Cloning and validation of *YELLOW STRIPE* 3 (*Ys3*) and its implication on iron metabolism in maize (*Zea mays* L.)

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

Der Mathematisch-Naturwissenschaftlichen Fakultät

Der Universität zu Köln

Vorgelegt von

María Gabriela Ronquillo-López

aus Ambato, Ecuador

Köln, 2013

Die vorliegende Arbeit wurde am Max-Planck-Institut für Pflanzenzüchtungsforschung in Köln in der Abteilung für Pflanzenzüchtung und Genetik (Prof. Dr. Maarten Koornneef) angefertigt.

Berichterstatter:

Prof. Dr. Maarten Koornneef
Prof. Dr. Marcel Bucher
Prof. Dr. Frank Hochholdinger

Prüfungsvorsitzender:

Prof. Dr. Wolfgang Werr

Tag der Disputation: 27. Juni.2013.

To my dear parents Cecilia and Galo,

to my sisters Paola and Cristina,

to my lovely nephews Hendrik, Niels, Cailan, and Conrad,

who have filled my days with happiness and love.

Dedicated to my family,

-Gabriela-

Abstract

In plants, iron (Fe) deficiency leads to chlorosis, reduced yield, and decreased nutritional quality. Graminaceous plants follow strategy II, in which chelating substances called "phytosiderophores" are produced and secreted into the rhizosphere. The Fe(III)-phytosiderophore complex is then taken up by the specific transporter, YELLOW STRIPE 1. The phytosiderophore in maize is 2'-deoxymugineic acid (DMA). In maize, the mutant *ys3* is characterized by an intervenial chlorosis due to a defect in phytosiderophore secretion. Understanding genome-wide gene regulation upon Fe stress in $ys3$-background plants will provide important insides about $Ys3$ and its implication on Fe homeostasis. Map based cloning located the *Ys3* gene in a 0.8 cM interval on chromosome 3 spanning 13.59 Mbps and containing 207 high confidence gene models. However, only 50 genes were present in the maize, rice, and sorghum genomes. Within this subset of candidate genes, GRMZM2G063306 was predicted to encode a DMA efflux transporter orthologous to *OsTOM1* and *HvTOM1*. The *Ys3* gene then was sequenced in plants of the *ys3* mutant and wt showing 10 SNPs and 3 InDels in the coding sequence. However, only 2 InDels and 2 synonymous SNPs were unique for the ys3 mutant. Isolation of one novel $ys3::Mu$ and four novel $ys3::Ac$ novel *ys3* mutations by direct

transposon tagging confirmed the candidate gene GRMZM2G063306. An additional F_1S_1 *ys3::Mu* individual showed a 6 bp insertion in exon 8, resulting in the insertion of two amino acids in the sequence. Increased expression levels of *Ys3_T01*, *Ys3_T02*, *Dmas*, and *Ys1* was shown in root tissue of the *ys3* mutant and wild type plants grown under Fe deficient conditions, in comparison with Fe sufficient conditions. Furthermore, a transcriptome profile of *ys3*×W22 F_2 individuals grown under deficient and sufficient Fe regimes revealed the immediate response of several Fe uptake and homeostasis genes along with bHLH transcription factors including GRMZM2G057413 (*ZmIro2*) and GRMZM2G350312 (*ZmIro3*) as well as novel candidate genes associated with transport, oxidation-reduction, and to the NAS family. In addition, phenotypic and ionomic analyses were carried out to complement the transcriptome profile and thus, provide a complete and deep overview of gene response during Fe stress. This study revealed that *Ys3* encodes a protein syntenic to rice and barley TOM1, which is in line with its predicted function as the specific transporter for phytosiderophore efflux in maize. Furthermore, it also provides important insides about *Ys3* and its implication on Fe homeostasis by investigating its response when grown under deficient and sufficient Fe regimes, which can later be used to improve Fe efficiency and thus, influence Fe content in grain to fight Fe deficiency in humans.

Zusammenfassung

In Pflanzen wird durch Eisenmangel Chlorose, Ernteausfälle und Nährwertreduktion verursacht. Zu den Gräsern gehörige Pflanzen führen Strategie II durch, wobei Chelatbildner, die so genannten Phytosiderophore, exprimiert und in die Rhizosphäre sekretiert werden. Der Fe (III) Phytosiderophor-Komplex wird dann von einem spezifischen Transporter, YELLOW STRIPE 1, aufgenommen. Das Phytosiderophor in der Maispflanze ist 2'-Deoxymugineic acid (DMA). Die Maismutante *ys3* weißt Defekte in der Phytosiderophor-Auscheidung auf, die eine Chlorose zwischen den Blattadern bewirken. Eine verbessertes Verständnis der Steuerung der Genexpression nach Eisenstress in *ys3*-Mutanten wird Aufschluss über *Ys3* und seine Funktion im Eisenhaushalt der Pflanze geben.

Map based cloning ergab, dass das *Ys3* Gen in einem 0.8 cM Intervall auf Chromosom 3 liegen muss und eine Region von 13.59 Mbps mit 207 sicher vorhergesagten Genen umspannt. Im Mais-, Reis- und Hirsegenom sind jedoch nur jeweils 50 dieser Genen vorhanden. In dieser Gruppe von Kandidatengenen wurde GRMZM2G063306 als DMA-Efflux Transporter identifiziert, der ortholog zu *OsTOM1* und *HvTOM1* ist. Das *Ys3*-Gen wurde daraufhin in *ys3*-Mutanten und Pflanzen des Wildtyps sequenziert und zehn Single Nucleotid Polymorphismen (SNPs) sowie drei InDels konnten in der kodierenden Sequenz entdeckt werden. Davon kamen jedoch lediglich zwei

InDels und zwei synonyme SNPs nur bei der *ys3*-Mutanten vor. Durch die Isolierung von einer neuartigen *ys3*::*Mu* und vier neuartigen *ys3*::*Ac* Mutationen durch direktes Transposon-Tagging konnte GRMZM2G063306 als Kandidatenlocus bestätigt werden. Ein zusätzliches F_1S_1 *ys3*::*Mu* Individuum wieß eine 6 bp Insertion in Exon 8 auf, die in eine Insertion von zwei Aminosäuren in die Sequenz resultierte. Im Wurzelgewebe der *ys3*-Mutante und von Pflanzen des Wildtyps konnte unter Eisenmangel eine Erhöhung der Expression von $Ys3\_T01$, *Ys3_T02*, *Dmas* und *Ys1* im Vergleich zu eisenreichen Bedingungen beobachtet werden. Zudem ergab eine Analyse des Transkriptoms von *ys*3×W22 F_2-Individuen, die unter Eisenmangel sowie eisenreichen Bedingungen gezogen wurden, eine sofortige Antwort verschiedener Gene, die in der Eisenaufnahme und im Eisenhaushalt involviert sind. Darunter waren auch verschiedene bHLH Transkriptionsfaktoren wie z.B. GRMZM2G057413 (*ZmIro2*) und GRMZM2G350312 (*ZmIro3*) sowie weitere, neue Kandidatengene, die mit Transport oder Redoxprozessen assoziiert sind oder zur NAS-Familie gehören. Desweiteren wurden zur Komplementierung des Transkriptomprofiles phänotypische und ionomische Analysen durchgefaührt und ermöglichten somit einen vollständige und tiefgehenden Überblick über die Genantwort während des Eisenstresses. Diese Studie ergab, dass das *Ys3*-Gen für ein Protein kodiert, das syntenisch zu TOM1 in Reis und Gerste ist, was mit der vorhergesagten Funktion als spezifischer Phytosiderophor-Transporter in Mais übereinstimmt. Zudem gaben Experimente unter eisenreichen und Eisenmangel Bedingungen Aufschluss über *Ys3* und seine Rolle im Eisenhaushalt. Diese Ergebnisse können in der Zukunft zu einer Verbesserung der Eiseneffizienz beitragen, den Eisengehalt in Getreide beeinflussen und somit Probleme des Eisenmangels im Menschen lösen.

Contents

List of figures

List of tables

Glossary

Following are listed the abbreviations used in this study. Gene names are in italics and capital letters and mutants are in lowercase.

3'	Three prime end of a DNA fragment
5'	five prime end of a DNA fragment
Ac	*Activator element*
B	Boron
BAC	Bacterial artificial chromosome
Ca	Calcium
CAPS	Cleaved amplified polymorphic sequences
CDS	Coding sequence
cM	Centimorgan
Cu	Copper
DEG	Differentially expressed gene
DMA	2'-deoxymugineic acid
Fe	Iron
FGS	Filtered gene set
GO	Gene ontology
Indel	Insertion / deletion
K	Potassium
Kbp	Kilo base pair
Mg	Magnesium
Mn	Manganese
Mu	*Mutator element*
Na	Sodium
NAM	Nested association mapping population
P	Phosphorus
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
S	Sulfur
SNP	Single nucleotide polymorphism
SPAD	Relative chlorophyll content
SSR	Simple sequence repeat or microsatellite
TOM1	*Transporter of mugineic acid family 1*
UTR	Untranslated region
WGS	Working gene set
ys3	*yellow stripe 3* mutant
Zn	Zinc

Acknowledgments

Finally, it is time to express my sincere gratitude to everyone, who have contributed in any way to this research, without any of you this PhD would have not been possible.
I am also very grateful to Prof. Dr. Maarten Koornneef for being an active supporter in the progress of my project. Thanks for your questions, discussions, and suggestions as well as for the funding during the last period of this project. Also, special thanks for being part of my defense committee.
I would also like to express my gratitude to Prof. Dr. Frank Hochholdinger for his support, advice, and encouragement in the development of this project and also for being part of my defense committee.
Special thanks to the members of my thesis committee Prof. Dr. Marcel Bucher and Prof. Dr.Wolfgang Werr. I am very glad you were able to be part of my defense.
My sincere gratitude to Dr. Benjamin Stich for opening the doors of his group and giving me the opportunity to work with the famous and challenging *ys*3 and allow me to discover the interesting world of maize genetics. Thanks you so much for your supervision, patience, and scientific discussions. All the experiences I obtained along this project will for sure support my future professional career.
I am very thankful with the current and former member of the Quantitative Crop Genetics group (AG Stich) for a nice working atmosphere. Special thanks to Andrea, Nele,and Nicole for their help either in the lab or the greenhouse. Many, many thanks to Isabell Scheibert "Isa" for being such a great friend and for the endless hours in the lab, the trips to Meckenheim, your dedication, and support. Thanks a lot to Enrico for his help with the hydroponics, those were a lot of liters of nutrient solution! My deepest thanks to Andi, for his help with R and LaTex! (it was VERY appreciated), the endless scientific/philosophical discussions about Fe and phytosiderophores, Family Guy, and your friendship.
My gratitude goes to Claude Urbany for his help and support, answering my question, the correction of this dissertation, your supervision and encourage-

ment when I needed the most! I am very sure that one day (I hope in a very close future) you would be an amazing supervisor! My thanks also go to Dr. Olof Persson for his support and help during these years as part of the IMPRS.
I also want to thank everyone at the MPIPZ who worked hard to contribute to this research. In this regard, thanks to the people at the MPIPZ Genome Center, gardeners, and technicians. Also, I would like to thank to Dr. Thorsten Kraska (University of Bonn) and the personnel at Campus Klein-Altendorf (University of Bonn) for their technical assistance.
Finally, I want to say thanks to my dear friends Cami and Joha Misas for always being there!. Muchas gracias por su apoyo, ayuda, y amistad! Las llevaré siempre en mis recuerdos!
Muchas gracias a mi familia en especial a mis papitos por su amor incondicional, consejo, y apoyo durante estos últimos años. A mis hermanas Paola y Crisitina. En especial, mi profundo agradecimiento va para Paola y Torben por abrirme las puertas de su casa cada vez y siempre! Gracias a mis sobrinos Niels, Hendrik, Cailan y Conrad por darle a mi vida muchos momentos de alegría y amor. Gracias a todos y cada uno de ustedes por formar parte de esta aventura! Finalmente lo logramos! Los amo mucho!
This research was funded by the Max Planck Society, the International Max Planck Research School (IMPRS), and the Deutsche Forschungsgemeinschaft (DFG).

Erklärung

Die vorliegende Arbeit wurde am Max-Planck-Institut für Züchtungsforschung in Köln-Vogelsang durchgeführt.

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit einschließlich Tabellen, Karten und Abbildungen , die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie noch nicht verffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Herr Prof. Dr. Maarten Koornneef betreut worden.

Köln, im Mai 2013

María Gabriela Ronquillo-López

Lebenslauf

Angaben zur Person

Name:	María Gabriela Ronquillo-López
Geburtsdatum:	25 Oktober, 1983
Geburtsort:	Ambato, Ecuador
Nationalität:	Ecuadorianisch

Schulbildung

07/2010	**Abitur** an der *Unidad Educativa de La Inmaculada*, Quito, Ecuador.

Studium

10/2009–aktuell	**Promotions-Studium** an der Universität zu Köln, Köln, Deutschland. Max-Planck-Institut für Pflanzenzüchtungsforschung, unter der Leitung von Prof. Dr. Maarten Koornneef und Dr. Benjamin Stich. *Stipendium* der Max-Planck-Gesellschaft (MPG)
01/2007-09/2009	**Postgraduierten Studium** zum "Master of Science / MSc. Plant Breeding and Plant Genetics" an der University of Wisconsin-Madison, Madison, USA, unter der Leitung von Prof. James Nienhuis. *Stipendium* des U. S. Department of Agriculture (USDA)
01/2002–12/2005	**Bachelor of Science "BSc. Studium Agrarwissenschaften"** an der Escuela Agrícola Panamericana (EAP) Zamorano, Tegucigalpa, Honduras. *Abschluss*: Bachelor of Science in Agrarwissenschaften. *Stipendium* der Swiss Agency for Development and Cooperation and Villaseca Fundation.

Praktika

05/2006-09/2006	**Wissenschaftliche Mitarbeiterin** an der University of Wisconsin-Madison, Madison, USA. Abteilung Pflanzenzüchtung und Genetik unter der Leitung von Prof. James Nienhuis.
	Stipendium der Pioneer Foundation
02/2004–02/2005	**Bachelor Arbeit** am EAP Zamorano, Tissue Culture und Micropropagation Labor, unter der Leitung von Prof. Dinnie de Rueda.

Publikationen

Ronquillo-López, M.G., C.R. Grau, and J. Nienhuis. 2010. Variation in reaction to *Fusarium spp.* identified in a common bean (*Phaseolus vulgaris* L.) population developed for field-based resistance to root rot and wilt. Crop Sci. 50:2303-2309.

Ronquillo-López, M.G., C.R. Grau, and J. Nienhuis. 2009. Heritability of resistance to *Fusarium oxysporum* f. sp. *Phaseoli* in common bean. BIC 52:96-97.

Chapter 1

Introduction

Maize (*Zea mays* L.)is one of the most extensively cultivated crops with 169.5 million hectares worldwide and producing 880.5 million tons in 2011 (www.fas.usda.gov/psdonline/). Maize is mainly used to feed animals, but also for human consumption and biofuel production. Approximately 15% of grain production is used for food (Awika, 2011). Maize has also been used as a model system for the identification and characterization of genes underlying a respective phenotype as well as for the study of genome-wide transcription profiles in several developmental stages (Curie et al., 2001; Li et al., 2010b; Sturaro et al., 2005; Wen et al., 2005; Sekhon et al., 2011; Li et al., 2010a; Davidson et al., 2011; Hansey et al., 2012).

1.1 Iron (Fe) importance and the effects of its deficiency

In spite of its consumption as one of the most important staple crops, maize lacks several important amino acids and micronutrients including lysine, tryptophan, zinc, and iron (Fe). The most common micronutrient malnourishment in humans are the lack of vitamin A, zinc, and Fe (Black, 2003). In humans, Fe deficiency and Fe deficiency anemia (IDA) are estimated to affect 25% and 50% of the world's population (Conte and Walker, 2011). IDA is mostly prevalent in developing countries and frequently exacerbated by infectious diseases (Stoltzfus, 2001; WHO, 2008). In plant production, iron (Fe) deficiency can lead to cholorosis, reduced yield, and a decreased nutritional quality (Curie et al., 2001). Fe deficiency in maize caused by growth on calcareous and high-pH soils can lead to a reduction of grain yield up to 20% (Godsey et al., 2003). Fe is the fourth element most found in the lithosphere comprising approximately 5% (Briat et al., 2006). In soils with a neutral pH value and in the presence of oxygen, Fe is mainly found in its oxidized form Fe (III), which has a low solubility and therefore a low bioavailabily for plants (Briat and Lobreaux, 1997; Curie and Briat, 2003; Thomine and Vert, 2013).

In fact, Fe plays an important role as a cofactor in many crucial metabolic pathways involving electron-transfer including photosynthesis, respiration, nitrogen fixation, hormone synthesis, and DNA synthesis (Briat

et al., 1995; Lobreaux et al., 1992; Conte and Walker, 2011), and thus, esencial for cell metabolism in living organisms.

1.2 Strategies for Fe uptake in plants

Dicotyledonous plants follow strategy I to take up iron (Römheld 1987; Fig. 1.1A). Plant species belonging to this class release protons via the plasma membrane H^+-ATPase, which is in *Arabidopsis thaliana* under the control of the $AHA2$ gene (Kobayashi and Nishizawa, 2012a). Furthermore, expression of $FERRIC\ REDUCTASE\ OXIDASE\ 2$ ($FRO2$) leads to a reduction of Fe (III) to Fe (II). Fe (II), which is more soluble than Fe (III) is later transported into the plant by the $IRON\ REGULATED\ TRANSPORTER\ 1$ ($IRT1$) (Curie and Briat, 2003; Guerinot, 2001; Walker and Connolly, 2008).

In contrast, graminaceous plants including rice, barley, and maize follow strategy II for Fe acquisition (Römheld 1987; Fig. 1.1B). These plant species produce chelating substances called phytosiderophores, which are low-molecular weight compounds. Phytosiderophore syntesis consist of a sequel of enzymatic reaction that are part of the *S*-adenosyl-L-methionine (SAM) pathway, in which nicotianamine synthase (NAS), nicotianamine aminotransferase (NAAT), and deoxymugineic acid synthase (DMAS) generate 2'-deoxymugineic acid (DMA) (Shojima et al., 1990; Kobayashi and Nishizawa, 2012a; Suzuki et al., 2006). Maize releases the phytosiderophore called DMA into the apical root area in response to iron deficient conditions. Recently, the

rice and barley *TRANSPOTER OF MUGINEIC ACID FAMILY* 1 (*TOM*1) were identified with a high-resolution microarray analysis to be specific transporters for DMA efflux (Nozoye et al., 2011). Thus, Fe(III)-phytosiderophore complexes are then transported into the plant by the specific transporter YELLOW STRIPE 1 (YS1) into the root plasmalemma (Curie et al., 2001; Curie and Briat, 2003; Guerinot, 2001; Lanfranchi et al., 2002; Von Wirén et al., 1994; Walker and Connolly, 2008).

Rice and barley *TOM*1 were upregulated in root tissues under Fe deficient conditions. In addition, tolerance to Fe deficiency was increased when *TOM*1 was overexpressed and decreased when it was repressed providing a strong evidence for its function (Nozoye et al., 2011). In maize, the *yellow stripe* 3 (*ys*3) mutant (Beadle, 1929; Motta et al., 2001; Wright, 1961) is characterized by an intervenial chlorosis due to a defect in phytosiderophore secretion rather than phytosiderophore biosynthesis (Lanfranchi et al., 2002). The wild type phenotype of *ys*3 plants can be restored by co-cultivation with wild type plants or by applying Fe chelates (Basso et al., 1994; Beadle, 1929; Curie and Briat, 2003; Motta et al., 2001). The *Ys*3 locus is located on chromosome 3 and is recessively inherited (Beadle, 1929; Motta et al., 1999). Recently, *ZmTOM*1 was identified by a semiquantitative reverse transcription (RT-PCR) analysis, in which unspliced introns were detected in the *ys*3 mutant, suggesting that *ZmTOM*1 might be involved with the *ys*3 phenotype. However, neither *Tom*1 nor *Ys*3 has been identified as the specific transporter responsible for the efflux of phytosiderophores (Nozoye et al.,

2011).

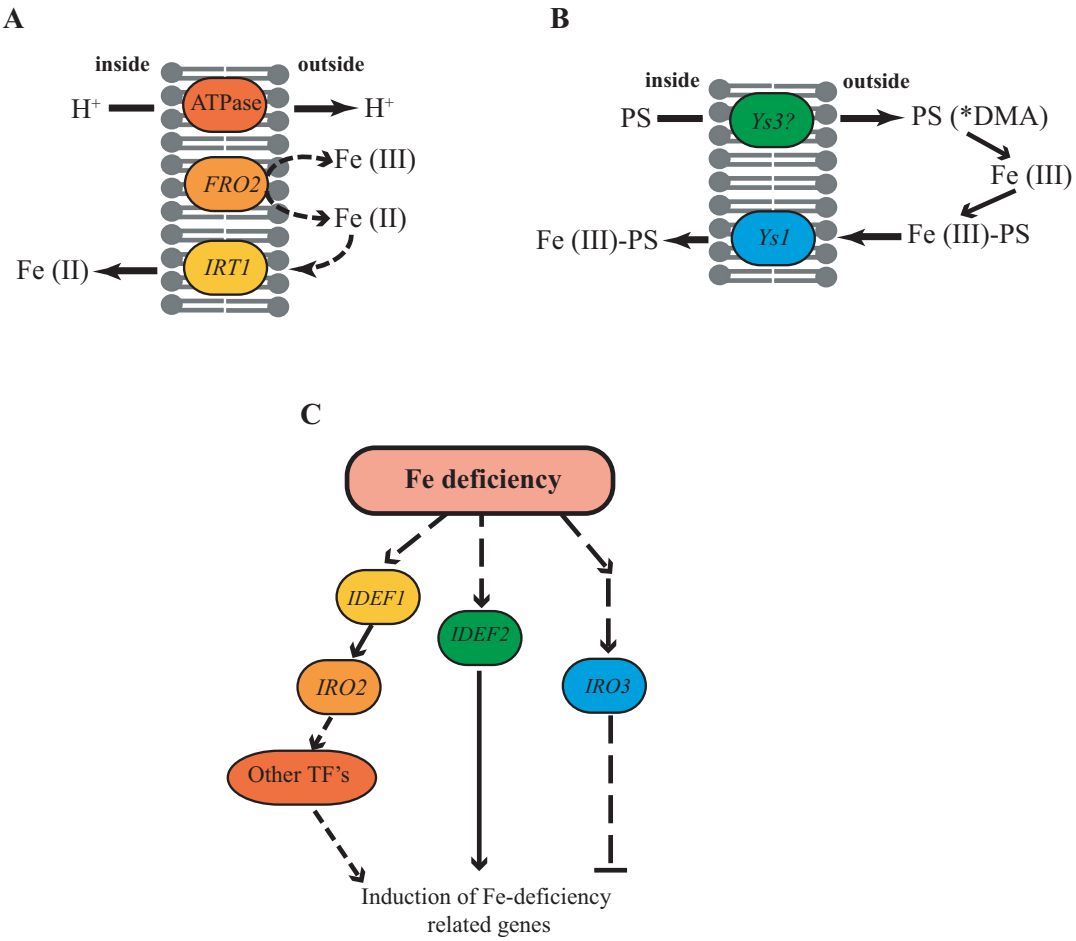

Figure 1.1: Iron acquisition strategies in **A)** dicotyledonous and **B)** graminaceous species. **C)** Regulation of Fe deficiency response in graminaceous species.

1.3 Fe homeostasis and storage

Plants have adaptated to different Fe constrain scenarios such as starvation or over-load by homeostatic responses that can control Fe uptake, its

mobilization between cells, and Fe storage. Unlike Fe starvation, Fe over-load can lead to oxidative stress by the formation of hydroxyl radicals also known as the Haber-Weiss or Fenton reaction (Guerinot and Yi, 1994; Lobreaux et al., 1992; Kobayashi and Nishizawa, 2012b; Thomine and Vert, 2013).

Fe over-load leads to the accumulation of ferritin in the cell. Ferritins are characterized by the storage of up to 4,500 Fe atoms in its central cavity, which is available to the cell in a non-toxic form (Guerinot and Yi, 1994; Lobreaux et al., 1992). The organelles where Fe is mostly accumulated include chloroplast, mithocondrion, and vacuole. In the chloroplast, 80%-90% of cellular Fe can be found due to its high requeriments for photosyntetis. Similarly, the mithocondrion is another compartment where Fe is required due to its intervetion in electron-transfer reactions as well as for the biogenesis of Fe-Sulfur clusters. However, in the vacuole, Fe is only accumulated to prevent cell toxity (Kobayashi and Nishizawa, 2012a). $AtFer1$ and $ZmFer1$ are induced in response to Fe overload and thus associated to Fe storage. However, studies in the $AtFer1$ suggested that its funcion is not only associated to storage, but also to the protection against oxidative stress (Kobayashi and Nishizawa, 2012a; Ravet et al., 2009).

Fe starvartion triggers the expression of several genes and transcription factors including $IDEF1$, $IDEF2$, $IRO2$, and $IRO3$. $IDEF1$ and $IDEF2$ have been shown to control genes at early stages in Fe deficiency (Fig. 1.1C). In contrast, $IRO2$ has been shown to regulate several genes in the methionine cycle and Fe uptake pathway. In addition, $IRO3$ regulates in a more complex

scenario controls genes related to Fe deficiency (Kobayashi and Nishizawa, 2012a).

1.4 Genome-wide transcriptome profile

Any insides on Fe maize homeostasis and its involvement on genome-wide gene regulation under different Fe regimes (deficient or sufficient conditions) has not been yet reported. The understanding of transcriptome regulation is important for interpreting gene expression differences and how these differences are associated to a specific phenotype (Wang et al., 2009; Sekhon et al., 2011; Hansey et al., 2012; Davidson et al., 2011; Li et al., 2010a).Transcriptome profiling by deep-sequencing technologies also called RNA-Seq allows to survey the entire transcriptome for gene expression differences, novel genes and isoforms identification, and sequence variants determination (Wang et al., 2009).

In maize, several studies have been published using the RNA-Seq technology (Davidson et al., 2011; Li et al., 2010a; Hansey et al., 2012; Eveland et al., 2010) in order to elucidate transcriptional networks associated to a specific phenotype or growing stage. However, none of these studies have explored thoroughly the response of the maize transcriptome under two different Fe regimes.

The main objectives of this project were to identify the $Ys3$ gene and understand its implication in gene expression on iron metabolism in maize. In that regard, the specific objectives were,

1. Identify the $Ys3$ gene by using map-based cloning,

2. Determine the function of the $Ys3$ gene by comparative sequencing in a broad germplasm set and transcription profiling,

3. Generate additional alleles of the $Ys3$ gene by transposon tagging,

4. Validate the $Ys3$ gene by sequencing additional alleles,

5. Determine gene response in $ys3\times$W22 F_2 individuals grown under two different Fe regimes by RNA-Seq,

6. Identify differential expressed genes, novel isoforms, and sequence variants in the RNA-Seq dataset.

Chapter 2

Material and Methods

2.1 Plant material

For genetic mapping of *Ys3*, a segregating F_2 population was developed from the cross between W22 and the *ys3* mutant with the germplasm bank code 311F (Fig. 3.1A). For causative mutation confirmation, the 26 parental inbreds of the nested association mapping (NAM) population were considered.

For direct transposon tagging, the four near isogenic *Ac* lines mon00178::*Ac*, bti03702::*Ac*, bti00220::*Ac*, and bti03526::*Ac* which contain *Ac* insertions near the predicted *Ys3* gene and the *Mutator* line 3820 *Mu*/B73 were crossed with the *ys3* mutant (Fig. 3.1B).

2.2 Phenotyping

2.2.1 *Cultivation in the greenhouse*

Individuals of the F_2 mapping population were grown in soil using a mix of Type ED73 soil mix (Einheitserde, Sinntal-Altengronau, Germany) and fine sand in the greenhouse. Three weeks after sowing, these individuals were classified based on their leaf intervenial chlorosis either as wild type (wt) plants, when they presented non visible signs of chlorosis or as mutant ($ys3$) plants, when they presented intervenial chlorosis. In each individual batch, $ys3$ plants were included as controls. Similarly, F_1 individuals derived from the crosses between $ys3$ and Ac as well as Mu were screened in the greenhouse for the $ys3$ phenotype. A total of 65,064 $Ys3 :: Mu$ and 66,355 $Ys3 :: Ac$ individuals were screened in a greenhouse at temperatures between 20 to 30°C and supplemental light.

2.2.2 *Cultivation in a hydroponic system*

Two replications, each with four seeds from each parental genotype (W22 and 311F) and hundred seeds from the $ys3$×W22 F_2 population were grown in a hydroponic system at different Fe concentrations to harvest clean root tissue. Seeds were sterilized by immersing them into a saturated $CaSO_4$ solution and heated at 60°C for 20 min. Then, seeds were transferred and germinated in petri dishes at room temperature in the dark until the primary root was developed. Afterwards, seedlings were transferred for 7 days into

a tip-removed 15 mL falcon tube for support and suspended into a 5 L pot containing a continuously aerated 100 μM Fe(III)-EDTA nutrient solution as described by Von Wirén et al. (1994). From day 14 to 28 plants were separated into two different Fe concentrations, 10 μM and 300 μM Fe(III)-EDTA. The cultivation between day 7 and 28 was performed in a growth chamber, in which the photoperiod, light intensity, relative humidity, and air temperature were 16 h, 170 μmol $m^{-2}a^{-1}$ in the leaf canopy, 60%, and 24°C, respectively. On day 28, root and leaf tissue was harvested from individual plants.

