAAAAGTCTGCT
GRMZM2G390374_ys3
CGGAAAGACTAAGTATAGGCGAGGTCCAGAGGTCACAATGAACCAGGTCAAAAGTCTGCT

GRMZM2G390374_B73
CAGCCCGAGAAGAAGAGGTAGTAAAAGGGAGACGAGTGTGTCGGCCAAGCTGACAAGCAT
GRMZM2G390374_ys3
CAGCCCGAGAAGAAGAGGTAGTAAAAGGGAGACGAGTGTGTCGGCCAAGCTGACAAGCAT

GRMZM2G390374_B73
GACACAGGCCCTCAAAATGTCCCCTATTACAAGAAGAATCTAAACTACTGGAAAGCTTGG
GRMZM2G390374_ys3
GACACAGGCCCTCAAAATGTCCCCTATTACAAGAAGAATCTAAACTACTGGAAAGCTTGG

GRMZM2G390374_B73
TCATGACGTCAGGTCCTAGGTGGCCGAGACGTCGATGCCAACGGTGGAGGTGGTGGAGAC
GRMZM2G390374_ys3
TCATGACGTCAGGTCCTAGGTGGCCGAGACGTCGATGCCAACGGTGGAGGTGGTGGAGAC

GRMZM2G390374_B73
CAGGACGAGAGGTGGAGACACGCAGGTGGAAGAGTGCCGGAGCGTGTAGAGGGCCCCGAC
GRMZM2G390374_ys3
CAGGACGAGAGGTGGAGACACGCAGGTGGAAGAGTGCCGGAGCGTGTAGAGGGCCCCGAC

GRMZM2G390374_B73
GCTGTCACATCGGGCGAGTGGGGTCCTGGTGGCCAGATCCTTCACAGAAAAACCAGAGGG
GRMZM2G390374_ys3
GCTGTCACATCGGGCGAGTGGGGTCCTGGTGGCCAGATCCTTCACAGAAAAACCAGAGGG

GRMZM2G390374_B73
GTCAAACTCAATGGAACAAGAATTGTCGGTGGTGAACCGACGGACGAAAAGGAGATTGTG
GRMZM2G390374_ys3
GTCAAACTCAATGGAACAAGAATTGTCGGTGGTGAACCGACGGACGAAAAGGAGATTGTG

GRMZM2G390374_B73
GGTGATGTGTGGGGCTACGAGAACGTCGTTGAGATAGAACGGTCCAGGGAGCACTGAGGC
GRMZM2G390374_ys3
GGTGATGTGTGGGGCTACGAGAACGTCGTTGAGATAGAACGGTCCAGGGAGCACTGAGGC

GRMZM2G390374_B73
ACCTACTGAGGTGACCGAAAGAGTGGAACCGTTTCCAAAAACGATCGAGTGGGTGTG
GRMZM2G390374_ys3
ACCTACTGAGGTGACCGAAAGAGTGGAACCGTTTCCAAAAACGATCGAGTGGGTGTG

GRMZM2G390374_B73
GGGAGGAGGTGGGTGTGAGCAAGATAGCATGCCTTCGGTGGGGGTGGTGTGGTAGGATGC
GRMZM2G390374_ys3
GGGAGGAGGTGGGTGTGAGCAAGATAGCATGCCTTCGGTGGGGGTGGTGTGGTAGGATGC

GRMZM2G390374_B73
ACCCGGAGTCGATGACCCAGTCAGAAGCTGATAGAGCTAGCGCCATGGTGCTGTATGCGGC
GRMZM2G390374_ys3
ACCCGGAGTCGATGACCCAGTCAGAAGCTGATAGAGCTAGCGCCATGGTGCTGTATGCGGC

GRMZM2G390374_B73
GGCGAGAGAGGCTGGGTCCTAGCCTCCAGCGAACGGGGACCAGGGTGTCGTGGTAGGGGT
GRMZM2G390374_ys3
GGCGAGAGAGGCTGGGTCCTAGCCTCCAGCGAACGGGGACCAGGGTGTCGTGGTAGGGGT

