TÜBİTAK

Turkish Journal of Botany

http://journals.tubitak.gov.tr/botany/

Research Article

Turk J Bot
(2023) 47: 511-528
© TÜBİTAK
doi:10.55730/1300-008X.2781

Genome-wide identification of QTLs associated with iron deficiency tolerance in soybean reveals GATA12 as the integrator of iron signaling in roots

Amir MAQBOOL[1], Ahmet ŞAHİN[1], Allah BAKHSH[2], Emre AKSOY[3*]

[1]Niğde Ömer Halisdemir University, Faculty of Agricultural Sciences and Technologies, Department of Agricultural Genetic Engineering, Niğde, Turkiye

[2]Centre of Excellence in Molecular Biology, University of the Punjab, Lahore, Pakistan

[3]Department of Biological Sciences, Middle East Technical University, Ankara, Turkiye

Received: 02.08.2023 • **Accepted/Published Online:** 11.10.2023 • **Final Version:** 22.11.2023

Abstract: Iron (Fe) is an essential micronutrient required for plant growth and development. However, its deficiency causes substantial yield losses, particularly in alkaline soils. Soybean (*Glycine max*) serves as an ideal model to study Fe deficiency chlorosis (IDC) due to its inefficient Fe uptake from the soil. Although the quantitative trait loci (QTL) associated with IDC tolerance were determined in soybean, the specific genes within these QTL regions remain unidentified. In this study, it was aimed to identify and analyze the expression levels of genes present in IDC-responsive soybean QTL under Fe deficiency. Through this investigation, 6593 genes were identified within 19 QTL linked to IDC tolerance in soybean, and among these, 607 genes exhibited differential expression under Fe deficiency conditions. Notably, the orthologs of 10 selected genes, referred to as the core group, were found to be induced in Fe-signaling mutants of *Arabidopsis thaliana*. These core group genes were enriched in metal transport and Fe-signaling pathways. Further examination of these genes in an IDC-sensitive soybean cultivar revealed their induction under Fe deficiency and high soil pH conditions. Particularly, GATA TRANSCRIPTION FACTOR12 (GATA12) stood out with significantly increased expression levels of approximately 5 and 20 times under Fe deficiency and high pH treatments, respectively. Coexpression network analyses of the core group genes highlighted the significance of the cluster containing GATA12 as a crucial integrator of Fe signaling between the epidermis-localized FER-LIKE IRON DEFICIENCY INDUCED TRANSCRIPTION FACTOR (FIT) and the stele-localized POPEYE (PYE) signaling networks. Overall, the findings suggest that GATA12 and its close paralogs may act in the transcriptional network linking the two Fe-signaling networks in the root stele and epidermis. This knowledge sheds light on the intricate mechanisms of Fe signaling in soybean and provides valuable insights for future studies on IDC tolerance and Fe-efficient cultivar development.

Key words: Arabidopsis, gene expression, iron deficiency, molecular networks, QTL, soybean

1. Introduction

Iron (Fe) is an essential micronutrient required by all living organisms to continue their metabolic activities properly (Schwarz and Bauer, 2020). Plants often face Fe deficiency in the field since Fe cannot be taken up by many species effectively due to its low bioavailability, even though enough Fe is present in the soil. These plants experience Fe deficiency, especially in calcareous soils, where the soil pH is high, and at different developmental stages with an increase in the sink demand (Schwarz and Bauer, 2020). Its deficiency in plants decreases photosynthesis and plant growth, which eventually limits yield. However, an excess amount of Fe can act as a catalyst in the Fenton reaction, which may result in the production of reactive oxygen species (ROS) that can damage the cells by inducing oxidative stress (Kermeur et al., 2023). Therefore, tight regulation of Fe homeostasis is necessary to avoid both its deficiency and toxicity. Plants have evolved different strategies to uptake Fe from the soil and translocate it from the roots into the shoots and/or seeds (Aksoy et al., 2018). Strategy I plants can uptake Fe in the form of Fe^{2+}, while Strategy II plants have evolved mechanisms to uptake the Fe^{3+}. Dicots and nongramineous monocots mobilize Fe from the soil by Strategy I. In this strategy, Fe is taken up by the plant roots in three steps, i.e., 1) the acidification of rhizosphere by proton extrusion mediated by H^+-ATPASE (AHA), 2) the reduction of nonsoluble Fe^{3+} into Fe^{2+} by the action of FERRIC CHELATE REDUCTASE/FERRIC REDUCTION OXIDASE (FCR/FRO), and 3) the uptake of Fe^{2+} by a divalent metal transporter, IRON-REGULATED TRANSPORTER1

* Correspondence: emreaks@metu.edu.tr

To maintain Fe homeostasis, plants must sense the bioavailable Fe in the growth medium and fine-tune their transcriptome accordingly. Fe uptake and long-distance translocation are primarily regulated by bHLH transcription factors (Gao et al., 2019). These factors fall into two major signaling groups centered on the FER-LIKE IRON DEFICIENCY INDUCED TRANSCRIPTION FACTOR (FIT) and the POPEYE (PYE) networks (Kobayashi, 2019). The FIT network, localized in the root epidermis, regulates Fe uptake from the rhizosphere, while the PYE network, specific to the root vasculature, is involved in Fe transport into and out of the vascular tissues (Mari et al., 2020). Although these two networks are proposed to function synchronously (Kobayashi et al., 2019), their interconnection has not been demonstrated yet. Apart from the bHLH transcription factors, different transcription regulators also play a role in the control of Fe uptake and distribution mechanisms. The studies conducted in the last 10 years have demonstrated the interaction of different transcription factors with the bHLH regulatory network to control the pathways involved in hormonal regulations, abiotic stress responses, and even plant growth and development (Gao et al., 2019; Kobayashi, 2019).

Soybean (Glycine max L.) is an important legume crop rich in protein and oil contents, and it is used for food, feed, and biofuel production. Unfortunately, soybean is one of the crops that is most sensitive to Fe deficiency since it is affected by even a small decrease in Fe bioavailability. Fe deficiency in soybean leads to Fe deficiency chlorosis (IDC) due to limited chlorophyll biosynthesis, which, in turn, causes the yellowing of younger leaves, a reduction in leaf area, and shoot and root dry weights. Consequently, the soybean yield is significantly reduced by Fe deficiency due to high soil pH. High annual economic losses due to IDC in soybean production led to the development of new strategies to cope with Fe deficiency. Soil amendments and foliar Fe sprays have been used traditionally to correct mild chlorosis (Schenkeveld, 2010). However, they are not economically feasible. Genetic engineering has shown promise in developing IDC-tolerant soybean genotypes, but transformation and regeneration efficiencies remain low, and public acceptance is limited. Hence, breeding Fe-efficient cultivars through breeding programs stands as the most practical solution to address IDC in soybean. Therefore, the best way to overcome the IDC problem in soybean is the development of Fe-efficient cultivars in breeding programs. IDC is a polygenic quantitative trait related to the regulation of Fe acquisition, translocation, and accumulation in soybean (Fourcroy et al., 2014). Extensive research has identified genetic markers associated with Fe efficiency in soybean (Merry et al., 2019), and several quantitative trait loci (QTL) have been linked to Fe deficiency symptoms, particularly the IDC score and SPAD measurement (Diers et al., 1992; Lin et al., 1997; Lin et al., 2000a, Lin et al., 2000b; Charlson et al., 2005; Peiffer et al., 2012). However, the specific genes within these QTL regions remain unidentified and uncharacterized. The identification of genes involved in IDC-related QTLs would enhance our ability to understand molecular mechanisms behind IDC tolerance and sensitivity in soybean and facilitate the development of IDC-tolerant genotypes in the future.

In silico studies of already available QTL data and prediction of genes and their functional characterization in response to Fe deficiency can help in broadening our ability to understand the mechanisms behind Fe deficiency tolerance in soybean. Therefore, this study aimed to identify the IDC-related genes present in the previously reported IDC-responsive QTLs in soybeans and determine their expression levels under limited Fe conditions.

2. Materials and methods

2.1. Identification of QTL linked with Fe deficiency tolerance in soybean

QTL regions linked with Fe deficiency symptoms, especially with the IDC score and SPAD measurement, were previously generated from 6 different experiments, representing a divergence in genetic background, experimental setup, field, and environmental conditions (Diers et al., 1992; Lin et al., 1997; Lin et al., 2000a; Lin et al., 2000b; Charlson et al., 2005; Peiffer et al., 2012) (Tables S1 and S2). The QTL data were obtained from the SoyBase (https://www.soybase.org/) and are summarized in Table 1. Among a total of 40 QTL regions generated from the 6 experiments, only 19 represented unique regions. Chromosomal distribution of the 19 QTL regions was generated using Geneious software (Kearse et al., 2012) according to the soybean chromosome sizes and distribution of the QTL regions compared to the size of the chromosomes (Nepal et al., 2015). The centromere position was determined by identifying 91–92 nucleotide tandem repeats within the centromeric region. A total of 6593 genes were found in the 19 QTL regions by screening through the SoyBase Genome Browser (Nepal et al., 2015).

For proper annotation of the identified genes, their Arabidopsis thaliana orthologs were obtained from a BLASTp search of each gene against the A. thaliana genome in The Arabidopsis Information Resource (TAIR) (Lamesch et al., 2012) with preset parameters, and the top query hit was selected as the ortholog, where the percent similarity was >75% and $E < 10^{-5}$.

2.2. In silico expression and gene set enrichment analysis (GSEA) of the IDC-related genes

The expression levels of single-copy Arabidopsis orthologs were determined using Genvestigator (Zimmermann et al., 2004) in 5 microarray experiments generated from the whole roots of wild-type Col-0 exposed to Fe deficiency. The Gene Expression Omnibus (GEO) datasets used in the in silico analyses included GSE40076 (Sivitz et al.,

Table 1. QTL data used in the study.

