Identification and Fine-Mapping of a Soybean Quantitative Trait Locus on Chromosome 5 Conferring Tolerance to Iron Deficiency Chlorosis

Ryan Merry, Karl Butenhoff, Benjamin W. Campbell, Jean-Michel Michno, Dechun Wang, James H. Orf, Aaron J. Lorenz,* and Robert M. Stupar*

R. Merry, K. Butenhoff, B.W. Campbell, J.-M. Michno, J.H. Orf, A.J. Lorenz, R.M. Stupar, Dep. of Agronomy and Plant Genetics, Univ. of Minnesota, Saint Paul, MN 55108; D. Wang, Dep. of Plant, Soil and Microbial Sciences, Michigan State Univ., East Lansing, MI 48824; K. Butenhoff, Bayer Crop Science, 704 12th Street, Kalona, IA 52247.

ABSTRACT Iron deficiency chlorosis (IDC) is an important nutrient stress for soybean [*Glycine max* (L.) Merr.] grown in high-pH soils. Despite numerous agronomic attempts to alleviate IDC, genetic tolerance remains the most effective preventative measure against symptoms. In this study, two association mapping populations and a biparental mapping population were used for genetic mapping of IDC tolerance. Quantitative trait loci (QTLs) were identified on chromosomes *Gm*03, *Gm*05, and *Gm*06. Heterogenous inbred families were developed to fine-map the *Gm*05 QTL, which was uniquely supported in all three mapping populations. Fine-mapping resulted in a QTL with an interval size of 137 kb on the end of the short arm of *Gm*05, which produced up to a 1.5-point reduction in IDC severity on a 1 to 9 scale in near isogenic lines.

Abbreviations: HIF, heterogeneous inbred families; IDC, iron deficiency chlorosis; LOD, logarithm of odds; NIL, near isogenic lines; QTL, quantitative trait locus; RIL, recombinant inbred line; SNP, single nucleotide polymorphism; URT, Uniform Regional Trials.

CORE IDEAS

- 'Fiskeby III' harbors a combination of abiotic stress traits, including iron deficiency chlorosis (IDC) tolerance.
- An IDC quantitative trait locus on chromosome *Gm*05 was identified in genome-wide association studies and biparental populations.
- Fine-mapping resolved a 137-kb interval containing strong candidate genes.

Iron deficiency chlorosis is an abiotic stress prevalent in soybean. Soils with a pH greater than 8 cause Fe to become unavailable for plant uptake. High soil pH is often accompanied in calcareous soils by bicarbonate release, which acts as a pH buffer and exacerbates IDC in soybeans because of the inability of the plant to acidify the rhizosphere (Hansen et al., 2003; Inskeep and Bloom, 1987). This lack of Fe availability reduces chlorophyll production, visually observed as interveinal chlorosis of leaves in soybean, which ultimately hinders photochemical capacity

Citation: Merry, R., K. Butenhoff, B.W. Campbell, J.-M. Michno, D. Wang, J.H. Orf, A.J. Lorenz, and R.M. Stupar. 2019. Identification and fine-mapping of a soybean quantitative trait locus on chromosome 5 conferring tolerance to iron deficiency chlorosis. Plant Genome 12:190007. doi:10.3835/plantgenome2019.01.0007

Received 23 Jan. 2019. Accepted 12 July 2019.
*Corresponding authors (lore0149@umn.edu; stup0004@umn.edu).

(Spiller and Terry, 1980). Minor symptoms can reduce grain yields by up to 20% and severe cases can result in total crop failure (Froechlich and Fehr, 1981; Niebur and Fehr, 1981). In the North Central United States, it has been estimated that soybean is grown on 1.8 million ha of high-pH soil with an accompanying estimated annual yield loss of 340,000 Mg (Hansen et al., 2003).

There are several agronomic treatments for IDC including foliar sprays, seed treatments, soil Fe chelate applications, and tolerant cultivar selection. Foliar sprays are used as a restorative measure after the onset of chlorosis. A comparison of different foliar applications indicated that although soybean chlorosis improved, no yield improvement was achieved (Chatterjee et al., 2017). Iron supplementation through seed treatments has had limited success. Although IDC-tolerant cultivars were found to respond positively to treating seeds with Fe, no yield increase was observed in susceptible cultivars (Karkosh et al., 1988). In studies comparing the effects of foliar Fe sprays, seed treatments, and Fe chelate soil application, it was concluded that cultivar selection was the most effective method for controlling IDC in soybean (Goos and Johnson, 2000; Kaiser et al., 2014).

Phenotyping of foreign soybean introductions that displayed some tolerance to IDC from the USDA soybean collection indicated that no cultivar existed that was completely tolerant to IDC (Fehr, 1982). Although there is a great deal of variation for IDC tolerance in breeding material, no completely tolerant cultivars were found during recent soybean variety trials in Minnesota (Minnesota Agricultural Experiment Station, 2017) and North Dakota (Kandel et al., 2017) when sufficient IDC pressure was present. Spatial variation and year-to-year variation in IDC severity (Inskeep and Bloom, 1987) make IDC tolerance a difficult trait to select for in breeding programs. Furthermore, little is known about the mechanism of Fe deficiency tolerance in soybean, which prevents indirect selection for trait improvement. Therefore, nearly all selection for IDC tolerance is done through visual ratings on breeding lines planted in IDC-prone nurseries (personal communication; Aaron Lorenz, Ted Helms, and George Graef; 2018).

Identifying QTLs for marker-assisted breeding for IDC tolerance would be beneficial, considering the difficulty and inefficiency of screening breeding populations on the basis of phenotype alone. Once QTLs have been identified, it would also be useful to identify the gene variants underlying the tolerance trait to illuminate the physiological and molecular mechanisms soybean uses to combat this abiotic stress. An early study on IDC tolerance indicated that it was controlled by a single major gene (Weiss, 1943); however, later studies firmly concluded that IDC tolerance is a quantitative trait controlled by many genes (Cianzio and Fehr, 1982; Mamidi et al., 2014). Relatively large-effect QTLs have been identified, including a single major-effect QTL required for soybean Fe efficiency reported on *Gm*03 (linkage group N) (Lin et al., 1997). Peiffer et al. (2012) used a backcrossing technique to develop near isogenic lines (NILs) of different recombination classes within the *Gm*03 QTL for fine-mapping. As a result, two genes encoding basic helix-loop-helix transcription factors, which were hypothesized to dimerize together, were identified as potential candidates. One of the two genes was found to contain a deletion in a predicted dimerization domain. Differences in the transcript abundance of genes induced by Fe deficiency had previously been identified (O'Rourke et al., 2009), making a transcription factor a likely candidate. This discovery led to the identification of additional transcription factors important in the Fe deficiency response (Li et al., 2018). Several minor effect QTLs for IDC tolerance have also been identified in soybean (Charlson et al., 2005; Lin et al., 1997,2000; Mamidi et al., 2011, 2014; Wang et al., 2008). To our knowledge, fine-mapping of these QTL has not been conducted, most probably because of the difficulties associated with phenotyping IDC and the lack of power to detect small differences in tolerances associated with these individual QTLs. Fine-mapping of additional IDC QTLs is an important step in developing markers for selection, as well as leading to gene discovery and furthering our understanding of IDC tolerance mechanisms in soybean.

Although no completely tolerant cultivars are present in the germplasm collection, there are a handful of cultivars with a high level of IDC tolerance. One such cultivar is Fiskeby III (PI 438471), which was identified as a potential cool climate cultivar (Holmberg, 1973) with multiple abiotic stress tolerances. Fiskeby III has presented an opportunity to map QTLs for tolerance to salt (Carter et al., 2006; Do et al., 2018), aluminum toxicity, ozone tolerance (Burton et al., 2016), and canopy wilt tolerance (Butenhoff, 2015). In this study, we used a previously developed biparental mapping population with Fiskeby III as a parent to fine-map an IDC tolerance QTL.