GRMZM2G390374_B73
CCCGTATCGCGGAACTGCAGGAGGCGGGCGGCGGCTCGGGGGGACCAGCCACGAGCAT
GRMZM2G390374_ys3
CCCGTATCGCGGAACTGCAGGAGGCGGGCGGCGGCTCGGGGGGACCAGCCACGAGCAT
**GRMZM2G390374_B73**

```
TGTTCCTGGATAATTTCCATGCAATGGGAGATGATCCAAGGAGAATACAGTAACATGTAGTA
```

**GRMZM2G390374_ys3**

```
TGTTCCTGGATAATTTCCATGCAATGGGAGATGATCCAAGGAGAATACAGTAACATGTAGTA
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
GAACAATGGTCAACAGGATCACTTGCCCAAGCGGCATCTGAATACGCATGAAGCTGAAGT
```

**GRMZM2G390374_ys3**

```
GAACAATGGTCAACAGGATCACTTGCCCAAGCGGCATCTGAATACGCATGAAGCTGAAGT
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
GGGCTATCGGAAGCATAGAGTAAACATCTGGATCTTGTCCCACGTAAATATCGAAGCACA
```

**GRMZM2G390374_ys3**

```
GGGCTATCGGAAGCATAGAGTAAACATCTGGATCTTGTCCCACGTAAATATCGAAGCACA
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
CGAAGCAAATGACCATAATGAACTGAAGTAGGTGTTGATACAAATTGGCTCAACGTGAAGCTGAAGT
```

**GRMZM2G390374_ys3**

```
CGAAGCAAATGACCATAATGAACTGAAGTAGGTGTTGATACAAATTGGCTCAACGTGAAGCTGAAGT
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
CGATACCGAGTTGGATCTGCTAGTGGTGTCCCATCCTTCGGACGAAGTTTTAAATGTATA
```

**GRMZM2G390374_ys3**

```
CGATACCGAGTTGGATCTGCTAGTGGTGTCCCATCCTTCGGACGAAGTTTTAAATGTATA
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
TCCATCGGTGTTGCAACAGTGCGTGTATCAGTAAGTCCAGAGCGATCAAGAAGATCTTGT
```

**GRMZM2G390374_ys3**

```
TCCATCGGTGTTGCAACAGTGCGTGTATCAGTAAGTCCAGAGCGATCAAGAAGATCTTGT
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
AAATAGCTAAGCGGTCCCAAATCAGACATCTGAAATTCCTTACTAAGATGTTGCTTGACA
```

**GRMZM2G390374_ys3**

```
AAATAGCTAAGCGGTCCCAAATCAGACATCTGAAATTCCTTACTAAGATGTTGCTTGACA
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
TGAGAAATATGTTCTGAGTTGTCACCTGTAATCAGCATATCATCTACATATAATAGTAGT
```

**GRMZM2G390374_ys3**

```
TGAGAAATATGTTCTGAGTTGTCACCTGTAATCAGCATATCATCTACATATAATAGTAGT
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
AAAGTACGACCTTTTGGAGAGACATGAACAAACAATGCAGGATCATGTTCACTAGAAAAA
```

**GRMZM2G390374_ys3**

```
AAAGTACGACCTTTTGGAGAGACATGAACAAACAATGCAGGATCATGTTCACTAGAAAAA
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
AAACCAGCAGCTGTGATGACAGAGATGAATCGCTCAAACCAAGCACGAGGAGCTTGTTTC
```

**GRMZM2G390374_ys3**

```
AAACCAGCAGCTGTGATGACAGAGATGAATCGCTCAAACCAAGCACGAGGAGCTTGTTTC
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
AGCCCATACAATACGCGACGAAGACGACAAACATGACCTGAGGGAGTATCAACCCCTGGA
```

**GRMZM2G390374_ys3**

```
AGCCCATACAATACGCGACGAAGACGACAAACATGACCTGAGGGAGTATCAACCCCTGGA
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
GGAGGATGCATATAAACTTCTTCTTGCAGATCCTGCAATTAGAAAAGCATATCTTGTGACATCC
```

**GRMZM2G390374_ys3**

```
GGAGGATGCATATAAACTTCTTCTTGCAGATCCTGCAATTAGAAAAGCATATCTTGTGACATCC
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
ATCTGGGAGATTGTCCAAGAAGATGAGGCTGCAACTGCAATCAAGGTTGCAAAGCAGATGTC
```