QTL number	QTL name	Parent1	Parent 2	Nearest sequence-based genetic marker				Reference
				5' end	Position	3' end	Position	
9	**Fe effic 1-1**	**A81356022**	**PI468916**	BARC-011635-00314	Gm08:6809647	BARC-028361-05840	Gm08:7723315	Diers et al., 1992
3	Fe effic 1-2	A81356022	PI468916	Satt129	Gm01:55166202	BARC-018211-03157	Gm01:56113247	Diers et al., 1992
2	Fe effic 1-3	A81356022	PI468916	Satt407	Gm01:53174848	Sat_160	Gm01:54121486	Diers et al., 1992
1	Fe effic 1-4	A81356022	PI468916	BARC-020113-04470	Gm01:52286178	BARC-039805-07589	Gm01:53429616	Diers et al., 1992
14	Fe effic 1-5	A81356022	PI468916	Satt217	Gm18:4713017	Sat_315	Gm18:5348175	Diers et al., 1992
-	Fe effic 2-1	A81356022	PI468916	-	-	-	-	Diers et al., 1992
-	Fe effic 3-1	Pride B216 Fe inefficient	A15 Fe efficient	Satt122	Gm14:17619136	Sat_424	Gm14:46281808	Lin et al., 1997
-	Fe effic 3-2	Pride B216 Fe inefficient	A15 Fe efficient	Satt063	Gm14:45993714	BARC-039667-07534	Gm14:47205629	Lin et al., 1997
-	Fe effic 3-3	Pride B216 Fe inefficient	A15 Fe efficient	BARC-017185-02246	Gm18:50785143	BARC-018441-03188	Gm18:52157617	Lin et al., 1997
4	Fe effic 3-4	Pride B216 Fe inefficient	A15 Fe efficient	Satt159	Gm03:3197845	BARC-016199-02307	Gm03:5664735	Lin et al., 1997
-	Fe effic 4-1	Pride B216 Fe inefficient	A15 Fe efficient	BARC-013927-01275	Gm14:16752071	BARC-056587-14511	Gm14:16256888	Lin et al., 1997
-	Fe effic 4-2	Pride B216 Fe inefficient	A15 Fe efficient	BARC-029827-06444	Gm20:24177772	Sat_104	Gm20:37789703	Lin et al., 1997
-	Fe effic 4-3	Pride B216 Fe inefficient	A15 Fe efficient	Sat_334	Gm12:9131428	BARC-049209-10821	Gm12:35692712	Lin et al., 1997
7	Fe effic 5-1	Anoka Fe inefficient	A7 Fe efficient	BARC-020479-04634	Gm05:32322642	BARC-014463-01558	Gm05:34473339	Lin et al., 1997
-	Fe effic 5-2	Anoka Fe inefficient	A7 Fe efficient	Sat_033	Gm03:32837825	Satt237	Gm03:38104512	Lin et al., 1997
-	Fe effic 6-1	Anoka Fe inefficient	A7 Fe efficient	BARC-029827-06444	Gm20:24177772	Sat_104	Gm20:37789703	Lin et al., 1997
-	Fe effic 6-2	Anoka Fe inefficient	A7 Fe efficient	Sat_033	Gm03:32837825	Satt237	Gm03:38104512	Lin et al., 1997
15	Fe effic 7-1	Pride B216 Fe inefficient	A15 Fe efficient	Satt012	Gm18:49174114	BARC-018441-03188	Gm18:52157617	Lin et al., 1997
5	Fe effic 7-2	Pride B216 Fe inefficient	A15 Fe efficient	BARC-044085-08610	Gm03:7805399	BARC-013561-01160	Gm03:33280183	Lin et al., 1997
-	Fe effic 8-1	Pride B216 Fe inefficient	A15 Fe efficient	Satt020	Gm14:41294144	Sat_424	Gm14:46281808	Lin et al., 1997
-	Fe effic 8-2	Pride B216 Fe inefficient	A15 Fe efficient	Satt012	Gm18:49174114	BARC-018441-03188	Gm18:52157617	Lin et al., 1997
-	Fe effic 8-3	Pride B216 Fe inefficient	A15 Fe efficient	Sat_334	Gm12:9131428	BARC-049209-10821	Gm12:35692712	Lin et al., 1997
-	Fe effic 8-4	Pride B216 Fe inefficient	A15 Fe efficient	BARC-044085-08610	Gm03:7805399	BARC-013561-01160	Gm03:33280183	Lin et al., 1997
10	Fe effic 9-1	Pride B216 Fe inefficient	A15 Fe efficient	Sat_270	Gm11:4234139	BARC-059851-16137	Gm11:24571364	Lin et al., 2000B
-	Fe effic 9-2	Pride B216 Fe inefficient	A15 Fe efficient	Satt063	Gm14:45993714	BARC-039667-07534	Gm14:47205629	Lin et al., 2000B

Table 1. (Continued).

-	Fe effic 9-3	Pride B216 Fe inefficient	A15 Fe efficient	Satt012	Gm18:49174114	BARC-018441-03188	Gm18:52157617	Lin et al., 2000B
17	Fe effic 9-4	Pride B216 Fe inefficient	A15 Fe efficient	BARC-044913-08839	Gm19:43329322	BARC-013505-00505	Gm19:45014383	Lin et al., 2000B
-	Fe effic 9-5	Pride B216 Fe inefficient	A15 Fe efficient	BARC-044085-08610	Gm03:7805399	BARC-013561-01160	Gm03:33280183	Lin et al., 2000B
18	Fe effic 10-1	Anoka Fe inefficient	A7 Fe efficient	BARC-029827-06444	Gm20:24177772	Sat_104	Gm20:37789703	Lin et al., 2000B
6	Fe effic 10-2	Anoka Fe inefficient	A7 Fe efficient	Sat_033	Gm03:32837825	Satt237	Gm03:38104512	Lin et al., 2000B
13	Fe effic 10-3	Anoka Fe inefficient	A7 Fe efficient	Sat_189	Gm14:16352945	BARC-039667-07534	Gm14:47205629	Lin et al., 2000B
12	Fe effic 11-1	Pride B216 Fe inefficient	A15 Fe efficient	BARC-013927-01275	Gm14:16752071	BARC-056587-14511	Gm14:16256888	Lin et al., 2000A
-	Fe effic 11-2	Pride B216 Fe inefficient	A15 Fe efficient	BARC-029827-06444	Gm20:24177772	Sat_104	Gm20:37789703	Lin et al., 2000A
11	Fe effic 11-3	Pride B216 Fe inefficient	A15 Fe efficient	Sat_334	Gm12:9131428	BARC-049209-10821	Gm12:35692712	Lin et al., 2000A
-	Fe effic 12-1	Anoka Fe inefficient	A7 Fe efficient	BARC-029827-06444	Gm20:24177772	Sat_104	Gm20:37789703	Lin et al., 2000A
-	Fe effic 12-2	Anoka Fe inefficient	A7 Fe efficient	Sat_033	Gm03:32837825	Satt237	Gm03:38104512	Lin et al., 2000A
8	Fe effic 13-1	P9254 Low Fe chlorosis resistance	A97-770012 High Fe chlorosis resistance	BARC-060051-16321	Gm05:41877400	BARC-018681-02991	Gm05:41347323	Charlson et al., 2005
16	Fe effic 13-2	P9254 Low Fe chlorosis resistance	A97-770012 High Fe chlorosis resistance	BARC-059657-15973	Gm19:40154846	Satt156	Gm19:40637071	Charlson et al., 2005
19	Fe effic 13-3	P9254 Low Fe chlorosis resistance	A97-770012 High Fe chlorosis resistance	BARC-038869-07359	Gm20:37596189	BARC-030259-06840	Gm20:38262192	Charlson et al., 2005
-	Fe effic 14-1	Anoka Fe inefficient	A7 Fe efficient	Sat_033	Gm03:32837825	BARC-040277-07706	Gm03:36368900	Peiffer et al., 2012

[a] All data were obtained from the SoyBase QTL database (https://soybase.org/search/qtllist_by_symbol.php#I).

2012), GSE24348 (Schuler et al., 2011), GSE21582 (Long et al., 2010), GSE15189 (Buckhout et al., 2009), and GSE10576 (Dinneny et al., 2008). Differentially expressed genes (DEGs) were identified via gene search analysis of perturbations with a minimum target log ratio of 0.5, maximum target p-value of 0.05, and tolerance percent of 0. Expression patterns of the Fe deficiency-related *Arabidopsis orthologous*genes were determined in the *bhlh100 bhlh101* (GSE40076 - (Sivitz et al., 2012)) and *pye1-1* (GSE21582 - (Long et al., 2010)) mutants using Genevestigator (Zimmermann et al., 2004).

The heat map construction and/or hierarchical clustering of the genes was completed in Genevestigator according to the Manhattan correlation distance and the average linkage rule (Eisen et al., 1998). The gene ontology (GO) enrichment was performed using PANTHER 14.0 according to Fisher's exact test and corrected by Bonferroni correction for multiple testing, where $p < 0.05$ (Mi et al., 2019).

To prove the relation of the core group genes with Fe deficiency, GSEA was done using GeneTrail (Schuler et al., 2011) as explained previously (Eroglu and Aksoy, 2017). GSEA was performed against datasets obtained from the GEO accessions of GSE40076 (Sivitz et al., 2012), GSE24348 (Schuler et al., 2011), GSE21582 (Long et al., 2010), GSE15189 (Buckhout et al., 2009), GSE10576 (Dinneny et al., 2008), E-MEXP2378 (Mizoguchi et al., 2010), and E-MEXP475 (Nishimura et al., 2007), and

NASCArray accessions 176 (Goda Backes et al., 2008), 139 (Kilian et al., 2007), and 140 (Nakashima et al., 2006). The gene set consisting of constitutively expressed genes in Arabidopsis was used as a negative control in the analysis (Czechowski et al., 2005). The false discovery rate was set at $p < 0.05$.

The genes that coexpressed together with the core group genes were identified using ATTED-II Network Drawer (Obayashi et al., 2007), where the coexpression option was set to 'add a few genes' and the PPI was set to 'Draw PPIs'.

2.3. Plant material

A Turkish soybean (*Glycine max* (L.) Merr.) variety known to be IDC-sensitive (Atakişi) was used in the study (Maqbool, 2018) and the seeds were generously provided by Prof. Sevgi Çalışkan from Niğde Ömer Halisdemir University, Niğde, Türkiye.