The objective of this study was to identify and fine-map QTLs controlling IDC tolerance in soybean. We first demonstrate that a QTL on *Gm*05 is an important contributor to genetic variation for IDC tolerance through association mapping by using plant introductions from the USDA Germplasm Resources Information Network soybean collection and elite breeding lines from Uniform Regional Trials (URTs). We then present results from a biparental linkage mapping population of Fiskeby III (PI 438471) by 'Mandarin (Ottawa)' (PI 548379), which also mapped a QTL to the same position on *Gm*05. Using this population, we created a series of NILs and narrowed the QTL interval to a 137-kb region containing 17 gene models. Results from this study form an important step towards the identification of genes controlling IDC tolerance in soybean.

METHODS

Association Mapping

Two independent panels of soybean genotypes were used for genome-wide association mapping: a panel of plant introductions (the PI Panel) and a panel of breeding lines

entered into the USDA Northern Uniform Regional Trials (the URT Panel). A total of 1737 accessions comprised the PI Panel, 1352 of which were evaluated for IDC tolerance at two locations in 2001 and 385 of which were evaluated for IDC tolerance at five locations in 2004. Phenotypic data for the PI Panel was downloaded from the Germplasm Resource Information Network website (https://npgsweb.ars-grin.gov/gringlobal/descriptordetail.aspx?id=51074, accessed 26 Aug. 2019). Briefly, the experimental design consisted of two randomized replications of the PI Panel grown under field IDC conditions, with symptoms scored on a 1 to 5 chlorosis severity rating scale, where 1 indicated no chlorosis and 5 indicated severe chlorosis (Cianzio et al., 1979). Final scores for the PI Panel were reported as adjusted means with the values of neighboring plots as covariates for adjustment.

A total of 854 advanced public breeding lines comprising the URT Panel were also screened for IDC tolerance. The IDC scores for the URT Panel were obtained from the final reports of the USDA Northern Region Uniform Soybean Tests between 1989 and 2017. Each year, breeding lines in the URT maturity groups 00 to II were planted into IDC conditions at one or two locations in a randomized complete block design with two replications. In most years, the plots consisted of single rows 0.91 m in length spaced 0.76 m apart. Experimental design parameters and the number of genotyped lines from each URT year are provided in Supplemental Table S1. The IDC scores were taken with the same phenotypic scale described for the PI Panel and reported as the mean of two scoring dates averaged across replications. Means of each breeding line were obtained from https://ars.usda.gov/mwa/lafayette/cppcru/ust (accessed 26 Aug. 2019) and compiled into a single dataset consisting of breeding line, location, year, and mean IDC score.

Genotypic data were downloaded from SoyBase (www.soybase.org, accessed 26 Aug. 2019) for the PI Panel and accessions were genotyped with the SoySNP50K iSelect SNP Beadchip (Song et al., 2013, 2015). To initiate URT line genotyping, remnant seeds of URT lines were requested from public soybean breeders who entered breeding lines in the Northern URT. Tissue was collected from 5 to 10 germinated seedlings per line and DNA was extracted via the hexadecyltrimethylammonium bromide method (Kisha et al., 1997). DNA was then genotyped with the SoySNP6K Infinium Beadchip (Illumina Inc. San Diego, CA) and the Illumina iScan platform. GenomeStudio Genotyping Module (Illumina Inc.) was used to call the single nucleotide polymorphism (SNP) alleles (Song et al., 2013). SoySNP6K genotype data were obtained for several individuals ['Burlison' (PI 533655), 'Kenwood' (PI 537094), 'Sturdy' (PI 542768), 'Hoyt' (PI 540552), 'Flyer' (PI 534646), and 'Resnik' (PI 534645)] by selecting the SoySNP6K SNP allele data from the SoySNP50K iSelect SNP Beadchip data (Song et al., 2013, 2015).

In both panels, SNPs with minor allele frequencies less than 0.025 were removed. A total of 42,294 and 5972 biallelic SNP markers for the PI Panel and the URT panel, respectively, were retained in the analysis after minor allele filtering. Single nucleotide polymorphism positions and all genomic positions within this study are in reference to the Wm82.a2.v1 genome assembly bp positions (Song et al., 2016). A genome-wide association analysis was implemented via the "GWAS" function within the *rrBLUP* package (rrBLUP version 4.6; Endelman, 2011) on the PI Panel and the URT Panel separately. In each analysis, the year, location, subpopulation structure (modeled as three principal components), and SNP effects were fitted as fixed effects into a mixed linear model. A principal component analysis implemented in the *FactoMineR* package via the *pca* function (Le et al., 2008) was performed separately on the genotypic data from the PI Panel and the URT Panel. The first three principal components accounted for 20 and 11% of the total SNP variation in the PI Panel and URT Panel, respectively. Random polygenic effects (u) were also included, and were modeled as $\sim MVN(0,\mathbf{G}\sigma_u^2)$, where $\mathbf{G}$ represents the genomic relationship matrix calculated internally by the "GWAS" function of the *rrBLUP* package. To declare statistically significant associations between SNPs and IDC scores, a genome-wide significance threshold of 5.93 and 5.07 ($-\log(p)$) was set for the PI Panel and URT Panel, respectively, with a Bonferroni correction for the number of tests (i.e., total number of SNPs tested) in each panel. This is a highly conservative form of multiple testing correction and thus any SNPs surpassing this threshold displayed very strong evidence of a true association with the observed IDC score. The reduced power resulting from a conservative threshold was not of concern because the objective of this study was to ultimately fine-map specific QTLs, not describe the full genetic architecture of IDC tolerance.

The "LD" function of the genetics package (Genetics version 1.3.8.1.2, Warnes et al., 2019) was used to calculate linkage disequilibrium between significant SNPs discovered near each other on the same chromosome. Single nucleotide polymorphisms associated with each other were considered to be a single QTL. To detect epistatic interactions among QTLs, a linear model was fitted with year, location, and the first three principal components as fixed effects and all interactions modeled between peak significant SNPs.

Biparental Mapping Population Development

The development of the biparental population used in this study has been described previously by Burton et al. (2016). Briefly, Fiskeby III (Maturity Group 000) was hybridized with the IDC susceptible cultivar Mandarin (Ottawa) (PI 548379, Maturity Group 0) in 2006 to create an $F_{5:6}$ population of 239 recombinant inbred lines (RILs). Mandarin (Ottawa), a selection from the Chinese landrace 'Mandarin', was a prominent parent in breeding crosses of early North American breeding programs. Through a pedigree analysis, Gizlice et al. (1994) estimated that Mandarin (Ottawa) contributed substantially to the U.S. cultivars released between 1947 and 1988. The

F_1 seed was harvested and planted in a winter nursery in Chile during 2006–2007. The F_2 seed was harvested and single-seed descent was used to advance the population to the F_5 generation. The F_5 seed was planted in the greenhouse at the USDA-ARS facility in Raleigh, NC, in the winter of 2010 and $F_{5:6}$ families were created by harvesting the seeds from individual F_5 plants. The $F_{5:6}$ families were then planted in the summer of 2010 in Minnesota in progeny rows. Seed was harvested and bulked by each progeny row to create $F_{5:7}$ families for use in this study.