**GRMZM2G390374_ys3**

```
ATCTGGGAGATTGTCCAAGAAGATGAGGCTGCAACTGCAATCAAGGTTGCAAAGCAGATGTC
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
ATATGCGCAACAGGTGCAAAAGTCTCATCATAATCAAGTCCTTAAGTCTGTTGAAAACCA
```

**GRMZM2G390374_ys3**

```
ATATGCGCAACAGGTGCAAAAGTCTCATCATAATCAAGTCCTTAAGTCTGTTGAAAACCA
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
CGGGCCACAAGACAAGCTTTATACCTCTGAATGGAACCATTTGATTTTGTCTTAACTTTG
```

**GRMZM2G390374_ys3**

```
CGGGCCACAAGACAAGCTTTATACCTCTGAATGGAACCATTTGATTTTGTCTTAACTTTG
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
AAGACCCATTTAATGGTATAGGAGAACAATCAGGAGGTTAATAGCAACATCTCCATGTA
```

**GRMZM2G390374_ys3**

```
AAGACCCATTTAATGGTATAGGAGAACAATCAGGAGGTTAATAGCAACATCTCCATGTA
```

```
******************************************************************************
```

**GRMZM2G390374_B73**

```
CCAGTGCGTTCAAGAGCAGCCAATTCATCATAATGCTAGCTGCCACTCAAGAATCCCA
```

**GRMZM2G390374_ys3**

```
CCAGTGCGTTCAAGAGCAGCCAATTCATCATAATGCTAGCTGCCACTCAAGAATCCCA
```

```
******************************************************************************
```
GRMZM2G390374_B73
ATATGTGCTACTGTAAAATGGCGCAGAAAAATGAGTTAACAGCCCAGTATAAAC