2.2.3 *Macro- and micronutrient quantification - Ionome analysis*

Eleven macro- and micronutrients including B, Ca, Cu, , Fe, K, Mg, Mn, Na, P, S, and Zn were analyzed by inductively coupled plasma optical emission spectroscopy (ICP-OES) at the Department of Soil Science, University of Hohenheim.

2.3 DNA extraction and genotyping

If no publicly available markers were available in genome regions of interest, new markers were developed from publicly available bacterial artificial chromosome (BAC) sequences. For microsatellite (SSR) markers, repeats were

identified using the MIcroSAtellite identification tool (MISA; Thiel et al., 2003) with a minimum of six repeats of dibasic motifs and four repeats of three to six base pair motifs. In addition, the maximum distance between two SSRs was 50 base pairs (Ingvardsen et al., 2010; Thiel et al., 2003).

Single nucleotide polymorphism (SNP) and cleaved amplified polymorphism sequence (CAPS) markers were identified by comparative sequencing of parental genotypes. The required primers for SSR, SNP, and CAPS markers were designed using Primer3 (Rozen and Skaletsky, 1999). Subsequently, primer sets were blasted against the BAC maize database at MaizeGDB and finally selected based on a non formation of self-dimers, pair-dimers, and hairpins using PrimerSelect^{TM} (DNASTAR® Lasergene v.8.02; Madison, WI, USA).

DNA was isolated from young leaves using a modified CTAB extraction protocol (Saghai Maroof et al., 1984). SSR genotyping was performed using a 4300 DNA Analyzer (LI-COR Inc., Lincoln, NE, USA) following standard protocols. CAPS screening was performed using *Bpm*I for digestion and verified on a 3% universal agarose gel (Bio-Budget Technologies GmbH, Krefeld, Germany). SNP markers were genotyped by Sanger sequencing on an Applied Biosystems 3130XL and 3730XL genetic analyzer using BigDye-terminator v3.1 chemistry (Weiterstadt, Germany). A haplomarker was determined with a minimum of four SNPs. Haplomarkers were visually scored for each DNA sample of the F_2 population and parental controls.

2.4 Initial genetic mapping

An initial genetic map was constructed based on 180 F_2 individuals derived from the cross of W22 $\times$ *ys*3 and the seven publicly available SSR markers bnlg1456, bnlg1957, umc1773, umc1449, umc1501, umc1908, and umc2002.

2.5 Recombinant selection and fine mapping

A total of 9,232 F_2 individuals were fingerprinted with the two flanking SSR markers on chromosome 3 (bnlg1957 and umc1773) and visually phenotyped as described before. A total of 76 recombinants were selected based on their genotypic and phenotypic information and later used for fine-mapping.

2.6 Synteny analysis

Synteny analysis of maize against the rice and sorghum genomes was performed based on physical coordinates of the fine mapping interval using the synteny tool of the maize sequence website (www.maizesequence.org). Additionally, individual genes were also searched against the Phytozome database (www.phytozome.net) in order to identify paralogous and homologous genes, functional annotations, and gene ancestry. For genes with non functional annotations, genomic and protein sequences were blasted against the UniProt database with a threshold $\leq e^{-10}$.

2.7 Genetic Map and phylogenetic tree construction

Genetic maps were calculated using JoinMap® (version 4.0; Van Ooijen, 2006). Phylogenetic tree construction was conducted using MEGA (version 4.0; Tamura et al., 2007).

2.8 Expression profile

Total RNA was extracted from pools of root tissue of parental genotypes W22 and *ys*3 and individuals of the *ys*3×W22 F_2 population grown under hydroponic conditions using the RNeasy® Plant Mini Kit (QIAGEN GmbH, Hilden, Germany). Total RNA was treated with DNase I (Ambion® Turbo DNA-$free^{TM}$, Invitrogen, Austin, TX, USA).

2.8.1 Quantitative RT-PCR

cDNA was synthesized using 1 μg of total RNA and the Transcriptor First Stand cDNA Synthesis Kit (Roche Applied Science, Penzberg, Germany) according to the manufacturer's instructions. The expression of $Ys3$ (DAA54978, DAA54979) and *Dmas* (DAA44291) was quantified in parental genotypes, W22 and 311F. Furthermore, expression response in candidates

genes identified by RNA-Seq analyses were queantified in individuals of the *ys*3×W22 F_2 population on a LightCycler® 480 (Roche Applied Science, Penzberg, Germany) using the DyNamo ColorFlash SYBR Green qPCR kit (Biozym, Hess, Oldendorf) and *Actin* (AFW81799) as internal control to calculate the relative transcript abundance of candidate genes (Table B1.).

2.8.2 Transcriptome sequencing (RNA-Seq)

Root tissue from our hydroponic study was classified into four groups based on the observed phenotype as either wilt type (wt) or *yellow stripe* 3 (*ys*3) and also by the Fe regime that was applied as either 10 *mu*M Fe-EDTA (10) or 300 *mu*M Fe-EDTA (300), respectively. The formed groups were identified as wt-10, wt-300, *ys*3-10, and *ys*3-300 and later used for further experiments. Total RNA from each group was cleaned up and concentrated using the RNeasy® MinElute® Cleanup Kit (QIAGEN GmbH, Hilden, Germany). RNA integrity and concentration was observed by loading the samples on an 1% agarose gel (Bio-Budget Technologies GmbH, Krefeld, Germany). Total RNA concentration was quantified using the Qubit ® RNA BR Assay Kit and measured with the Qubit ® Fluorometer. Ribosomal RNA (rRNA) was depleted using the RiboMinus^{TM} Plant Kit for RNA-Seq (Life Technologies GmbH, Darmstadt, Germany) and concentrated using the RiboMinus^{TM} Concentration Module (Life Technologies GmbH, Darmstadt, Germany), following the manufacturer's instructions. Depleted rRNA was analyzed with

the Agilent 2100 Bioanalyzer using a Plant RNA Pico Array (Agilent Technologies, B öblingen, Germany).

2.8.3 Library construction and sequencing

A total of 4 libraries per replication were constructed based on the previous described four groups using the TruSeq^{TM} RNA Sample Preparation Kit and paired-end sequenced using the Illumina® HiSeq 2000. On replication one, sequencing was performed on a single lane per library, whereas on replication two, samples were bar-coded and sequenced on two lanes. Library construction was performed by the MPIPZ Genomic Center. Raw sequencing data was processed with Illumina software CASAVA (ver. 1.8.2).

2.8.4 Transcriptome profiling

Raw RNA-Seq reads were analyzed with R/Bioconductor software using the *ShortRead* package. An indexed reference was created using the ZmB73 RefGen_v2 assembly and ZmB73_5a Working Gene Set (WGS, http://ftp.maizesequence.org/current/assembly/) with Bowtie2 v.2.0.0-beta 6 (Langmead et al., 2009). High-quality reads were mapped against the reference using TopHat version 2.0.3 (Trapnell et al., 2012) with the following settings inner distance between mate pair was set to 300 bp, maximum intron length was set to 50000 bp, standard deviation for the distribution on inner distances between mate pairs was 40 bp, library type to fr-unstranded, and the remained parameters were set to default conditions. Generated BAM

files were sorted and indexed using SAMtools v.1.18 (Li et al., 2009).

Transcript assembly, gene abundance, and identification of differentially expressed genes was performed using four different procedures. The first two procedures consisted in the use of $Cufflinks$ to assemble transcripts using sorted SAM files and by either providing any reference annotation file (ZmB73_5aWGS) or not. Transcript assembly performed by $Cufflinks$ reference annotation based transcript (RABT) is known to better identify novel transcripts based on a reference annotation (Roberts et al., 2011). $Cufflinks$ analyses used version 2.0.2 (Trapnell et al., 2012) with all parameters set as default.

Moreover, sorted BAM files along with an annotation reference were used to obtain count tables to investigate differential gene expression. Count tables were generated by using R/Bioconductor packages $EasyRNASeq$, *biomaRt*, $GenomicRanges$, and $GenomicFeatures$ (Delhomme et al., 2012). Determination of differentially expressed genes based on count tables used R/Bioconductor packages $DeSeq$ and $edgeR$, respectively (Anders and Huber, 2010; Robinson et al., 2010). For $DeSeq$, dispersion estimation was calculated using the blind method and a fit-only sharing-mode for each library per replication. Similarly, individual libraries per replication were analyzed with $edgeR$ using a dispersion coefficient of 20%. identification of differentially expressed genes was performed in four comparison, in which differences in Fe regimes response was tested in $ys3$ and wt F_2 individuals as comparison 1 and 2, respectively. In addition, differences in genotypes were also analyzed

between *ys*3 and wt individuals grown under 10 μM and 300μM Fe-EDTA as comparison 3 and 4, respectively. The threshold of significance was set to a false discovery rate (FDR) lower than 0.05 for all procedures. Significant differentially-expressed-genes throughout all the four bioinformatic-biometric tests were further analyzed.

2.8.5 Gene Ontology (GO-term) enrichment and pathway analysis

GO-term analysis was performed based on significant differentially-expressed-genes (DEG) identified by all four comparisons (FDR $\leq$ 0.05) and reference annotation (ZmB73 RefGen_5a) using *agriGO* Analysis Kit (Du et al., 2010). Visualization of differences in transcript expression within specific pathways used MapMan software v.3.5.1R2 (Thimm et al., 2004; Benke et al., 2011). Significant DEGs were converted into transcripts by the addition of "_T01" at the end of every gene name. Thus, pathway analysis used the list of DEGs and custom mapping and pathway files, which consisted of a list of gene identifiers and a diagram that showed transcripts involved in Fe uptake and homeostasis.

2.8.6 Variant Calling and data adjustment

Detection of polymorphisms between *ys*3 and wt F_2 individuals used BAM files generated by grouping all FASTAQ files of a given genotype, regardless their Fe regime nor replication using TopHat (version 2.0.3) (Trap-

nell et al., 2012).

Furthermore, BAM files were sorted and indexed using SAMtools (version 1.4). Variant calling was performed by using the mpileup function with default settings along with the ZmB73 RefGen_v2 assembly as reference.Further analysis used SAMtools/BCFtools version 1.4 (Li et al., 2009)in order to produce a VCF files. Only polymorphisms that presented a phred score $\geq$ 20 and an average read depth $\geq$ 10 were kept to further analysis. Coding sequences were annotated based on the filtered VCF file using Variant Effect Predictor (VEP) perl script version 2.8 and API and DB version 70 (McLaren et al., 2010).

Chapter 3

Results

3.1 Cloning and Validation of $Ys3$

3.1.1 Map based cloning of $Ys3$

The initial genetic map was constructed using the phenotypic and genotypic information of a subset of 180 out 9,232 F_2 individuals derived from the cross between 311F, also known as $ys3$ mutant due to its intervenial chlorosis caused by a defect in DMA secretion, and the inbred line W22, which presents no chlorosis (Fig. 3.1A). A total of 50 out of 180 F_2 individuals showed the $ys3$ phenotype. The segregation ratio between wt and $ys3$ phenotypes was 2.6:1 and, thus, not significantly (α=0.05) different from the expectation of a recessively inherited gene.

The marker order of the genetic map, with the exception of bnlg1456

and umc1449, was consistent with the location on the maize B73 RefGen_v2 physical map. The *ys*3 gene was mapped between SSR markers bnlg1957 and umc1773 on chromosome 3 (Fig. 3.2A). These markers were used as flanking markers for searching recombinants in the entire *ys*3× W22 F_2 population. The genetic map distance between the flanking markers bnlg1957 and umc1773 and *ys*3 was 1 cM and 4.1 cM, respectively (Fig. 3.2A). The physical distance between both flanking markers was 25,725,000 bp, 38,130,413 bp, and 61,167,870 bp based on the BAC-based Maize B73, B73 RefGen_v1, and _v2, respectively. This region was found to be in vicinity to the centromere (Fig. 3.2C and D).

Out of 9,232 F_2 individuals derived from the cross between W22 and the *ys*3 mutant, 76 showed recombinations between the flanking markers bnlg1957 and umc1773. In this region, nine additional markers were developed to further fingerprint the recombinant genotypes including 1 CAPS, 2 SSRs, and 6 haplomarkers (Table B1). The closest markers flanking *ys*3 were HAP-84 and HAP-129 with a genetic map distance of 0.7 cM and 0.1 cM, respectively (Fig. 3.2B).

The first synteny analysis attempt based on the Maize B73 RefGen_v1 showed 17 positional candidate genes in the fine mapping interval, which consisted of 1.22 Mbp. After sequencing one by one (data not shown), only two genes presented unique polymorphisms when comparing between W22

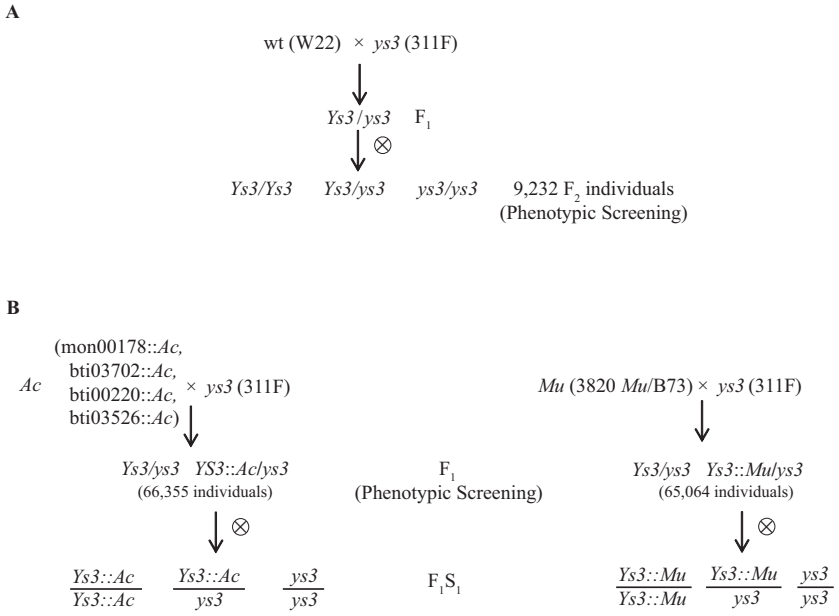

Figure 3.1: Diagram of individuals used for **A** construction of the genetic map of $ys3$ and **B** direct transposon tagging, where a total of four Ac NILs and one Mu genotype were used.

and the NAM parental inbreds against the $ys3$ mutant. However, both genes shared no homology to genes in rice and sorghum, and, thus, both were removed from further analyses.

3.1.2 Identification of a $Ys3$ candidate gene

After the release of the Maize B73 RefGen_v2, the region between markers HAP-84 and HAP-129 corresponds to 13.59 Mbp including 207 high confidence gene models in the maize B73 RefGen_v2 (Fig. 3.2C). A synteny analysis revealed that among the 207 candidate genes, 56 were shared between maize and rice, 64 between maize and sorghum, and 50 between all three species (Table B2). All genes involved in Fe uptake are known to

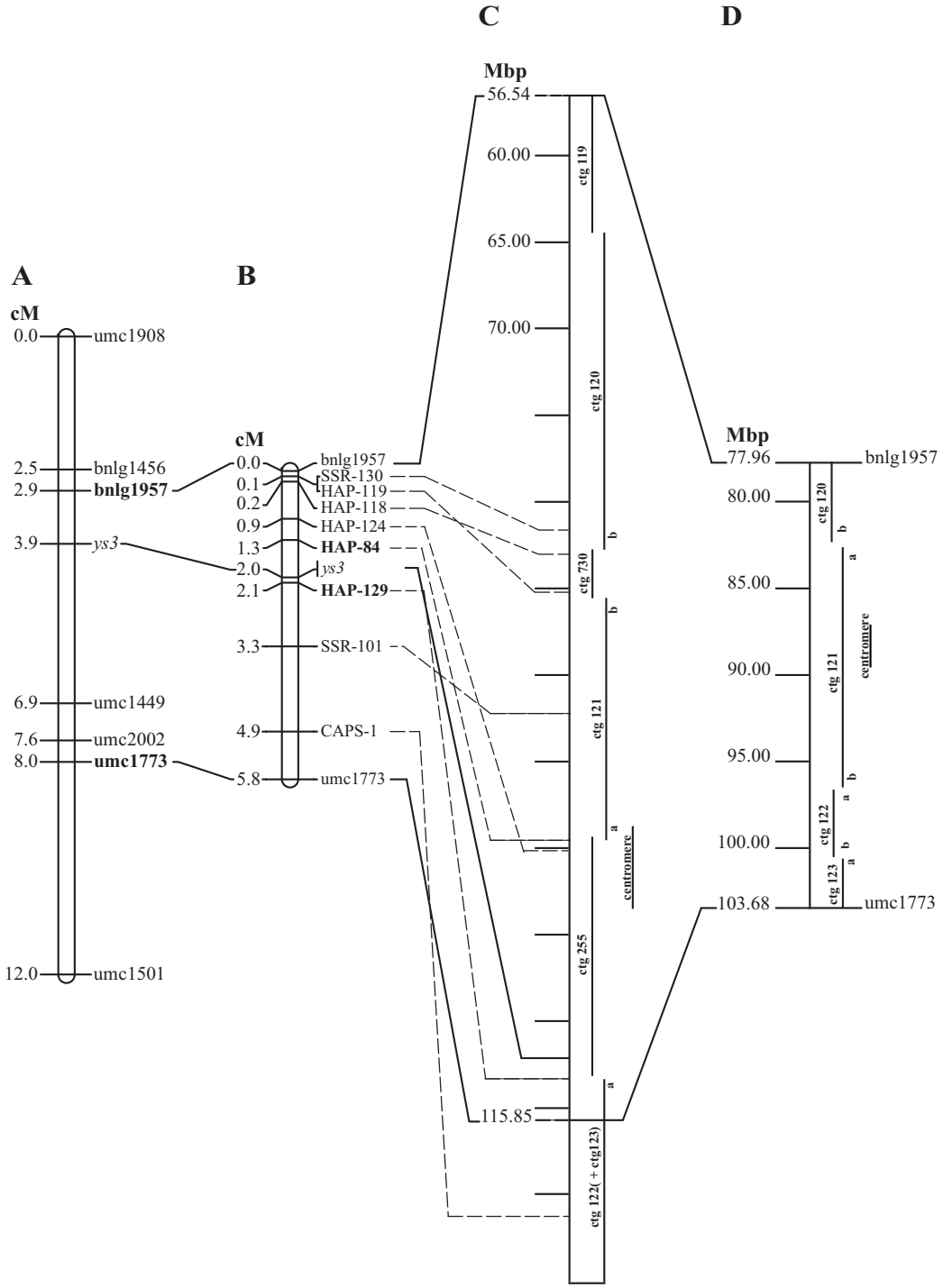

Figure 3.2: **A** Genetic map of *ys*3, where marker order and genetic map distance were calculated based on 180 F_2-individuals. **B** Genetic fine map of *ys*3, where marker order and genetic distance were calculated based on 76 recombinant F_2-individuals selected from a total of 9,232 F_2-individuals. **C** Maize B73_v2 physical map and **D** BAC-based Maize B73 physical map.

be conserved among graminaceous species. On that regard, only candidate genes that showed its orthologous in rice and sorghum were further analyzed.

Thirteen of these 50 genes encoded transmembrane proteins, but only four of these genes were characterized as transporters. Among these four genes, GRMZM2G063306 encodes a DMA efflux transporter orthologous to Os*TOM*1 and Hv*TOM*1 with an average of 75% and 62% identity of the amino acid sequences compared to rice and barley TOM1 proteins.

GRMZM2G063306 was predicted to enconde two transcripts (GRMZM2G063306_T01 and GRMZM2G063306_T02) and have a Major Facilitator Superfamily (MFS) domain. GRMZM2G063306_T01 is predicted to contain 7 exons encoding a transcript and protein of 657 bp and 186 aa, respectively. The protein is predicted to contain five transmembrane domains. In contrast, GRMZM2G063306_T02 is predicted to contain 13 coding exons with a transcript and protein length of 1,289 bp and 337 aa, respectively. For the protein, eight transmembrane helices are predicted (Fig. 3.3A). In addition, orthologous proteins from sorghum, brachypodium, switchgrass, and foxtail millet have been identified showing an identity on the protein level between 72% and 83% (Table 3.1, Fig. 3.4).

Table 3.1: $Ys3$ (GRMZM2G063306) synteny in graminaceous species.

Gene name	**Specie**	**Length (AA)**	**GRMZM2G063306_T01**	**GRMZM2G063306_T02**
			% of identity	
GRMZM2G063306_T01	Maize	186	-	93.00
GRMZM2G063306_T02	Maize	337	93.00	-
LOC_Os11g04020	Rice	473	78.00	71.00
HvTOM1	Barley	460	73.00	51.00
Sb08g008410	Sorghum	418	88.00	78.00
BRADI4G26380_T01	Brachypodium	497	73.00	70.00
BRADI4G26380_T02	Brachypodium	407	73.00	50.00
Pavirv00037536m	Switchgrass	473	83.00	79.00
Pavirv00003023m	Switchgrass	475	85.00	80.00
Pavirv00003022m	Switchgrass	482	85.00	80.00
Si021907m	Foxtail millet	483	83.00	80.00
Si021962m	Foxtail millet	468	83.00	75.00

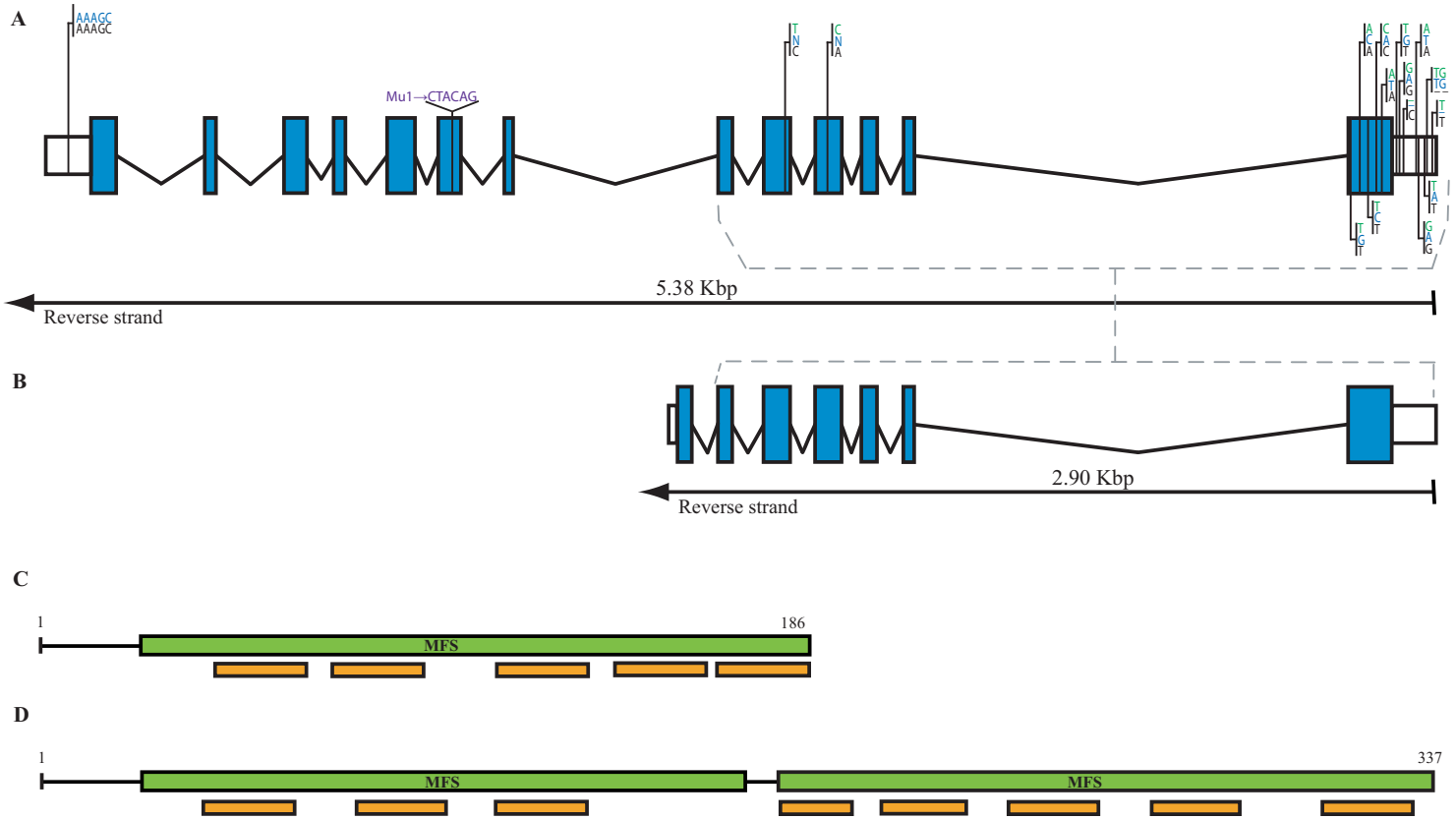

Figure 3.3: **A** Gene model of GRMZM2G063306_T02 (*Ys3*) including SNPs and InDels, where green represents B73, blue represents W22, black represent 311F, and purple represents the additional F_1S_1 *Ys3::Mu*. Any missing nucleotide is represented by N. Blue boxes represents exons and white boxes UTRs. **B** Gene model of GRMZM2G063306_T01 (*Ys3*). Grey lines represent the same polymorphism pattern in exon 1 to 6. **C** Putative conserved domains present in GRMZM2G063306_P01 (186 aa), where green boxes represent the Major Facilitator Superfamily (MFS) domain and orange boxes represent the transmembrane domains. **D** Putative conserved domains present in GRMZM2G063306_P02 (337 aa).

3.1.3 Polymorphisms in the $Ys3$ gene

The coding sequence (CDS) of $Ys3$ was sequenced in the parental inbred lines 311F and W22 of our study as well as in the parental individuals of the NAM population in order to identify unique causative mutations responsible for the $ys3$ phenotype. Sequence comparison between 311F and W22 showed 10 SNPs and 3 InDels. However, only 2 InDels located in the 5´-untranslated region (5´-UTR) and the 2 synonymous SNPs located in the CDS were unique for 311F when comparing them with B73 and W22 (Fig. 3.3). Furthermore, the parental individuals of the NAM population, HP301 and Oh43, also presented the same InDels like 311F, but not the two synonymous SNPs located in exon 4 and 5.

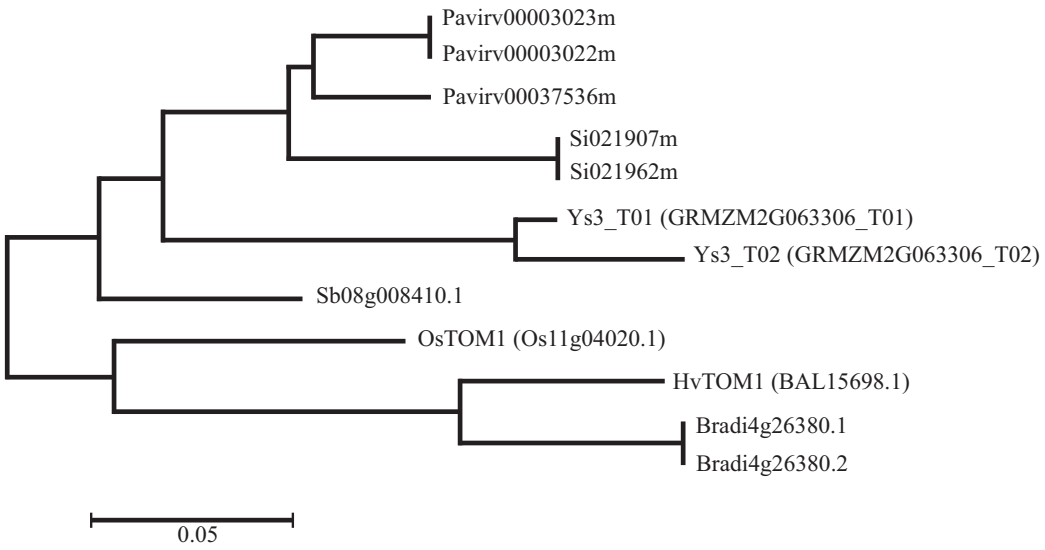

Figure 3.4: Phylogenetic tree of YS3 and its orthologous.

3.1.4 Confirmation of the $Ys3$ gene by an independent allele

A novel $ys3$ mutant allele was generated by crossing a Mu genotype with the $ys3$ mutant (Fig. 3.1B). One ($Ys3$::Mu) out of 65,064 and 4 ($Ys3$::Ac) out of 66,355 F_1 individuals were identified to carry a $ys3$ mutation due to their $ys3$ like phenotype. Sequencing of $Ys3$ in one F_1S_1 $Ys3$*::Mu* individual showed a 6 bp insertion in exon 8 (Fig. 3.3A).

3.1.5 Expression of $Ys3$ in maize roots under different Fe regimes

Expression levels of $Ys3$ was quantified in root tissue of the 311F mutant and W22 grown under deficient (10 μM Fe-EDTA) and sufficient (300 μM Fe-EDTA) iron conditions. Expression levels of $Ys3$ (GRMZM2G063306_T01 and GRMZM2G063306_T02), $Dmas$, and $Ys1$ in roots of 311F and W22 grown under deficient conditions were strongly induced in comparison with sufficient conditions.