Biparental Mapping, Design, and Phenotyping

The 239 RILs were planted in an IDC nursery near Danvers, MN in 2011, 2012, and 2013. The soil type consisted of 50% Bearden–Quam (Bearden: fine-silty, mixed, superactive, frigid Aeric Calciaquolls; Quam: fine-silty, mixed, superactive, frigid Cumulic Endoaquolls), 30% Byrne (fine-loamy, mixed, superactive, frigid Calcic Hapludolls), and 10% each of Quam and Malachy (coarse-loamy, mixed, superactive, frigid Aquic Calciudolls). All soil series shared the common characteristics of being poorly drained (which facilitates bicarbonate release) and prone to IDC symptoms. The RILs were planted into single-row plots arranged in a randomized complete block design with three replications. Plots consisted of 25 seeds planted in rows 0.91 m in length and spaced 0.76 m apart. Only two replications were planted in 2013 because of a shortage of seed. Seven check lines were included to control for environmental variation. Check lines included Fiskeby III (RIL parent, resistant), Mandarin (Ottawa) (RIL parent, susceptible), 'MN0095' (resistant check), 'Sheyenne' (resistant check), 'Dawson' (resistant check), 'Corsoy 79' (susceptible check), 'Bicentennial' (susceptible check used in 2011), 'Parker' (susceptible check used in 2012), and 'Lambert' (susceptible check used in 2013).

Iron deficiency chlorosis was scored for each replication at the third trifoliate (V3) stage and again at the sixth trifoliate (V6) stage (Fehr and Caviness, 1977) on a 1 to 9 chlorosis severity rating scale on a whole-plot basis. The following guidelines were used: 1 = healthy, green plants within the plot, with no signs of chlorosis; 2 = several chlorotic leaves visible within the plot; 3 = 10 to 20% of leaves chlorotic; 4 = 20 to 30% of leaves chlorotic; 5 = 30 to 50% leaves chlorotic; 6 = >50% of leaves chlorotic; 7 = >50% of leaves chlorotic with <50% necrosis occurring; 8 = >50% of leaves chlorotic with >50% necrosis; 9 = complete necrosis and plant death. This nine-point scale was used in this experiment for increased resolution and thus better ability to detect small differences among the NILs.

For estimation of broad-sense heritability on an entry-mean basis (Holland et al., 2010) of IDC scores, the following model was fitted:

$$y_{ijk} = \mu + g_i + e_j + r_{k(j)} + ge_{ij} + \varepsilon_{ijk} \quad [1]$$

where y_{ijk} is the IDC score of the plot, μ is the model intercept, g_i is the effect of the i^{th} RIL, e_j is the effect of the j^{th} environment, $r_{k(j)}$ is the effect of the k^{th} replicate (block) nested within the j^{th} environment, ge_{ij} is the interaction effect between the i^{th} genotype and the j^{th} environment, and ε_{ijk} is the residual. To calculate heritability, all effects were considered to be random for the estimation of variance components via restricted maximum likelihood. For purposes of QTL mapping, RIL least-squares means were calculated via the same linear model, with the exception of the RIL effect being fitted as a fixed effect.

DNA from the RIL plants were extracted with the Qiagen DNeasy Plant Mini Kits in a QiaCube workstation (Qiagen, Hilden, Germany). DNA was extracted from a sample of 15 to 20 primary root tips for each RIL. In total, 1536 SNP markers were genotyped in the RIL population via the Illumina GoldenGate assay (Hyten et al., 2010). The analysis was completed following the protocol described by Fan et al. (2006) on the Illumina BeadStation 500G (Illumina Inc.). Linkage map construction for this population has been described by Burton et al. (2016). The following criteria were used to exclude markers in Illumina's GenomeStudio software, version 2011.1: monomorphism, a call frequency less than 95%, a GenTrain score less than 0.25, cluster separation less than 0.20, and an AB_T Mean less than 0.20 or between 0.80 and 1.00. The software package R/qtl, version 1.32–10 (Broman et al., 2003) was used to evaluate the quality of the linkage map by checking for duplicate and missing markers, marker order compared with the USDA consensus map (Hyten et al., 2010), and segregation distortion. Upon completion of marker allele calls and map quality evaluation, 366 SNP markers were retained for QTL analysis. The final map length was 1811 cM, with an average marker spacing of 5.2 cM.

Composite interval mapping in the R/qtl statistical package (Broman et al., 2003) was used for QTL mapping. Composite interval mapping was performed with five marker covariates, a window size of 10 cM, and a walking speed of 2 cM. Logarithm of odds (LOD) thresholds for declaring QTL significance were set to ensure an experiment-wise error rate of 0.05. The thresholds for the 1, 5, and 10% significance levels were 4.57, 3.77, and 3.46, respectively. These thresholds were determined from 1000 permutation tests (Doerge and Churchill, 1996). Haley–Knott regression, implemented in R/qtl, was used to estimate the amount of phenotypic variation that significant QTLs accounted for, as well as marker effect estimates on the 1 to 9 rating scale.

Fine-Mapping of the *Gm*05 QTL

Twenty $F_{5:6}$ families from the Fiskeby III × Mandarin (Ottawa) population were identified as being heterogeneous for a significant QTL found on *Gm*05. Individuals from each RIL family were genotyped with a custom panel of 19 SNP assays (Supplemental Table S2, designed from SNPs from the SoySNP50K iSelect SNP Beadchip) on a MassARRAY platform (Sequenom, Agena Biosciences, San Diego, CA) in the winter of 2014–2015 to identify heterozygous plants and develop heterogeneous inbred families (HIFs) (Tuinstra et al., 1997). Six $F_{5:6}$ families were found to still have heterozygous individuals. Heterozygous plants were identified and advanced for three

generations to the F_9 generation. The $F_{9:10}$ NILs were developed for each HIF family from single F_9 heterozygous plants. The $F_{10:11}$ seed was harvested from NILs and used for phenotyping (Fig. 1). To further resolve the QTLs, an additional panel of 27 SNP assays (Supplemental Table S2, designed with SNPs found in whole-genome resequencing of Fiskeby III and Mandarin (Ottawa)) were used to find a critical recombination point in HIF 609.

Phenotyping of the fine-mapping population was done at the Danvers, MN, site in 2016 and 2017. The $F_{10:11}$ NILs in 2016 and $F_{11:12}$ NILs in 2017 from each HIF were planted in an augmented block design with five NIL pairs in each incomplete block alongside two checks: Fiskeby III as a tolerant check and Mandarin (Ottawa) as a susceptible check. These were planted in the center of each incomplete block. Each pair of NILs appeared in the same incomplete block together but pairs were randomized across the incomplete blocks. Three replications of each NIL pair were planted and evaluated in complete blocks, so that incomplete blocks were nested inside complete blocks. Plot dimensions were identical to those used for the initial QTL mapping. Ratings were conducted on the same 1 to 9 scale used for initial QTL mapping. The statistical significance of SNP effects was determined via the following linear model:

$$y_{ijklm} = \mu + z_i m + e_j + r_{k(j)} + Bx_l + h_m + \varepsilon_{ijklm} \quad [2]$$

where y_{ijklm} is the IDC score of the plot, μ is the intercept, z_i is an indicator variable indicating whether line i carries the Mandarin (Ottawa) allele or the Fiskeby III allele at the SNP locus being tested, m is the SNP effect, e_j is the effect of the j^{th} year, $r_{k(j)}$ is the effect of the k^{th} complete block nested within the j^{th} year, B is an incomplete block effect modeled as a quantitative covariate using the mean IDC score of the checks in each incomplete block (x_l); h_m is the effect of the m^{th} heterogeneous inbred family; and ε_{ijklm} is the residual. The significance of the SNP effect was calculated and used to fine-map the position of the QTLs. To estimate the effects of individual NILs, the model described above for testing SNP effects was used, with the exception of $z_i m$ being replaced by g_i, which represents the effect of the i^{th} NIL. The NIL effects were used to calculate the NIL-adjusted means. Significant differences ($P < 0.05$) between paired NILs were determined from the LSD obtained from the "LSD.test" function of the "agricolae" package in R (De Mendiburu Delgado, 2009).