GRMZM2G390374_ya3
ATATGTGCTACTGTAAAATGGCGCAGAAAAATGAGTTAACAGCCCAGTATAAAC

GRMZM2G390374_B73
GCGGAGAGCTCAATGGCAACCAAGAAGTATTGAATATGCAAGAGTTGCTATAGGCA

GRMZM2G390374_ya3
GCGGAGAGCTCAATGGCAACCAAGAAGTATTGAATATGCAAGAGTTGCTATAGGCA

GRMZM2G390374_B73
CGACTAGACGGATGATTTGGATGACTTGGAAATCTGGTCGCGTCGTCCTTCTACCTGCAA

GRMZM2G390374_ya3
CGACTAGACGGATGATTTGGATGACTTGGAAATCTGGTCGCGTCGTCCTTCTACCTGCAA

GRMZM2G390374_B73
ACTAACAGGCGACCAGCCCGACTAATCGCGATTAGTCGGACGACTTGTAATGCACAAGCA

GRMZM2G390374_ya3
ACTAACAGGCGACCAGCCCGACTAATCGCGATTAGTCGGACGACTTGTAATGCACAAGCA

GRMZM2G390374_B73
ACTGACTAGTATTGGTAAGAATTTGGTGTCTTGGTGTGCTTGCAAGCAGCCCACAATCTC

GRMZM2G390374_ya3
ACTGACTAGTATTGGTAAGAATTTGGTGTCTTGGTGTGCTTGCAAGCAGCCCACAATCTC

GRMZM2G390374_B73
AAGATTAGGCACTGAAGCTGAATATAAAGCTATGGCCAACACAACTGTAAGAATTGATGTG

GRMZM2G390374_ya3
AAGATTAGGCACTGAAGCTGAATATAAAGCTATGGCCAACACAACTGTAAGAATTGATGTG

GRMZM2G390374_B73
GGTGACGTCCCTTTTACAAAGACGCTAGTATTGGTCGTTCTCCCTCCTACCTGCAAGGAGGTTAGCA

GRMZM2G390374_ya3
GGTGACGTCCCTTTTACAAAGACGCTAGTATTGGTCGTTCTCCCTCCTACCTGCAAGGAGGTTAGCA

GRMZM2G390374_B73
TGACACTGTTGGAGCAACCTATCTTACAGCAAATCCATTGTTTTATAGTCATATGAAGCA

GRMZM2G390374_ya3
TGACACTGTTGGAGCAACCTATCTTACAGCAAATCCATTGTTTTATAGTCATATGAAGCA

GRMZM2G390374_B73
TCTTGAGATTGGTTTTTATTTTATGAGAGGGAGTAGCTAGAGGACTTTCGGAGGTCAGAA

GRMZM2G390374_ya3
TCTTGAGATTGGTTTTTATTTTATGAGAGGGAGTAGCTAGAGGACTTTCGGAGGTCAGAA

GRMZM2G390374_B73
TCATATCTACAAAAGATCAAGTTGTTGATGGCTTTACCAAGATAATAACTATACAGATTA

GRMZM2G390374_ya3
TCATATCTACAAAAGATCAAGTTGTTGATGGCTTTACCAAGATAATAACTATACAGATTA

GRMZM2G390374_B73
CAGAAGATGGCTTTATCTAGAGACAATCTTAACTTGTACAAGTTATGATTGAGGGGGGTG

GRMZM2G390374_ya3
CAGAAGATGGCTTTATCTAGAGACAATCTTAACTTGTACAAGTTATGATTGAGGGGGGTG

GRMZM2G390374_B73
TTATAATAGAATGTGCTTGTCAGGCTATGGCCGTATCAGATCAGGGGTCCACATGGTTGT

GRMZM2G390374_ya3
TTATAATAGAATGTGCTTGTCAGGCTATGGCCGTATCAGATCAGGGGTCCACATGGTTGT

GRMZM2G390374_B73
TCTATAATAGAAGAAGACCTTAACCCGACTCCGATTAAAGATAGGATAGAATCAGGTTTG

GRMZM2G390374_ya3
TCTATAATAGAAGAAGACCTTAACCCGACTCCGATTAAAGATAGGATAGAATCAGGTTTG

GRMZM2G390374_B73
ATCTATCTCAGAGATAGCAGAGTTAGAATTGGAGTCGGTTAAGTTAGTTGTACCTGAATT

GRMZM2G390374_ya3
ATCTATCTCAGAGATAGCAGAGTTAGAATTGGAGTCGGTTAAGTTAGTTGTACCTGAATT

GRMZM2G390374_B73
GGAATCCTTGTTATGTTGTAATCTTTATGGCAATATTCTGCTCTACATATACTATACAT

GRMZM2G390374_ya3
GGAATCCTTGTTATGTTGTAATCTTTATGGCAATATTCTGCTCTACATATACTATACAT

GRMZM2G390374_B73
CCTCGGCCTAATAGGTCCGGGTGAACAGCTTAAATCGGATGTTGTGTTTTCTTACACTAT

GRMZM2G390374_ya3
CCTCGGCCTAATAGGTCCGGGTGAACAGCTTAAATCGGATGTTGTGTTTTCTTACACTAT

GRMZM2G390374_B73
GACCCATGATAGATCACTACACCGATCCTTGGAGGGGCTCCTGAAGAATACTTACATCC

GRMZM2G390374_ya3
GACCCATGATAGATCACTACACCGATCCTTGGAGGGGCTCCTGAAGAATACTTACATCC

GRMZM2G390374_B73
TCTGTTTCAAATTATAATTTATTTTGTTTTTAGATACATGAATTTTGGTAGGTGCATAG

GRMZM2G390374_ya3
TCTGTTTCAAATTATAATTTATTTTGTTTTTAGATACATGAATTTTGGTAGGTGCATAG

GRMZM2G390374_B73
CAAAATTGAGATAGAATACAGGATACGAGGACCCAGCAATCCTTGAAGAATAGATAGATATTA

GRMZM2G390374_ya3
CAAAATTGAGATAGAATACAGGATACGAGGACCCAGCAATCCTTGAAGAATAGATAGATATTA

81
Supplemental Figure 4: Genomic alignment of GRMZM2G502560 (B73 vs ys3)

GRMZM2G502560 Alignment B73 vs ys3

CLUSTAL O (1.2.1) multiple sequence alignment