2.4. Plant growth and stress application

In the experiment, a custom-made hydroponics culture system was used with some modifications (Kurtz, 2017). A cheesecloth was stretched over the mouth of 0.5-L polypropylene cups and the seeds placed on this setup were covered with polypropylene transparent cups with a 25-mm opening on the top for 10 days while they were germinating. The soybean seeds were surface sterilized by washing with 70% ethanol for 5 min, followed by washing with 0.25% sodium hypochlorite for 2 min. Then, they were rinsed 5 times with sterile distilled water. For the stratification of the seeds, they were kept at 4 °C in the dark for 2 days in sterile distilled water. Subsequently, the seeds were transferred to the hydroponics 1 h before the lights were turned on and germinated on ½ × Hoagland's solution at the following mineral element concentrations: 7.5 mM N, 3.0 mM K, 0.5 mM P, 2.5 mM Ca, 1.0 mM Mg, 1, 0 mM S, 50 µM Fe, 7.4 µM Mn, 0.96 µM Zn, 1.04 µM Cu, 7.13 µM B, and 0.01 µM Mo (Hoagland et al., n.d.). Fe was applied as sodium Fe-EDTA (Carl Zeiss Str., Roedermark, Germany). Three plants were transferred to each cup (as technical replicates) and 4 cups were used for each application as biological replicates. The cups were wrapped with aluminum foil to prevent algal growth. Under these conditions, the plants were grown for 2 weeks under Fe deficiency (0 µM Fe-EDTA, pH: 5.7), high pH (50 µM Fe-EDTA, pH: 9.0), or the control conditions (50 µM Fe-EDTA, pH: 5.7) in a controlled growth chamber at 25/20 °C (day/night) under a long day cycle (16:8 h light/dark photoperiod), 70%–80% relative humidity, and 200 µmol m^{-2} s^{-1} light provided by high-pressure sodium lamps placed over the plant canopy to prevent excessive light, and thus reduce the photoinhibition on the photosystem. The plants were sampled at the V2 stage since the IDC symptoms are best observed at this developmental stage in soybean. During the study, growth media were replaced with new ones every 48 h, thus eliminating the diminishing element and oxygen deficiencies in the environment. The experiment was performed in 5 replicates according to the randomized block design, where the location of the cups was randomly changed every 2 days.

2.5. Physiological and biochemical measurements

All of the aboveground measurements were taken from the top second trifoliate leaves that completed their full development. Chlorophyll index measurements were taken at noon using the SPAD-502 chlorophyll meter (Konica Minolta, Tokyo, Japan) from at least 3 independent measurements from each leaf and evaluated by calculating the mean and standard deviation. After the measurements, the second trifoliate leaves were used for total chlorophyll determination following extraction in 80% acetone (Aksoy et al., 2013). The absorbance readings of the extracts were determined at 470, 646.8, and 663.2 nm and compared to 80% acetone.

Plant roots and shoots were photographed with a digital camera at the end of the stress application. Root and shoot lengths were determined by ImageJ software (http://rsbweb.nih.gov/ij/) (Aksoy et al., 2013). After imaging, the fresh weights (FWs) of roots and shoots were recorded.

FCR activity was analyzed from the roots following the method in (Aksoy and Koiwa, 2013). Briefly, after the physiological measurements were completed, the plant roots were washed 3 times in distilled water, dried with tissue paper, and weighed. Then, the root samples were placed in the FCR activity solution (0.1 mM of Fe (III)-ethylenediaminetetraacetic acid (EDTA) and 0.3 mM of ferrozine in distilled water). The samples were kept in the solution for 12 h in the dark at room temperature. Then, the absorbance of the solution was read at 562 nm against a blank without the roots. The enzyme activity was calculated with a molar extinction coefficient of 28.6 mM^{-1} cm^{-1} and presented as normalized to the root FW.

2.6. Total RNA extraction and gene expression analyses

Total RNA was isolated from 200 mg of root sample in a 1-step method using TRIzol reagent (Thermo Fisher Scientific Inc., Waltham, MA, USA) (Chomczynski and Sacchi, 1987). After elimination of the genomic DNA contamination using a RapidOut DNA removal kit (Thermo Fisher Scientific) according to the manufacturer's instructions, RNA integrity and quantity were determined using 1% agarose gel electrophoresis and a microspectrophotometer, respectively. Complementary DNA was produced by the RevertAid First Strand cDNA synthesis kit (Thermo Fisher Scientific) by using random hexamers according to the manufacturer's instructions. Finally, the expression levels of the genes *(DIN1, MTPA2, AIP1, IREG2, ST2A,* and *GATA12)* were determined by real-time quantitative polymerase chain reaction (RT-qPCR) (Qiagen, USA) using the SYBR Green PCR kit (Roche Diagnostics, Basel, Basel-Stadt, Switzerland)

and the primers given in Table 2 following our previous study (Aksoy et al., 2013). Relative gene expression was calculated according to the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) after data normalization with *ELF1B* (Glyma.02g44460) reference gene (Jian et al., 2008). RT-qPCR experiments were designed in 2 biological and 3 technical replicates.

2.7. Statistical analysis

MINITAB 18.0 (Minitab, LLC, State College, PA, USA) was used in the statistical analyses. The physiological and biochemical data obtained in the study were compared at a 5% significance level by subjecting the variance analysis of the mean values (followed by the Tukey HSD post hoc test) according to the randomized block design. The gene expression data were analyzed using the Student's t-test at a 5% significance level. Statistical test results of the individual bioinformatic analyses are provided under the appropriate sections below.

3. Results

3.1. QTLs related to Fe deficiency tolerance in soybean

The analysis of 40 QTL regions linked with Fe deficiency tolerance found that only 19 represented unique regions in the soybean genome, demonstrating a conserved high-level genetic effect of IDC tolerance in soybean (Figure 1). Distributed on 10 soybean chromosomes, these regions encompass a total size of 101.763 megabases, representing 0.88% of the entire soybean genome (1.15 Gb) (Shultz et al., 2006). All of the genes in each QTL region were identified and this list was named the master gene list (Table S3). Accordingly, a total of 6593 genes were found in the master gene list and 347 genes on average occupied each QTL region. Genes with multiple copies in the soybean were eliminated from future analyses, since *Glycine* species contain paleopolyploid genomes that were established by an ancient genome duplication event that occurred 50,000 years ago (Shoemaker et al., 2006). As the soybean genome has not been well-annotated yet, a BLASTp search was performed of the proteins encoded by each gene against the *A. thaliana* genome to find its Arabidopsis ortholog and as a result, 3822 unique genes were found.

3.2. Identification of loci linked with Fe deficiency tolerance in soybean

To understand if the selected Arabidopsis orthologs were related to the Fe deficiency, their in silico expression levels under Fe deficiency were checked. Out of the 3822 unique genes, 607 were selected because their expression levels were differentially altered by more than 1.5 times (where $p < 0.05$) by one or more of the Fe deficiency treatments on the Arabidopsis (Figure 2). As the list of DEGs was long, how many of these genes were also affected by Fe deficiency signaling was further investigated. To this end, the expression levels of 607 genes were further evaluated in microarray experiments of Fe deficiency signaling mutants, i.e., *bhlh100bhlh101* and *pye1-1* (also known as *pye*). Thus, 68 genes were found to be differentially altered specifically in the *bhlh100bhlh101* mutant, while 13 DEGs were exclusively identified in the *pye1-1* mutant (Figure 3a, Table S4). It is interesting to note that DEGs identified in both *bhlh100bhlh101* or *pye1-1* suggested that they are unique genes that might be under the control of these transcription factors. Ten DEGs were found to be affected by both bHLH100-bHLH1001 and PYE, suggesting their importance in Fe uptake and/or translocation in plants. These core set of 10 genes include DNA J PROTEINC77 (DJC77), GLUTAMINE-DEPENDENT ASPARAGINE SYNTHASE1/DARK INDUCIBLE6 (ASN1/DIN6), UDP-GLUCOSYL TRANSFERASE72E1 (UGT21E1), METAL TOLERANCE PROTEIN3 (MTP3), IRON-REGULATED PROTEIN2/FERROPORTIN2 (IREG2/FPN2), SULFOTRANSFERASE2A (ST2A), GATA TRANSCRIPTION FACTOR12 (GATA12), CONSERVED IN THE GREEN LINEAGE AND DIATOMS27 (CGLD27), an adenine nucleotide alpha hydrolases-domain containing protein kinase (AT1G77280) and ALUMINUM-INDUCED PROTEIN1 (AT4G27450) (Table 3). Their gene expression levels were upregulated under Fe deficiency in at least one of the microarray

Table 2. RT-qPCR primers used in the gene expression analyses.

AGI number	SoyBase Number	Gene Code	Forward primer (5' – 3')	Reverse primer (5' – 3')
AT3G47340	Glyma.11g170300	*ASN1 / DIN6*	CATACTTGCTGTGCTTGGTTG	GCCCACTCCAGTCAGGACCA
AT3G58810	Glyma.11g135500	*MTP3 / MTPA2*	GGAAAGGACACCAAGTGAGA	CCTGCACTCCCTTGATATTTCTG
AT4G27450	Glyma.12g150500	*AIP1*	CCCTGCATCTAACAGTGGCT	GGAAGGTGGATCAGGACGAG
AT5G03570	Glyma.03g042400	*IREG2 / FPN2*	AGAGCCTCTTCTTGCCCAAC	AATTCCCACATCCTGGCACC
AT5G07010	Glyma.11g158900	*ST2A*	GGTCACTACCCACATTGGCA	AATGGGCCAAAACCAACCAC
AT5G25830	Glyma.14g145700	*GATA12*	GGCATGAATGCGCGAAACTA	CTCATAGTGGAGCCCGCC
AT5G12110	Glyma.02g44460	*ELF1B*	GTTGAAAAGCCAGGGGACA	TCTTACCCCTTGAGCGTGG