Moreover, a two-fold decrease in expression was observed in the $ys3$ mutant for the $Ys3$ and $Dmas$ genes when comparing with W22 under Fe deficient conditions. However, a two-fold increase of $Ys1$ was observed in root tissue of 311F in contrast to W22 under deficient conditions. Furthermore, under sufficient conditions 311F showed a four and two-fold increase for both

transcripts of $Ys3$ and the same induction for *Dmas* and $Ys1$ (Fig. 3.5).

3.2 Genome-wide transcription profiling in $Ys3$ and $ys3$ background

3.2.1 Phenotypic characterization

A greater performance was observed in F_2-individuals that presented the wt phenotype, in contrast to $ys3$ individuals across all harvesting coefficients including shoot weight (SW), dry shoot weight (DW), shoot length (SL), root weight (RW), and water content (WC). Furthermore, a similar trend was observed in individuals grown under 300 μM Fe-EDTA (high), in comparison with individuals grown 10 μM Fe-EDTA (low). However, significant differences (α=0.05) were observed between wt and $ys3$ individuals grown under low Fe conditions , but not at high Fe conditions in SW and WC (Fig. 3.6A and D). For DW, significant differences (α=0.05) were observed between low and high Fe conditions, but not between wt and $ys3$ individuals (Fig. 3.6B). Moreover, significant differences (α=0.05) were observed within $ys3$ and wt individuals grown under low Fe condition, as well as high Fe concentration. However, no difference was observed between wt individuals grown under low Fe conditions and individuals grown under high Fe condition (Fig. 3.6C). No significant difference was observed between wt and $ys3$ individuals nor among Fe conditions in RW (Fig. 3.6E).

The relative chlorophyll content (SPAD) was measured in leaves 3

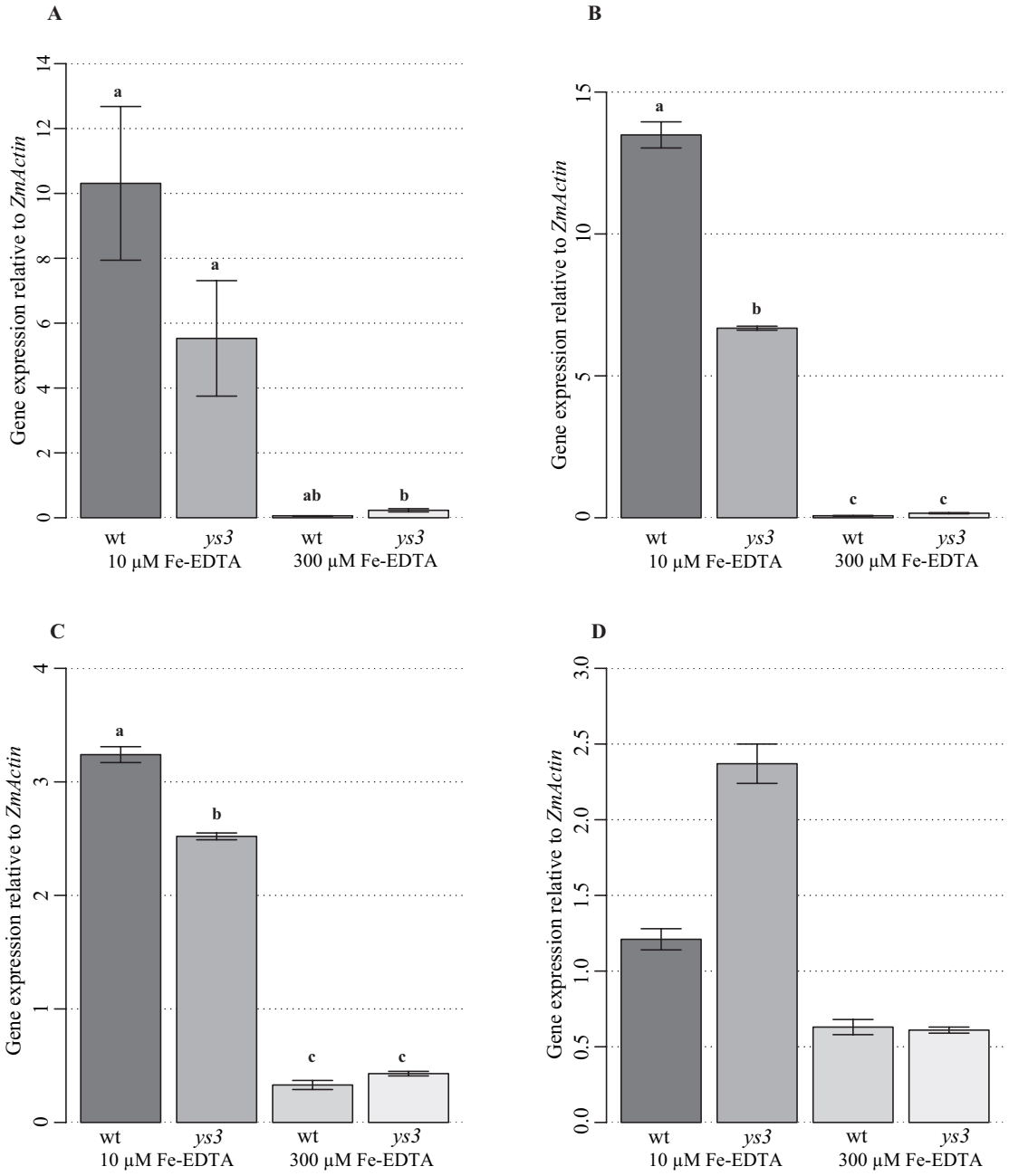

Figure 3.5: **A** and **B** Expression of two possible transcripts of $Ys3$ (GRMZM2G063306_T01 and GRMZM063306_T02), **C** *Dmas*, and **D** $Ys1$ relative to *Actin*1 in root tissue of wt and $ys3$ mutant plants under Fe deficient (10 μM Fe-EDTA) and sufficient (300 μM Fe-EDTA) conditions. Error bars represent SE and were calculated based on two biological replicates in **A**, **B**, and **C** and on four technical replicates in **D**. Letters represent significant differences ($p < 0.05$) between treatments.

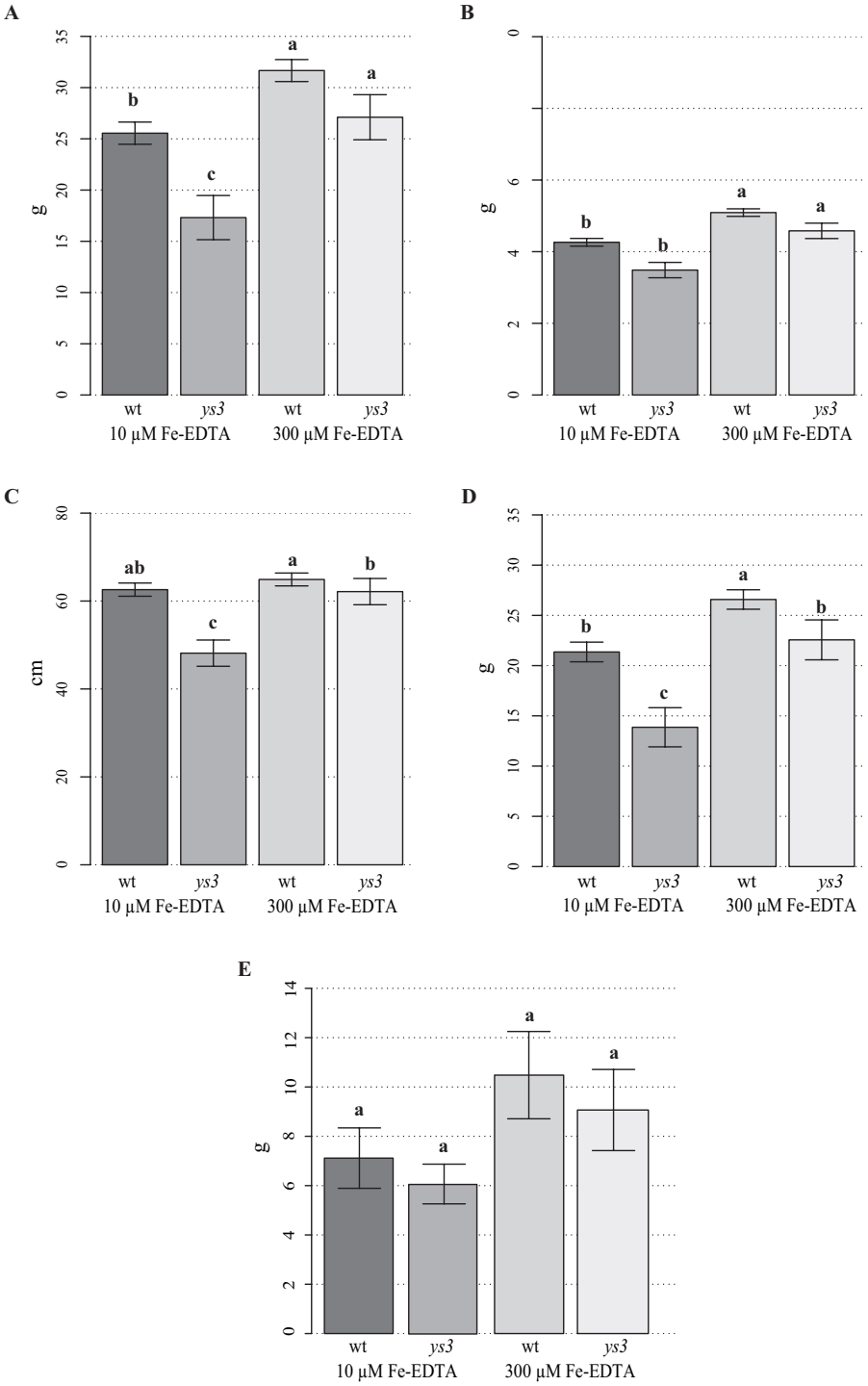

Figure 3.6: Harvesting coefficients including **A)** shoot weight, **B)** dry shoot weight, **C)** shoot length, **D)** water content, and **E)** root weight. All the harvesting coefficients were measured in F_2 -individuals derived from the cross between W22 and *ys*3, in which plants that showed the *ys*3 and wt phenotype were grouped and grown under Fe deficient (10 μM Fe-EDTA) and sufficient (300 μM Fe-EDTA) conditions, respectively. Error bars represent SE and were calculated based on two biological replicates. Letters represent significant differences ($p < 0.05$) between treatments.

(SB3), 4 (SB4), and 5 (SB5) before the application of both Fe regimes and in leaves 3 (SA3), 4 (SA4), 6 (SA6) after Fe treatment. However, chlorophyll content was measured only in leaf 4 before and after Fe treatment and leaf 6 after Fe treatment for both replicates. Significant difference (α=0.05) in SB4 was observed only in wt individuals grown under high Fe conditions, in comparison with *ys*3 individuals grown at low and high Fe and wt individuals grown at low Fe conditions (Fig. 3.7 B). Significant differences (α=0.05) were also observed between wt and *ys*3 individuals grown under low and high Fe conditions, respectively. No significant difference was observed between wt plants at low Fe and *ys*3 plants at high Fe conditions (Fig. 3.7E). Significant differences were observed (α=0.05) in SA6 within wt and *ys*3 individuals at low Fe conditions. However, no significant difference was observed between wt and *ys*3 individuals grown under high Fe conditions. In fact, *ys*3 individuals consistently showed lower SPAD values at both regimes, although more drastic symptoms were observed at low Fe conditions.

3.2.2 Micronutrient response to Fe deficient and sufficient regimes - Fe content

A total of eleven micronutrients were analyzed by inductively coupled plasma optical emission spectroscopy (ICP-OES). Fe content was greater in wt plants than in *ys*3 plants for both Fe regimes (Fig. 3.8). Furthermore, Fe content was highly and positively correlated with harvesting coefficients SA4, SA6, SWB, DW, WC, and RW (Fig. 3.9, Table B4.). In contrast, zinc (Zn)

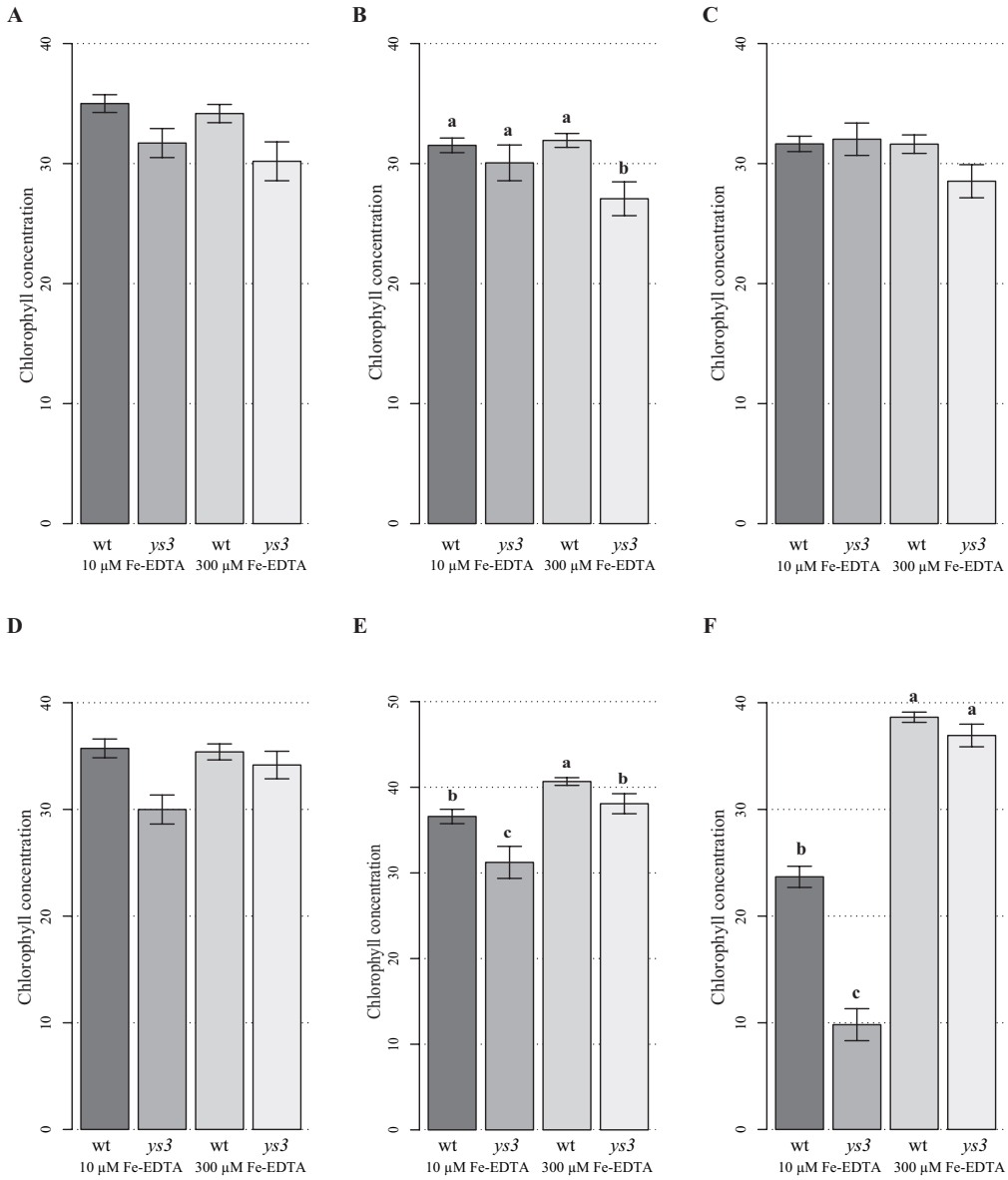

Figure 3.7: Relative chlorophyll concentration (SPAD) measured in leaf **A** 3, **B** 4, and **C** 5 before Fe treatment and leaf **D** 3, **E)**4, and **F** 6 after Fe treatment. SPAD was measured in F_2 -individuals derived from the cross between W22 and *ys*3, in which plants that showed the *ys*3 and wt phenotype were grouped and grown under Fe deficient (10 μM Fe-EDTA) and sufficient (300 μM Fe-EDTA) conditions, respectively. Error bars represent SE and were calculated based on two biological replicates in **B**, **E**, and **F** and on technical replicates in **A**, **C**, **D**, and **F**. Letters represent significant differences ($p < 0.05$) between treatments.

content was high but negatively correlated to the same harvesting coefficients as Fe. Potassium (K) was also high and positively correlated to all harvesting coefficients, expect for SB4 and SL. In addition, manganese (Mn) an sodium (Na) were high and positively correlated with SB3. Fe content was positively correlated with K content. However, it was negatively correlated with Zn, Copper (Cu), and sulfur (S) (Fig. 3.9, Table B5, and Fig. A7).

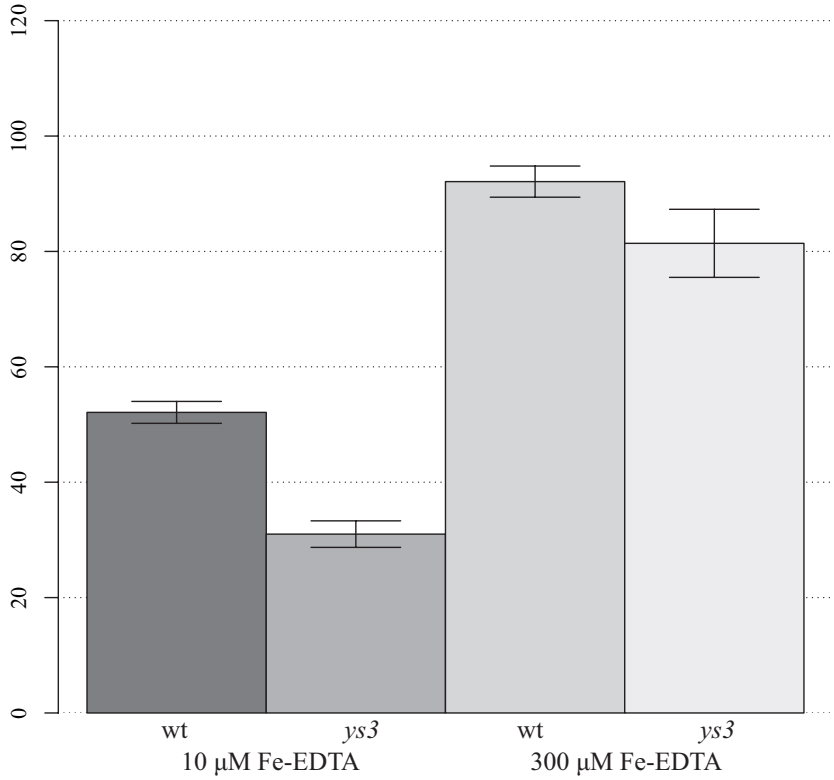

Figure 3.8: Barplot of Fe content measured in leaf 6 of $ys3\times$W22 F_2 - individuals grown under Fe deficient (10 μM Fe-EDTA) and sufficient (300 μM Fe-EDTA) regimes, respectively. Error bars represent SE and were calculated based on technical replicates

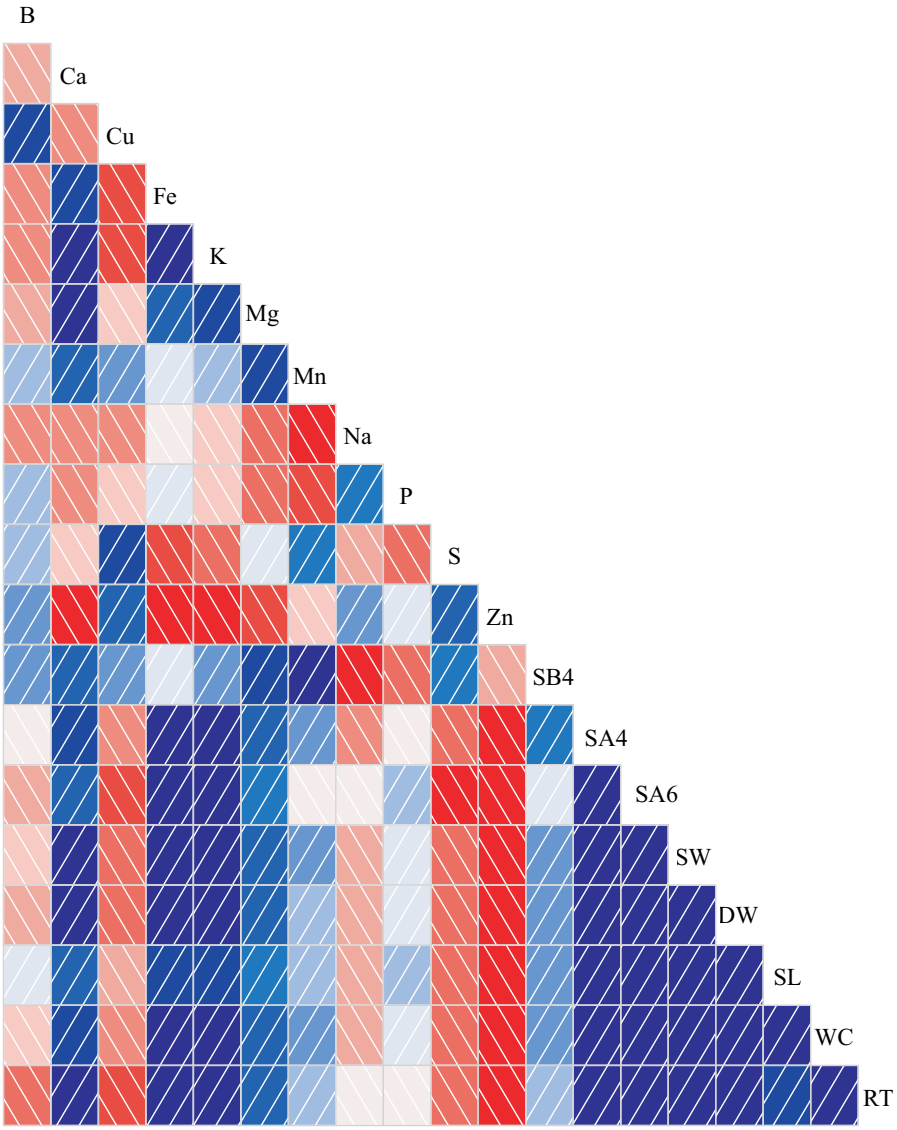

Figure 3.9: Correlation matrix between harvesting coefficients and eleven micronutrients measured in $ys3\times$W22 F_2 -individuals grown under Fe deficient (10 μM Fe-EDTA) and sufficient (300 μM Fe-EDTA) regimes, respectively. Positive correlation is expressed by gradients of blue color, in contrast to negative correlation that is associated with gradients of red color.

3.2.3 Maize root transcriptome sequencing in response to deficient and sufficient Fe conditions

RNA-Seq reads from root tissue coming from two biological replications grown under deficient and sufficient Fe regimes were used to determine differences in gene response between *ys*3 and wt F_2 individuals. Individuals from the *ys*3×W22 F_2 population were used in this experiment in order to avoid any sequencing noise due to the lack of genetic background homogeneity between the wt and *ys*3-mutant. Approximately 459 million high quality reads were obtained from both replications among all four libraries. A coverage estimation of 119x, 133x, 158x, 153x was observed in *ys*3-*10*, $ys3-300$, wt-10, and wt-300 libraries, respectively and calculated based on the total amount of reads obtained from both replications, an average gene length of 2.5 Kbp, and a total number of 32,540 genes, based on the filtered gene set (FGS). A total coverage of 564x was calculated across all treatments (Table 3.2).

An average of 36,544 genes were expressed in our experiment. A 5 % more genes were expressed in the first replication in contrast to the second replication (Table 3.2).

Table 3.2: Summary of RNA-Seq experiment. Total number of expressed genes was calculated based on Maize B73 WGS5a.

Treatment	**Rep.**	**No. of reads**	**Total reads**	**Coverage**	**Expressed genes**
wt-10	1	93,684,352	128,431,291	157.87	36,574
wt-10	2	34,746,939			34,233
wt-300	1	86,717,678	124,795,237	153.05	38,805
wt-300	2	38,077,559			34,717
ys3-10	1	63,396,867	97,075,623	119.33	37,488
ys3-10	2	33,678,756			34,792
ys3-300	1	57,878,378	108,401,017	133.25	37,023
ys3-300	2	50,522,639			38,720
Overall		458,703,168		563.86	36,544

3.2.4 Gene response to Fe deficient and sufficient conditions - Differentially expressed genes (DEGs)

In this study, the determination of DEGs was based on the use of four different procedures including $cuffdiff$ ($Cufflinks$ v. 2.0.2) with and without a reference annotation based transcript (RABT) assembly, $DeSeq$, and $edgeR$. Four comparisons were performed in order to investigate the differences in gene expression between the $ys3$-mutant and wt plants in response to two different Fe regimens. Gene expression differences in response to two different Fe regimens, 10 μM and 300 μM, were determined for $ys3$ plants in Comparison 1 and wt plants in Comparison 2. In addition, expression differences were also determined between wt and $ys3$ plants grown under 10 μM Fe-EDTA as Comparison 3 and 300 μM Fe-EDTA as Comparison 4, respectively.

DEG analysis was performed in both replications separately, due to a moderate biological coefficient of variation found in this experiment. The final list of DEGs was determined based on overlapping genes coming from the four different bioinformatic-biometric procedures on both replications (Fig. A1). A total of 190 DEGs were identified based on coincidences across the four procedures. The number of DEGs were similar among the four procedures, except for *edgeR*. However, *DeSeq* was the most stringent test based on the number of identified DEGs across all comparisons (Table B6-9).

Comparison 1, 2, 3, and 4 identified a total of 115, 58, 49, and 38 DEGs, respecitvely (Fig. 3.10). Comparison 3 and 4 identified more stress related genes and lower number of candidate genes. Therefore, comparison 1 and 2 were further analyzed. Thus, observed candidate genes were associated with Fe uptake and homeostasis including $Fer1$, $NAAT$, $Ys1$, $Idi4$, $Nramp3$, and Mtk (Fig. 3.11, Table 3.3, and Table B6-9). In addition, novel candidate genes showing bHLH domains were also identified including GRMZM2G057413 and GRMZM2G350312, whose orthologs were found in *A. thaliana* as well as rice and involved in the regulation of $FRO2$, a Fe reductase in response to Fe deficiency (Long et al., 2010). In addition, other identified candidate genes presented domains that are involved with oxidation-reduction, transport, response to ROS, and NAD synthesis (Table 3.3).

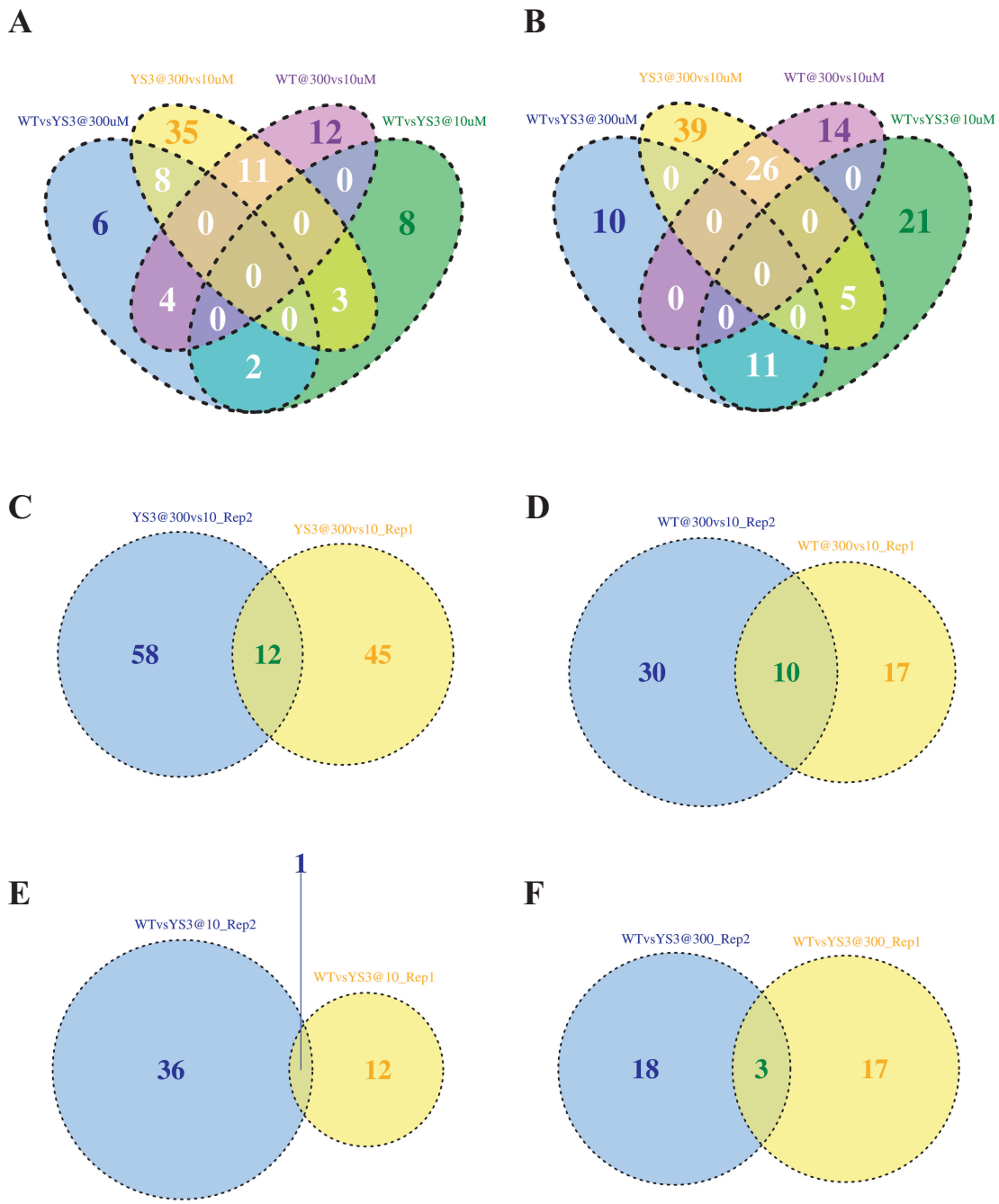

Figure 3.10: Overall representation of differentially expressed genes (DEG) across replication **A** one and **B** two. DEGs across comparison **C** 1, **D** 2, **E** 3, and **F** 4 using a FDR <0.05, in which red color represent up- and blue down gene regulation.