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAGATGTTGTTTCCTCCTTCCTTCTTTATATTTAACAAAAGATTATATATATACCTTCGA</td>
<td>AAGATGTTGTTTCCTCCTTCCTTCTTTATATTTAACAAAAGATTATATATATACCTTCGA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTTCGTAAGATAGGATAGGAGTAAAGTAAATATATACAGTCTGAGGCGATATAATA</td>
<td>CTTCGTAAGATAGGATAGGAGTAAAGTAAATATATACAGTCTGAGGCGATATAATA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCCATTATTATTTCTCTCACTCTCAAAAGAGCTGTGCTGGACAGCAAGAATGTAACAGCG</td>
<td>GCCATTATTATTTCTCTCACTCTCAAAAGAGCTGTGCTGGACAGCAAGAATGTAACAGCG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>TATTTTTAACTGCAGAGCTGTGTTCACCCCTCATGTGCACATGCTGTTACTCCGAAGCTT</td>
<td>TATTTTTAACTGCAGAGCTGTGTTCACCCCTCATGTGCACATGCTGTTACTCCGAAGCTT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATACAATTGAGCTAGCAAATATATTAGGATACTGCAGTCAATTTCTATAGCTGTCTTGTT</td>
<td>ATACAATTGAGCTAGCAAATATATTAGGATACTGCAGTCAATTTCTATAGCTGTCTTGTT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GACCTCTCATGGTTTTGGCTTGTTCACCCCTCATGTGCACATGCTGTTACTCCGAAGCTT</td>
<td>GACCTCTCATGGTTTTGGCTTGTTCACCCCTCATGTGCACATGCTGTTACTCCGAAGCTT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTTGGGTATTCTACCTGCAACACTTATTAATTCAAGCTACAAAACAGTTGTCGAGTTAGT</td>
<td>CTTGGGTATTCTACCTGCAACACTTATTAATTCAAGCTACAAAACAGTTGTCGAGTTAGT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCGCTTTTGTATTCGTTGTTGCATTGCAACCATCCCCGCCGAAGGTGAGCCGAAGGTAAT</td>
<td>GCGCTTTTGTATTCGTTGTTGCATTGCAACCATCCCCGCCGAAGGTGAGCCGAAGGTAAT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTTTTTTTTTACCTTCGAAAAGAAGACTTCCGGCAATGCACAACTTCCCATCTGCATTATC</td>
<td>TTTTTTTTTTACCTTCGAAAAGAAGACTTCCGGCAATGCACAACTTCCCATCTGCATTATC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTGAGCAGGATTGTAGGCACACAGTGATGACGAAGACAGAGACAGCAATATACACAACCG</td>
<td>GTGAGCAGGATTGTAGGCACACAGTGATGACGAAGACAGAGACAGCAATATACACAACCG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCCATTATTATCACGGATAATTAATCAACAGTGAGTGACAGAAGAAGAGAATACAGTGAC</td>
<td>GCCATTATTATCACGGATAATTAATCAACAGTGAGTGACAGAAGAAGAGAATACAGTGAC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCCATTATTATCACGGATAATTAATCAACAGTGAGTGACAGAAGAAGAGAATACAGTGAC</td>
<td>GCCATTATTATCACGGATAATTAATCAACAGTGAGTGACAGAAGAAGAGAATACAGTGAC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCCATTATTATCACGGATAATTAATCAACAGTGAGTGACAGAAGAAGAGAATACAGTGAC</td>
<td>GCCATTATTATCACGGATAATTAATCAACAGTGAGTGACAGAAGAAGAGAATACAGTGAC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTTTTTTTTTACCTTCGAAAAGAAGACTTCCGGCAATGCACAACTTCCCATCTGCATTATC</td>
<td>TTTTTTTTTTACCTTCGAAAAGAAGACTTCCGGCAATGCACAACTTCCCATCTGCATTATC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GRMZM2G502560-B73</th>
<th>GRMZM2G502560_ys3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCCATTATTATCACGGATAATTAATCAACAGTGAGTGACAGAAGAAGAGAATACAGTGAC</td>
<td>GCCATTATTATCACGGATAATTAATCAACAGTGAGTGACAGAAGAAGAGAATACAGTGAC</td>
</tr>
</tbody>
</table>
GRMZM2G502560-B73  ATTACCCTCTCTCTCTCTCCTCTCATCTTGGATTCGTCGATCCATATATGACA
GRMZM2G502560_ys3  ATTACCCTCTCTCTCTCTCCTCTCATCTTGGATTCGTCGATCCATATATGACA

GRMZM2G502560-B73  GTCAGGGACGGGGGAGAGAGAGAGTGACAGGGGCCGGTAGTAGTATAGATTACATCCA
GRMZM2G502560_ys3  GTCAGGGACGGGGGAGAGAGAGAGACAGGGCCGG-----TAGTA-----------GTAT