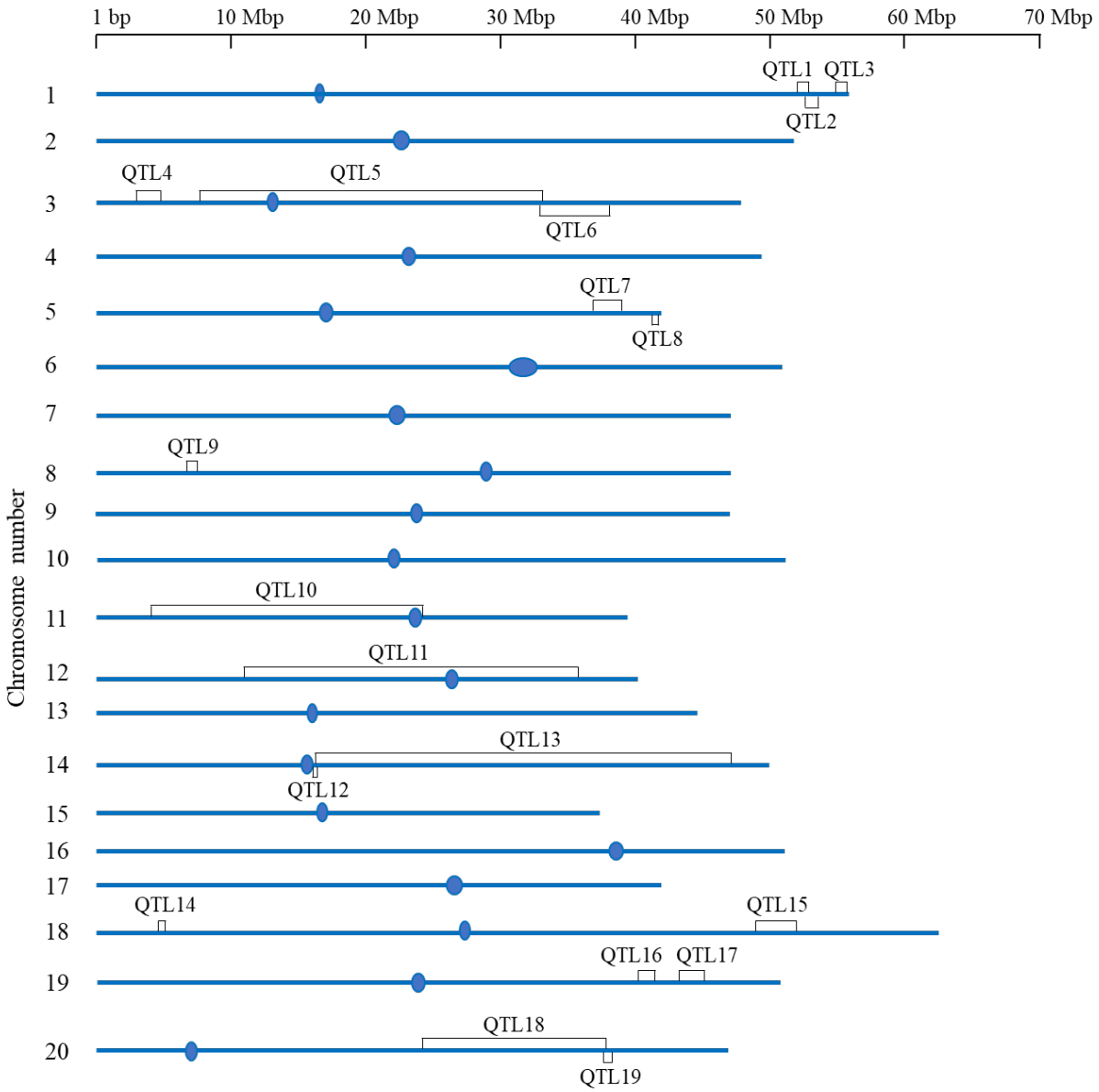

Figure 1. 19 QTL regions are distributed on ten soybean chromosomes. The figure was generated using Geneious software (Kearse et al., 2012) according to the soybean chromosome sizes and distribution of QTL regions compared to the size of the chromosomes (Nepal et al., 2015). The centromere position was determined by identifying 91–92 nucleotide tandem repeats within the centromeric region.

experiments performed from the roots, and all of them showed a high level of similar expression patterns under Fe deficiency in a hierarchical clustering analysis (Figure 3b), suggesting their close evolutionary relationship with each other. As the core group genes were not linked with Fe deficiency before, we performed a GSEA of 10 genes under available Fe deficiency microarray experiments (Table S5). As expected, the core group genes were explicitly enriched as a group in Fe deficiency, while they were not enriched in salinity, abscisic acid, or drought stresses. Overall, these analyses revealed that the core set of 10 genes was upregulated under Fe deficiency and might be under the control of bHLH transcription factors in plant roots for providing tolerance to Fe deficiency in soybean.

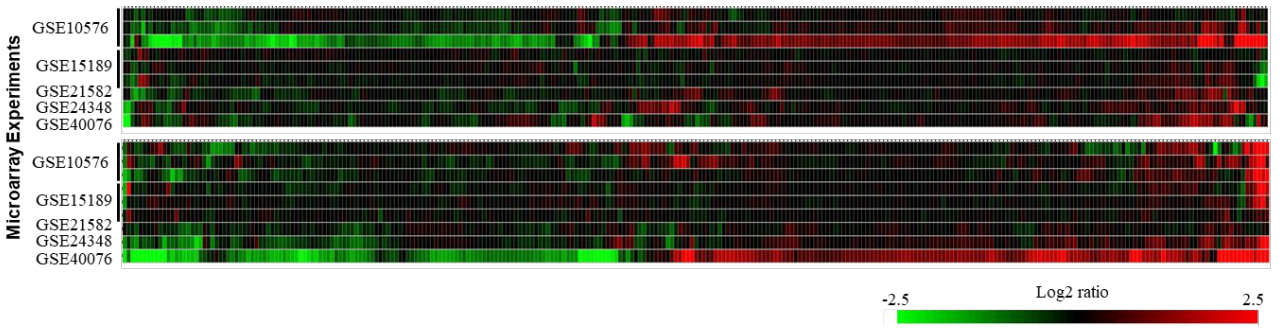

Figure 2. Expression levels of Arabidopsis orthologs under Fe deficiency. The heat map of 607 unique genes was generated from the microarray experiments, including the Fe deficiency treatment to the whole roots by Genevestigator (Zimmermann et al., 2004). GEO datasets used in the heat map construction included GSE40076 (Sivitz et al., 2012), GSE24348 (Schuler et al., 2011), GSE21582 (Long et al., 2010), GSE15189 (Buckhout et al., 2009), and GSE10576 (Dinneny et al., 2008).

3.3. Core group gene association with metal ion transport and Fe deficiency signaling

For the functional characterization of the core group genes, the GO enrichment was utilized (Table 4). Accordingly, the top 10 most enriched GOs include metal ion transport, specifically of Fe and zinc, in the biological processes, whereas Fe and zinc ion transmembrane transporter activity was enriched in the molecular functions. Some additionally enriched GOs include asparagine metabolic process ($p = 1.82 \times 10^{-03}$), (glutamine-hydrolyzing) asparagine synthase activity ($p = 1.82 \times 10^{-03}$), and sulfotransferase activity ($p = 6.91 \times 10^{-03}$), providing a clue for the involvement of sugar and sulfur metabolisms in Fe deficiency response and/or tolerance in plants.

To infer the potential functions of the core group genes, genes that coexpressed together with them were determined (Figure 4). The coexpression gene network identified both Fe-signaling components (i.e., PYE and FIT/FRU) as linked with the core group genes. The network also included several transcription factors and metal transporters with known functions in metal translocation and Fe homeostasis. Interestingly, there were some genes involved in sugar and amino acid metabolism in the network. As the network analysis presented coexpressing genes with known functions in Fe uptake, translocation, and signaling as well as the ones involved in other metabolic pathways, the network was divided into 6 clusters according to the network topology (Figure 4 and Table S6) and a GO enrichment was performed for the genes included in each of the 6 clusters (Table S7). As known from previous bioinformatics studies (Schwarz and Bauer, 2020), clusters 1 and 2 included the genes involved in the Fe-signaling networks regulated by PYE and FIT in the root stele and epidermis, respectively. Surprisingly, cluster 3 which was centered around GATA12, represented a knob between the 2 Fe-signaling networks, indicating the potential missing link between the 2 most important Fe-signaling components of Fe homeostasis in plant roots. Cluster 4 contained the genes with GO enrichment in amino acid metabolism, while clusters 5 and 6 contained the genes with GO enrichment in ethylene response and gibberellic acid-mediated signaling, respectively. Taken together, these data indicated that the core group genes are associated with metal ion transport and Fe deficiency signaling in plants (Table 4).

3.4 Core group genes induced by Fe deficiency and high pH in soybean roots

Finally, to prove that the core group genes identified in the bioinformatic analyses were induced by Fe deficiency in the soybean roots, Fe deficiency-sensitive soybean plants (Atakişi variety) were exposed to Fe deficiency for 2 weeks. Except for the FCR activity, the root and shoot lengths, root, and shoot FWs, soil plant analysis development (SPAD) value, and chlorophyll content were significantly decreased in the plants exposed to Fe deficiency, whereas the FCR activity increased by 2 times under Fe deficiency (Figure 5). As plants cannot uptake divalent elements such as Fe^{2+} efficiently in alkaline soils (Olsen and Brown, 1980), also investigated was the effect of a high pH (9.0) on the growth and development of the Atakişi variety. Except for the shoot length and FCR activity, the other parameters increased significantly with a high pH, as expected. On the other hand, the high pH did not affect the shoot length and FCR activity, suggesting that the induction of the FCR enzyme is unique to the Fe deficiency treatment.

The expression levels of some genes in the core group were determined in the plant roots exposed to Fe deficiency or high pH (Figure 6). The expression of *DIN1, MTPA2, AIP1, IREG2, ST2A,* and *GATA12* increased significantly under Fe deficiency and a high pH compared to the control

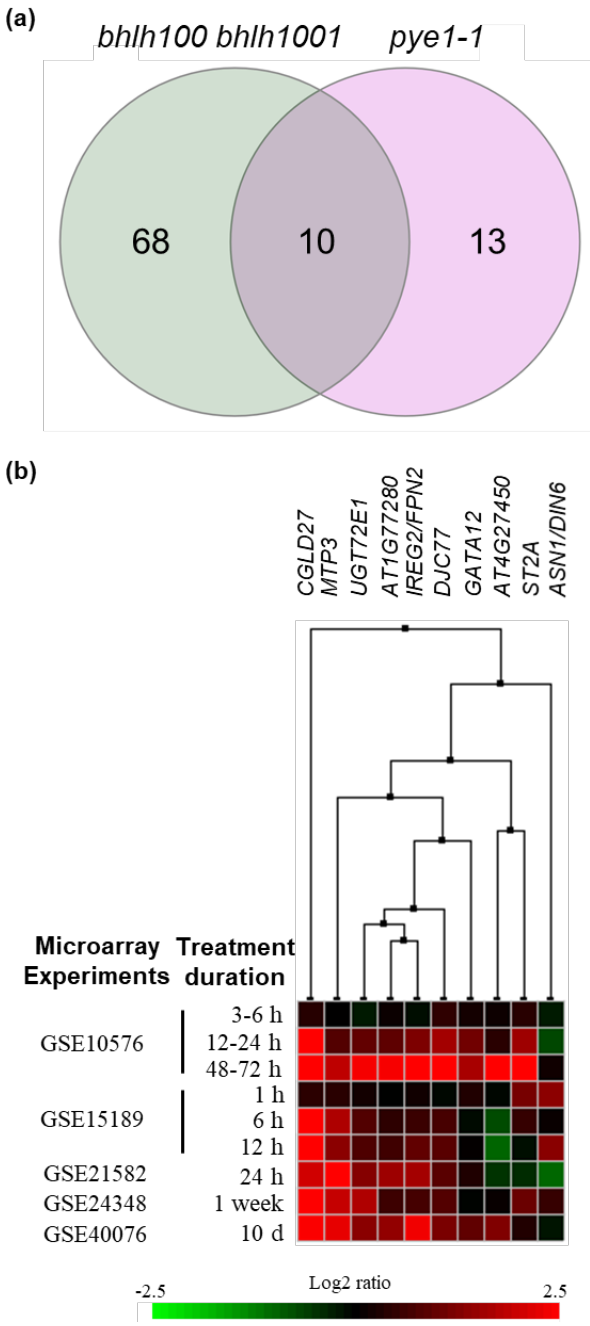

Figure 3. Genes linked with Fe signaling in the Arabidopsis roots. Expression patterns of Fe deficiency-related Arabidopsis orthologous genes were determined in the *bhlh100 bhlh101* (GSE40076 – [90]) and *pye1-1* (GSE21582 – [92]) mutants using Genevestigator [21]. (**a**) Venn diagram of the DEGs in the *bhlh100bhlh1011* and *pye1-1* mutants. (**b**) Hierarchical clustering of the 10 DEGs common in *bhlh100bhlh1011* and *pye1-1* under Fe deficiency microarrays. GEO datasets used in the hierarchical clustering construction included GSE40076 (Sivitz et al., 2012), GSE24348 (Schuler et al., 2011), GSE21582 (Long et al., 2010), GSE15189 (Buckhout et al., 2009), and GSE10576 (Dinneny et al., 2008).