Table 3.3: List of candidate genes identified by maize root transcriptome profile using *ys*3×W22 F_2-individuals, which were grown under tow different Fe regimes (10 and 300 μM Fe-EDTA). Identification of candidate genes was based on the use of four procedures including $cuffdiff$ with and without RABT assembly, $DeSeq$, and $edgeR$ with a FDR <0.05.

Comparison[a]	**Gene**	**Location**	**Log2fold mean**	**Predicted function**
1	GRMZM2G308463	2: 230,640,837-230,641,721	-3.1	Oxidation-reduction process
	GRMZM2G161746	4: 166,243,312-166,245,736	-3.3	Enzymatic reaction of N-methyltransferase
	GRMZM2G057413	3: 148,031,503-148,032,613	-5.7	bHLH transcription factor
	GRMZM2G400602	3: 174,361,366-174,363,882	-4.0	MFS domain - transporter
	GRMZM2G085381	4: 3,256,234-3,258,478	-4.8	Bx1-NAD synthesis
	GRMZM2G350312	1: 65,657,063-65,660,290	-2.7	bHLH transcription factor
	GRMZM2G104563	2: 172,852,270-172,853,086	-3.1	MFS domain - transporter
	GRMZM2G155546	6: 118,073,100-118,075,002	2.7	Oxidation-reduction process
	GRMZM2G325575	4: 183,588,190-183,591,209	2.4	Ferritin-1 $Fer1$
	GRMZM5G866024	3: 222,670,492-222,671,716	-4.8	Membrane protein
	GRMZM2G150952	3: 213,005,290-213,017,220	-3.2	ATP-binding
2	GRMZM2G308463	2: 230,640,837-230,641,721	-4.1	-
	GRMZM2G161746	4: 166,243,312-166,245,736	-3.9	-
	GRMZM2G103342	3: 146,522,696-146,524,904	2.5	Peroxidase-12, response to ROS
	GRMZM2G124061	6: 66,895,501-66,897,831	-5.2	von Willebrand factor, type A
	GRMZM2G057413	3: 148,031,503-148,032,613	-3.8	-
	GRMZM2G430902	6: 67,454,716-67,457,087	-4.1	Chloride transport, transmembrane transport
	GRMZM2G137440	6: 67,258,224-67,260,551	-3.1	von Willebrand factor, type A domain,
	GRMZM2G035599	2: 144,880,257-144,881,689	-2.8	Cell death and cellulose biosynthesis
	GRMZM2G066840	2:230,858,935-230,860,296	-4.3	Oxidation-reduction process
	GRMZM2G104563	2: 172,852,270-172,853,086	-2.6	-
3	GRMZM2G011523	3: 23,955,988-23,956,940	-5.1	Forkhead domain
4	GRMZM2G011523	3: 23,955,988-23,956,940	3.8	-
	GRMZM2G134618	3: 58,561,607-58,562,993	5.4	Nucleotide binding

[a]Comparison 1 and 2representing 10 μM vs. 300μM Fe-EDTA in *ys*3 and wt individuals respectively. Comparison 3 and 4 representing wt vs. *ys*3 at 10 μM and 300 μM Fe-EDTA, respectively.

3.2.5 GO-term enrichment analysis of differentially expressed transcripts

A GO-enrichment analysis was performed using the 190 DEGs identified across all four comparisons (Table B6-9). Identified GO-terms were grouped based on their biological process as stress-related response, ion homeostasis, metabolic processes, as well as defense response to biotic and abiotic stimulus and developmental process (Fig. 3.12). In addition, sulfur, carboxylic acid, aromatic compounds, and nitrogen processes were also enriched in this analysis. Furthermore, methionine and aspartate metabolic processes were also identified as well as GO-terms in response to ion, cation, chemical, and especially iron and di-, tri-valent inorganic cation homeostasis (Fig. 3.12).

3.2.6 Polymorphism identification and annotation

A total of 204,123 variations including single nucleotide polymorphisms (SNP) and insertion and deletions (InDel) were identified in coding regions of the Maize B73 RefGen_v2. Furthermore, 194,189 and 9,934 variations were annotated as SNPs and InDels, respectively. In the SNP dataset, silent and missense mutations were the most common with 56.9% and 41.6%. Non-sense muations and splice sites were present with a 0.4% and 1.1% (Fig. 3.13A). In contrast, in the InDel dataset missense mutations were the most present with 96.6%, followed by splice site with 3.4% (Fig. 3.13B).

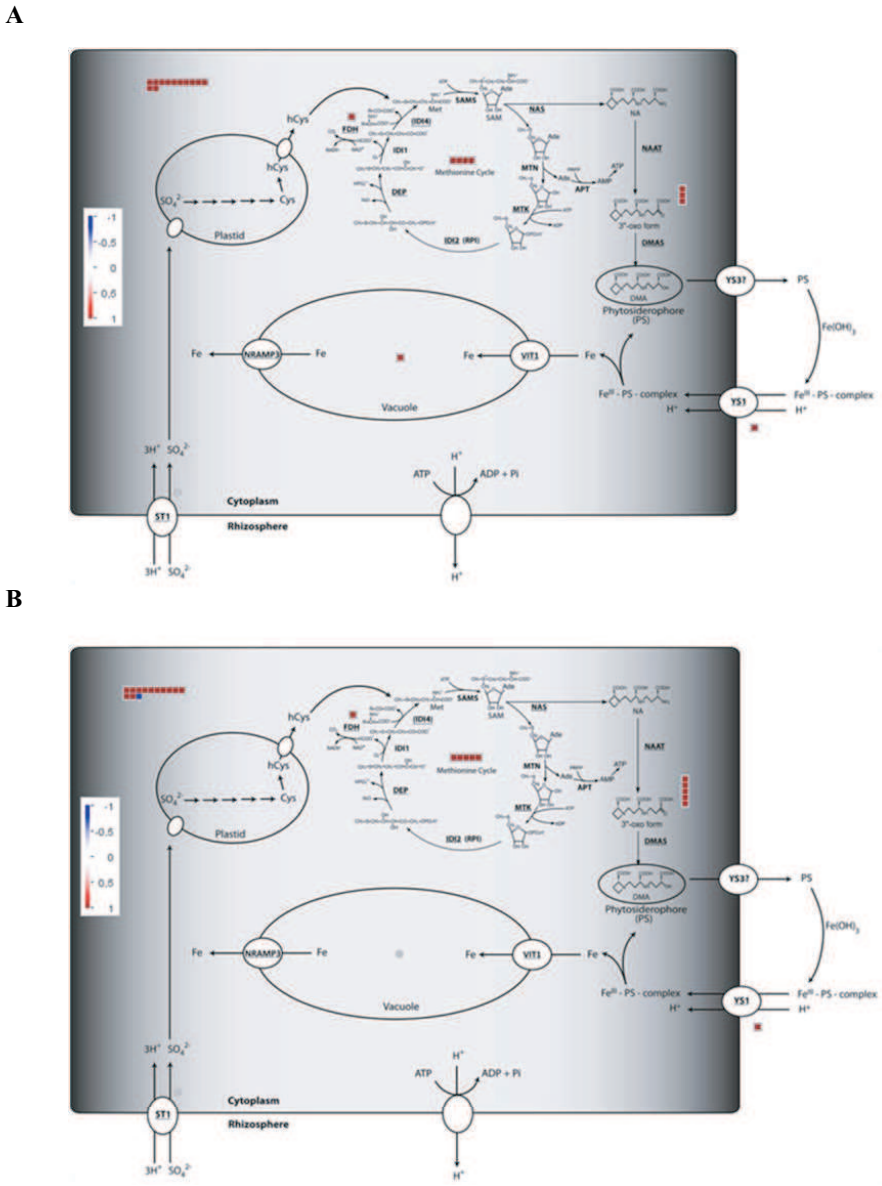

Figure 3.11: Pathway analysis using differentially expressed genes that were identified in W22×*ys*3 F_2 -individuals that presented **A** the *ys*3 and **B** wt phenotype and were grown under Fe deficient (10 μM Fe-EDTA) and sufficient (300 μM Fe-EDTA) conditions, respectively. Pathway analysis was performed using MapMap v.3.5.1R2.

3.2.7 Correlation between RNA-Seq and qRT-PCR

Expression response measured by RNA-Seq and qRT-PCR in thriteen genes followed the same pattern. The correlation coefficient and the proportion of the explained variance was 42% (Fig. 3.14).

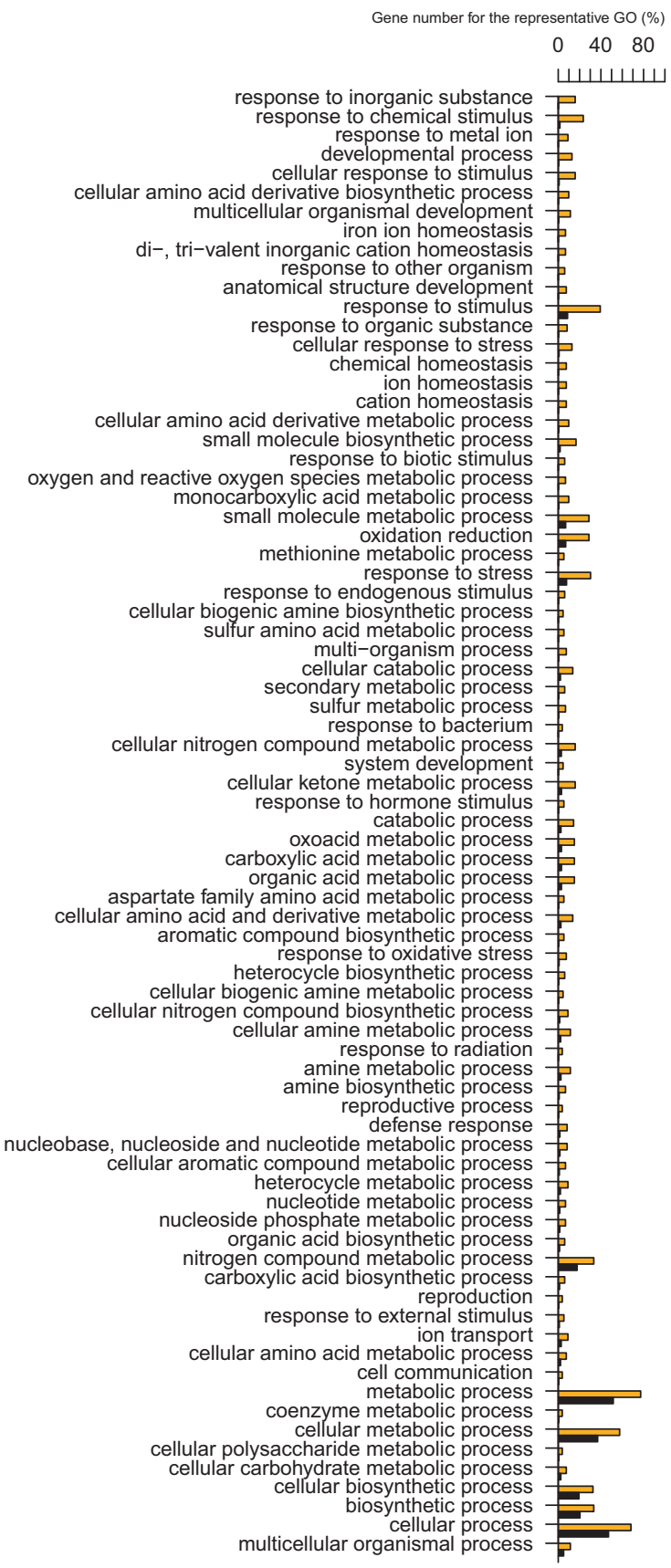

Figure 3.12: GO enrichment analysis.

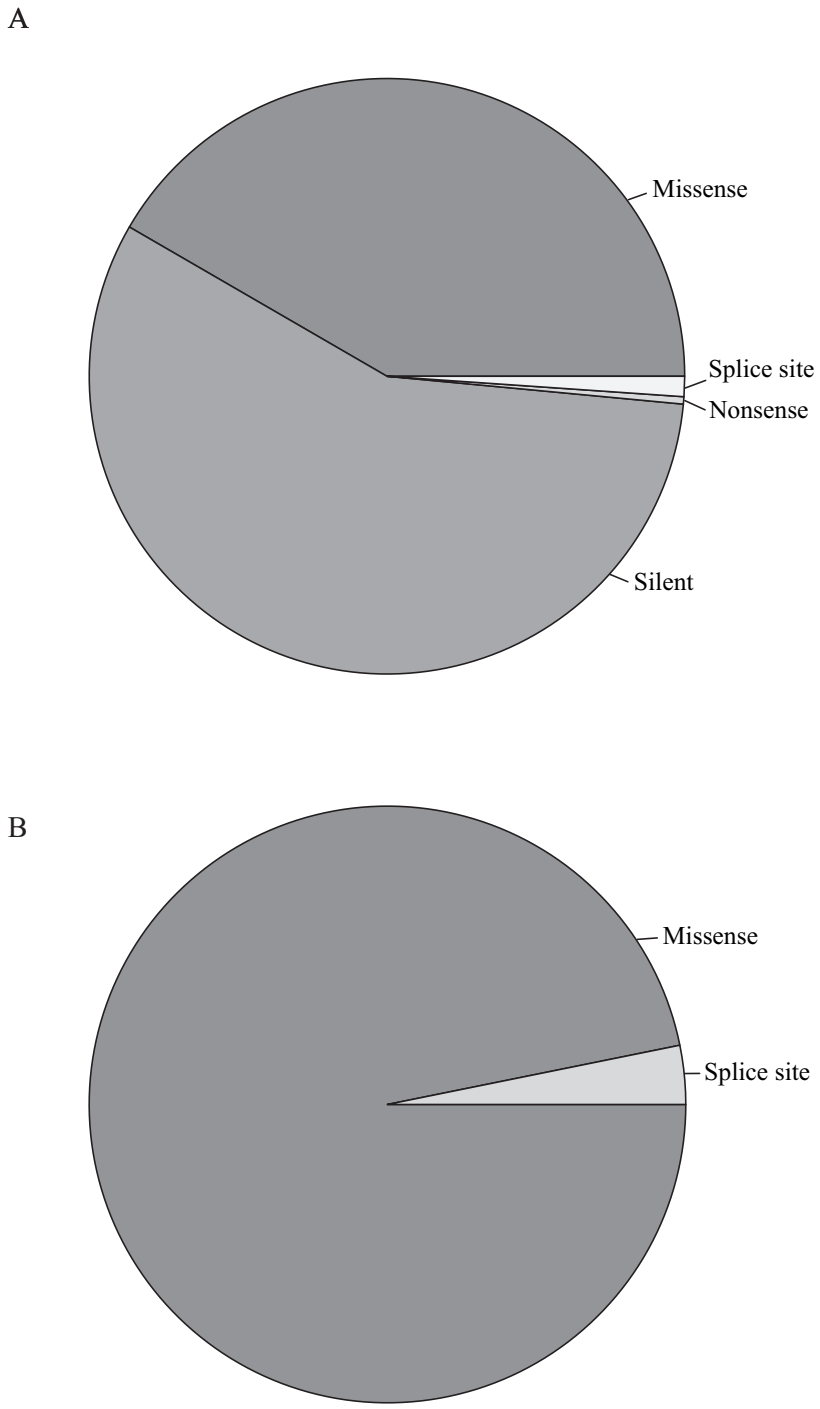

Figure 3.13: Distribution of mutations based on **A** single nucleotide polymorphisms (SNPs) and **B** insertions and deletions (InDels) that were identified and annotated using Samtools (v. 0.1.18) and Variant Effector Predictor (VEP, v.2.8), respectively.

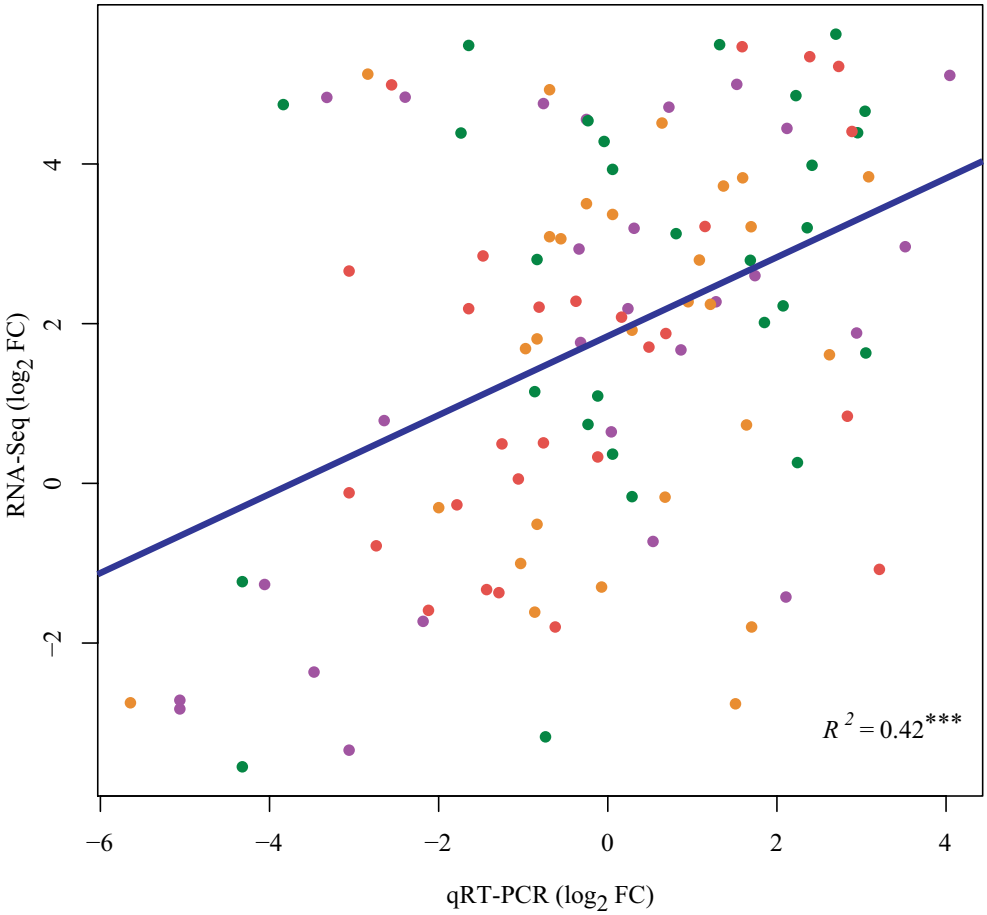

Figure 3.14: Correlation of expression response in thirteen candidate genes between RNA-Seq and qRT-PCR. Significance level with a p <0.05.

Chapter 4

Discussion

4.1 Mapping of $Ys3$ and chromosome walking

In the W22 × *ys*3 F_2-mapping-population, the *ys*3 phenotype was observed to segregate in a 3:1 ratio characteristic for recessive inheritance. This is consistent with the first description of the *ys*3 mutant (Beadle, 1929; Basso et al., 1994). The initially published genetic location of *ys*3 was on the long arm of chromosome 3 above the phenotypical and molecular markers umc102, Pgd2, and vp1 with a genetic distance of 2.87 cM, 9.29 cM, and 14.29 cM, respectively (Motta et al., 1999; Wright, 1961). Furthermore, on the IBM2 2008 Neighbors consensus map, the location of *ys*3 was estimated to be in between SSR markers umc1773 and umc2002 (Schaeffer et al., 2008).

However, in our study, *ys*3 was mapped to a region slightly upstream between SSR markers bnlg1957 and umc1773 and in vicinity to the cen-

tromere. The difference in genetic map position between our study and the IBM 2008 Neighbors map might be due to the estimation of loci positions on the consensus map.

When this projected started, no publicly available markers were found in the genome region between SSR marker, umc1773 and bnlg1957. Therefore, 70 new SSR markers were developed based on BAC sequences (Table B3). However, only two SSR markers were found polymorphic and showed consistent band patterns when comparing W22 and *ys*3 mutant. This observation could be explained due to the proximity of our region to the centromere (Schnable et al., 2009). The high amount of repetitive sequences observed in the maize genome especially in centromeric regions might lead to inconsistent band patterns that hamper polymorphism identification. Our observation is in accordance to Ingvardsen et al. (2010), who reported that 60-70% of developed SSR markers from a near region on chromosome 3 were discarded due to the repetitiveness nature of the sequence around the mapping area.

In addition, one CAPS and four haplomarkers were also developed and screened for polymorphism initially in our parental genotypes and later in the F_2-recombinants. In the present study, haplomarkers were defined with a minimum of four SNPs to tag a specific haplotype. Haplotype markers are a cost efficient genotyping strategy that allows the proper identification of parental and heterozygous alleles, increasing the genetic informativeness and quality (Ching et al., 2002).

The region between both flanking markers (bnlg1957 and umc1773)

has rapidly changed along with the release of the Maize B73 sequence. According with the BAC based Maize B73 map, this region was composed by four contigs (ctg) including ctg 120 to 123. However, when the Maize B73 RefGen_v1 was released, the same region presented several changes including the addition of ctg 255 coming from chromosome 5 between ctgs 120 and 121 and the merge of ctg 123 with ctg 122 (Wei et al., 2009). Furthermore, after the latest release of the maize sequence (Maize B73 RefGen_v2) the same region presented new changes. In this respect, ctg 730 has been placed between ctgs 120 and 121, the switch of ctg 255 between ctgs 121 and 122, and the reorder of ctg 121. These changes have had a considerable impact in the physical size of our fine mapping region. For instance, based on the BAC-based Maize B73 our fine mapping region consisted of 25.73 Mbp. Furthermore, based on the Maize B73 RefGen_v1 and _v2, the same region spans 38.13 Mbp and 61.17 Mbp, respectively. These continuous changes in the *ys*3 region have hampered chromosome walking and therefore the application of map-based cloning.

Marker positions in the rough genetic map as well as in the fine mapping were partially in agreement with the current physical map (Maize B73 RefGen_v2). Similarly, marker incongruence has been observed when the maize B73 RefGen_v2 physical map was compared with the predicted genetic positions on the ISU Integrated IBM 2009, a high level of disagreement was observed in our fine mapping region with an average of disagreement of 17.5 Mbp (Andorf et al., 2010). These observed disagreements due to

repetitive nature of the sequence close to the centromete strongly suggest that in our fine mapping region there are still probes that might need to be reordered and oriented and gaps to be closed. The resequencing of B73 with long fragments might improve probe positioning and orientation and gap closing.

A total of 26 SNPs were observed in 2,061 bp sequenced in the wt and *ys*3 genotypes, representing one SNP change occurring every 79.3 bp. Moreover, in a study using 36 elite maize inbred lines, a slightly high SNP frequency was observed with one SNP change occurring every 60.8 bp (Ching et al., 2002). In contrast, one SNP change was observed every 133 bp in 19 accessions of *B. distachyon* (Luo et al., 2012). In this respect, a proportional high number of SNPs was observed in our study, indicating a potential high degree of diversity in the *ys*3 region.

Gene density in our fine mapping region was found to be one gene per 65.7 Kbp. Consistently, a gene density of one gene per 67 Kb was also observed in a neighbor region in chromosome 3 (Ingvardsen et al., 2010). However, a higher gene density was observed in non-repeat regions of the Maize B73 RefGen_v1, which was determined to be one gene per 13.07 Kbp (Schnable et al., 2009; Vicient, 2010). In maize, gene density and recombination frequency were found to be non randomly distributed with a high density observed at the end of the chromosomes, in contrast to centromeric regions. Therefore, the observed low gene density in our region might be explained by its pericentromeric location.

4.2 Identification of a candidate gene for $Ys3$

Selection of candidate genes in our fine mapping interval were based on the assumption that the gene responsible for phytosiderophores efflux needed to present orthologous in other graminaceous species. In fact, graminaceous species follow the same mechanism for Fe acquisition. In that regard, only 24% of the positional candidate genes were found to be in synteny with maize, rice, and sorghum. In contrast, 70% of the candidate genes in the $bm6$ region in the short arm of chromosome 2 were syntenic to maize and rice (Chen et al., 2012). Low levels of synteny between maize, sorghum, and rice have been reported in pericentrometic regions of maize and therefore expected in our study (Schnable et al., 2009).

The gene GRMZM2G063306 was identified as an ortholog to $OsTOM1$ and $HvTOM1$. $OsTOM1$ was first described as the responsible for phytosiderophore efflux in rice and identified based on a high resolution microarray analysis and confirmed by transgenic approaches including GUS-tissue localization and function analysis by repression and overexpression of $TOM1$ (Nozoye et al., 2011). Furthermore, additional orthologous to GRMZM2G063306 were also identified in other graminaceous species including sorghum, brachypodium, switchgrass, and foxtail millet. The fact that several orthologous were identified for GRMZM2G063306 and that its described function was related to DMA efflux, not only confirmed that the mechanism for Fe acquisition is well conserved across graminaceous species, but also indicated that GRMZM2G063306 may be the $Ys3$.

The sequence and location of GRMZM2G063306 has rapidly changed along with the release of the maize genome sequence. In the BAC-based Maize B73 sequence, this gene was located on chromosome 5 and only one transcript was predicted with a positive strand orientation. However, in the Maize B73 RefGen_v1, this gene was located on chromosome 3 in positions 83,194,573-83,196,907 and was part of the working gene set (WGS), but not in the filtered gene set (FGS). In our study, only genes from the FGS, which are evidence-based gene models, and located in our fine mapping region were sequenced one by one to identify causative mutations. Furthermore, Zm$TOM1$ was located outside of our fine mapping region, due to the positioning of ctg 255 between ctg 120 and 121, and therefore, not taken on account as positional candidate gene. In the maize RefGen_v2, $Ys3$ was located in chromosome 3 in positions 112,044,581-112,047,482. This gene was predicted to have a negative strand orientation and a total of two transcripts, $Ys3$_T01 and $Ys3$_T02 that were characterized by the presence of a Major Facilitator Superfamily (MFS) domain. In addition, three and eight transmembrane helices were identified in $Ys3$_T01 and _T02 using an *in silico* transmembrane identification tool, strongly suggesting that $Ys3$ was a membrane protein and may be potentially responsible for the phytosiderophore efflux (Hirokawa et al., 1998).

A total of thirteen members of the $Ys3$ gene family were identified for $Ys3$ showing a homology percentages ranging from 15 to 66% on the protein level. However, based on the physical location of $Ys3$ on choromosome 3

and its sorghum ortholog on chromosome 8, the $Ys3$ gene was identified to be part of the maize subgenome 2. However, the maize subgenome 1 region of $Ys3$ might be located either on chromosome 1 or 10. The number of genes retained in the maize subgenome 1 was low with only 126 genes (Schnable et al., 2011). In our study, two members of the $Ys3$ gene familiy have been identified in both chromosomes. Nevertheless, GRMZM2G415785 and GRMZM2G101928 showed a low homology percentage of 41 and 26% based on the protein level, respectively, suggesting that both genes showed some similarity, but failed to be identified as a second copy for $Ys3$.

$Ys3$ was sequenced and compared with parental genotypes 311F, W22, and B73 in order to identify any presumably polymorphism responsible for the $ys3$ phenotype. A total of two InDels in the 5´-untranslated region (5´-UTR) and two SNPs in the coding sequence (CDS) were observed in the comparison. However, to ensure that these polymorphisms were unique for the $ys3$ mutant and causative for its phenotype, all parental genotypes of the NAM population were sequenced. The parental genotypes, HP301 and Oh43 showed the same InDel pattern in the 5´-UTR, indicating that these polymorphisms might not be responsible for the $ys3$ phenotype. Furthermore, based on public RNA-Seq data a comparison of expression levels between HP301, Oh43, and B73 showed no expression of $Ys3$ in shoot-apical-meristem, root, tassel, shoot, and ear tissue. This lack of expression of $Ys3$ might be caused due to that these genotypes were not subjected to any Fe deficient nor sufficient conditions (qTELLER, unpublished). Moreover, sig-

nificant expression levels of $Ys3$ were observed mainly in root tissue of the B73 inbred line (Sekhon et al., 2011).