Genomic Coordinates

All genomic coordinates used in this study are in reference to the Wm82.a2.v1 genome assembly bp positions (Song et al., 2016).

RESULTS

Genome-Wide Association Mapping of IDC Tolerance

Phenotypic variation for IDC was found in all years and locations for the PI Panel and URT Panels. In total, 1336 lines were included in the PI Panel. Iron deficiency chlorosis scores over the 2-yr study period in the PI Panel ranged from 1 to 5, with an average score of 3.07. The URT Panel included 854 lines over the course of the 19-yr study, with IDC scores ranging from 1 to 5 and an average score of 3.11.

A total of 42,294 and 5972 biallelic SNP markers were retained in the analysis after minor allele filtering in the PI Panel and URT Panel, respectively, and used in genome-wide association studies. In the PI Panel, four significant SNPs were found in association with each other ($p < 0.001$) in a 369-kb region on *Gm*03, three significant SNPs were found to be in association with each other ($p < 0.001$) in a 1.1-Mb region on *Gm*05, and seven significant SNPs were found to be in association with each other ($p < 0.001$) in a 24-kb region on *Gm*06 (Fig. 1a). These findings indicated three distinct QTLs. A significant epistatic interaction was found between the *Gm*03 and *Gm*06 QTL ($p < 0.001$). Two significant SNPs were found in the URT Panel at the genome-wide significance level ($-\log(p) > 5.07$, Fig. 2b), both on *Gm*05 and in high association with each other ($p < 0.001$). This *Gm*05 QTL in the URT Panel colocalized within the *Gm*05 region discovered in the PI Panel. Marker bp positions and $-\log(p)$ scores are summarized in Table 1. Of the SNPs associated with the QTLs discovered in association mapping in the PI Panel on chromosomes *Gm*03 and *Gm*06, only one SNP from each QTL was present on the SoySNP6K Infinium Beadchip used in the URT Panel. The minor alleles of these SNPs had frequencies of 0.03 (ss715585486) and 0.029 (ss715595697) in the URT population. Though these minor allele frequencies are above the minimal threshold for inclusion in the analysis, their low frequencies greatly reduced detection power in the URT Panel despite these QTLs having strong effects.

Recombinant Inbred Line Analyses for IDC Phenotypic Variation and QTL Mapping

Genetic mapping of IDC symptoms was performed for a RIL population derived from Fiskeby III and Mandarin (Ottawa) via a 1 to 9 chlorosis severity rating scale. As expected, Fiskeby III showed greater tolerance to IDC than Mandarin (Ottawa) in all 3 yr of evaluation. The average IDC scores for Fiskeby III and Mandarin (Ottawa) over the 3-yr study period were 3.25 and 5.19, respectively. A high amount of phenotypic variation was observed in the RIL population, which is typical for IDC. The mean IDC score over the 3-yr period for the RILs was 3.98 but ranged from 1.29 to 7.15 following a normal distribution. If we compare the years, 2012, a year with record heat and drought, had the lowest average IDC score and a maximum IDC score of 4.97 averaged across repetitions, whereas 2013, which had a cool and extremely wet spring, had a much higher average and maximum IDC score and a minimum score nearly equal to the average score in 2012 (4.01) (Table 2). Significant differences in IDC score were found for RIL genotypes ($p < 0.001$) as well as years ($p < 0.001$). The heritability of IDC tolerance on an entry-mean basis averaged over the

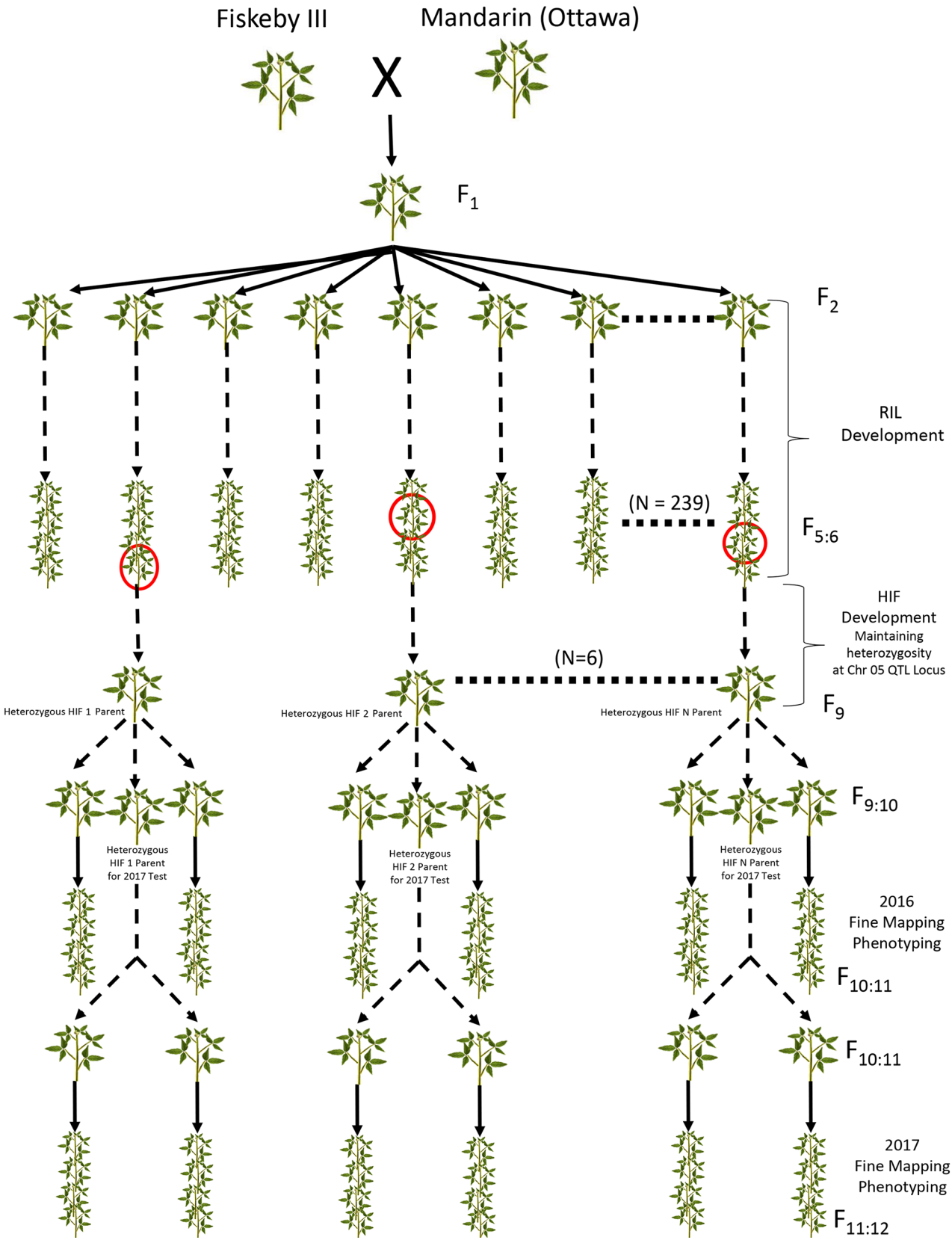

Fig. 1. Population development for the biparental mapping, near isogenic line (NIL) development, and fine-mapping populations of soybean. Here, 239 recombinant inbred lines (RILs) were generated from the initial hybridization of Fiskeby III and Mandarin (Ottawa). After quantitative trait locus (QTL) mapping in the $F_{5:7}$ generation, six RILs with residual heterozygosity in individuals were found and used to generate heterogeneous inbred families (HIFs), depicted by red circles. Heterozygous plants within HIFs were advanced to the F_9 generation. Homozygous NIL pairs were identified in the $F_{9:10}$ generation, and $F_{10:11}$ seed was harvested for planting in the 2016 iron deficiency chlorosis (IDC) nursery. Further inbreeding was conducted from $F_{9:10}$ heterozygous individuals to develop $F_{11:12}$ NIL pairs for phenotyping in the 2017 IDC nursery.