Table 3. Expression levels of the ten core genes in Fe deficiency microarrays.

AGI number	SoyBase number	Gene name	Gene code	Relative expression								
				GSE40076 (Sivitz et al., 2012)	**GSE24348** (Schuler et al., 2011)	**GSE21582** (Long et al., 2010)	**GSE15189** (Buckhout et al., 2009)			**GSE10576** (Dinneny et al., 2008)		
				10 days	**1 week**	**24 h**	**1 h**	**6 h**	**24 h**	**3–6 h**	**12–24 h**	**48–72 h**
AT1G77280	Glyma.11g108000	Adenine nucleotide alpha hydrolases-domain containing protein kinase	-	**2.80**	**1.57**	**2.99**	–1.01	1.40	**1.53**	1.06	**1.93**	**5.21**
AT2G42750	Glyma.12g130000	*DNA J PROTEIN C77*	*DJC77*	**2.34**	**1.93**	**1.87**	–1.06	**1.65**	**1.83**	1.39	**1.64**	**6.29**
AT3G47340	Glyma.11g170300	*GLUTAMINE-DEPENDENT ASPARAGINE SYNTHASE 1/ DARK INDUCIBLE 6*	*ASN1/DIN6*	–1.12	**2.11**	**–2.06**	**1.84**	1.04	**2.53**	–1.21	**–1.65**	1.23
AT3G50740	Glyma.03g032600	*UDP-GLUCOSYL TRANSFERASE 72E1*	*UGT72E1*	**2.61**	**3.06**	**2.51**	1.12	**1.81**	**1.68**	–1.19	**2.02**	**5.52**
AT3G58810	Glyma.11g135500	*METAL TOLERANCE PROTEIN 3/METAL TOLERANCE PROTEIN A2*	*MTP3/MTPA2*	**4.95**	**3.63**	**8.78**	1.35	**3.53**	**2.57**	1.02	**1.82**	**1.91**
AT4G27450	Glyma.12g150500	*ALUMINUM INDUCED PROTEIN 1*	*AIP1*	**2.41**	1.13	–1.42	–1.09	**–2.57**	**–2.18**	1.09	1.49	**13.86**
AT5G03570	Glyma.03g042400	*FE-REGULATED PROTEIN 2/ FERROPORTIN 2*	*IREG2/FPN2*	**5.33**	**1.54**	**3.11**	1.14	**1.53**	**1.94**	–1.09	**2.54**	**6.72**
AT5G07010	Glyma.11g158900	*SULFOTRANSFERASE 2A*	*ST2A*	1.27	**3.78**	–1.37	**2.67**	1.38	–1.09	1.27	**2.91**	**11.28**
AT5G25830	Glyma.14g145700	*GATA TRANSCRIPTION FACTOR 12*	*GATA12*	**1.96**	–1.00	1.24	1.25	–1.07	1.03	1.15	**2.23**	**3.20**
AT5G67370	Glyma.01g191600	*CONSERVED IN THE GREEN LINEAGE AND DIATOMS 27*	*CGLD27*	**5.74**	**8.30**	**4.20**	1.46	**7.22**	**6.60**	1.19	**7.07**	**4.67**

Relative expression levels in bold indicate a statistical significance differential expression (cut-off value of 1.5-times) compared to the control condition ($p < 0.00$)

Table 4. GO enrichment of the ten core genes.

GO term	Fold enrichment	raw p-value
Biological process		
Detoxification of zinc ion	>100	3.65E-04
Stress response to zinc ion	>100	7.29E-04
Asparagine metabolic process	>100	1.82E-03
Cellular response to Fe ion starvation	>100	2.91E-03
Fe ion transmembrane transport	>100	3.28E-03
Response to Fe ion starvation	>100	4.37E-03
Stress response to metal ion	>100	6.18E-03
Zinc ion transport	>100	9.80E-03
Fe ion transport	88.44	1.12E-02
Cellular Fe ion homeostasis	70.30	1.41E-02
Molecular function		
Asparagine synthase (glutamine-hydrolyzing) activity	>100	1.82E-03
Sulfotransferase activity	>100	6.91E-03
Fe ion transmembrane transporter activity	>100	8.72E-03
Zinc ion transmembrane transporter activity	>100	9.80E-03
Transferase activity	80.64	1.27E-02
Transition metal ion transmembrane transporter activity	74.10	3.36E-04
Divalent inorganic cation transmembrane transporter activity	68.54	1.49E-02
Ligase activity	38.08	2.63E-02
Metal ion transmembrane transporter activity	23.33	3.20E-03
Inorganic anion transmembrane transporter activity	21.93	4.50E-02

group. Under a high pH, the induction of ST2A and GATA12 reached 10 and 20 times, respectively. Overall, this indicated that the core group genes were induced by Fe deficiency and high pH in soybean roots and therefore their functions in Fe uptake and/or translocation in plants should be investigated in future studies.

4. Discussion

Herein, the identification of 19 unique regions containing 6493 genes were presented for the first time in soybean (Figure 1) and 3822 unique ortholog genes were obtained in *A. thaliana*. In silico expression analysis of 3822 unique genes identified 607 of them as differentially expressed under Fe deficiency. To gain a more detailed understanding of their regulation under Fe deficiency, their expression levels in the *bhlh100bhlh101* and *pye1-1* mutants were evaluated, since POPEYE (PYE) is localized to the stele, while basic helix-loop-helix transcription factors bHLH100 and bHLH1001 are localized to the root epidermis and function as a dimer with other bHLH transcription factors including the major Fe deficiency signaling regulator FIT (Gao et al., 2019).

The in silico analysis revealed 10 DEGs affected by both bHLH100-bHLH1001 and PYE (Table 3). GSEA showed that the core group of genes was linked with Fe deficiency response, while the hierarchical clustering demonstrated their similar expression patterns in different Fe deficiency experiments. Among these genes, the molecular functions of METAL TOLERANCE PROTEIN3 (MTP3) and IRON-REGULATED PROTEIN2/FERROPORTIN2 (IREG2/FPN2) have already been shown under Fe deficiency. The core group of genes was enriched in GOs linked with Fe tolerance in previous studies (Schwarz and Bauer, 2020). Some additionally enriched GOs include sulfur metabolism. The positive regulatory effect of hydrogen sulfide (Chen et al., 2020) and the involvement of sulfur metabolism in the alleviation of Fe deficiency responses have been demonstrated in plants. Interestingly, these studies demonstrated the necessity of sulfur for proper Fe deficiency tolerance in plants (Robe et al., 2020), and the contribution of hormones, especially auxin and gibberellic acid, to the tolerance mechanisms (Chen et al., 2020).

The coexpression gene network analysis identified 6 clusters linked with the core group of genes (Figure 5).

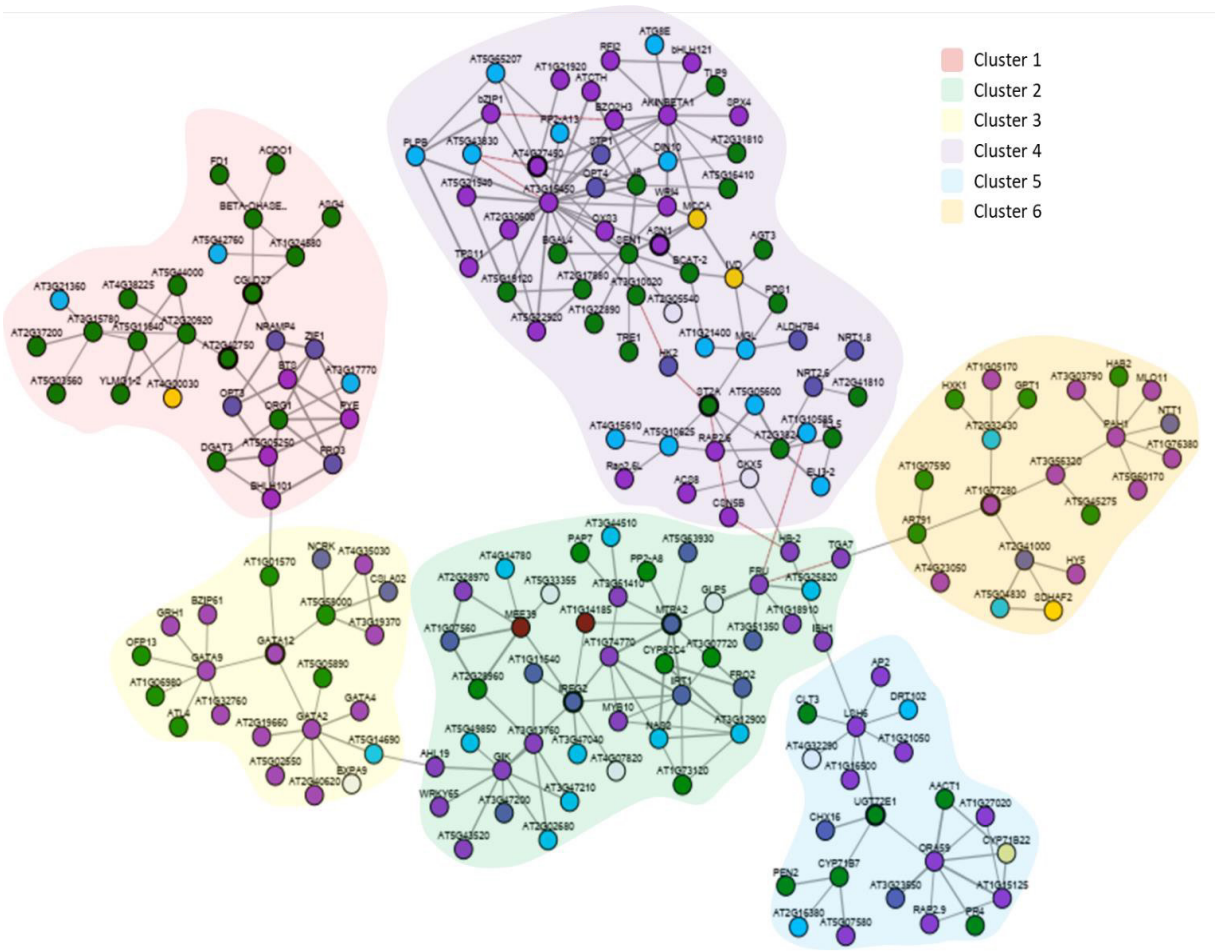

Figure 4. Coexpression network analysis of the core group genes. The coexpression network drawer of Atted II (Obayashi et al., 2007) was used to generate the network map. Dark circles present the 10 core group genes. Colors indicate the predicted subcellular localizations according to WoLF PSORT (Horton et al., 2007). Green: chloroplast, yellow: mitochondria, purple: nucleus, pale blue: cytosol, red: vacuole, dark blue: plasma membrane, pale grey: extracellular, and beige: peroxisome.