The remaining two SNPs in the CDS were silent mutations. However, there is still a gap of approximately 100 bp in this gene between positions 112,044,481-112,044,581. In our study, several attempts to amplify this gap failed, preventing any chance of sequencing and determining any possible polymorphism in the CDS. Silent or synonymous mutations have been assumed to be neutral during protein formation. However, silent mutations can induce structural changes in the mRNA structure and therefore influence stability and protein formation or function (Czech et al., 2010). The formation of rare codons due to silent mutations can affect protein translation or production of an accurate protein sequence (Czech et al., 2010; Hunt et al., 2009). In the $ys3$ mutant, the CDS of Zm$Ys3$ showed the formation of several rare codons. For instance, the identified SNP in exon 4 resulted in the formation of a rare codon, which codes for isoleucine (ATA). Nevertheless, there was no difference in the protein sequence of YS3. Therefore, the $ys3$ phenotype might be caused due to an expression change caused by possible polymorphisms either in the promoter region or in the 100 bp gap, which have not been analyzed in this study.

4.3 Expression of $Ys3$ in maize roots under different Fe regimes

In the present study, several genes involved in iron acquisition showed an increase in expression under low Fe availability. Similarly, expression of both $Ys3$ transcripts was mainly found in root tissue of W22 and $ys3$ plants grown under Fe deficiency. Similarly, $OsTOM1$ and $HvTOM1$ were expressed in Fe deficient roots showing the same expression pattern (Nozoye et al., 2011). However, expression levels of $Ys3$ in the $ys3$ mutant were much lower than in W22 plants. The reduced expression of $Ys3$ in the $ys3$ mutant might be caused by the formation of rare codons due to silent mutations or polymorphisms in the promoter region.

The expression pattern of $Ys3$ is in line with a role of $Ys3$ in Fe metabolism. Moreover, *Dmas* expression in roots of the $ys3$ mutant grown under Fe deficiency showed an increased fold change, compared with sufficient Fe conditions, confirming the active production of phytosiderophores, but the inactive efflux of the latter (Lanfranchi et al., 2002; Motta et al., 2001). In addition, the $Ys1$ gene involved in the uptake of Fe(III)-phytosiderophore complex showed also an increased abundance in plants under Fe deficient conditions (Curie et al., 2001; Von Wirén et al., 1994). Therefore, genes which are involved in the Fe uptake pathway might trigger an upregulation response against Fe starvation conditions to overcome any deficiency.

4.4 Validation of *Ys3* by Transposon Tagging

Verification assays using the mutator (*Mu*) and activator (*Ac*) systems were performed to generate new mutations allelic to the *ys3* mutant. The mutation frequency observed in the *Ys3*::*Mu* and *Ys3*::*Ac* individuals was 1.53×10^{-5} and 7.53×10^{-4}, respectively. The *Mu* element is commonly used due to its high copy number and transposition frequency compared to the *Ac* element (Kolkman et al., 2005; Lisch, 2002). In addition, transposition in the *Mu* is performed randomly, in contrast to the *Ac* element which preferentially transposes within 10cM of the donor element (Brutnell, 2002). In our study, the *Ac* system showed a higher mutation frequency compared to the *Mu*. This difference in mutation frequency might be caused by the location of the *Ac* near the *Ys3* gene and therefore, the selected *Ac* lines carried only *Ac* insertions close to the *Ys3* gene. In the case of the *Mu* element, the higher and random transposition may lead to insertions of the *Mu* in several genes, reducing the likelihood to identify unique *Mu* lines carrying *Mu* insertions close to the *Ys3* gene.

In the present study, an additional *Ys3*::*Mu* mutation allelic to the *ys3* mutant was identified by its *yellow stripe* like phenotype. Marker analyses with SSR markers umc1773 and bnlg 1957 confirmed that the additional mutation was indeed a F_1-individual derived from the cross between the *Mu* donor and the *ys3* mutant (data not shown).

Therefore, a set of 10 F_1S_1 individuals and parental genotypes, which were demonstrated to be *ys3*, because the wt phenotype could be restored

by supplying Fe, were sequenced in order to identify any footprint left by the *Mu* element. *Mu* insertions are usually inserted in the 5´-UTRs or exons and leave a 9 bp target site duplication (TSD) (Dietrich et al., 2002; Liu et al., 2009). However, different sizes of the TSDs have been observed in additional mutations allelic to the *gl*8 gene (Dietrich et al., 2002). In the present study, an additional F_1S_1 *Ys*3::*Mu* individual showed an insertion of six bp TSD in exon 8 of *Ys*3. Thus, confirming *Ys*3 as the responsible gene for the *ys*3 phenotype.

4.5 Maize root transcriptome on *ys*3 background F_2-individuals grown under different Fe regimes

4.5.1 Phenotypic response

Performance of the *ys*3 F_2-individuals was weaker across all harvesting coefficients, SPAD, and micronutrient content in both Fe regimes, in contrast with wt F_2-individuals. However, these results were expected as it is well known that the *ys*3-mutant, parent of the F_2 population, showed a limited growth when cultivated under Fe limiting conditions (Lanfranchi et al., 2002). In fact, the response of *ys*3 individuals in terms of SPAD content was lower and followed the same pattern across SB3, SB4, SB5, SA3, SA4, and

SA6, in comparison with wt individuals. However, it was clear in SA6 at Fe deficient conditions that *ys*3 individuals were highly susceptible to Fe deficiency and thus, indicating its limitations on Fe homeostasis. SPAD content in leaves SB3 and SA3 was similar before and after the application of both Fe regimes. On the contrary, SPAD in leaves SB4 and SA4 showed an increase in content after the application of both Fe regimes, showing a clear Fe remobilization across the plant. However, *ys*3 individuals showed a limited increase in SPAD content at a Fe deficient regime, in contrast to a sufficient regime. Nevertheless, the response of *ys*3 individuals at a Fe sufficient regime remained lower than in wt individuals, indicating that *ys*3 individuals were not able to handle properly the Fe found in the hydroponic solution, despite of its availability, showing its inefficient ability to handle Fe. Furthermore, SPAD content was consistently higher at sufficient conditions (300μM Fe-EDTA) than at deficient conditions, indicating that SPAD content showed a dependency on Fe availability (Lanfranchi et al., 2002).

Moreover, *ys*3 and wt individuals showed a lower Fe content when grown under Fe deficient conditions. As observed in harvesting coefficients, *ys*3 individuals presented lower Fe content than wt individuals in both Fe regimes. In fact, *ys*3 individuals showed 41% less Fe accumulation than wt individuals at Fe deficient conditions, strongly indicating that *ys*3 was a Fe-ineffficient mutant, which might have an impaired Fe homeostasis functioning. Similar results were observed when Fe content was measured in leaves of the *ys*3-mutant, in which a reduction between 30 - 40% was obsereved in

comparison to its isogenic wt (Motta et al., 2001).

Zn content was higher in *ys*3 individuals at both Fe regimes, and even higher at Fe deficient conditions. The fact that phytosiderophores mobilize other cations including Zn, Mn, and Cu, besides Fe, resulted in the unspecific uptake of these micronutrients and even suggesting a competition among cations due to a concentration effect (Römheld, 1991; Treeby et al., 1989; Kanai et al., 2009). Furthermore, the increased accumulation of Zn was expected in this experiment due to the strong negative correlation between Fe and Zn content (Fig. 3.9, Fig. A7, and Table B5). In this regards, exudates of barley root shown mobilization of higher amounts of cations in the following order Cu < Fe < Zn < Mn (Treeby et al., 1989). However, phytosiderophores shown a preference for Fe uptake, specially when influenced by Fe nutritional conditions (Römheld, 1991).

In this study, no significant differences were observed in Cu and Mn content in *ys*3 and wt individuals nor in both Fe regimes, indicating a preference in Fe and Zn uptake specially at Fe deficient conditions. Moreover, no significant correlations were observed between Fe, Cu, Mn, and Mg content (Fig. A7 and Table B5). On the contrary, in a study using F_4-individuals from the B84×Os6-2 population significant correlations were shown between these cations (Sorić et al., 2012). The difference between our study and the later might be explained by the difference in the setting of both experiments. For instance, the B84×Os6-2 F_4-individuals were screened in a field and Fe supplemented with fertilizers, in contrast to our experiment in which plants

were grown in a controlled hydroponic system that allowed a proper application of macro and micronutrients and where plants were subjected to two different Fe regimes.

Furthermore, Fe content was highly correlated with SA6, DW, and RW indicating that these harvesting coefficients can be used to estimate Fe content in plants. In addition, K and Zn content were also found highly correlated to these harvesting coefficients (Fig. 3.9, Table B5).

4.5.2 Maize root transcriptome profile in a *ys*3 background F_2 population

Total number of expressed genes across libraries and their coverage was slightly different between both replications. These observed differences might be explained by the fact that both replications were set differently. For instance, on replication one individual libraries were sequenced in individual lanes, while in replication two all the libraries were multiplexed. Although the total number of reads varied among the sequenced libraries, the total number of expressed genes slightly varied between both reps, indicating that either single-lane or multiplexing can be used for transcriptome sequencing and decision in either of these methods might depend on the depth of sequencing (coverage) and sequencing costs (Wang et al. 2009, Table 3.2).

In this study, novel transcript discovery was attempted for *ys*3 and wt F_2-individuals using *de novo* assembly of unmapped RNA-Seq reads. However, the analysis was unable to be carried out due to high memory usage

requirement by *Velvet* (Zerbino and Birney, 2008). *De novo* assembly of unmapped RNA-Seq reads was used in 21 inbred lines including B73 in order to identify additional novel genes specific for determined genotype. The *Velvet/Oases* (Zerbino and Birney, 2008; Schulz et al., 2012) platform was used to assemble the unmapped reads, which yielded a total of 1,321 high-confidence novel transcripts. Although, only 654 novel transcripts were present across all the 21 inbred lines, 757 transcripts were identified as unique in a subset of non-B73 inbred lines (Hansey et al., 2012).

Another advantage within the RNA-Seq technology is the identification of variants (SNP and InDels) based on RNA-Seq reads. Among the identified variants, the most common were silent and missense mutations and in a minor proportion non-sense mutations and splice site (Fig. 3.12). A similar trend was observed among the identified variants in the Maize HapMap2, in which 103 pre-domesticated and domesticated *Zea mays* varieties were sequenced (Chia et al., 2012). Thus, based on the identified variants within the *ys*3×W22 F_2-population, changes in the protein sequence are expected and even might explain the observed phenotypic differences between *ys*3 and wt individuals and their response to deficient and sufficient Fe regimes.

4.5.3 Differential expressed genes (DEG) identification

In an initial analysis using the data from both replications for each library, the calculated biological coefficient of variation (BCV) was on average approximately 40%. Furthermore, when DEG analysis was performed

using $cuffdiff$ with and without RABT assembly, $Deseq$, and $edgeR$, the number of DEG was lower and in some comparisons even nule. Common observed values for BCV in well-controlled RNA-Seq experiments are 0.4 for human data, 0.1 for data on genetically identical model organisms, and 0.01 for technical replicates (Robinson et al., 2010). In this study, consideration of any possible causes that might explain the relative high BCV include the differences in the experimental setting of both replications. For instance, in replication one plants share the same pot and therefore, roots of these plants grown together and took the shape of the pot. However, in the case of replication two, plants were split as individuals when growing under hydroponic conditions. Thus, this difference might have influenced and increased the BCV. In terms of sequencing conditions, in replication one pooled root tissue from each library was sequenced in individual lanes as in replication two, root samples were multiplexed. Therefore, number of reads and coverage of reads throught the genome was also influenced and thus, might lead to an increase in BCV. Hence, every library from each replications was analyzed as individual and later on put together based on the coincidences between $cuffdiff$, $Deseq$, and $edgeR$ (Fig. A1).

The number of identified DEGs among these procedures were slightly similar, except for $edgeR$. In fact, $Deseq$ was the most stringent test, in which the number of DEGs ranged from 44 - 129 across all four comparison (Fig. A2). Furthermore, $cuffdiff$ with and without the use of RABT-assembly yielded a similar number of DEGs ranging from 101 to 181 across

the comparisons (Fig. A3-4). However, *edgeR* using a BCV of 20% identified 113 to 966 DEGs across all comparison (Fig. A5.). Based on the observed differences among the four procedures used for the identification of DEGs, the use of either $cuffdiff$ or *Deseq* is suggested due to the better handle in the identification of DEGs especially when datasets lack replications.

In the case of *edgeR*, the identification of DEGs is based on a given arbitrary BCV. Using our data, we simulated the identification of DEGs using different BCVs including 20, 25, and 30%. Based on the results of identified DEGs with a BCV of 20%, we observed a reduction of 56% and 84% of the number of genes across all comparison with the use of a higher BCV. These findings indicate that *edgeR* might not be a good option for the identification of DEGs when replicatios are lacking, unless previous knowledge on BCV is known with certantly.

4.5.4 Candidate gene identification and its response to Fe metabolism

Four comparison were analyzed in this study based on Fe response among the *ys*3 and wt F_2-individuals grown under 10 vs. 300 μM Fe-EDTA as comparison 1 and 2, respectively and on genotype differences including *ys*3 vs. wt grown under 10 and 300 μM Fe-EDTA as comparison 3 and 4, respectively. A total of 190 DEGs were identified across all comparison. However, individual comparisons such as 1, 2, 3, and 4 yielded a total of 115, 57, 49, and 38 DEGs, respectively (Fig. 3.10). A GO-enrichment analysis was per-

formed using the total number of identified DEGs in order to unravel their involvement into a specific pathway (Fig. 3.11). In this regard, the identified DEGs showed an involvement on ion, cation, and chemical homeostasis including Fe and other di−, tri−valent cations, which suggest the activation of responsive mechanisms to establish and maintain ion homeostasis in order to cope deficiencies. In addition, the importance of methionine, sulfur, and nitrogen processes were also highlighted and thus, indicating its association with phytosiderophore biosynthesis (Ma et al., 1995). Other important pathways related to stress response were induced including response to oxidative stress (ROS) and defense responses, which might be triggered as a response to Fe deficiency in attempt to maintain cellular homeostasis (O′Rourke et al., 2007). The fact that responses such as iron homeostasis, methionine, sulfur, and stress related responses were identified, evidence the immediate response triggered by deficient and sufficient Fe regimes inside of the plant in order to establish, regulate or maintain cellular homeostasis. Overall, comparison 3 and 4 yielded the lower amount of DEGs and identified more stress−response related genes (Table B8-9). However, comparisons 1 and 2 identified Fe−related genes as well as a higher number of candidate genes than any other comparison. Therefore, these comparisons were further discussed (Table B6-7).

Comparison 1 and 2 identified several genes involved in the methionine cycle (*ZmMTK*, *ZmMTN*, *ZmFDH*, *ZmIDI*4, *ZmPRPPs*, and *ZmRPI*) as well as phytosiderophore synthesis (*ZmDmas* and *ZmNaat*1)

and Fe homeostasis ($ZmYs1$, $ZmNramp3$, and $Fer1$).

In addition, four other genes that were orthologs to $OsNAS1$ *and* $OsNAS2$ were also identified in comparison 1 and 2 including GRMZM2G124785, AC233955.1_FG003, GRMZM2G704488, and GRMZM2G312481. $OsNAS1$ *and* $OsNAS2$ were identified previously to be highly expressed in Fe-deficient roots and involved in Fe long-distance transport (Inoue et al., 2003). However, in maize $NAS2$ was also identified to be positively expressed in Fe-deficient root, but lacking any *in vitro* activity due to the duplication of its structure, which is very similar to $ZmNAS1$ (Mizuno et al., 2003). Moreover, GRMZM2G124785 has 601 AA and presented 97% similarity to $NAS2$, which was also expressed in comparison 1 and 2 (Table B6-7). Furthermore, AC233955.1_FG003, GRMZM2G704488, and GRMZM2G312481 have 327 AAs and showed an average similarity of 94% with $NAS1$. The fact that a high homology is observed between these NAS isoforms, strongly indicates that an inaccurate association of the reads to a specific isoform might be observed and thus, hampering the detection of true DEGs between these isoforms.

Furthermore, GRMZM2G057413 and GRMZM2G350312 were also identified to be upregulated in Fe deficient conditions in both comparison. These transcription factors presented a helixloophelix domains (bHLH). Moreover, GRMZM2G057413 ($ZmIro2$) is orthologue to $OsIRO2$, which is known to be involved in the regulation of Fe deficiency inducible genes (Ogo et al., 2006; Kobayashi and Nishizawa, 2012a). Similarly, GRMZM2G350312 ($ZmIro3$)

is an orthologue to $AtPYE$, which has been characterized by its involvement in maintaining iron homeostasis under Fe deficient conditions, root development, as well as stress response (Long et al., 2010). In that regard, $OsIRO2$ regulates the expression of several Strategy II genes including $OsNAS1$, $OsNAS2$, $OsNAAT1$, $OsDMAS$, $TOM1$, $OsYSL15$ as well as other genes involved in the methionine cycle (Kobayashi and Nishizawa, 2012a). In contrast, the involvement of PYE has only been described in Strategy I, *A. thaliana* and not in maize (Long et al., 2010). However, expression levels of its orthologous $OsIRO3$ measuered in overexpression lines showed a more complex regulation response, which needs to be further characterized (Kobayashi and Nishizawa, 2012a). Nevertheless, the expression of these transcription factors might improved the gene expression response triggered by Fe deficient conditions in maize, and thus, improving tolerance to Fe deficiency.

In this study, these transcription factors were strongly expressed in Fe deficient conditions and even showed a greater expression levels on $ys3$ F_2-individuals based on qRT-PCR validation . However, read coverage of $ZmIRO2$ showed an elongation after the end of transcription of approximately 6.2 Kbp (Fig. A6), indicating that the annotation of this gene model might need to be improved.

GRMZM2G325575 ($Fer1$) was also identified in comparison 1 and 2. However, it was found to be upregulated in plants grown under Fe sufficient conditions (Table 3.3). Induction of $Fer1$ has been observed in *A.*

thaliana and maize as a response to Fe overload regimes (Kobayashi and Nishizawa, 2012a). Ferritins are known to be the source of Fe storage as well as a mediator between Fe homeostasis and ROS response (Ravet et al., 2009). In addition, regulation of ferritins is performed by an iron-dependent regulatory sequences (IDRS), which repress the expression of $ZmFer1$ when cultivated under Fe deficient conditions (Briat et al., 2006). Furthermore, $ys3$ individuals shown a slightly low expression levels of $Fer1$ as observed in wt individuals. Moreover, lower Fe content, which was measured in leaf tissue was observed to be lower in $ys3$ individuals grown under Fe deficient and sufficient regimes, which suggests that $ys3$ individuals might struggle during Fe homeostasis and therefore hampering Fe allocation in the plant.

Other additional candidate genes that were identified to be expressed under Fe deficient conditions including GRMZM2G400602, GRMZM2G104563, and GRMZM5G866024, which presented a described function as transporters. Although, the major transporters involved in Strategy II have been unraveled in graminaceous species (Curie et al., 2001; Nozoye et al., 2011, 2013), the expression of these novel transporters, which were validated by qRT-PCR, indicated the possible involvement of these in intracellular Fe transport.

Conclusions

This study identified the $Ys3$ gene in maize by a map based cloning approach that confirmed its function as a specific transporter for phytosiderophore efflux in maize. Furthermore, upregulation of both transcripts produced by $Ys3$ was shown when plants were grown under Fe deficient conditions, thus, confirming the involvement of $Ys3$ in Fe metabolism. In addition, insides on genome-wide gene regulation associated with Fe deficiency using a $ys3$-mutant background was also investigated using RNA-Seq technology. Furthermore, phenotypic and ionomic analyses were also performed in order to complement our transcriptome profile and provide a complete and better understanding of gene response during Fe deficiency. Several genes involved in Fe uptake and homeostasis were identified as well as novel genes associated with Fe deficiency response, transport, and oxidation-reduction.

The results of this research provide important insides about $Ys3$ and its implication on Fe homeostasis by investigating its response when grown under deficient and sufficient Fe regimes, which can later be used to improve Fe efficiency and thus, influence Fe content in grain to tackle Fe deficiency in humans.

Bibliography

Anders, S. and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol*, 11(10):R106.

Andorf, C. M., Lawrence, C. J., Harper, L. C., Schaeffer, M. L., Campbell, D. A., and Sen, T. Z. (2010). The locus lookup tool at maizegdb: identification of genomic regions in maize by integrating sequence information with physical and genetic maps. *Bioinformatics*, 26:434–436.

Awika, J. M. (2011). *Major Cereal Grains Production and Use around the World*, chapter 1, pages 1–13. American Chemical Society.

Basso, B., Bagnaresi, P., Bracale, M., and Soave, C. (1994). The yellow stripe 1 and 3 mutants of maize: nutritional and biochemical studies. *Maydica*, 39:97–105.

Beadle, G. W. (1929). Yellow Stripe a Factor for Chlorophyll Deficiency

in Maize Located in the Pr pr Chromosome. *The American Naturalist*, 63:189–192.

Benke, A., Stich, B., and Donini, P. (2011). An analysis of selection on candidate genes for regulation, mobilization, uptake, and transport of iron in maize. *Genome*, 54(8):674–683.

Black, R. (2003). Micronutrient deficiency-an underlying cause of morbidity and mortality. *BULLETIN-WORLD HEALTH ORGANIZATION*, 81:79.

Briat, J.-F., Cellier, F., and Gaymard, F. (2006). Ferritins and iron accumulation in plant tissues. In Barton, L. and Abadia, J., editors, *Iron Nutrition in Plants and Rhizospheric Microorganisms*, pages 341–357. Springer Netherlands.

Briat, J. F., Fobis-Loisy, I., Grignon, N., Lobreaux, S., Pascal, N., Savino, G., Severino, T., von Wirén, N., and van Wuytswinkel, O. (1995). Cellular and molecular aspects of iron metabolism in plants. *Biology of the Cell*, 84:69–81.

Briat, J. F. and Lobreaux, S. (1997). Iron transport and storage in plants. *Trends in Plant Science*, 2:187–193.

Brutnell, T. P. (2002). Transposon tagging in maize. *Functional & Integrative Genomics*, 2:4–12.

Chen, Y., Liu, H., Ali, F., Scott, M., Ji, Q., Frei, U., and Lübberstedt, T. (2012). Genetic and physical fine mapping of the novel brown midrib gene bm6 in maize (*Zea mays* L.) to a 180kb region on chromosome 2. *Theoretical and Applied Genetics*, 125:1223–1235.

Chia, J.-M., Song, C., Bradbury, P. J., Costich, D., de Leon, N., Doebley, J., Elshire, R. J., Gaut, B., Geller, L., Glaubitz, J. C., et al. (2012). Maize hapmap2 identifies extant variation from a genome in flux. *Nature genetics*, 44(7):803–807.

Ching, A., Caldwell, K., Jung, M., Dolan, M., Smith, O., Tingey, S., Morgante, M., and Rafalski, A. (2002). SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. *BMC Genetics*, 3:19.

Conte, S. S. and Walker, E. L. (2011). Transporters contributing to iron trafficking in plants. *Molecular plant*, 4(3):464–476.

Curie, C., Panaviene, Z., Loulergue, C., Dellaporta, S. L., Briat, J. F., and Walker, E. L. (2001). Maize *yellow stripe*1 encodes a membrane protein directly involved in Fe(III) uptake. *Nature*, 409:346–349.

Curie, C. C. and Briat, J. F. (2003). Iron Transport and Signaling in Plants. *Annual Review of Plant Biology*, 54:183–206.

Czech, A., Fedyunin, I., Zhang, G., and Ignatova, Z. (2010). Silent mutations

in sight: co-variations in trna abundance as a key to unravel consequences of silent mutations. *Molecular Biosystems*, 6(10):1767–1772.

Davidson, R. M., Hansey, C. N., Gowda, M., Childs, K. L., Lin, H., Vaillancourt, B., Sekhon, R. S., de Leon, N., Kaeppler, S. M., Jiang, N., et al. (2011). Utility of RNA sequencing for analysis of maize reproductive transcriptomes. *The Plant Genome*, 4:191–203.

Delhomme, N., Padioleau, I., Furlong, E. E., and Steinmetz, L. M. (2012). easyrnaseq: a bioconductor package for processing rna-seq data. *Bioinformatics*, 28(19):2532–2533.

Dietrich, C. R., Cui, F., Packila, M. L., Li, J., Ashlock, D. A., Nikolau, B. J., and Schnable, P. S. (2002). Maize mu transposons are targeted to the 5 untranslated region of the gl8 gene and sequences flanking mu target-site duplications exhibit nonrandom nucleotide composition throughout the genome. *Genetics*, 160:697–716.

Du, Z., Zhou, X., Ling, Y., Zhang, Z., and Su, Z. (2010). agrigo: a go analysis toolkit for the agricultural community. *Nucleic acids research*, 38(suppl 2):W64–W70.

Eveland, A. L., Satoh-Nagasawa, N., Goldshmidt, A., Meyer, S., Beatty,

M., Sakai, H., Ware, D., and Jackson, D. (2010). Digital gene expression signatures for maize development. *Plant physiology*, 154(3):1024–1039.

Godsey, C. B., Schmidt, J. P., Schlegel, A. J., Taylor, R. K., Thompson, C. R., and Gehl, R. J. (2003). Correcting Iron Deficiency in Corn with Seed Row - Applied Iron Sulfate. *Agronomy Journal*, 95:160–166.

Guerinot, M. L. (2001). Improving rice yields - ironing out the details. *Nature*, 19:417–418.

Guerinot, M. L. and Yi, Y. (1994). Iron: nutritious, noxious, and not readily available. *Plant Physiology*, 104:815.

Hansey, C. N., Vaillancourt, B., Sekhon, R. S., de Leon, N., Kaeppler, S. M., and Buell, C. R. (2012). Maize (*zea mays* l.) genome diversity as revealed by rna-sequencing. *PLoS One*, 7(3):e33071.

Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). Sosui: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*, 14(4):378–379.

Hunt, R., Sauna, Z. E., Ambudkar, S. V., Gottesman, M. M., Kimchi-Sarfaty, C., et al. (2009). Silent (synonymous) SNPs: should we care about them? *Methods Mol. Biol.*, 578:23–39.

Ingvardsen, C., Xing, Y., Frei, U., and Lüberstedt, T. (2010). Genetic and physical fine mapping of *scmv*2, a potyvirus resistance gene in maize. *Theoretical and Applied Genetics*, 120:1621–1634.

Inoue, H., Higuchi, K., Takahashi, M., Nakanishi, H., Mori, S., and Nishizawa, N. K. (2003). Three rice nicotianamine synthase genes, osnas1, osnas2, and osnas3 are expressed in cells involved in long-distance transport of iron and differentially regulated by iron. *The Plant Journal*, 36(3):366–381.

Kanai, M., Hirai, M., Yoshiba, M., Tadano, T., and Higuchi, K. (2009). Iron deficiency causes zinc excess in Zea mays. *Soil Sci. Plant Nutr.*, 55:271–276.

Kobayashi, T. and Nishizawa, N. (2012a). Iron uptake, translocation, and regulation in higher plants. *Annual Review of Plant Biology*, 63:131–152.

Kobayashi, T. and Nishizawa, N. K. (2012b). Iron uptake, translocation, and regulation in higher plants. *Annual Review of Plant Biology*, 63:131–152.

Kolkman, J. M., Conrad, L. J., Farmer, P. R., Hardeman, K., Ahern, K. R., Lewis, P. E., Sawers, R. J. H., Lebejko, S., Chomet, P., and Brutnell, T. P. (2005). Distribution of Activator (Ac) throughout the maize genome for use in regional mutagenesis. *Genetics*, 169:981–995.

Lanfranchi, S., Basso, B., and Soave, C. (2002). The *yellow stripe* 3 mutant of maize is defective in phytosiderophore secretion. *Maydica*, 47:181–184.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S. L., et al. (2009). Ultrafast and memory-efficient alignment of short dna sequences to the human genome. *Genome Biol*, 10(3):R25.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., et al. (2009). The sequence alignment/map format and samtools. *Bioinformatics*, 25(16):2078–2079.

Li, P., Ponnala, L., Gandotra, N., Wang, L., Si, Y., Tausta, S. L., Kebrom, T. H., Provart, N., Patel, R., Myers, C. R., et al. (2010a). The developmental dynamics of the maize leaf transcriptome. *Nature genetics*, 42:1060–1067.

Li, Q., Yang, X., Bai, G., Warburton, M. L., Mahuku, G., Gore, M., Dai, J., Li, J., and Yan, J. (2010b). Cloning and characterization of a putative gs3 ortholog involved in maize kernel development. *TAG*, 120:753–763.

Lisch, D. (2002). *Mutator* transposons. *Trends in Plant Science*, 7:498–504.

Liu, S., Yeh, C.-T., Ji, T., Ying, K., Wu, H., Tang, H. M., Fu, Y., Nettleton, D., and Schnable, P. S. (2009). *Mu* transposon insertion sites and meiotic

recombination events co-localize with epigenetic marks for open chromatin across the maize genome. *PLoS Genetics*, 5:e1000733.

Lobreaux, S., Massenet, O., and Briat, J.-F. (1992). Iron induces ferritin synthesis in maize plantlets. *Plant molecular biology*, 19:563–575.

Long, T. A., Tsukagoshi, H., Busch, W., Lahner, B., Salt, D. E., and Benfey, P. N. (2010). The bhlh transcription factor popeye regulates response to iron deficiency in arabidopsis roots. *The Plant Cell*, 22(7):2219–2236.

Luo, N., Yu, X., Liu, J., and Jiang, Y. (2012). Nucleotide diversity and linkage disequilibrium in antioxidant genes of *Brachypodium distachyon*. *Plant Science*, 197:122–129.