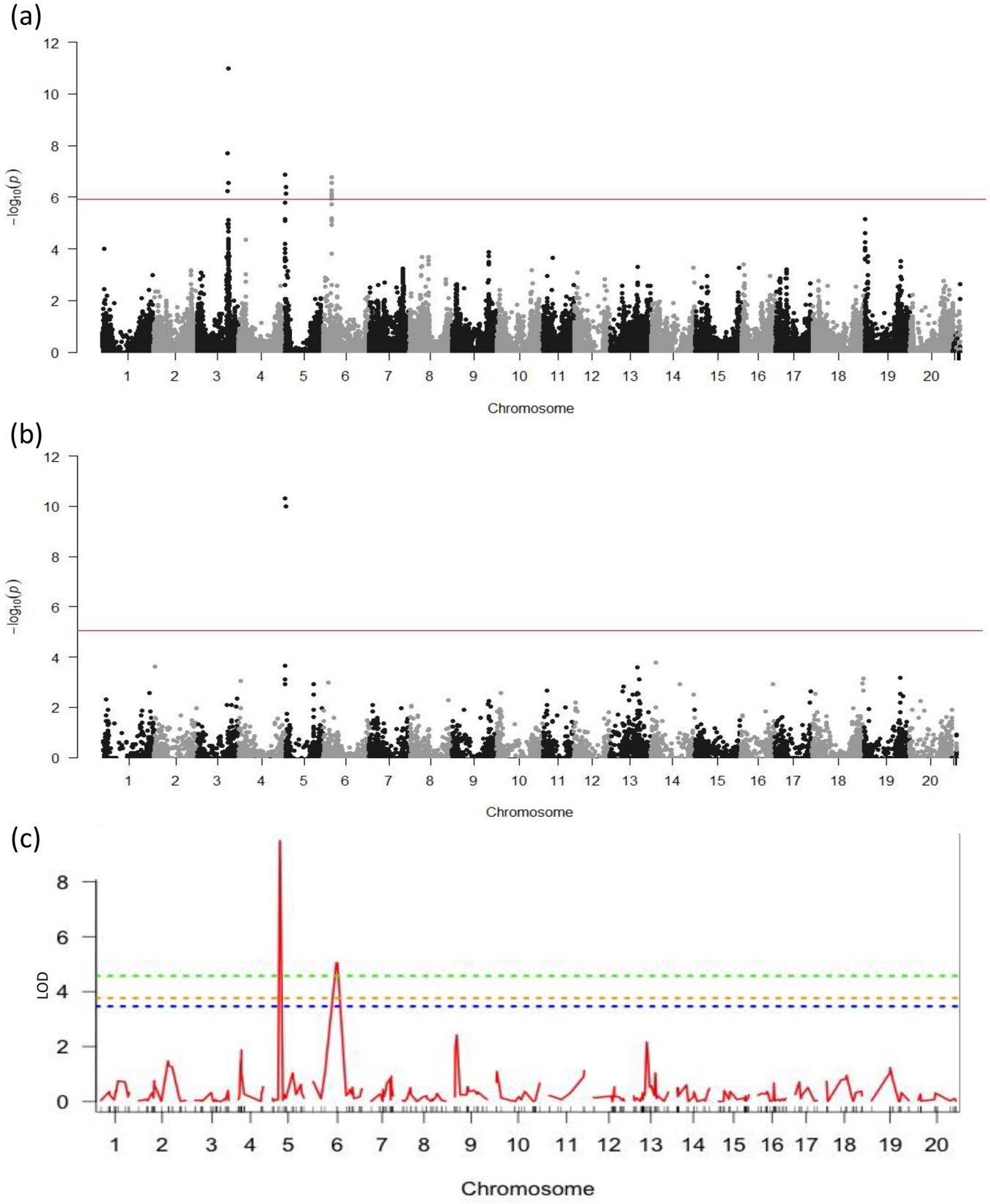

Fig. 2. Single nucleotide polymorphism (SNP) significance from genome-wide association studies of the PI Panel and Uniform Regional Trials (URT) panels and quantitative trait locus (QTL) mapping in the biparental population of soybean. (a, b) Chromosomes 1 to 20 are ordered on the *x*-axis; −log(*p*) is shown on the *y*-axis. Dots represent individual SNPs. The red line indicates the significance thresholds of −log(*p*) = 5.93 and 5.07 for the PI Panel (a) and URT Panel (b), respectively. (c) Marker logarithm of odds (LOD) scores across the genome for the Fiskeby III × Mandarin (Ottawa) QTL mapping population. The green dashed line represents the 0.01 significance threshold (LOD = 4.57), the yellow dashed line represents the 0.05 significance threshold (LOD = 3.77), and the blue dashed line represents the 0.10 significance threshold (LOD = 3.46).

Table 1. Significant single nucleotide polymorphisms (SNPs) discovered in association mapping of the PI Panel and Uniform Regional Trials (URT) panels of soybean, showing SNP ID, chromosomes, bp position, and –log(p) value for significant SNPs discovered in association mapping in the PI and URT panels.

SNP ID†	Chromosome	Position	–Log(p)‡	
			PI Panel	URT Panel
ss715585469	3	34,532,040	10.985	–
ss715585427	3	34,245,545	7.691	–
ss715585486	3	34,612,476	6.554	–
ss715585425	3	34,243,241	6.247	–
ss715592610	5	201,448	6.880	10.306
ss715592592	5	308,102	–	10.000
ss715592489	5	1,323,486	6.389	–
ss715592488	5	1,325,450	6.125	–
ss715595695	6	9,762,865	6.766	–
ss715595694	6	9,747,938	6.558	–
ss715595697	6	9,767,612	6.268	–
ss715595696	6	9,765,765	6.141	–
ss715595698	6	9,767,799	6.055	–
ss715595700	6	9,772,068	6.024	–
ss715595699	6	9,768,760	5.960	–

† The SNPs, chromosomes, and bp positions were aligned with the Wm82.a2.v1 genome assembly.

‡ The SNPs were considered significant when –log(p) was greater than 5.93 in the PI Panel and 5.07 in the URT Panel.

3-yr period was 0.61, which is similar to the heritabilities for IDC reported in previous studies (Lin et al., 1997, 2000; Charlson et al., 2005). Variation caused by year had the highest impact (54%), whereas genotypes and genotypes nested within years accounted for 9.8 and 9.6%, respectively. Variance caused by repetition and repetitions nested within year was negligible and error variance accounted for 26.4% of the total variation.