Clusters 1 and 2 included the genes involved in the Fe-signaling networks regulated by PYE and FIT/FRU in the root stele and epidermis, respectively, as determined in previous bioinformatic studies (Schwarz and Bauer, 2020). Although these spatially separated signaling networks have been known for a long time, their interdependence has not been shown before. The cluster analysis herein showed that cluster 3 represented a link between the 2 Fe-signaling networks, indicating the potential missing link between the 2 most important Fe-signaling components of Fe homeostasis in plant roots. Interestingly, cluster 3 was centered around GATA12. By binding to the main sequence of G-A-T-A, GATA transcription factors are classified in the IV group of the zinc finger transcription factor family and function in different pathways (Reyes et al., 2004). Thirty GATA transcription factors have been identified in the Arabidopsis genome and divided into 4 subgroups preserved within themselves. While GATA12 was found in the first subgroup of GATAs together with GATA9, GATA2, and GATA4, the clustering in this study showed a coexpression of these 4 GATA transcription factors together in cluster 3. It has been reported that GATA transcription factors are controlled by different hormones, and thus play important roles in plant growth and stress responses (Behringer and Schwechheimer, 2015). In a recent study, it was shown that GATA12 was expressed especially in the root and stem, localized to the nucleus, and negatively affected dormancy (Ravindran et al., 2017). In the same study, it was shown that the activation of GATA12 was dependent on the presence of gibberellic acid and GATA12 was activated as a result of the interaction of the DELLA transcription factor RGL2, which is the main

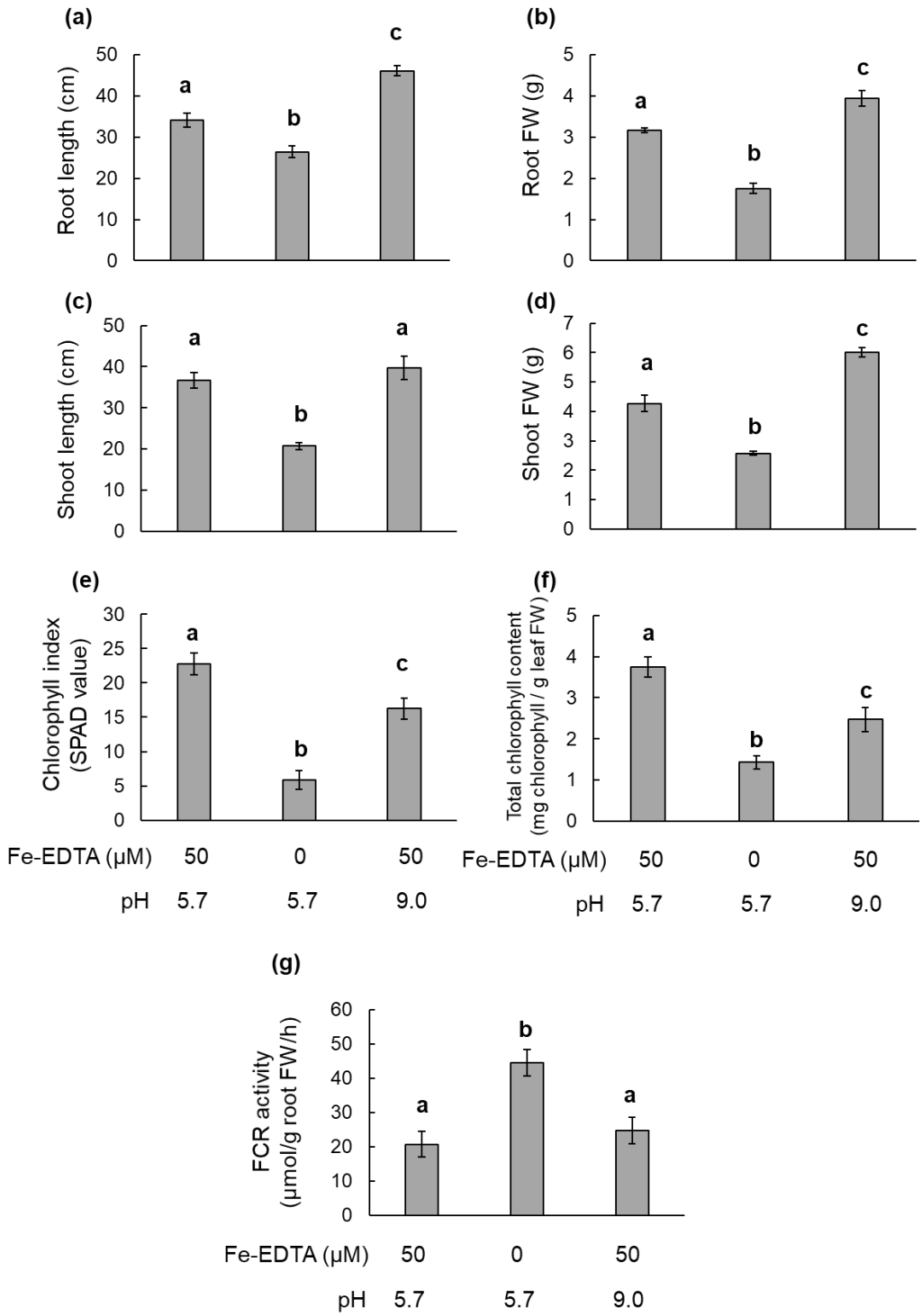

Figure 5. Physiological and biochemical responses of soybean under Fe deficiency and high pH. The soybean plants were grown under Fe deficiency (0 μM Fe-EDTA, pH: 5.7) and high pH (50 μM Fe-EDTA, pH: 9.0) compared to the control conditions (50 μM Fe-EDTA, pH: 5.7) for 2 weeks. (**a**) Root length. (b) Root FW. (**c**) Shoot length. (**d**) Shoot FW. (**e**) SPAD value. (**f**) Chlorophyll content. (**g**) FCR enzyme activity. Values indicate the means ± SEM (n ≥ 3). Different letters indicate significant changes among the treatments ($p < 0.05$).

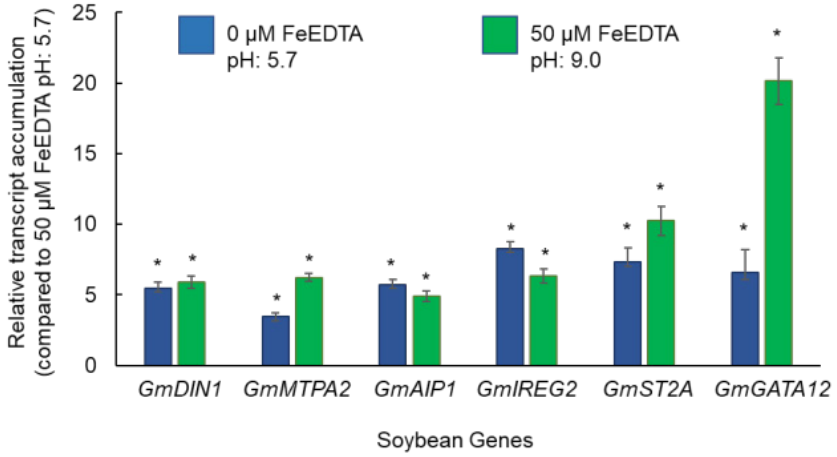

Figure 6. Expression levels of *DIN1, MTPA2, AIP1, IREG2, ST2A*, and *GATA12* in the soybean roots under Fe deficiency and high pH. The gene expression levels were determined in the roots of soybean plants grown under Fe deficiency (0 µM Fe-EDTA, pH: 5.7) and high pH (50 µM Fe-EDTA, pH: 9.0) compared to the control conditions (50 µM Fe-EDTA, pH: 5.7) for 2 weeks. Values indicate the means of relative gene expression ± SEM (n = 3). * Significant change compared to the control ($p < 0.05$).

regulator of gibberellic acid signaling, with DOF6, the stress transcription factor. This suggests that the DELLA proteins are one of the possible regulators of GATA transcription factors. In another study, it was shown that the DELLA proteins post-translationally regulate the Fe-signaling regulator FIT in the roots (Wild et al., 2016). In that study, it was reported that the local distribution of DELLA suppressors responsible for gibberellin signaling was regulated by Fe and contributed to the modulation of the root system architecture. DELLA interacts with FIT activity to inhibit it, but DELLA becomes unstable, and unable to release the FIT to activate Fe uptake genes in epidermis cells under Fe deficiency. Taken together, the coexpression network analysis conducted in the present study, together with the RT-qPCR results, highlight the importance of the interaction of GATA transcription factors with DELLAs in regulating the stability of FIT (and possibly other bHLH transcription factors involved in Fe signaling) for proper Fe uptake from the rhizosphere in plants.

Cluster 1 was centered around DJC77 (AT2G42750), which is a J domain-containing cochaperone (J protein) that interacts with the chloroplast-localized heat shock protein HSP70 to regulate the binding of HSP70 to folding proteins under stress conditions (Chiu et al., 2013). Interestingly, *DJC77* is highly expressed in the leaves, especially under selenium treatment (Van Hoewyk et al., 2008), and was shown to be mobile from leaves to the roots via the phloem (Thieme et al., 2015), suggesting its potential regulatory effect under mineral imbalances.