Ma, J. F., Shinada, T., Matsuda, C., and Nomoto, K. (1995). Biosynthesis of phytosiderophores, mugineic acids, associated with methionine cycling. *Journal of Biological Chemistry*, 270(28):16549–16554.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., and Cunningham, F. (2010). Deriving the consequences of genomic variants with the ensembl api and snp effect predictor. *Bioinformatics*, 26(16):2069–2070.

Mizuno, D., Higuchi, K., Sakamoto, T., Nakanishi, H., Mori, S., and Nishizawa, N. K. (2003). Three nicotianamine synthase genes isolated

from maize are differentially regulated by iron nutritional status. *Plant physiology*, 132(4):1989–1997.

Motta, A., Basso, B., DellOrto, M., Briat, J.-F., and Soave, C. (2001). Ferritin synthesis in response to iron in the fe-inefficient maize mutant *ys*3. *Plant Physiology and Biochemistry*, 39:464–435.

Motta, A., Soave, C., and Basso, B. (1999). An integrated map of a portion of chromosome 3l, in proximity of the *ys*3 locus. *Maize Genetics Conference Abstracts*, 41:78.

Nozoye, T., Nagasaka, S., Kobayashi, T., Takahashi, M., Sato, Y., Sato, Y., Uozumi, N., Nakanishi, H., and Nishizawa, N. K. (2011). Phytosiderophore efflux transporters are crucial for iron acquisition in graminaceous plants. *The Journal of biological chemistry*, 286:5446–5454.

Nozoye, T., Nakanishi, H., and Nishizawa, N. K. (2013). Characterizing the Crucial Components of Iron Homeostasis in the Maize Mutants *ys*1 *and ys*3. *PLOS ONE*, 8(5):e62567.

Ogo, Y., Itai, R. N., Nakanishi, H., Inoue, H., Kobayashi, T., Suzuki, M., Takahashi, M., Mori, S., and Nishizawa, N. K. (2006). Isolation and characterization of iro2, a novel iron-regulated bhlh transcription factor in graminaceous plants. *Journal of experimental botany*, 57(11):2867–2878.

O′Rourke, J., Charlson, D., Gonzalez, D., Vodkin, L., Graham, M., Cianzio, S., Grusak, M., and Shoemaker, R. (2007). Microarray analysis of iron deficiency chlorosis in near-isogenic soybean lines. *BMC Genomics*, 8.

Ravet, K., Touraine, B., Boucherez, J., Briat, J.-F., Gaymard, F., and Cellier, F. (2009). Ferritins control interaction between iron homeostasis and oxidative stress in arabidopsis. *The Plant Journal*, 57(3):400–412.

Roberts, A., Pimentel, H., Trapnell, C., and Pachter, L. (2011). Identification of novel transcripts in annotated genomes using rna-seq. *Bioinformatics*, 27(17):2325–2329.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140.

Römheld, V. (1987). Different strategies for iron acquisition in higher plants. *Physiologia Plantarum*, 70:231–234.

Römheld, V. (1991). The role of phytosiderophores in acquisition of iron and other micronutrients in graminaceous species: An ecological approach. *Plant and Soil*, 130(1-2):127–134.

Rozen, S. and Skaletsky, H. (1999). Primer3 on the www for general users and for biologist programmers. *MMethods Mol. Biol.*, 132:365–386.

Saghai Maroof, M. A., Soliman, K. M., Jorgensen, R. A., and Allard, R. W. (1984). Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance,chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences*, 81:8014–8018.

Schaeffer, M., Sanchez-Villeda, H., and Coe, E. (2008). IBM Neighbors 2008.

Schnable, J. C., Springer, N. M., and Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of the National Academy of Sciences*, 108:4069–4074.

Schnable, P., D, W., Fulton, R., Stein, J., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T., Minx, P., Reily, A., Courtney, L., Kruchowski, S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, W., Rock, S., Belter, E., Du, F., Kim, K., Abbott, R., Cotton, M., Levy, A., Marchetto, P., Ochoa, K., Jackson, S., Gillam, B., Chen, W., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., and Cordes, M. (2009). The B73 Maize Genome: Complexity, Diversity, and Dynamics. *Science*, 326:1112–1115.

Schulz, M. H., Zerbino, D. R., Vingron, M., and Birney, E. (2012). Oases:

robust de novo rna-seq assembly across the dynamic range of expression levels. *Bioinformatics*, 28(8):1086–1092.

Sekhon, R. S., Lin, H., Childs, K. L., Hansey, C. N., Buell, C. R., de Leon, N., and Kaeppler, S. M. (2011). Genome-wide atlas of transcription during maize development. *The Plant journal*, 66:553–563.

Shojima, S., Nishizawa, N.-K., Fushiya, S., Nozoe, S., Irifune, T., and Mori, S. (1990). Biosynthesis of phytosiderophores in vitro biosynthesis of 2-deoxymugineic acid from l-methionine and nicotianamine. *Plant Physiology*, 93(4):1497–1503.

Sorić, R., Ledenčan, T., Zdunić, Z., Jambrović, A., Brkić, I., Lončarić, Z., Kovačević, V., and Šimić, D. (2012). Quantitative trait loci for metal accumulation in maize leaf. *Maydica*, 56(4).

Stoltzfus, R. J. (2001). Iron-Deficiency Anemia : Reexamining the Nature and Magnitude of the Public Health Problem Defining Iron-Deficiency Anemia in Public Health Terms : A Time for Reflection. *Journal of Nutrition*, 131:565–567.

Sturaro, M., Hartings, H., Schmelzer, E., Velasco, R., Salamini, F., and Motto, M. (2005). Cloning and characterization of glossy1, a maize gene

involved in cuticle membrane and wax production. *Plant physiology*, 138:478–489.

Suzuki, M., Takahashi, M., Tsukamoto, T., Watanabe, S., Matsuhashi, S., Yazaki, J., Kishimoto, N., Kikuchi, S., Nakanishi, H., Mori, S., et al. (2006). Biosynthesis and secretion of mugineic acid family phytosiderophores in zinc-deficient barley. *The Plant Journal*, 48(1):85–97.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, 24:1596–1599.

Thiel, T., Michalek, W., Varshney, R. K., and Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeumvulgare* L.). *Theoretical and Applied Genetics*, 106:411–422.

Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L. A., Rhee, S. Y., and Stitt, M. (2004). mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal*, 37(6):914–939.

Thomine, S. and Vert, G. (2013). Iron transport in plants: better be safe than sorry. *Current opinion in plant biology*.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., and Pachter, L. (2012). Differential gene and transcript expression analysis of rna-seq experiments with tophat and cufflinks. *nature protocols*, 7(3):562–578.

Treeby, M., Marschner, H., and Rmheld, V. (1989). Mobilization of iron and other micronutrient cations from a calcareous soil by plant-borne, microbial, and synthetic metal chelators. *Plant and Soil*, 114(2):217–226.

Van Ooijen, J. W. (2006). *JoinMap® 4 - Software for the calculation of genetic linkage maps in experimental populations*. Kyazma B.V., Wageningen, Netherlands.

Vicient, C. (2010). Transcriptional activity of transposable elements in maize. *BMC Genomics*, 11:601.

Von Wirén, N., Mori, S., Marschner, H., and Romheld, V. (1994). Iron inefficiency in maize mutant *ys1* (*Zea mays* L. cv Yellow-Stripe) is caused by a defect in uptake of iron phytosiderophores. *Plant Physiol.*, 106:71–77.

Walker, E. L. and Connolly, E. L. (2008). Time to pump iron: iron-deficiency-signaling mechanisms of higher plants. *Current Opinion in Plant Biology*, 11:530–535.

Wang, Z., Gerstein, M., and Snyder, M. (2009). Rna-seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10:57–63.

Wei, F., Zhang, J., Zhou, S., He, R., Schaeffer, M., Collura, K., Kudrna, D., Faga, B. P., Wissotski, M., Golser, W., Rock, S. M., Graves, T. a., Fulton, R. S., Coe, E., Schnable, P. S., Schwartz, D. C., Ware, D., Clifton, S. W., Wilson, R. K., and Wing, R. A. (2009). The physical and genetic framework of the maize B73 genome. *PLoS Genetics*, 5:1–9.

Wen, T.-J., Hochholdinger, F., Sauer, M., Bruce, W., and Schnable, P. S. (2005). The roothairless1 gene of maize encodes a homolog of sec3, which is involved in polar exocytosis. *Plant physiology*, 138:1637–1643.

WHO (2008). *Worldwide prevalence of anaemia 1993-2005*.

Wright, J. (1961). A new yellow stripe on chromosome 3. *Maize Newsletter*, 35:111.

Zerbino, D. R. and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de bruijn graphs. *Genome research*, 18(5):821–829.

Appendix I

Supplementary figures

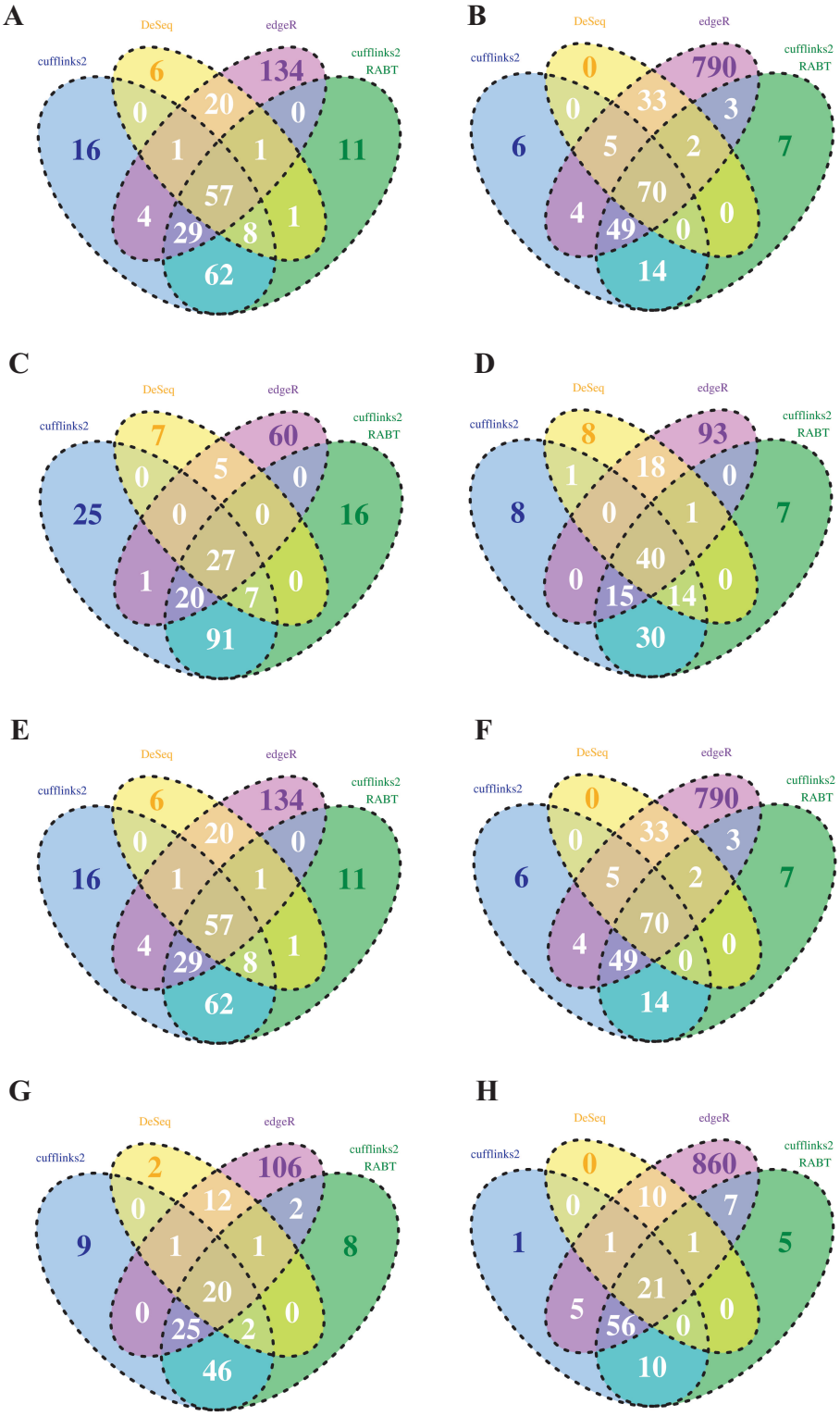

Fig. A1 Venn Diagram of overlapping differentially expressed genes identified using a FDR $\leq$0.05 with *Cuffdiff* with and without RABT-assembly, *Deseq*, and *edgeR* in **A, B)** comparison 1, **C, D)** 2, **E, F)** 3, and **G, H)** 4. **A, C, E, G)** represent replication 1 and **B, D, F, H)** replication 2.

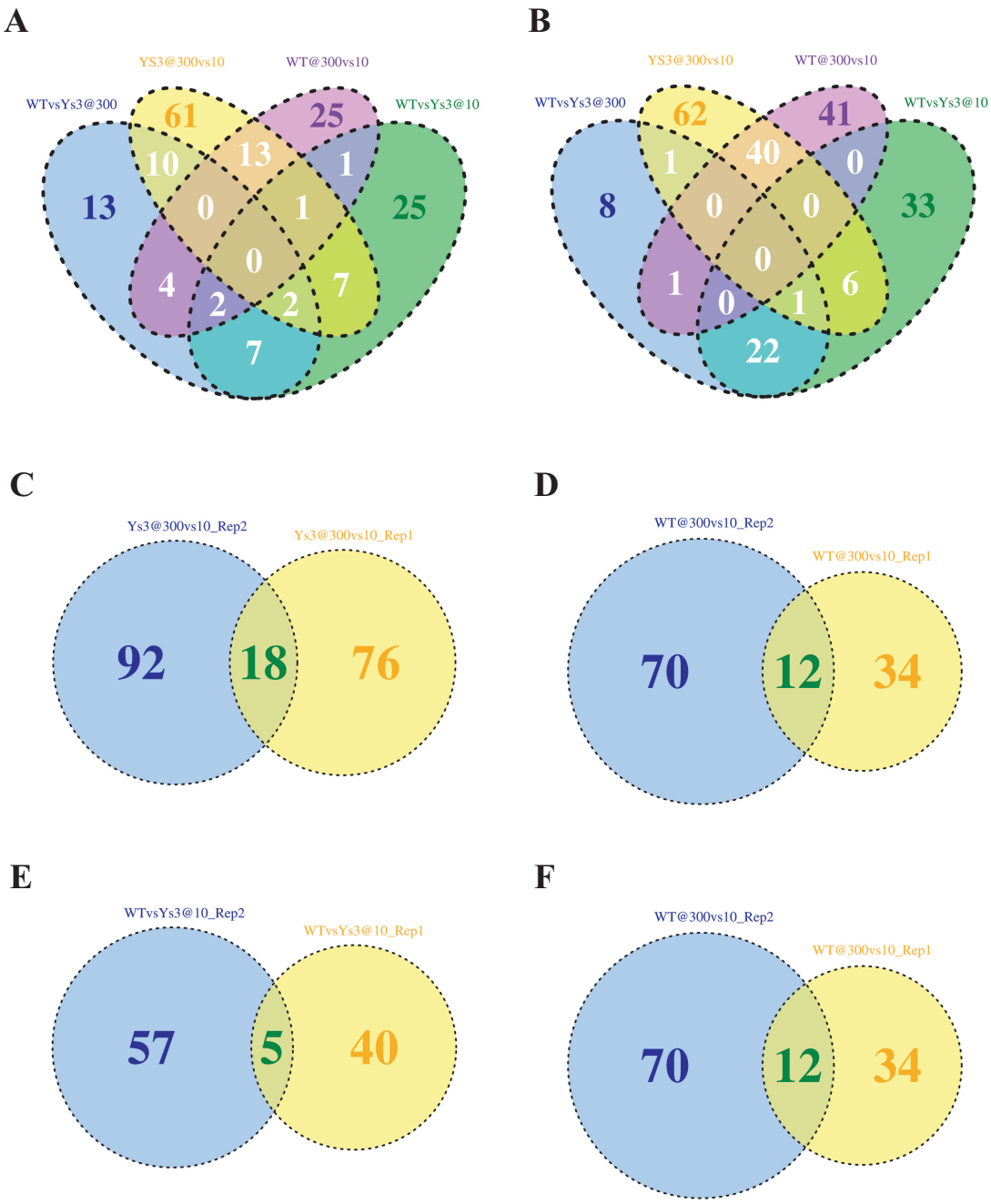

Fig. A2 Overall representation of differentially expressed genes (DEG) across replication **A** one and **B** two. DEGs were determined using *Deseq* across comparison **C** 1, **D** 2, **E** 3, and **F** 4 using a FDR ≤ 0.05.

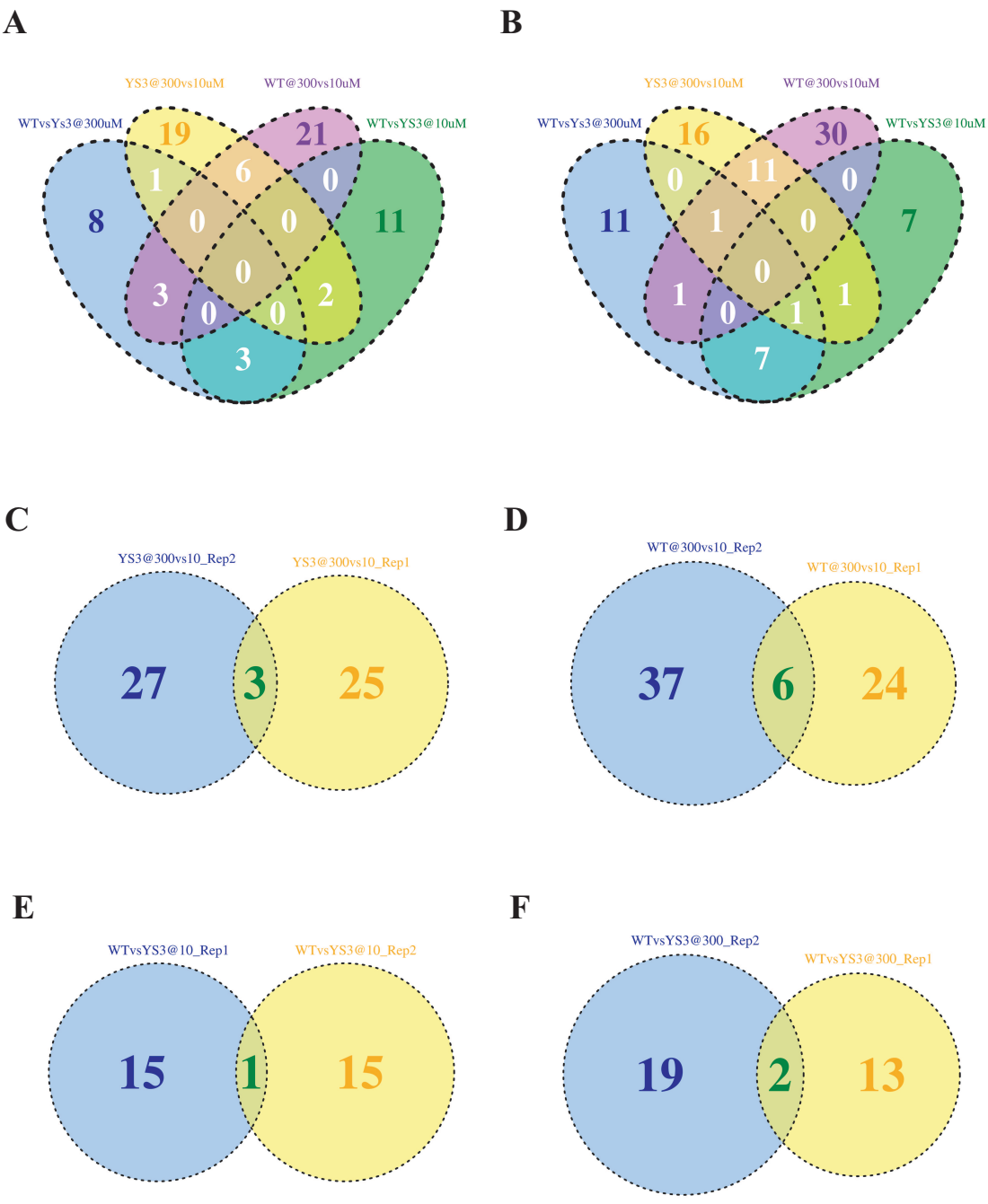

Fig. A3 Overall representation of differentially expressed genes (DEG) across replication **A** one and **B** two. DEGs were determined using $Cuffdiff$ across comparison **C** 1, **D** 2, **E** 3, and **F** 4 using Bonferroni correction with ≤ 0.05.

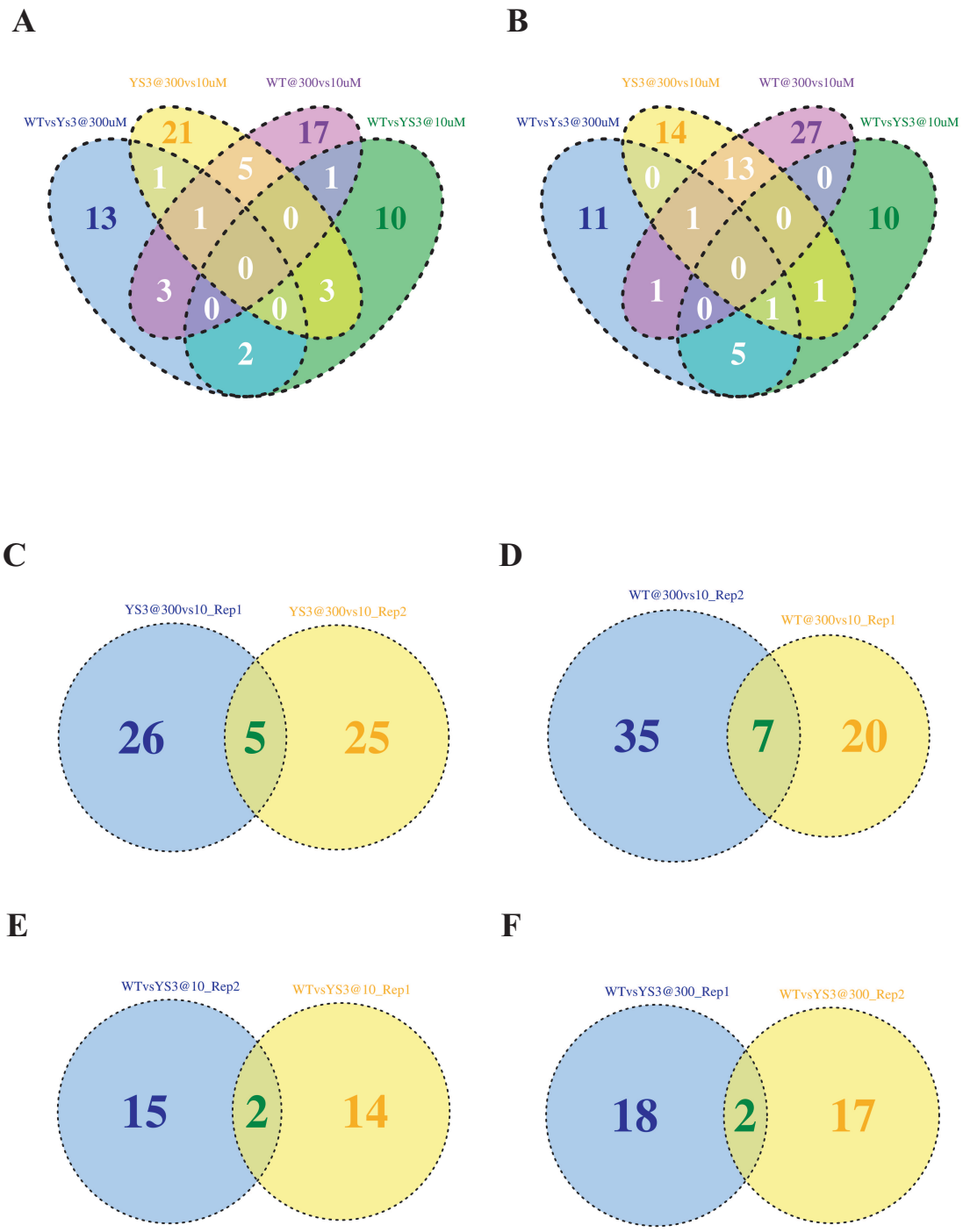

Fig. A4 Overall representation of differentially expressed genes (DEG) across replication **A** one and **B** two. DEGs were determined using *Cuffdiff* with RABT-assembly across comparison **C** 1, **D** 2, **E** 3, and **F** 4 using Bonferroni correction with ≤ 0.05.

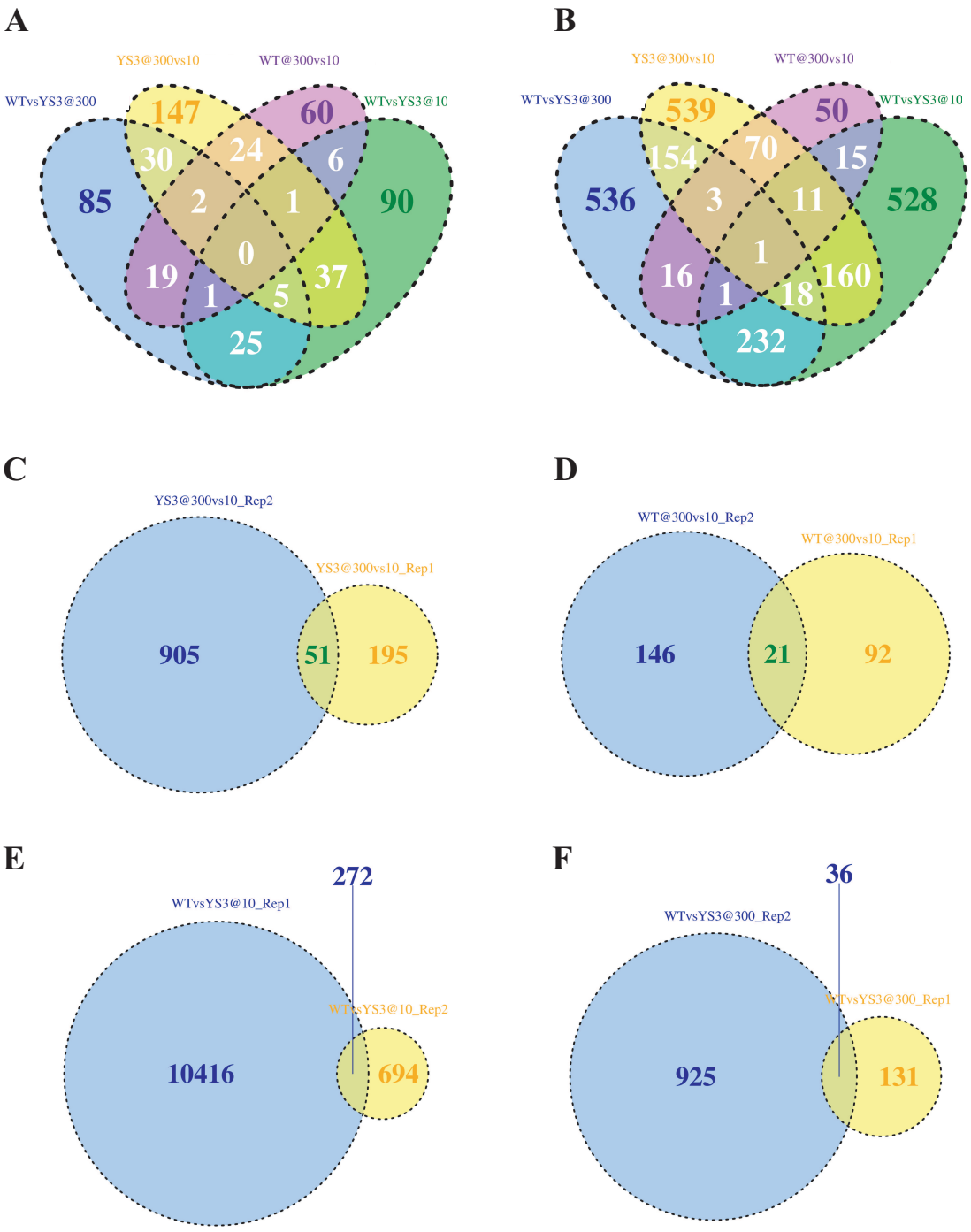

Fig. A5 Overall representation of differentially expressed genes (DEG) across replication **A** one and **B** two. DEGs were determined using *edgeR* across comparison **C** 1, **D** 2, **E** 3, and **F** 4 using a FDR ≤ 0.05.