Two QTLs for IDC tolerance were detected, one on *Gm*05 and another on *Gm*06 (Fig. 2c). The *Gm*05 QTL was found to have a peak LOD score of 8.5 at 0 cM and spanned 0 to 2 cM on a 1.5-LOD support interval. The QTL on *Gm*05 displayed an additive effect of 0.25 and explained 16.2% of the phenotypic variation for IDC in this study. The *Gm*06 QTL had a peak LOD at 67 cM, spanned 57 to 81 cM based on a 1.5-LOD support interval, and had a peak LOD score of 5.5 (Fig. 2c). The *Gm*06 QTL had an additive effect of 0.16 and explained 7.7% of the phenotypic variation for IDC in this study (Table 3). The Fiskeby III alleles contributed the IDC tolerance in both cases (Table 3).

Table 2. Descriptive statistics of iron deficiency chlorosis (IDC) scores in soybean for quantitative trait locus mapping of the Fiskeby III × Mandarin (Ottawa) biparental population.

Year	Mean score	SD	Median score†	Min. score†	Max. score†	Skewness	Excess kurtosis
2011	3.88	0.66	3.84	2.30	5.78	0.27	–0.08
2012	2.95	0.77	2.88	1.29	4.97	0.30	–0.58
2013	5.69	0.66	5.67	4.01	7.15	0.02	–0.72
All years	3.98	1.28	3.84	1.29	7.15	0.35	–0.62

† This score is the recombinant inbred line average of three repetitions.

Fine-Mapping of the *Gm*05 QTL

The QTL discovered on *Gm*05 in both the association mapping and biparental mapping experiments was fine-mapped to provide higher resolution of the causative locus. Six HIFs were generated from RILs in the biparental mapping population and were further inbred. In each inbreeding generation, heterozygosity was maintained within the QTL region to allow for new recombination within the locus while increasing the homogeneity of the genetic background within each HIF. Near isogenic line pairs generated from HIFs were analyzed for differences in IDC tolerance in Danvers, MN in 2016 and 2017 via the 1 to 9 rating scale.

The $F_{10:11}$ NIL pairs generated from the $F_{9:10}$ HIFs showed significant variation for IDC tolerance in the 2016 field season ($p < 0.001$) (Fig. 3). A single NIL pair (722) showed a statistically significant difference in IDC score, which differentiated for a 450-kb region on *Gm*05. Other NIL pairs that differentiated at this locus showed similar results, with the Fiskeby III haplotype exhibiting lower mean IDC scores, although these differences were not statistically significant (Fig. 3). The average IDC score reduction of NILs contrasting in the 450-kb region in 2016 was 0.55 points on a 1 to 9 scale, similar to the homozygous effect of the QTL estimated by biparental mapping, and ranged from reductions of 0.49 to 0.82 points. Moreover, two NIL pairs (642, 693) that were monomorphic for this region had very similar IDC scores (Fig. 3). These results indicated that the QTL region was narrowed to a 450-kb region on the end of the short arm of *Gm*05 (Fig. 4b).

In the 2017 field season, five $F_{11:12}$ NIL pairs were phenotyped in Danvers, MN, to further resolve the 450-kb region (Fig. 3). Significant differences were again found between NILs ($p = 0.014$). Significant differences within pairs were only found for HIF 609, which differentiated a ~160-kb region; the Fiskeby III haplotype at this region exhibited superior IDC tolerance. Near

Table 3. Summary of results for composite interval mapping of iron deficiency chlorosis (IDC) based on least-square mean IDC scores. For the quantitative trait locus (QTL) on chromosomes 5 and 6, the associated peak marker, logarithm of odds (LOD) position, LOD score, 1.5-LOD support interval, percentage of phenotypic variation explained by the QTL, and additive effect of a single copy of the Fiskeby III allele for each QTL are identified.

Chromosome	Peak marker	Marker position†	LOD peak position	LOD score	1.5-LOD support interval	% Variation explained	Additive effect
		bp	cM		cM		
5	BARC–044481–08709	21,750	0	8.5	0–2	16.2	–0.25
6	BARC–014557–01578	11,851,152	67	5.5	57–81	7.7	–0.16

† Marker position coordinates are based on the Wm82.a2.v1 genome assembly.

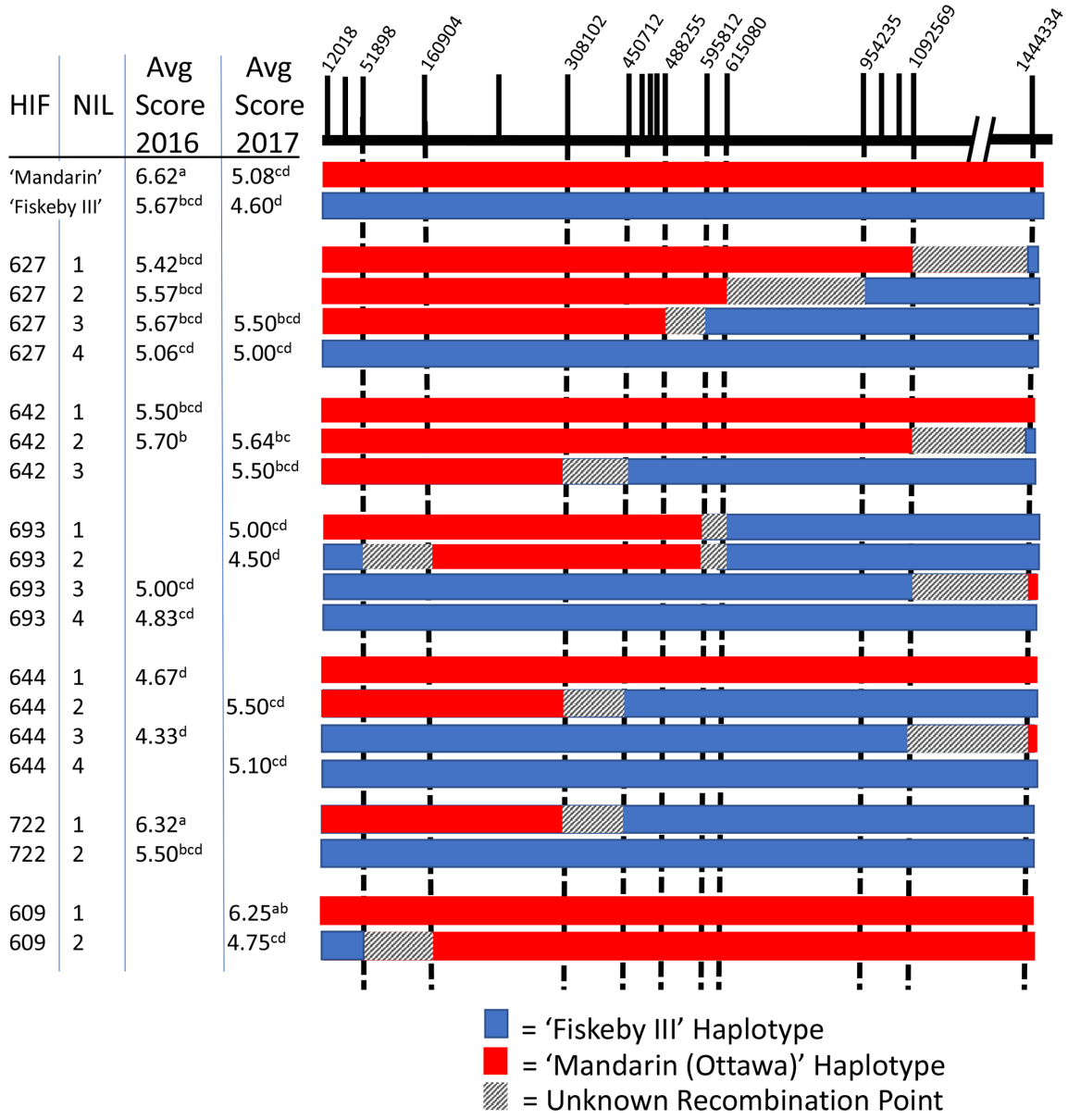

Fig. 3. Near isogenic line (NIL) pair haplotype classes generated in heterogeneous inbred families (HIFs) of soybean. The family identification number is indicated in the HIF column and the NIL identification number within each family is indicated in the NIL column. The iron deficiency chlorosis (IDC) score was averaged for each NIL in 2016 ($F_{10:11}$, Avg Score 2016) and 2017 ($F_{11:12}$, Avg Score 2017). Average IDC scores with different letters indicate significant differences between NIL pairs by LSD. Custom Sequenom MassARRAY single nucleotide polymorphism markers and bp positions of markers flanking crucial recombination points in the *Gm05* quantitative trait locus region are indicated. Positions are in reference to the Wm82.a2.v1 genome assembly.