GRMZM2G502560-B73  TCTTACATATACCACCACCATCATAACCAGATCATGCTGGCAGCTTCACCAACTCGTGGT
GRMZM2G502560_ys3  GA

GRMZM2G502560-B73  GCACCACTAC--CATACCCTCTCGTCTGATCCAAACGGAGGAAGGAGGAAGAAGAGCTAGC
GRMZM2G502560_ys3  GCACCACTATACATACCCTCTCGTCTGATCCAAACGGAGGAAGGAGGAAGAAGAGCTAGC

GRMZM2G502560-B73  TATCC
GRMZM2G502560_ys3  TATCC

89
Supplemental Figure 5: Genomic alignment of GRMZM2G105766 (B73 vs ys3)

GRMZM2G105766 Alignment B73 vs ys3

CLUSTAL O (1.2.1) multiple sequence alignment

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************

GRMZM2G105766_B73
GRMZM2G105766_ys3

************************************************************
**Supplemental Figure 6:** Genes presented on chromosome 3 between (118,288,000-119,675,000) that were not targeted for sequence analysis.

<table>
<thead>
<tr>
<th>Gene Identifier</th>
<th>Annotation</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRMZM2G587694</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G173996</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G587701</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>AC212472_3_FG003</td>
<td>No data available</td>
<td>Bundle sheath cells</td>
</tr>
<tr>
<td>AC212472_3_FG001</td>
<td>No data available</td>
<td>Bundle sheath cells, Embryo</td>
</tr>
<tr>
<td>AC212472_3_FG002</td>
<td>No data available</td>
<td>Anthers, Pollen</td>
</tr>
<tr>
<td>GRMZM5G875963</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G156907</td>
<td>No data available</td>
<td>Anthers, Mature tassel</td>
</tr>
<tr>
<td>GRMZM2G577933</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G403475</td>
<td>glycosylhydrodrolase activity!</td>
<td>Embryo, Un-pollinated ear</td>
</tr>
<tr>
<td>AC183510_4_FG001</td>
<td>No data available</td>
<td>Bundle sheath cells</td>
</tr>
<tr>
<td>GRMZM2G519927</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G444597</td>
<td>No data available</td>
<td>Seedling roots</td>
</tr>
<tr>
<td>GRMZM2G144004</td>
<td>No data available</td>
<td>Seedling roots</td>
</tr>
<tr>
<td>GRMZM2G569191</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G569189</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G444621</td>
<td>Calmodulin</td>
<td>Silk</td>
</tr>
<tr>
<td>GRMZM2G144011</td>
<td>No data available</td>
<td>Embryo, Developing leaf</td>
</tr>
<tr>
<td>GRMZM2G569195</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G010955</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G569202</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G144010</td>
<td>Transposable element</td>
<td>Mesophyll cells, Embryo</td>
</tr>
<tr>
<td>GRMZM2G490255</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G313137</td>
<td>No data available</td>
<td>Embryo, Ear, Seed</td>
</tr>
<tr>
<td>GRMZM2G357722</td>
<td>Transposable element</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G560567</td>
<td>Transposable element</td>
<td>No data available</td>
</tr>
<tr>
<td>AC212014_3_FG002</td>
<td>No data available</td>
<td>Leaf meristem</td>
</tr>
<tr>
<td>GRMZM2G579601</td>
<td>Transposable element</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G169850</td>
<td>Transposable element</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM5G807081</td>
<td>No data available</td>
<td>No data available</td>
</tr>
<tr>
<td>GRMZM2G342819</td>
<td>No data available</td>
<td>Mesophyll cells</td>
</tr>
<tr>
<td>GRMZM2G390371</td>
<td>No data available</td>
<td>Anthers, Pollen</td>
</tr>
<tr>
<td>GRMZM5G845752</td>
<td>No data available</td>
<td>Bundle sheath cells</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


Yan, Hanchi, Weiyun Huang, Chuangye Yan, Xinqi Gong, Sirui Jiang, Yu Zhao, Jiawei Wang, and Yigong Shi. “Structure and Mechanism of a Nitrate Transporter.” *Cell Reports* 3, no. 3 (2013): 716–23. doi:10.1016/j.celrep.2013.03.007.