Cluster 4 included bHLH121, also known as UPSTREAM REGULATOR OF IRT1 (URI), which was recently shown to act as a crucial regulator of the Fe deficiency signaling pathway (Gao et al., 2020). The phosphorylated form of URI accumulates under Fe deficiency and forms heterodimers with bHLH38/39/100/101. These transcription factors in turn heterodimerize with the FIT and drive the transcription of *IRT1* and *FRO2* to increase Fe uptake. It also interacts with PYE, BTS, and BTSL1. Surprisingly, URI clustered with STA2, ASN1, and AIP1 in cluster 4 instead of clustering together with FIT, FRO2, IRT1, and other bHLHs (Figure 5). Moreover, *FIT* was found to be epistatic to *bHLH121*, and *FIT* overexpression partially rescued the severe phenotype of the *bhlh121* mutant under Fe deficiency (Lei et al., 2020). The expression of *URI* was not affected by Fe availability. This suggests URI as an upstream transcription factor regulating the Fe-signaling pathway depending on the availability of nutrients, especially of Fe, via amino acid metabolism, since cluster 4 was enriched in GOs linked to amino acid metabolism. Taken together, these observations indicate that the core group genes are associated with metal ion transport and Fe deficiency signaling in plants. Moreover, they expand our understanding of the Fe deficiency signaling network toward the integration of plant hormonal signals.

Under Fe deficiency, the root and shoot lengths as well as the FWs decreased significantly compared to the control conditions (Figure 6), which is in agreement with previous studies done on soybean (Kurtz, 2017; Maqbool, 2018). As Fe is required for chlorophyll biosynthesis, the total

chlorophyll content and the chlorophyll index decrease significantly under Fe deficiency in Fe-sensitive soybean genotypes (Maqbool, 2018). Hence, the observations herein are in agreement with the literature. FCR was determined as the major factor affecting the Fe uptake in dicots and the application of Fe deficiency enhances the FCR activity in plants (Jiménez et al., 2019). Herein, a two-fold increase was observed in the FCR activity in the roots of the Atakişi plants under Fe deficiency, proving that the plants experienced Fe deficiency in the hydroponics setup. Dicots require the protonation of the rhizosphere for proper solubilization and uptake of Fe. Therefore, they cannot uptake Fe efficiently in alkaline soils (Olsen and Brown, 1980). To understand how high pH affects plant growth, Atakişi plants were also grown under a high pH (9.0) for two weeks. As expected, high pH decreased the chlorophyll content while the root length increased together with root and shoot FWs (Figure 6). The effects of a high pH and Fe deficiency are opposite in terms of the root growth (Gheshlaghi et al., 2020), which was also observed in the current study. While the high pH and Fe deficiency caused a decrease in the chlorophyll content in plants (Maqbool, 2018), Fe deficiency showed a more remarkable reduction compared to the high pH application. Interestingly, the high pH did not increase the FCR activity, while Fe deficiency did, suggesting that different molecular mechanisms would be affected by individual stresses. This observation confirms the previously reported mildly acidic pH optimum of *FRO2*-mediated Fe chelate reduction (Tsai et al., 2018).

The molecular networks affected by alkalinity (Tsai and Schmidt, 2020) and Fe deficiency have already been studied in plants. However, genes involved in individual networks have not been compared to determine the common genes. Interestingly, the expression levels of some core group genes were significantly induced under both Fe deficiency and high pH, indicating their potential involvement in the common set of genes regulated under both stresses. The identification of these core genes involved in IDC-related QTLs in soybean would enhance our ability to understand molecular mechanisms behind IDC tolerance and sensitivity in soybean and facilitate the development of IDC-tolerant genotypes in the future. Among these genes, *GATA12* stands out with its induced expression levels around 5 and 20 times under Fe deficiency and high pH, respectively (Figure 7). Since cluster 3, which was centered around GATA12, represented a knob between the 2 Fe-signaling networks, indicating the potential missing link between the 2 most important Fe-signaling components of Fe homeostasis in plant roots. Therefore, the potential regulatory roles of the GATA transcription factors in Fe signaling should be investigated in future studies.

5. Conclusions

Several molecular studies have been performed with IDC-tolerant and sensitive soybean genotypes to find the QTLs linked with IDC in soybean. However, none of these studies have identified the potential genes in these QTL regions and their interactions in Fe uptake in soybean. Herein, 10 core group genes were identified that have not been linked with Fe deficiency responses before. Their interactions with other well-known Fe-signaling components represent the interaction of different nutrient metabolisms and hormonal regulations. Among these 10 core genes, GATA12 stands out as an important integrator of Fe signaling between the epidermis-localized FIT and the stele-localized PYE signaling networks. Overall, these results suggest that the core group genes are induced by Fe deficiency and high pH in soybean roots and GATA12 should be characterized under Fe deficiency in plants.

Supplementary Materials

The supplementary materials are available online at https://shorturl.at/syCLW

Funding

No external funding was received for this research.

Acknowledgments

The authors thank Prof. Sevgi ÇALIŞKAN for providing the soybean seeds used in the study. This article is based upon work from COST Action PLANTMETALS, CA19116, supported by COST (European Cooperation in Science and Technology).

Conflict of interest

The authors have no conflicts of interest to declare.

References

Aksoy E, Koiwa H (2013). Determination of ferric chelate reductase activity in the *Arabidopsis thaliana* root. Bio-protocol 3 (15): e843-e843. https://doi.org/10.21769/BioProtoc.843

Aksoy E, Jeong IS, Koiwa H (2013). Loss of function of Arabidopsis C-terminal domain phosphatase-like1 activates iron deficiency responses at the transcriptional level. Plant Physiology 161 (1): 330-345. https://doi.org/10.1104/pp.112.207043

Aksoy E, Yerlikaya BA, Ayten S, Abudureyimu B (2018). Iron uptake mechanisms from the rhizosphere in plants. Turkish Journal of Agriculture - Food Science and Technology 6 (12): 1673-1683. https://doi.org/10.24925/turjaf.v6i12.1673-1683.1326

Behringer C, Schwechheimer C (2015). B-GATA transcription factors-insights into their structure, regulation, and role in plant development. Frontiers in Plant Science 6: 90. https://doi.org/10.3389/fpls.2015.00090

Buckhout TJ, Yang TJ, Schmidt W (2009). Early iron-deficiency-induced transcriptional changes in Arabidopsis roots as revealed by microarray analyses. BMC Genomics 10: 14. https://doi.org/10.1186/1471-2164-10-147

Charlson DV, Bailey TB, Cianzio SR, Shoemaker RC (2005). Molecular marker Satt481 is associated with iron-deficiency chlorosis resistance in a soybean breeding population. Crop Science 45 (6): 2394-2399. https://doi.org/10.2135/cropsci2004.0510.

Chen J, Zhang NN, Pan Q, Lin XY, Shangguan Z et al. (2020). Hydrogen sulphide alleviates iron deficiency by promoting iron availability and plant hormone levels in *Glycine max* seedlings. BMC Plant Biology 20 (1): 122. https://doi.org/10.1186/s12870-020-02601-2

Chiu CC, Chen LJ, Su PH, Li H (2013). Evolution of chloroplast J proteins. PloS One 8 (7): e70384. https://doi.org/10.1371/journal.pone.0070384

Chomczynski P, Sacchi N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate phenol chloroform extraction. Analytical Biochemistry 162 (1): 156-159.https://doi.org/10.1016/0003-2697(87)90021-2

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiology 139 (1): 5-17. https://doi.org/10.1104/pp.105.063743

Diers BW, Cianzio SR, Shoemaker RC (1992). Possible identification of quantitative trait loci affecting iron efficiency in soybean. Journal of Plant Nutrition 15 (10): 2127-2136. https://doi.org/10.1080/01904169209364462

Dinneny JR, Long TA, Wang JY, Jung JW, Mace D et al. (2008). Cell identity mediates the response of Arabidopsis roots to abiotic stress. Science 320 (5878): 942-945. https://doi.org/10.1126/science.1153795

Eisen MB, Spellman PT, Brown PO, Botstein D (1998). Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America 95 (25): 14863-14868. https://doi.org/10.1073/pnas.95.25.14863

Eroglu S, Aksoy E (2017). Genome-wide analysis of gene expression profiling revealed that COP9 signalosome is essential for correct expression of Fe homeostasis genes in Arabidopsis. Biometals 30 (5): 685-698. https://doi.org/10.1007/s10534-017-0036-8

Fourcroy P, Sisó-Terraza P, Sudre D, Savirón M, Reyt G et al. (2014). Involvement of the ABCG 37 transporter in secretion of scopoletin and derivatives by Arabidopsis roots in response to iron deficiency. New Phytologist 201 (1): 155-167. https://doi.org/10.1111/nph.12471

Gao F, Robe K, Bettembourg M, Navarro N, Rofidal V et al. (2020). The transcription factor bHLH121 interacts with bHLH105 (ILR3) and its closest homologs to regulate iron homeostasis in Arabidopsis. The Plant Cell 32 (2): 508-524. https://doi.org/10.1105/tpc.19.00541

Gao F, Robe K, Gaymard F, Izquierdo E, Dubos C (2019). The transcriptional control of iron homeostasis in plants: a tale of bHLH transcription factors? Frontiers in Plant Science 10: 6. https://doi.org/10.3389/fpls.2019.00006

Gheshlaghi Z, Luis-Villarroya A, Álvarez-Fernández A, Khorassani R, Abadía J (2020). Iron deficient *Medicago scutellata* grown in nutrient solution at high pH accumulates and secretes large amounts of flavins. Plant Science 110664. https://doi.org/10.1016/j.plantsci.2020.110664

Goda H, Sasaki E, Akiyama K, Maruyama-Nakashita A, Nakabayashi K et al. (2008). The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. Plant Journal 55 (3): 526-542. https://doi.org/10.1111/j.1365-313X.2008.03510.x

Horton P, Park K-J, Obayashi T, Fujita N, Harada H et al. (2007). WoLF PSORT: protein localization predictor. Nucleic Acids Research, 35 (2): W585-W587. https://doi.org/10.1093/nar/gkm259

Hu R, Fan C, Li H, Zhang Q, Fu YF (2009). Evaluation of putative reference genes for gene expression normalization in soybean by quantitative real-time RT-PCR. BMC Molecular Biology 10: 93. https://doi.org/10.1186/1471-2199-10-93

Jian B, Liu B, Bi Y, Hou W, Wu C et al. (2008). Validation of internal control for gene expression study in soybean by quantitative real-time PCR. BMC Molecular Biology 9 (1): 1-14. https://doi.org/10.1186/1471-2199-9-59

Jiménez MR, Casanova L, Saavedra T, Gama F, Suárez MP et al. (2019). Responses of tomato (*Solanum lycopersicum* L.) plants to iron deficiency in the root zone. Folia Horticulturae 31 (1): 223-234. https://doi.org/10.2478/fhort-2019-0017

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M et al. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12): 16471649. https://doi.org/10.1093/bioinformatics/bts199

Kermeur N, Pédrot M, Cabello-Hurtado F (2023). Iron availability and homeostasis in plants: a review of responses, adaptive mechanisms, and signaling. In: Couée, I. (editor) Plant Abiotic Stress Signaling. Methods in Molecular Biology. Vol 2642. New York, USA: Humana Press, pp. 49-81. https://doi.org/10.1007/978-1-0716-3044-0_3

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S et al. (2007). The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant Journal 50 (2): 347-363. https://doi.org/10.1111/j.1365-313X.2007.03052.x

Kobayashi T (2019). Understanding the complexity of iron sensing and signaling cascades in plants. Plant and Cell Physiology 60 (7): 1440-1446. https://doi.org/10.1093/pcp/pcz038

Kurtz S (2017). Hydroponic Testing of Iron Deficiency Chlorosis in Soybeans. PhD, Iowa State University, Ames, IA, USA.