isogenic lines from HIFs 627, 644, and 693 also showed improved IDC scores when the Fiskeby III haplotype was present within the 160-kb region, although this was not statistically significant. Heterogeneous inbred family 642 showed only minor differences in IDC score and was homogeneous for the region. An additional 27 SNP markers developed from the whole-genome resequencing data from Fiskeby III and Mandarin (Ottawa) in the 160-kb region showed that HIF 609 had a recombination point between bp positions 133,185 and 137,207 on *Gm*05, roughly narrowing the QTL to 137 kb in size (Fig. 4c and 4d). The average score reduction of NILs with the Fiskeby III haplotype in this region was 0.73 points on a 1 to 9 scale. Heterogeneous inbred family 609, with the smallest introgression region, showed an IDC score reduction of 1.5 points with the Fiskeby III haplotype.

A combined analysis was performed on the 2016 and 2017 populations to validate significance between NIL pairs. Markers from bp positions 12,018 to 308,102 showed significant differences ($p < 0.01$) for IDC score between Fiskeby III and Mandarin (Ottawa) alleles (Table 4), indicating that the causative gene is near the end of the chromosome. In all cases, the Fiskeby III marker alleles contributed IDC tolerance, with an average score reduction of 0.61 points on a 1 to 9 scale.

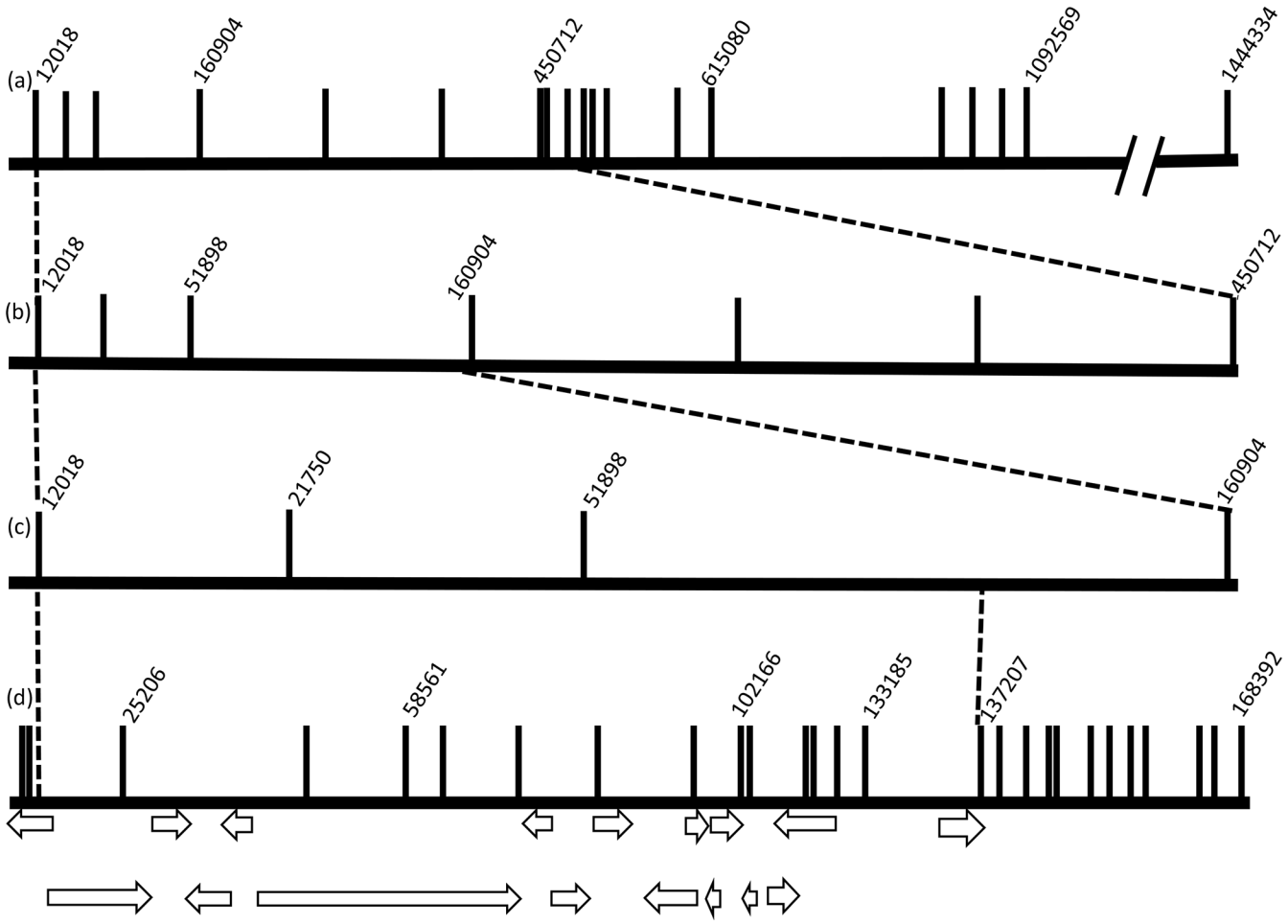

Fig. 4. Fine-mapping of the *Gm*05 quantitative trait locus (QTL) region through heterogeneous inbred families (HIFs) of soybean. (a) Custom Sequenom MassARRAY single nucleotide polymorphism (SNP) markers in the *Gm*05 QTL region. All numbers indicate bp position in reference to the Wm82.a2.v1 assembly. (b) Region fine-mapped in 2016. (c) Region fine-mapped in 2017. (d) Region fine-mapped to HIF 609 with an additional 27 SNPs. A critical recombination point was discovered between SNPs at bp positions 133,185 and 137,207. The final fine-mapped interval resides between bp positions 0 and 137,207. Arrows indicate gene models in the fine-mapped region.

DISCUSSION

Resolving the IDC QTL Region on *Gm*05

A QTL for soybean IDC tolerance was discovered on *Gm*05 through association mapping for both the PI Panel and URT Panel, as well as biparental mapping. In the association mapping, significant SNPs were found between bp positions 201,448 to 1,325,450 on *Gm*05, with the peak marker at bp position 201448 in both populations. In the biparental mapping, the most significant marker genome-wide was found on *Gm*05 at bp position 21750, which overlapped with the association mapping interval. Results from biparental and association mapping indicate that the QTL is located on the end of the chromosome, which coincides with another association mapping study that also used elite breeding lines (Mamidi et al., 2014). The *Gm*05 QTL is unique in that it is strongly supported by identification in all populations in this study. Unlike the QTLs on *Gm*03 and *Gm*06 in the PI panel, the *Gm*05 QTL is an important source of genetic variation among elite breeding lines, as it was the only significant QTL discovered in the URT Panel. Because Mandarin (Ottawa), the susceptible parent in the biparental population, is prominent in the ancestry of elite lines released from 1947 to 1988 (Gizlice et al., 1994), many modern elite cultivars may have inherited the susceptible allele from Mandarin (Ottawa) at this locus.