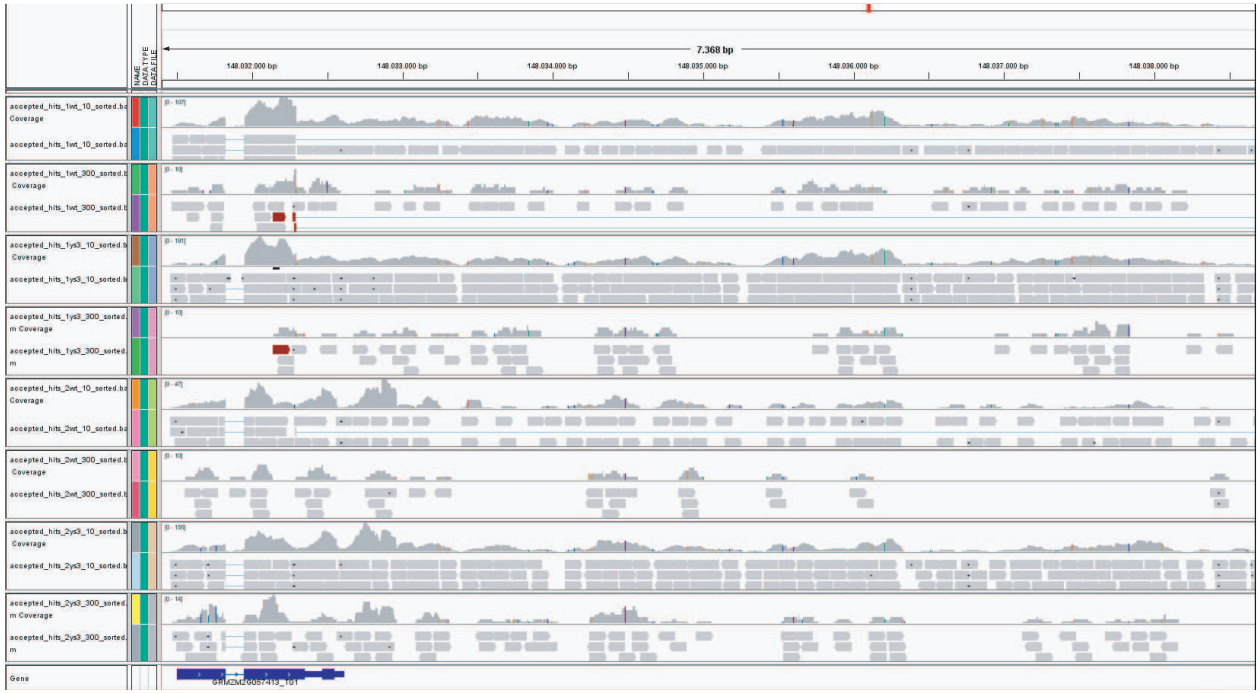

Fig. A6 Read coverage for GRMZM2G057413 visualized with IGV (v.2.0).

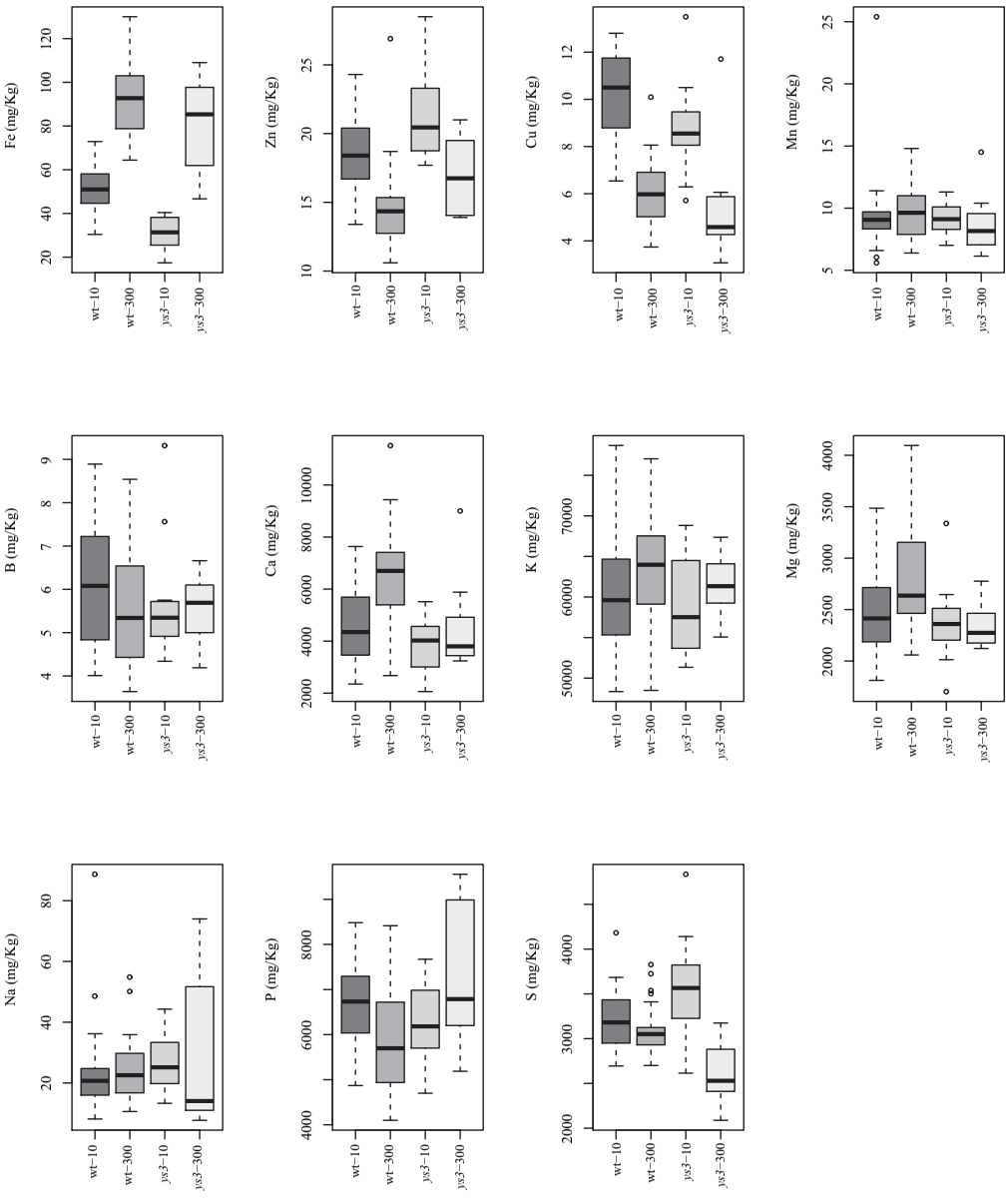

Fig. A7 Response of 11 micronutrients measured by inductively coupled plasma optical emission spectroscopy (ICP-OES) in individuals of the *ys*3×W22 F_2 population.

Appendix II

Supplementary tables

Table B1 List of primers used for fine mapping of *ys*3 and qRT-PCR of TOM1. Haplotype information included for parental genotypes W22 and 311F.

Gene / marker	Type	Genotype	Haplotype	Forward primer (5'-3')	Reverse primer (5'-3')
HAP-84	Fine Map	W22	C,T,A,C,G,C,C,C,A,C,A,A	ATCCGCAGACAGTTCACTCA	GTGCAAAATGACCATCTACCC
		311F	A,A,T,T,A,T,T,T,G,G,G,T		
HAP-118	Fine Map	W22	A,A,T,A	GCTTGATGTGGAGCATTGAA	TAGGTGAGGAATTGGGCAAG
		311F	G,G,C,G		
HAP-119	Fine Map	W22	T,C,G,A,A	GATTCCGGCACAAATGAAAG	ATGGCAAAGGATGGAACAAA
		311F	C,A,A,C,G		
HAP-124	Fine Map	W22	Present	GTGAGGCACCCCTATCTCTG	TCGGATCTAAAACGGAGAAGTT
		311F	Absent		
HAP-129	Fine Map	W22	Present	CCAAAAGAGGGCCACACTAA	TTTCCAGCACTCGTCTCCTT
		311F	Absent		
SSR-101	Fine Map			CCCCATGACTCCACATCAAT	CCACAACCATGACCAACTCA
SSR-130	Fine Map			TTCAACTCAGGCCACACAAA	TCGTCTACTCTCTGGATGGACA
CAPS-1	Fine Map			CAGCCATAGCCACTCTCCTC	CCCCTGCTGTCCCTTATGTA
Ys3	Fine Map	W22	G,G,T,T,G	CCAACCGTCTTTTGTCCTTG	TTATTTTGTGCGAGGGGAGA
		311F	A,A,G,G,A		
TOM1_T01	qRT-PCR			TGGACTGTACGCTGGTTTTC	AAGCAATCTTGTGGTGATGG
TOM1_T02	qRT-PCR			CGACTGCAATGTCCTTCTTTA	CGGTACTGCTAGGAATGGTTT
Dmas1	qRT-PCR			CTTCACGCCCGAGGACTT	ATGGTGGCGAAGGAGAGC
Actin1	qRT-PCR			ACCTCACCGACCACCTAATG	GCAGTCTCCAGCTCCTGTTC

Table B2 Synteny analysis of all genes present in the fine mapping region of maize with rice and sorghum genomes.

Zea mays genes	B73 RefGen_v2 location	Homologues		Predicted function	Transmembrane domain
		Oryza sativa	*Sorghum bicolor*		
GRMZM5G878970	99710199-99728406	Os12g14110	Sb08g009810	Uncharacterized Protein	No
GRMZM2G409893	100248087-100285775	Os01g33090	Sb03g021850	ATP Binding	Yes
GRMZM2G401179	101037267-101041120	Os05g38980	Sb03g022070	Ion Binding	Yes
GRMZM2G349191	101324993-101327116	Os04g58700	Sb06g033400	ATP Binding	Yes
GRMZM2G045714	101957769-101964970	Os01g33784	Sb03g022580	Lipid Metabolic process	Yes
GRMZM2G084208	102213870-102214796	Os11g47330	Sb05g027180	ATP Binding	No
GRMZM2G082437	102488454-102489626	Os01g34350	Sb03g022880	RNA polymerase activity	No
GRMZM2G145201	102518470-102521623	Os01g34350	Sb03g022880	RNA polymerase activity	No
GRMZM5G837123	102939883-102940838	Os12g22800	Sb08g015131	Uncharacterized Protein	No
GRMZM2G461983	103099092-103100756	Os02g14500	Sb06g010606	Uncharacterized Protein	No
GRMZM2G353957	103172508-103175313	Os12g30150	Sb08g014910	ATP Binding	Yes
GRMZM2G054852	103325154-103326907	Os01g74480	Sb01g012530	Cu and Zn Ion Binding	No
GRMZM2G125832	103798946-103802578	Os12g30040	Sb08g014900	Transmembrane Helix	Yes
GRMZM2G416913	103815531-103817630	Os11g03700	Sb08g001570	Zn and Ion Binding	No
GRMZM2G083643	103899326-103900331	Os12g30000	Sb08g014880	Uncharacterized Protein	No
GRMZM2G179071	105058467-105058857	Os03g51984	Sb01g009310	Methylation	No
GRMZM2G156813	105345433-105347329	Os12g29950	Sb08g014860	Transmembrane transport	Yes
GRMZM2G163464	105481151-105482225	Os01g67870	Sb03g043150	Uncharacterized Protein	Yes
GRMZM2G306193	106123061-106131455	Os01g58680	Sb03g037240	DNA Binding and Repair	No
GRMZM2G014805	106336887-106353717	Os12g30540	Sb08g015020	Ubiquitin-dependent Protein	No
GRMZM2G016457	106385610-106413663	Os12g18630	Sb08g011730	Chloroplast Inner Membrane	No
GRMZM5G847787	106409358-106411423	Os07g22590	Sb05g007530	Nucleic acid Binding	No
GRMZM2G444195	106768084-106769467	Os05g36160	Sb09g021580	DNA Binding	No
GRMZM2G008123	106951826-106958399	Os06g20610	Sb08g006830	Multicellular Organismal Development	No
GRMZM2G539377	107180671-107184887	Os10g36710	Sb01g017320	ATP Binding	Yes

Table B2 continued

Zea mays genes	B73 RefGen_v2 location	Homologues		Predicted function	Transmembrane domain
		Oryza sativa	*Sorghum bicolor*		
AC198518.3_FG003	107351263-107362589	Os12g27994	Sb08g013180	Uncharacterized Protein	No
AC198518.3_FG002	107406722-107407300	Os12g29480	Sb08g014080	SAM domain	No
GRMZM2G384695	107473593-107474909	Os12g29500	Sb08g014181	Uncharacterized Protein	No
GRMZM2G352159	107571179-107577553	Os12g29520	Sb08g014320	Auxin Signalin	No
GRMZM2G404132	108141894-108146866	Os12g29580	Sb08g014350	ATP binding	No
GRMZM5G826456	108422620-108424952	Os03g57430	Sb0011s0139	Uncharacterized Protein	No
GRMZM2G179349	108706693-108708062	Os12g29220	Sb08g013840	Sugar Transport	Yes
AC207392.3_FG001	108723852-108724739	Os08g45170	Sb07g023510	Carboxyl-terminal Proteinase	No
GRMZM2G319465	108824593-108836300	Os12g28270	Sb0011s0118	Hydrolase Activity	No
GRMZM2G439578	108858232-108871663	Os12g28270	Sb0011s0118	Hydrolase Activity	No
GRMZM2G061732	109627068-109679539	Os12g19304	Sb08g010620	NAD Biosynthetic Process	No
GRMZM2G061804	109702193-109702942	Os12g19304	Sb08g010620	NAD Biosynthetic Process	No
GRMZM2G404025	109816194-109817284	Os01g58024	Sn03g020190	Cellular Response to Sulfate Starvation	Yes
AC214360.3_FG001	110185330-110189107	Os02g36220	Sb05g022320	Phytoalexin Metabolic Process	No
AC225185.3_FG004	110670070-110673574	Os12g17310	Sb0011s0054	Plasma Membrane	No
GRMZM2G004468	111057348-111082507	Os06g27890	Sb07g014650	Defense Response	No
GRMZM2G314702	111187331-111190741	Os03g43990	Sb01g013680	Response to Light Stimulus	No
GRMZM2G033521	111343931-111347964	Os12g17900	Sb08g010420	Leaf Senescence	No
GRMZM2G134681	111478344-111479324	Os01g71310	Sb03g045410	Cytokinin Metabolic Process	No
GRMZM2G309220	111511297-111515898	Os08g28010	Sb02g026186	Uncharacterized Protein	No
GRMZM2G411536	111553578-111570930	Os12g24080	Sb08g012560	Cellular Protein Modification Process	No
GRMZM2G135322	111842679-111866759	Os12g18760	Sb06g015870	Ubl Conjugation Pathway	Yes
GRMZM2G063306	112042104-112047482	Os11g04020	Sb08g008410	DMA Efflux Transporter	Yes
GRMZM2G086882	112198031-112206945	Os12g13380	Sb08g008360	Response to Cd and Cu binding	No

Table B3 List of primers used for mapping *ys*3.

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SSR-01	AC190770	GAGGGTATGAGGGAGCAACA	CTTACGCCGCTGAAAATAGG
SSR-02	AC190770	CATCTCCCATCCAACGGTAA	CCGGAAGTCTGCAAATAGGA
SSR-03	AC190770	GGGAGAGAATACCGAACAGG	GTCTCAACGCAGGCATAGGT
SSR-04	AC190770	GAGCATGGAGGAGGGTATGA	GCTTACGCCGCTGAAAATAG
SSR-05	AC191518	TCAGAAACTCGACGTGCAAC	AAGTGGGGCTTTCCTCAGAT
SSR-06	AC191518	AGCAAACTGGCCTCAGATGT	AGGAGGGAGGAAACCAGAGA
SSR-07	AC191518	GCCCCTCACTACGTGGATT	GTGTAGGCAAAGGATCATACCC
SSR-08	AC191518	TCCTTTTGCGAAGCCCAAT	ATGTGGGGAGGGGGATTGT
SSR-09	AC191518	TAGCAAACTGGCCTCAGATGTA	ACCAGAGAAAGAAGGGAATGTG
SSR-10	AC183510	GTTTGTCGAGCGAGCGTATG	AGATCGTCGTGGTCCTATCG
SSR-11	AC183510	TTTGAGTTGACGGCCAGAC	ATCTTCGTGTCGGTGCGTA
SSR-12	AC183510	CCGTCAACAAGACAACAATCA	CAATCGACGGGAATAAGCAT
SSR-13	AC183510	TCTGTTCGGCCTAAATGGTC	AATCCTGCTTATGGGTGTGG
SSR-14	AC183510	GGAAACCAACGTCAATGCTT	ATGTACTCGCGGAACAGAGG
SSR-15	AC211204	CATCTTGTTGAGGGCTTCGT	ATAGGGGGCTATGGGTAGGG
SSR-16	AC211204	CCATGACCTTACCTTCCTGTC	CATTGCCGAGGGATGTTATA
SSR-17	AC211204	GTAGGCCACGAGAGGGTTG	CACCCCCATTTTCATCAGTT
SSR-18	AC211204	GGATACCGTTGCGTTCATCT	CAGGGGGACTAGGCACAAT
SSR-19	AC195233	ACCTACCACCTGAGCCAGTC	CGAGTGTAGGTCAGCCAAGA
SSR-20	AC195233	GGCGACGACGATGATGATA	TATTGGGAGGGGGTTACACA
SSR-21	AC183510	ACTTGGCTTGCGCTTCTATC	GGCTTGGATCACCGACAC
SSR-22	AC182625	AGCACGAGATCGGGAACAG	TTATGGAGGCTTTCGGGAAC
SSR-23	AC185252	TCGGAGAAACGAACATCATT	TGAGTCGGAGTACTCTTCGG
SSR-24	AC185252	AGCAAGTGGATTAAAGCAGCAG	TCGAGCAAGAAGCAAGAAGC
SSR-25	AC185252	GCAAGTGGATTAAAGCAGCA	AGCAAGAAGCCAAGAACCAT
SSR-26	AC185252	AAGTTGGTGCAAAGGAATGG	TGAGGGTGGGAAAAACAAAG
SSR-27	AC185252	ACCAATGCAGTCAAAAACACAG	CAAGTGGCTAACTCAAAAACCC
SSR-28	AC185252	CAATGCAGTCAAAAACACAGC	GCCCATTACACAACGCTATTAC
SSR-29	AC185252	ATGTATCCGTGCATTTGTCA	CCGACGATCATGTTGTAGTG
SSR-30	AC185252	GGAGCATGTGTGGTTCTCTG	ATGGGCTGCTTTCGTTTTT

Table B3 continued

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SSR-31	AC205890	CGAACATGCAAGGAATGGAC	AGAGAGTCGGGGGTTGATTT
SSR-32	AC205890	ACCTTCTTGGGCCTGATTTA	AGACCCCTCCTATCTCAAAAGA
SSR-33	AC205890	TCCCTTCCCACTCTGCTCT	TCTACGGCTGTTGCTGCTC
SSR-34	AC205890	ATCCCTTCCCACTCTGCTCT	GCAGCGGTTTTCTTTGTGA
SSR-35	AC205890	CGCTCTGGGACAGGTTCT	CTGTGTTGGGTTGTGTTGTG
SSR-36	AC205890	GCGTGGGCTCATCTCTATGT	GAGACGTTTGGGGGTGGT
SSR-37	AC208041	GAGGTGCAATAGCGGTGTTT	GCAAGACCAGCCCATACAGT
SSR-38	AC213775	TACCCCATGACTCCACATCA	TCCACAACCATGACCAACTC
SSR-39	AC213775	ATGCTCACTGCCAATCACAA	CGGGCCAGTATGAATGAACT
SSR-40	AC190717	GGGCTTGTATTGTGCGTGTA	GAGGAATTGAGCGGAGAAGA
SSR-41	AC190717	TCCTCCGTCAGCTTAGAGTG	CAGATCATTGCCCTTGACAT
SSR-42	AC205321	GGCTCCCAATATCCTTACCC	GCGACAGCCGTATTTGTTG
SSR-43	AC202151	TCAGCCAATCATCAGCGTAG	CGAGAGGAGATGCACCAATAG
SSR-44	AC202151	CTATTGGTGCATCTCCTCTCG	GCCAACCTATTCAACCGAAG
SSR-45	AC211203	CCGGAACCTCCTCCTCTACT	AACACAAGCCTCTGCCTCAT
SSR-46	AC211203	CAGAGCGGGTATCCTTTGTTAT	AGGTCGGTGGTTCATAGACG
SSR-47	AC226735	CGAGAGGACGAGTAGGCAAC	CTCACCCTGCATTGGATTTT
SSR-48	AC226435	TCCGACCCATATACCACGAT	AACAACATCTACCATCACAACACC
SSR-49	AC190963	GCAGCAGCTACGATTTGTGA	CCTAACCGCCTCAGTGTGTT
SSR-50	AC190963	CCTCTCACTGCTGGTTCTCC	CTTCACCCCCTGCTTTACAA
SSR-51	AC194101	GCTTCCTAGCCCCAGTTACC	CATCCCTCTTGTATCGTTTTTCTT
SSR-52	AC194101	TCACAAGGGCACACAAGGTA	CTCAGATTTTCATTCCGCTCTT
SSR-53	AC214839	CCCTTTTCCTTTCCACAGTTAC	TGCGCTATTATTCCCCCTAA
SSR-54	AC207816	GACCTGCCCTTCAAAAACAC	TGCTACACACACATGCTATTCG
SSR-55	AC207816	CAAGATGCCTGACGAGTGAA	CTTCTGATCCCTATACTGCTCCA
SSR-56	AC213020	ATGATGCGTAGATGCGTGAG	ATTTTGGTGTAGTGTCCTCTTGC
SSR-57	AC195122	GGCTTGTGTGGGAATGTCAC	CCGAGAACAATGCTGAAGGTT
SSR-58	AC195122	AGCCACTACCCCTCCTGTCT	CGTCGGTTCTCTTCACTCGT
SSR-59	AC216054	GCTAGGCACATAGGATGAGGAT	GCAGCCAGAGTAGGGTCAAC
SSR-60	AC196434	AGCCTTCTTTATCCCACAAGC	CACCCCTTTTCCTTTTATCCA

Table B3 continued

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SSR-61	AC196434	ATCGAAGGAGAGCAGTCACC	CCGTTTCAACAGGGATTCAT
SSR-62	AC213775	CTCACTGCCAATCACAACAAA	CACCTGACCAGCACATCATT
SSR-63	AC225179	AGGAGGTCGTAGCAGTCGTC	TCACCGTATCTAGGGGCAAC
SSR-64	AC213874	CTGTATCCCGCGCCTATAAA	CACGAACCAAATAGCCCCTA
SSR-65	AC213874	GTCAGTCCTCCGGTCCTGTC	AGACCACCGCTCTTGTTGCT
SSR-66	AC217048	TCTTGTCCTTGTCTCCTCATTGT	TCAAGTGATCTCCAAAGGCTCT
SSR-67	AC194190	GAGTCCCGTGAAGGTAAGTTG	AATACCCTGCCTGATGTTCC
SSR-68	AC194190	CACTAATGACTCCGAAAAACCT	AAACAATACCCTGCCTGATG
SNP-01	AC194939	ACTCAACAAAGCGTCGTCAA	GGCCTCCACATGAAACTTCT
SNP-02	AC202151	CAAGGGCTCATACGCAAAAC	ACCACCACGCTCAACAGATT
SNP-03	AC207793	TGATTTCAGCCATAAGTTCCTG	TTGATTGGAATGCACAAGTAAGTA
SNP-04	AC198201	ATCCATGGGCATGAATCCTA	TCTGGTTTCAACGGCACATA
SNP-05	AC198201	GCTATTTAGGCATCGGTGGA	CTCTCCTTTCGCAACCACTC
SNP-06	AC198201	TCCTTCCTTCAAACCACAC	CCGAAGATGGTGTGACTCAG
SNP-07	AC201960	CCAAACCAAAAACCATCACC	TGGACTCACCAACCTCAGATACT
SNP-08	AC213020	AATTGGTCCTAACGCACGAG	TCCCCCTTCCTGATTCTTTT
SNP-09	AC194664	ATTATGTGGCTGTGGATTTGC	TGATGGGACCTTTCTTTTGC
SNP-10	AC194664	AGGAGAGGCATGAAACATAAGC	CTCAGGTACACTCCCTTTTTGG
SNP-11	AC194664	GGATGGCTAGTGAGGCAAAC	TGACACAATCTTCCCGTTGA
SNP-12	AC206180	ATCTTCACCGCCACCTCTTA	AGCCGACATGCTTTTCTTTC
SNP-13	AC206180	AGGTAAAACTACTCGTCATCAGCA	CTTAGTGACATTGGGCAGCA
SNP-14	AC216054	GGTTCTTCACCCCTTTTTCC	ATGCAGGCCACTGTTTCTTT
SNP-15	AC225179	GTTGCGGCTCATAGTGTTGA	GGTTGAAGGGAGTTCTGACG
SNP-16	AC204430	AGCCGAGAGGAGATGTAGCA	TTGTCACCGTCCTGATAAACC
SNP-17	AC204430	AACCATCACAACCACGACATC	CAAATCTGGGCAGCAAAGTT
SNP-18	AC183889	CATGAATCCCATCACATTGC	CTTGTCTCTGGCCTTCCTTG
SNP-19	AC183889	AAAGTGGGGAGAGGAGGAGA	TGGCATCACACTATGGAGGA
SNP-20	AC195274	TCTACTAAAACACGGAAGTCATCG	ACCGTCTCTCGTCGCATAAT
SNP-21	AC195274	CGCCTTCGCTCATTGTAAGT	TTTGCGGGACCATAGACAGT
SNP-22	AC182620	GGAAGAGGGGGTAAGAGGAA	CAAGGACCAACAAGGACCAT

Table B3 continued

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SNP-23	AC191382	GGCCCACTCTATTCCCTTTT	TGCACCAGATTAAGTCGTCAA
SNP-24	AC210176	TGCTGGGGTTAGGGTTGTTA	CCTCTTAGTGCTGGCGTCAT
SNP-25	AC204595	CTACCGCACCTGTCATTCCT	CTACAGCAACAAAATCCCACAA
SNP-26	AC204595	GGATGGAATAGGAGGGGAAG	TGCTACCGTTTTGACTTGTGA
SNP-27	AC212774	GTCAACTCCGTAACAGCATCAG	CATTGAAGGAGATCGTGGTGT
SNP-28	AC212774	ACTTCTGCCACCGATTTGTC	ACACTCGAACTGAGCCACCT
SNP-29	AC208790	CCATCACCTTCAACAAAATGC	GACGTGTTCCTCCAGTCCAT
SNP-30	AC208790	TGACACGTTCTTTAGCGGTTT	TTCCATAGGATTTGCCAGGT
SNP-31	AC194954	AGATCCGAAGAACGACGAAG	GTCTCTACCTGCTTGCTCTGG
SNP-32	AC194954	CCACTTGACGTTGGGAAAAT	TCTTTAGAGGAGGACTTCGGTTC
SNP-33	AC190717	TGAAGGCAGCAAATGAGAAA	CGGATCGTGAATTAAGCACATA
SNP-34	AC190717	CTCAGGGATCAAGGGCTATG	TGCGTGGGTATTTCTTAGGC
SNP-35	AC194939	GAGGACGAGGCTAACGAGTG	GGTTCAAACACCCGCATAAT
SNP-36	AC194939	TGCTTGCTTGAGGAGGAGTT	ATACGAGGGGAGGAGGTGTT
SNP-37	AC202151	ATCTGTTGAGCGTGGTGGTT	GAGCTGTCGATGATGAGGTGT
SNP-38	AC202151	TTGAATAGGTTGGCAGGTAGGT	ATCCCCATGACTTGTTGAGC
SNP-39	AC211203	CACAGACAGCAGCAGGAGAG	CTGAGGAACTCGGTGGTCAT
SNP-40	AC211203	GGATGGATCACTGGGGACTA	TCTACCGCCAAGGTCAGAAT
SNP-41	AC207793	CAGTGGCTGTGAAGAACACCT	GCGATGAAACACCAAGTCAA
SNP-42	AC207793	TTTCCTTTGCCTTCGGTAGA	CATAAGCAGGATGGGGATTTT
SNP-43	AC226735	GCGTCGTCAAGTCATTCTTC	TGAACTGATCGGAGATGGAG
SNP-44	AC201960	TCGAGTTCCAAACACTGACAAC	GCCGTGTACGATGCTGTAGA
SNP-45	AC201960	TTCCTCCCTCTTGTTGTTGC	GAACATTGAACAGTGCGTGACT
SNP-46	AC190963	ATCTTCGGCTGTTCTGCTGT	GTGGCTAGGATCTTGGATGC
SNP-47	AC190963	TGTTGCTAGTTCGGTGATGG	CACAAGGTTTTCTTTAGGATAGGC
SNP-48	AC213020	TCTCGTTGCCGCTTATTTTT	ATCCACCCAAGGTATGTCAGAT
SNP-49	AC207816	GGAAAATCAGTGCCCAAAGTT	GCCAGATGAGAGGACAGAGG
SNP-50	AC207816	AGGGAGAATGGGGACATAACTT	CTAACATCTCGTTCCTCGTGTG
SNP-51	AC214819	TCCCCTTTTCTTCGTCAATG	CTTATCCAGTCCGTCGCAAT
SNP-52	AC194101	CATCGCCAACGTCTAACTCTAC	GCTTAACCGTACACAGATGACTTC