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C et al. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Research 40 (D1): D1202-D1210. https://doi.org/10.1093/nar/gkr1090

Lei R, Li Y, Cai Y, Li C, Pu M et al. (2020). bHLH121 functions as a direct link that facilitates the activation of FIT by bHLH IVc transcription factors for maintaining Fe homeostasis in Arabidopsis. Molecular Plant. https://doi.org/10.1016/j.molp.2020.01.006

Lin S F, Baumer J, Ivers D, Cianzio S, Shoemaker R (2000a). Nutrient solution screening of Fe chlorosis resistance in soybean evaluated by molecular characterization. Journal of Plant Nutrition 23 (11-12): 1915-1928. https://doi.org/10.1080/01904160009382153

Lin S F, Grant D, Cianzio S, Shoemaker R. (2000b). Molecular characterization of iron deficiency chlorosis in soybean. Journal of Plant Nutrition 23 (11-12): 1929-1939. https://doi.org/10.1080/01904160009382154

Lin S, Cianzio S, Shoemaker R (1997). Mapping genetic loci for iron deficiency chlorosis in soybean. Molecular Breeding 3 (3): 219-229. https://doi.org/10.1023/A:1009637320805

Livak K J, Schmittgen TD (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. Methods 25 (4): 402-408. https://doi.org/10.1006/meth.2001.1262

Long TA, Tsukagoshi H, Busch W, Lahner B, Salt DE et al (2010). The bHLH transcription factor POPEYE regulates response to iron deficiency in Arabidopsis roots. Plant Cell 22 (7): 2219-2236. https://doi.org/10.1105/tpc.110.074096

Maqbool A (2018). Investigation of physiological, biochemical and molecular responses of soybean cultivars under iron deficiency. MSc. Niğde Ömer Halisdemir Üniversitesi, Niğde, Türkiye.

Mari S, Bailly C, Sebastien T (2020). Handing off iron to the next generation: how does it get into seeds and what for? Biochemical Journal 477 (1): 259-274. https://doi.org/10.1042/BCJ20190188

Merry R, Butenhoff K, Campbell BW, Michno J, Wang D et al. (2019). Identification and fine-mapping of a soybean quantitative trait locus on chromosome 5 conferring tolerance to iron deficiency chlorosis. The Plant Genome 12 (3): 1-13. https://doi.org/10.3835/plantgenome2019.01.0007

Mi H, Muruganujan A, Ebert D, Huang X, Thomas PD (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Research 47 (D1): D419-D426. https://doi.org/10.1093/nar/gky1038

Mizoguchi M, Umezawa T, Nakashima K, Kidokoro S, Takasaki H et al. (2010). Two closely related subclass II SnRK2 protein kinases cooperatively regulate drought-inducible gene expression. Plant and Cell Physiology 51 (5): 842-847. https://doi.org/10.1093/pcp/pcq041

Nakashima K, Fujita Y, Katsura K, Maruyama K, Narusaka Y et al. (2006). Transcriptional regulation of ABI3- and ABA-responsive genes including RD29B and RD29A in seeds, germinating embryos, and seedlings of Arabidopsis. Plant Molecular Biology 60 (1): 51-68. https://doi.org/10.1007/s11103-005-2418-5

Nepal MP, Benson BV (2015). CNL disease resistance genes in soybean and their evolutionary divergence. Evolutionary Bioinformatics 11: 49-63. https://doi.org/10.4137/EBO.S21782

Nishimura N, Yoshida T, Kitahata N, Asami T, Shinozaki K et al. (2007). ABA-hypersensitive germination1 encodes a protein phosphatase 2C, an essential component of abscisic acid signaling in Arabidopsis seed. Plant Journal 50 (6): 935-949. https://doi.org/10.1111/j.1365-313X.2007.03107.x

Obayashi T, Kinoshita K, Nakai K, Shibaoka M, Hayashi S et al. (2007). ATTED-II: A database of co-expressed genes and cis elements for identifying co-regulated gene groups in Arabidopsis. Nucleic Acids Research 35 (1): D863-D869. https://doi.org/10.1093/nar/gkl783

Olsen RA, Brown JC (1980). Factors related to iron uptake by dicotyledonous and monocotyledonous plants. I. pH and reductant. Journal of Plant Nutrition 2 (6): 629-645. https://doi.org/10.1080/01904168009362804

Onyango DA, Entila F, Dida MM, Ismail AM, Drame KN (2019). Mechanistic understanding of iron toxicity tolerance in contrasting rice varieties from Africa: 1. Morpho-physiological and biochemical responses. Functional Plant Biology 46 (1): 93-105. https://doi.org/10.1071/FP18129

Peiffer GA, King KE, Severin AJ, May GD, Cianzio SR et al. (2012). Identification of candidate genes underlying an iron efficiency quantitative trait locus in soybean. Plant Physiology 158 (4): 1745-1754. https://doi.org/10.1104/pp.111.189860

Ravindran P, Verma V, Stamm P, Kumar PP (2017). A novel RGL2-DOF6 complex contributes to primary seed dormancy in *Arabidopsis thaliana* by regulating a GATA transcription factor. Molecular Plant 10 (10): 1307-1320. https://doi.org/10.1016/j.molp.2017.09.004

Reyes JC, Muro-Pastor MI, Florencio FJ (2004). The GATA family of transcription factors in Arabidopsis and rice. Plant Physiology 134 (4): 1718-1732. https://doi.org/10.1104/pp.103.037788

Robe K, Gao F, Bonillo P, Tissot N, Gaymard F et al. (2020). Sulphur availability modulates *Arabidopsis thaliana* responses to iron deficiency. PloS One 15 (8): e0237998. https://doi.org/10.1371/journal.pone.0237998

Schenkeveld W (2010). Iron fertilization with FeEDDHA The fate and effectiveness of FeEDDHA chelates in soil-plant systems. PhD. Wageningen University and Research, Wageingen, The Netherlands.

Schuler M, Keller A, Backes C, Philippar K, Lenhof HP et al. (2011). Transcriptome analysis by GeneTrail revealed regulation of functional categories in response to alterations of iron homeostasis in *Arabidopsis thaliana*. BMC Plant Biology 11: 87. https://doi.org/10.1186/1471-2229-11-87

Schwarz B, Bauer P (2020). FIT, a regulatory hub for iron deficiency and stress signaling in roots, and FIT-dependent and-independent gene signatures. Journal of Experimental Botany 71 (5): 1694-1705. https://doi.org/10.1093/jxb/eraa012

Shoemaker RC, Schlueter J, Doyle JJ (2006). Paleopolyploidy and gene duplication in soybean and other legumes. Current Opinion in Plant Biology 9 (2): 104-109. https://doi.org/10.1016/j.pbi.2006.01.007

Shultz JL, Kurunam D, Shopinski K, Iqbal MJ, Kazi S et al. (2006). The Soybean Genome Database (SoyGD): a browser for display of duplicated, polyploid, regions and sequence tagged sites on the integrated physical and genetic maps of *Glycine max*. Nucleic Acids Research 34 (1), D758-D765. https://doi.org/10.1093/nar/gkj050

Sivitz AB, Hermand V, Curie C, Vert G (2012). Arabidopsis bHLH100 and bHLH101 control iron homeostasis via a FIT-independent pathway. PloS One 7 (9): e44843. https://doi.org/10.1371/journal.pone.0044843

Peiffer GA, King KE, Severin AJ, May GD, Cianzio SR et al. (2012). Identification of candidate genes underlying an iron efficiency quantitative trait locus in soybean. Plant Physiology 158 (4): 1745-1754. https://doi.org/10.1104/pp.111.189860

Thieme CJ, Rojas-Triana M, Stecyk E, Schudoma C et al. (2015). Endogenous Arabidopsis messenger RNAs transported to distant tissues. Nature Plants 1 (4): 1-9. https://doi.org/10.1038/nplants.2015.25

Tsai HH, Schmidt W (2020). pH-dependent transcriptional profile changes in iron-deficient Arabidopsis roots. BMC Genomics 21 (694): 1-11. https://doi.org/10.1186/s12864-020-07116-6

Tsai HH, Rodriguez-Celma J, Lan P, Wu YC, Vélez-Bermúdez IC et al. (2018). Scopoletin 8-hydroxylase-mediated fraxetin production is crucial for iron mobilization. Plant Physiology 177 (1): 194-207. https://doi.org/10.1104/pp.18.00178

Van Hoewyk D, Takahashi H, Inoue E, Hess A, Tamaoki M et al. (2008). Transcriptome analyses give insights into selenium-stress responses and selenium tolerance mechanisms in Arabidopsis. Physiologia Plantarum 132 (2): 236–253.

Wild M, Davière JM, Regnault T, Sakvarelidze-Achard L, Carrera E et al. (2016). Tissue-specific regulation of gibberellin signaling fine-tunes Arabidopsis iron-deficiency responses. Developmental Cell 37 (2): 190-200. https://doi.org/10.1016/j.devcel.2016.03.022

Xie M, Chung CYL, Li MW, Wong FL, Wang X et al. (2019). A reference-grade wild soybean genome. Nature Communications 10 (1): 1-12. https://doi.org/10.1038/s41467-019-09142-9

Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruisse W (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. Plant Physiology 136 (1): 2621-2632. https://doi.org/10.1104/pp.104.046367