Table B3 continued

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SNP-53	AC186661	CGCTTGGCTTTACTGTTTTGA	ATGGTCGATGCCCTAACTCA
SNP-54	AC196661	ATCCGCAGACAGTTCACTCA	GTGCAAAATGACCATCTACCC
SNP-55	AC182620	TATTGCCCGAAAACAAAAGG	GCCAGGAAGGTACACGCTAC
SNP-56	AC182620	GGATGGCACGTAGGTTCTTC	CCATTTGGGTTCGAGATGTAA
SNP-57	AC211410	GGGAGTTCAGTCGCTTTCTC	CCTTCGGTTGTCCTTATGCT
SNP-58	AC211410	GATTTCTTGCGGGACTTGTT	TTTCCGAGTGTCTTCCTGTG
SNP-59	AC208987	ACTTGAGGAGGCGTGAGAAA	CCAAAGGCTACGGTTGAAGA
SNP-60	AC208987	CGACAATAGCTCCACAACCA	TTTCTCACGCCTCCTCAAGT
SNP-61	AC208987	GGTTTCCACGCAAAGATGTC	GCCGCTGCTCAAGAACTATC
SNP-62	AC208041	CAATCTGTATGATGGAAACGAC	TCTGTATGCGAGAAGTGAGGT
SNP-63	AC208041	AAAGAGAGCACGTCTATTGAGA	CAAGAATGGTAGATGGACAAAC
SNP-64	AC208041	CGAGCCCATTTTTCTTGTGA	CATCCATGCCACTTATTTCCA
SNP-65	AC208041	AGGGAGAGGTTCTTGGACTTG	ATTTGTGGAACTGGCTTTCG
SNP-66	AC208041	ATTTGGACGGCACATAGAGC	CTTGATGAATGGGAGGACTCA
SNP-67	GRMZM2G094100	AATGCAAAATGGGGAATGAG	ATGTAGGCAAGGGGGAAAAT
SNP-68	GRMZM2G042855	CCAGAGGGGTTTATTCAAAGG	GACCAAGATCAACACGAGACAA
SNP-69	GRMZM2G042855	GAGGAAAACATTGACGCTGAG	GATTCTAGGGTATCACTTGGCATT
SNP-70	GRMZM2G409893	GCTCTCCCTCCCTAAGAAGC	CACACATGGACCCTTTTGAA
SNP-71	GRMZM2G470427	AGCCGTGCTTTTAGGGAAAT	CCGTAGGCCACCACTACAAT
SNP-72	GRMZM2G470427	TGTCCAACCCAAAGAAGAATG	GCACGTAAACAGCCACAGAA
SNP-73	GRMZM2G000076	GAGTTGGGGTTGGAGTAGCA	CAGATGAAAGGAGAGGGACAGT
SNP-74	GRMZM2G000076	CTACAATCAGTGCCCCCAGT	CAGTGCCAGCTTTCAAGGAT
SNP-75	AC207793	ACAATCCCTTGGGTGATGTC	CAATCCCGTCAAGAAGTCGT
SNP-76	GRMZM2G170037	CTAGTACGCAGCCAAGTCAGA	TCCTAAAGTTCGGGGAAAAA
SNP-77	GRMZM2G170037	GGGATATTGCTTTTCGGATCT	TCTATTTTGCACCACATTCCAC
SNP-78	GRMZM2G042855	TTTCGGCAAATGAAGACCAT	GCGTCAATGTTTTCCTCCTG
SNP-79	GRMZM2G042855	TTTGAGAAGTCCCAGATAAGCA	CGCAACTCTCCTGTGTGTTT
SNP-80	GRMZM2G042855	CGACCCTTCCTGCTCCTATT	GAACCAACCTGCTGATGCTC
SNP-81	AC194101	CCGGTTATGAGAGCAACTGG	CGAACAACAGGGCTGAAAGT
SNP-82	GRMZM2G395853	TGGAGTTAATGTGTTTGGTAGGC	AAAATGACTGGTTGGGATGC

Table B3 continued

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SNP-83	GRMZM2G395853	GTGCATTTGGCCTGATATTCT	AAAGGTACAGAAGCCGTTGC
SNP-84	GRMZM5G878970	GACCAGCAATCGCAGAAATC	AAAGCCAGCACTTCAGTCAGTT
SNP-85	GRMZM5G878970	TCGGCACTTTCTGAGTTCCT	AAGCCAAGCTGACCACAAAT
SNP-86	GRMZM2G409893	ACCTCAACCTCCGCAGCAAC	TTTGAACCAGCAGCCACACC
SNP-87	GRMZM2G409893	TGACTTGAGGCCAAACAATG	GCACGCGATACTACCACCTT
SNP-88	GRMZM2G094038	CCACCACCTCCATCAAATCT	CAATCGCACACACATACACG
SNP-89	GRMZM2G094038	TGGGAATTTTGGATGAAGAGAG	CAACAGGGAATAAGTAGGGGTTC
SNP-90	GRMZM2G055545	CATAAGGGGAGTGAGACAGAAAA	TGCAGAAGGTGTACCAGTTGA
SNP-91	GRMZM2G055545	GGAACCTGAGCCTGCTGACC	CGCTGCTGATCGAGATGGATT
SNP-92	GRMZM2G055545	TTTCGCTCCCATTACAGTCC	CAAGAAGCCAGTAGAACAAATCG
SNP-93	GRMZM2G055545	GATGTCCGTGTTCTGGCTCA	ACCAAAAAGGCACCAACAGC
SNP-94	AC207793	TCCTCCACGTAACTAGCATTTTC	CTCAGTTTATCGCAGCATCG
SNP-95	AC207793	ACTAAGGGCCACAGCGTCAC	GCTGCAACCGATCTTCTTCC
SNP-96	AC207793	AGAGTCGTGCAGGGTTGAAG	TACGAATGCGGGGTGTTTAG
SNP-97	GRMZM2G316106	TAACGAGGTAATCCATGCTTAG	CACCCTCTTCTCTTATTCAGGT
SNP-98	GRMZM2G074488	CGAGACAAGCCAAAAACACA	GGCATCAGAAGGGATAAGACC
SNP-99	GRMZM2G074488	CGAGACAAGCCAAAAACACA	GGGAGAAAACATTCCAGTGC
SNP-100	GRMZM2G823484	TTACTGAAACCAGTCCCAAAGA	GCTGCCAAATATGAACAGAAAG
SNP-101	GRMZM2G823484	CCTAGAAAAGCAAGAACGACT	GAGTAAACAAGGCATAAACAGC
SNP-102	GRMZM2G823484	TCAGTAGTTGAGGCTGCTGCT	TCTGAAAATCAAGGTCCGCTAT
SNP-103	GRMZM2G823484	GATGAAGGGCCAGGAGAAAC	GACAGCCATGAACCAGAAGC
SNP-104	GRMZM2G823484	TCACCCTGCATTTACTATGTGG	GTCCTTTGCATTTTTCTCTTGG
SNP-105	GRMZM2G824275	TTCTCTTGGTTTCGCTGCTC	CTGGTTGCTTTTAGGTTGTTCG
SNP-106	GRMZM2G803397	GTTCTGCATTCCTCGTCCAG	CATTCCGTCAGGTTTGTCATC
SNP-107	GRMZM2G701635	TGCTTGCATCTGTGATTTGTG	ATTGGCGGTTTGGCTACTTT
SNP-108	GRMZM2G067231	CCCACGCATCTTCTTCCTTA	CAGCATCAGCCAACCCTAAT
SNP-109	GRMZM2G067231	TTTATGCCGCTGAAGAGATG	TCTACCCCTTGCCTAAGTCG
SNP-110	GRMZM2G067223	AGTGCTGTGCTACTACCCCTCT	CATGGCAGGACACTTGTAGG
SNP-111	GRMZM2G067223	CGACCTCCAAGTGAGTACCAA	ACTCCTAGTGCGATGAACAACA
SNP-112	GRMZM2G156533	CTTCTTCAACCGCAACAAGC	AATCGGACACGAAAAACACG

Table B3 continued

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SNP-113	GRMZM2G156533	CACTGTCCCGTGTTTTTCGT	TCCTCCACGTAACTAGCATTTTCT
SNP-114	GRMZM2G701633	ACGCGGTACAAAACACGACT	ATGCGAGCAAAATCACCTTC
SNP-115	GRMZM2G373272	CTAACGGGCTTGAAACACAT	TTGAACAAGAACCTGGCTTT
SNP-116	GRMZM2G169702	AACCCAAATTCGATCTGTTCTC	TCGGTCAGAGGTGGAAGAAT
SNP-117	GRMZM2G855629	CACCTCACCCCTAATCCTCA	CACCTCACCCCTAATCCTCA
SNP-118	GRMZM2G428672	ACGCCCCTGTTTTCTTCTTC	ACGCCCCTGTTTTCTTCTTC
SNP-119	GRMZM2G373292	TTTGAACAACCAAGCAACGA	TTTGAACAACCAAGCAACGA
SNP-120	GRMZM2G373292	CCCTTGCAGTGTAGGTCTGATT	AATGCCTCAAAGGTTATCATCG
SNP-121	GRMZM2G373277	GCATCCAGTTTCAGTTCCAATC	TGGCATGTGTCAGCTTATTCAC
SNP-122	GRMZM2G373277	GCCCCTCTAACCCCTCTAAC	CCCTACGACATTCATTTTGGT
SNP-123	GRMZM2G373277	CCCCTCTAACCCCTCTAACC	GTAAAGACCCCCTACGACATTC
SNP-124	GRMZM2G373277	TCATACCATCGCCTCTTATCCT	ATCACAATTCAGCCTCATGTTG
SNP-125	GRMZM2G074373	GCGGAAACAGAAACCAAACTAA	AGGCAATACCCACACACACC
SNP-126	GRMZM2G074373	CTAAGTGACAGGGACCAAGGATA	ACACAAACAGATGGCTGAAAAG
SNP-127	GRMZM2G074373	TGTTTTTGCTAGTGAGTGGGTA	CTTTCATTGGATACACAGACCA
SNP-128	GRMZM2G074373	GTTTTTGCTAGTGAGTGGGTAA	TCTTTCATTGGATACACAGACC
SNP-129	GRMZM2G074373	GCATGTGCATCGTTGAGTAAT	AGGCCAGTGAACAGAAAATATG
SNP-130	GRMZM2G074373	TCGGAGAAGCTATTCATGTTGA	TACTACACGCACACCCCTACAC
SNP-131	GRMZM2G094081	GCCGCCAACTACTACCTCAC	TCATCGCCTGCTTCCTAGAT
SNP-132	GRMZM2G094081	ACAAACCAAGGAGGGCATTT	TGTTTCCTGTAGATCGGACCA
SNP-133	GRMZM2G126309	CCATGTTATCTGTATTCCCCCTAA	CAACCCCACTTCCAAAAATG
SNP-134	GRMZM2G126309	GATCACCCAAAAGATGCTAA	GATCCGAGCAAGTGTATCAT
SNP-135	GRMZM2G067231	CCAAGGGCTTCCACTACTTCT	ATGGCGTAGCTCAACAAACC
SNP-136	GRMZM2G156533	GATCGAGCTTTGTCCCAACC	TCCCTACACACTTTCATTCACACA
SNP-137	GRMZM2G316106	CTGAAGTGAGATGGTGATCTTGA	AAAAGAAAGAGGCCACATGAAG
SNP-138	GRMZM2G316106	CCCAGGTATATTGTTACTACTACGACA	TCGGCCTACGTGACTCTTTA
SNP-139	GRMZM2G316106	TCATGGCAATAAAAGACGATGC	CGAAAAGAGCGTGACATCCA
SNP-140	GRMZM2G316106	GAATGGATGTCACGCTCTTT	CATACTACACTCACCCCAACAA
SNP-141	GRMZM2G373277	GTAACTGTGCTTGCTGTTCAAT	CAGCTTATTCACCCCATGTT
SNP-142	GRMZM2G373277	TCGTGGAGCCATACACTTTTC	GCCTCATGTTGCTCTTGGTT

Table B3 continued

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SNP-143	GRMZM2G067231	AGCACACATAGCCCCAGTTC	CGTAGCTCAACAAACCAGGAG
SNP-144	GRMZM2G701633	CGAGAAAACAAGGCAGGGTA	AACTTGCAGATTCGGTCCTG
SNP-145	GRMZM2G701633	CGGATGGATCACTGGGGACT	TGCTGCTTGCCTCGCTTCTA
SNP-146	GRMZM2G701633	AGCAGGGACAAGATGAGGAGA	CAGATGCGAGCAAAATCACC
SNP-147	GRMZM2G701633	GCTCGCATCTGTGATTTGTG	AGAAGGGATGGGGTTTATGG
SNP-148	GRMZM2G701633	TCCAATCTCCCTCTCAAACG	GGGGTCCTCACTGTCAATACTC
SNP-149	GRMZM2G701633	CTGGAATAGGATGCGTCGTC	TGAAAGGAGGGAGGTGACAG
SNP-150	GRMZM2G864903	CACAGATAGCCACACAAGCAA	GCAGTGAACGGTAGTCATTCTG
SNP-151	GRMZM2G864903	CAGAATGACTACCGTTCACTGC	AGTGGGATCAGGGTGGTATG
SNP-152	GRMZM2G864903	TTACTGCGTCATTGTCATGT	GGGGGATATGCTAGGTTTAC
SNP-153	GRMZM2G864903	TGAGGGAGGAGGTGCTGAGA	GGCAAAAGGAAGTGGCAGTG
SNP-154	GRMZM2G063306	CCAACCGTCTTTTGTCCTTG	TTATTTTGTGCGAGGGGAGA
SNP-155	GRMZM2G063306	GGTAGTTATTATTGGTGGTTAGAGC	TACTCACACTGAAGCGACTGA
SNP-156	GRMZM2G063306	CGATTTTTGGTCTCGTTCTTC	GATGCTCCCATTTTTGTTCA
SNP-157	GRMZM2G063306	AACCCAGGATATTAAGAGCATCTG	GGAACTACATACCCCAAGAGCA
SNP-158	GRMZM2G063306	CAGGCGCTATCCATACCACT	GAACCGTGGGAACTGATGAT
SNP-159	GRMZM2G063306	TGCACCACCCTAGTACCTGT	TACGCTTGCGGAGTAAAGTT
SNP-160	GRMZM2G063306	GGAGAGTGAGGGACTTTACGC	CTTGGTGGTTTAGCAGCACA
SNP-161	GRMZM2G063306	GTAGCGTGTCCCTGTCCATT	CCCAGAAGAACGAGACCAAA
SNP-162	GRMZM2G063306	TTTGCAGGCTTACTCAATCG	TGCTCTTAATATCCTGGGTTGA
SNP-163	GRMZM2G063306	CATGGTATCCGCTGTTTCAA	AAGTTAGGAGCCCCACAGGT
SNP-164	GRMZM2G063306	ATTGTGCCAGGGTTTGACTC	GGAAAGGCCAGCAACAGATA
SNP-165	GRMZM2G063306	GGTCCCCGACACTTATGATCA	TGGTGGGCAGGTAATTTGAGT
SNP-166	GRMZM2G063306	TCTCCCCTCGCACAAAATAA	GGCAAAGCTGAAAAGGAACA
SNP-167	GRMZM2G063306	GATCATAATTGCCGCCCTCT	AAGAAGGGAAACAGGGACGA
SNP-168	GRMZM2G063306	GAACCCAAGCAATGGAATAATG	CTCCCACACCCCAAAATGTA
SNP-169	GRMZM2G063306	TACTGCCTTCCATTCCCAAC	CCTCGCATCCACCTTTATTC
SNP-170	GRMZM2G063306	AGGGTGTGTGTGTGCTGCTA	TATTCCATTGCTTGGGTTCG
SNP-171	GRMZM2G063306	GTAGCGTGTCCCTGTCCATT	AATTCCGTTGTATGCCACTCA
SNP-172	GRMZM2G063306	TGTGCTGCTAAACCACCAAG	GTTTCGGTACTGCTAGGAATGG

Table B3 continued

Name	Origin	Forward primer (5'-3')	Reverse primer (5'-3')
SNP-173	GRMZM2G063306	AGGGTGTGTGTGCTGCTA	TTCGGTACTGCTAGGAATGGTT
SNP-174	GRMZM2G063306	CATGCTTGCACCAATAAAGG	TGCTCTTAATATCCTGGGTTGA
SNP-175	GRMZM2G063306	GCATCCCCTACAAGGAGTTCT	TgtGAAAGGAACAGAGAGTGATG
SNP-176	CTG730	CCGAGCAACTCATCACTTCA	TGCCCTGTGCTATCAAATACC
SNP-177	CTG730	GCTTGATGTGGAGCATTGAA	TAGGTGAGGAATTGGGCAAG
SNP-178	CTG730	GATTCCGGCACAAATGAAAG	ATGGCAAAGGATGGAACAAA
SNP-179	CTG730	CCATCGTCATCAGCTTCTCC	GAGACTGCTCCGATTTCCAA
SNP-180	CTG730-tidp7081	GGACTAGACCACTCGCATCC	TTTATGAGGCGACGAAGACC
SNP-181	CTG255	AGCATGGGTTGAATGACTCC	TTTTCGGCTACCACATCACA
SNP-182	CTG255	TTGCGAAAGTGTTAGCCGTA	GAGATCATGGTTGTGGTTTCAA
SNP-183	CTG255	GTGAGGCACCCCTATCTCTG	TCGGATCTAAAACGGAGAAGTT
SNP-184	CTG255	AACGACAGCGCAAATAGGTT	ATAATGCCACTTGCCACACA
SNP-185	CTG255	CTGCGGGTAGAGGAACTTGT	TTTTCAATGGCTTCGTGAGA
SNP-186	CTG255	CGGCCTATGTGTCAAAACCT	CGGTCTGATGCAAGGTATCC
SNP-187	CTG255	AAGCTGGACCTTCTGGAACA	CTTGACGAATAGTTGGGGACA
SNP-188	CTG255	CCAAAAGAGGGCCACACTAA	TTTCCAGCACTCGTCTCCTT
SNP-189	CTG255-idp9014	GGGTCTACTTCGGCTCTCG	GGGTCTACTTCGGCTCTCG
SNP-190	CTG255-idp4360	CAAATCCACATAACCCATTTGC	ACAGAGCTCAAGGATGACCC
SNP-191	CTG255-tidp7141	AACATGCTCGTGATGTTTGG	GGTTGAACTAGCAGAAGCCG
SNP-192	GRMZM2G063306	GTAGCGTGTCCCTGTCCATT	TGCGTGTCCCTTAAAAGTTCA
SNP-193	GRMZM2G063306	ATTTTCAGTCCTCAACCTAATGC	ACCCTCAATTGATTGAAAAACG
SNP-194	GRMZM2G063306	GCAATTAAATTAAAGGCTATATGTTTC	TCACACTGAAGCGACTGAAAA
SNP-195	GRMZM2G063306	TTTGTCGGTCGATTTTTGGT	AGATGCTCCCATTTTTGTTCA
SNP-196	GRMZM2G063306	GCTTGGGGCCTAGGTCTTAT	TGTGGACTGGCCTGTAGATG
SNP-197	GRMZM2G063306	TGTGCTGCTAAACCACCAAG	TGCGTGTCCCTTAAAAGTTCA

Table B4 Correlation between harvesting coefficients and 11 micronutrients measured in the *ys*3×W22 F_2 population. Upper numbers represent correlation coefficients, and bottom numbers represent significant levels.

	B	Ca	Cu	Fe	K	Mg	Mn	Na	P	S	Zn
SB4	0.31	0.61	0.39	0.08	0.30	0.79	0.99	-0.98	-0.69	0.44	-0.32
	0.69	0.39	0.61	0.92	0.70	0.21	0.01	0.02	0.31	0.56	0.68
SA4	-0.13	0.85	-0.48	0.89	0.92	0.71	0.38	-0.48	-0.01	-0.60	-0.96
	0.87	0.15	0.52	0.11	0.08	0.29	0.62	0.53	0.99	0.40	0.04
SA6	-0.40	0.70	-0.78	0.98	0.92	0.46	-0.05	-0.03	0.26	-0.86	-0.94
	0.60	0.30	0.22	0.02	0.08	0.54	0.95	0.97	0.74	0.14	0.06
SW	-0.24	0.86	-0.58	0.94	0.95	0.69	0.30	-0.37	0.02	-0.66	-0.98
	0.76	0.14	0.42	0.06	0.05	0.31	0.70	0.63	0.98	0.34	0.02
DW	-0.33	0.87	-0.65	0.97	0.97	0.69	0.26	-0.31	0.01	-0.68	-1.00
	0.67	0.13	0.35	0.03	0.03	0.31	0.74	0.69	0.99	0.32	0.00
SL	0.03	0.69	-0.41	0.84	0.82	0.53	0.24	-0.42	0.22	-0.70	-0.89
	0.97	0.31	0.59	0.16	0.18	0.47	0.76	0.58	0.78	0.30	0.11
WC	-0.22	0.85	-0.57	0.94	0.95	0.69	0.30	-0.38	0.03	-0.66	-0.98
	0.78	0.15	0.43	0.06	0.05	0.31	0.70	0.62	0.97	0.34	0.02
RT	-0.60	0.86	-0.83	0.99	0.99	0.67	0.14	-0.10	-0.06	-0.67	-0.97
	0.40	0.14	0.17	0.01	0.01	0.33	0.86	0.90	0.94	0.33	0.03

Table B5 Correlation coefficients between Fe and other micronutrients measured in the *ys*3×W22 F_2 population. Upper coefficients represent correlation coefficients, and bottom coefficients represent significant levels.

Micronutrients	Fe		
	Overall	10 M Fe-EDTA	300 M Fe-EDTA
B	-0.51	0.36	0.10
	0.49	0.02	0.52
Ca	0.80	0.36	0.30
	0.20	0.02	0.05
Cu	-0.82	0.74	0.36
	0.18	¡.0001	0.02
K	0.97	-0.10	0.09
	0.03	0.54	0.58
Mg	0.58	0.03	0.03
	0.42	0.87	0.87
Mn	0.05	0.31	0.31
	0.95	0.04	0.04
Na	-0.06	-0.08	0.39
	0.94	0.62	0.01
P	0.10	0.15	0.30
	0.90	0.32	0.05
S	-0.78	0.03	0.50
	0.22	0.83	0.00
Zn	-0.97	-0.19	-0.04
	0.03	0.21	0.82

Table B6 List of candidate genes from comparison 1 (*ys*3 grown at 10 vs. 300 μM Fe-EDTA) with a FDR$\leq$0.05.

Gene	Locus	Gene	Locus	Gene	Locus	Gene	Locus
GRMZM2G308463		GRMZM2G380784		AC225344.3_FG006		GRMZM2G085924	
GRMZM2G161746		GRMZM2G428035		GRMZM5G874955		GRMZM2G704488	*OsNas2*
GRMZM2G057413	*Iro2*	AC203989.4_FG001		GRMZM2G030036	*Nas2*	AC193786.3_FG005	
GRMZM2G300965		GRMZM2G104563		GRMZM2G133475		GRMZM2G066049	
GRMZM2G124061		GRMZM2G349895		GRMZM2G036629		GRMZM2G563190	
GRMZM2G400602		GRMZM2G106393		GRMZM2G124785	*Nas*	GRMZM2G141473	
GRMZM2G098875		GRMZM2G164974		GRMZM2G096958	*Naat1*	GRMZM2G083091	
GRMZM2G070087		GRMZM2G030159		GRMZM2G029951		GRMZM2G312712	
GRMZM2G026780		GRMZM2G155546		GRMZM2G038153		GRMZM2G102760	
GRMZM2G137440		GRMZM2G421491		GRMZM5G851266		GRMZM2G309109	
GRMZM2G119219		GRMZM2G430902		GRMZM2G122853		GRMZM2G063756	
GRMZM2G384311		GRMZM2G325575		GRMZM2G036711		GRMZM2G464137	*MTK*
GRMZM2G106511		GRMZM5G866024		GRMZM2G355572		GRMZM2G103342	
GRMZM2G138640		GRMZM2G010251		GRMZM5G878558		GRMZM2G028685	
GRMZM2G074672		GRMZM2G040689		GRMZM2G040638		GRMZM2G065030	*PRPP*
GRMZM2G013448		GRMZM2G171096		GRMZM2G410338		GRMZM2G171111	*MTN*
GRMZM2G150952		GRMZM2G038487		GRMZM2G147716		GRMZM2G165998	
GRMZM2G118731		GRMZM2G534430		GRMZM2G025441		GRMZM2G035599	*RPI*
GRMZM2G066840		GRMZM2G099467		GRMZM2G032198		GRMZM2G107639	
GRMZM2G030444		GRMZM2G046532		GRMZM2G020054		GRMZM2G028041	
GRMZM2G327890		GRMZM2G106413		GRMZM2G090487		GRMZM2G332660	
GRMZM2G095725		GRMZM2G035198		GRMZM2G137839		GRMZM2G157760	
GRMZM2G138450		GRMZM2G115190		GRMZM2G137352		AC233955.1_FG003	*Nas*
GRMZM2G316362		GRMZM2G001035		GRMZM2G067265	*Idi4*	GRMZM2G060952	*Dmas1*
GRMZM2G085381		GRMZM2G038677		GRMZM2G391272		GRMZM2G178190	*Nramp1*
GRMZM2G097141		GRMZM2G115839		GRMZM2G048474		GRMZM2G167549	
GRMZM2G350312	*Iro3*	GRMZM2G374213		GRMZM2G072071		GRMZM2G336824	
GRMZM2G147399		GRMZM2G156599	*Ys1*	GRMZM2G135960			
GRMZM2G131421		GRMZM2G049811	*FDH*	GRMZM2G410175			

Table B7 List of candidate genes from comparison 2 (wt grown at 10 vs. 300 μM Fe-EDTA) with a FDR$\leq$0.05.

Gene	Locus	Gene	Locus
GRMZM2G308463		GRMZM2G124785	*NAS*
GRMZM2G161746		GRMZM2G096958	*Naat1*
GRMZM2G127087		GRMZM2G412604	*Naat1*
GRMZM2G103342		GRMZM2G049811	*FDH*
GRMZM2G057208		GRMZM2G704488	*OsNas2*
GRMZM2G110369		GRMZM2G067265	*Idi4*
GRMZM2G124061		GRMZM2G165998	
GRMZM2G057413	*Iro2*	GRMZM2G156599	*Ys1*
GRMZM2G140455		GRMZM2G464137	*MTK*
GRMZM2G389903		GRMZM2G171111	*MTN*
GRMZM5G883985		GRMZM2G028685	
GRMZM2G039757		GRMZM2G036629	
GRMZM2G430902		AC233955.1_FG003	*NAS*
GRMZM2G028306		GRMZM2G312481	*OsNas2*
GRMZM2G137440		GRMZM2G131907	
GRMZM2G074672		GRMZM2G017959	
GRMZM2G055834		GRMZM2G180930	
GRMZM2G117971		GRMZM2G029951	
GRMZM2G010731		GRMZM5G878558	
GRMZM2G035599	*RPI*	GRMZM2G060952	*Dmas1*
GRMZM2G302171		GRMZM2G165098	
GRMZM5G866024		GRMZM2G026802	
GRMZM2G306345		GRMZM2G325575	*Fer1*
GRMZM2G400602		GRMZM2G313020	
GRMZM2G152079		GRMZM2G150952	
GRMZM2G066840		GRMZM2G065030	*PRPP*
GRMZM2G104563		GRMZM2G177942	
GRMZM2G030036	*Nas2*	GRMZM2G163406	
GRMZM2G038153			

Table B8 List of candidate genes from comparison 3 (wt vs. *ys*3 grown at 10 μM Fe-EDTA) with a FDR$\leq$0.05.

Gene	Gene
GRMZM2G070087	GRMZM2G086163
GRMZM2G018716	GRMZM2G039996
GRMZM2G011523	GRMZM2G051806
GRMZM2G136032	GRMZM2G009719
GRMZM2G118731	GRMZM2G059700
GRMZM2G309071	GRMZM2G019806
GRMZM2G134618	GRMZM2G409722
GRMZM2G147014	GRMZM5G847982
GRMZM2G506270	GRMZM2G178209
GRMZM2G360234	GRMZM2G098346
GRMZM2G158378	GRMZM2G038365
GRMZM2G083156	GRMZM2G038677
GRMZM2G038487	GRMZM2G305856
AC197705.4_FG001	GRMZM2G400602
GRMZM2G099420	GRMZM2G147716
GRMZM2G133475	AC204711.3_FG003
GRMZM2G074017	GRMZM2G179024
GRMZM2G140817	GRMZM2G178074
GRMZM2G152417	GRMZM2G147399
GRMZM2G428216	GRMZM2G131421
GRMZM2G157269	GRMZM2G563190
GRMZM2G127404	GRMZM2G162250
GRMZM2G447795	GRMZM2G032107
GRMZM5G853245	GRMZM2G163514
GRMZM2G312712	

Table B9 List of candidate genes from comparison 4 (wt vs. *ys*3 grown at 300 μM Fe-EDTA) with a FDR$\leq$0.05.

Gene
GRMZM2G011523
GRMZM2G384311
GRMZM2G534430
GRMZM2G134618
GRMZM2G476762
GRMZM2G040689
GRMZM2G030159
GRMZM2G103342
GRMZM2G300965
GRMZM2G028306
GRMZM2G127087
GRMZM2G170613
GRMZM2G013448
GRMZM2G164974
GRMZM2G119879
GRMZM2G153488
GRMZM2G131421
GRMZM2G306345
GRMZM2G180244
GRMZM5G888204
AC197705.4_FG001
GRMZM2G099420
GRMZM2G137535
GRMZM2G098875
GRMZM2G059700
GRMZM2G441656
GRMZM2G009719
GRMZM2G086163
GRMZM2G074017
GRMZM2G447795
GRMZM2G305856
GRMZM2G122543
GRMZM2G157269
AC190933.3_FG004
GRMZM2G022386
GRMZM2G162250
GRMZM2G035268
GRMZM2G103972