Fine-mapping of the *Gm*05 QTL through HIFs allowed for higher resolution of the QTL interval. In the 2016–2017 combined HIF analysis, significant markers were found from the beginning of *Gm*05 to bp position 308,102, again indicating that the causative gene is located on the end of the short arm of *Gm*05. Direct NIL comparisons showed results similar to those of Tuinstra et al. (1997) in that NILs with the tolerant haplotype showed an improvement; however, the differences were not always statistically significant from their paired NIL with the susceptible haplotype. This is not surprising, considering the difficulty of accurately scoring IDC and the strong influence of genotype × environment effects on this trait. Despite this, at least one NIL pair in each year had statistically significant differences in IDC score. Because significance was found in the combined analysis, it is possible to

further narrow the region by using individual HIFs that showed significant differences in IDC score between NILs. Near isogenic lines from HIF 609 showed significant differences in IDC score and differed between Fiskeby III and Mandarin (Ottawa) haplotypes in a ~137-kb region on the end of *Gm*05, representing the smallest resolved region that may contain the causative gene. The region fine-mapped to HIF 609 contains 17 protein coding gene models in the Wm82.a2.v1 soybean genome annotation (Fig. 4). It is not clear whether this QTL is sufficiently valuable to be used in marker-assisted selection programs. However, further cloning of the IDC tolerance gene in this region will enhance our understanding of this trait.

Understanding the causative gene will allow for increased knowledge of the physiological mechanisms of tolerance, as was proposed for previous studies of the IDC tolerance gene on *Gm*03 (Peiffer et al., 2012). Furthermore, identifying the causative gene will provide a basis for the development of novel strategies for crop improvement, including mining for novel variation, or creating novel variation through gene editing technologies and transgenesis (Jaganathan et al., 2018). There are several candidate genes within the fine-mapped *Gm*05 locus that have annotations suggesting potential functions in IDC tolerance, such as Glyma.05g001400, which is a FxxhVQxhTG (a VQ-containing motif) domain, known to regulate abiotic stress responses (Jing and Lin, 2015). The VQ motif proteins interact with WRKY transcription factors, which have been shown to regulate phosphate acquisition in *Arabidopsis thaliana* (L.) Heynh. (Devaiah et al., 2007), as well as Fe homeostasis in rice (*Oryza sativa* L. ssp. *indica*) (Ricachenevsky et al., 2010). The region also contains a Multidrug and Toxin Efflux (MATE) family protein gene (Glyma.05g001700). Multidrug and Toxin Efflux proteins *Gm*FRD3a (Glyma.15g274600) and *Gm*FRD3b (Glyma.09g102800) have been shown to be involved in Fe transport in the xylem of soybean, and increased expression of these genes improved Fe efficiency (Rogers et al., 2009). Although Glyma.05g001700 shares only slight protein homology with *Gm*FRD3a (E-value = 0.32, 24% identity) and *Gm*FRD3b (E-value = 0.75, 33% identity) over the complete protein length, it is possible that it plays a role in Fe transport within the roots, where it has been shown to be expressed (Severin et al., 2010). Lastly, it is possible that the source of the IDC tolerance, Fiskeby III, may contain a causative gene or genes in the mapped interval that is not present in the soybean reference cultivar Williams 82 genome (Schmutz et al., 2010). Further fine-mapping, genome analysis, and gene expression comparisons between NILs in the 609 family may help elucidate the gene and molecular mechanism(s) responsible for the Fiskeby III *Gm*05 IDC tolerance QTL.

Additional IDC QTLs Identified in This Study

The QTL on *Gm*05 was found in all mapping populations and methods in this study. However, two other IDC QTLs were identified within different subsets of the analysis. Through association mapping, a QTL on *Gm*03 was discovered in the PI Panel. This QTL overlaps with the known locus required for Fe efficiency in soybean (Peiffer et al., 2012), which can explain 70% of phenotypic variation for IDC tolerance in some populations (Lin et al., 1997). The NILs Clark (PI 548533) and IsoClark (PI 547430) have been used to characterize the genetic mechanism of the locus (O'Rourke et al., 2009). The Fe-inefficient IsoClark does not show a response to Fe deficiency stress, whereas the Fe-efficient Clark activates a suite of pathways to overcome the deficiency. Because none of the differentially expressed genes reside within the *Gm*03 locus, Peiffer et al. (2012) hypothesized the tolerance gene to be a transcription factor or regulator. The epistatic interaction with this QTL and the *Gm*06 QTL may be explained by this hypothesis. Although this QTL appears important in unadapted germplasm, it should be noted that no significance was found in this region in the URT Panel. The low minor allele frequency of marker ss715585486 within the QTL in the URT Panel suggests that the Fe-inefficient allele is present in low frequency among elite breeding lines.

Another QTL was identified on *Gm*06 in association mapping in the PI Panel and biparental mapping; however, it was not identified in the URT Panel. In the PI Panel, the peak significant marker (ss715595695) was at bp position 9,762,865 on *Gm*06; the peak marker in

Table 4. Combined marker analysis of near isogenic lines (NIL) in 2016–2017. In each year, NILs were pooled and genotype data from each NIL were used to calculate the iron deficiency chlorosis (IDC) score average associated with the alleles., marker IDC score averages for 'Mandarin (Ottawa)' and 'Fiskeby III' alleles, and significance of each marker are reported below.

SNP ID†	Position†	'Mandarin (Ottawa)' average	'Fiskeby III' average	Significance
ss715592634	12,018	5.784	5.182	****
ss715592632	21,750	5.784	5.182	***
–	51,898	5.784	5.182	***
ss715592615	160,904	5.781	5.146	***
ss715592597	292,492	5.781	5.146	***
ss715592592	308,102	5.766	5.181	***
ss715592569	450,712	5.611	5.494	NS‡
ss715592558	469,490	5.579	5.515	NS
ss715592553	484,990	5.579	5.515	NS
ss715592552	487,599	5.611	5.494	NS
ss715592551	488,255	5.611	5.494	NS
ss715592539	595,812	5.621	5.491	NS
ss715592538	615,080	5.588	5.512	NS
ss715592510	954,235	5.500	5.315	NS
ss715592509	954,941	5.516	5.548	NS
ss715592507	984,321	5.516	5.548	NS
ss715592503	1,049,939	5.500	5.315	NS
ss715592500	1,092,569	5.513	5.584	NS
ss715592479	1,444,334	5.368	5.581	NS

*** Significant at the 0.001 probability level.

**** Significant at the 0.0001 probability level.

† Marker ID (SSID) and marker bp positions are derived from the Wm82.a2.v1 genome assembly.

‡ NS, not significant.

biparental mapping (BARC-014557–01578) was at bp position 11,851,152. Despite the peak positions being roughly 2 Mb apart, this is likely to be a single QTL on *Gm*06 if we consider that the large support interval in the biparental mapping population was caused by low marker coverage in the region. Similar to the *Gm*03 QTL found in this study, the minor allele frequency of the *Gm*06 QTL in the URT Panel was low (0.029), which is most probably why it was not detected in that population. Interestingly, the Fe-inefficient allele appeared to be nearly fixed in the URT Panel. This QTL shows a very small effect (–0.16 of a point per allele copy), which limits its value in a breeding program and potentially explains why it has perhaps not been under heavy selection.

CONCLUSION

An IDC tolerance QTL on *Gm*05 was found to be a significant source of variation in both the PI Panel and elite breeding material. Two additional QTLs identified on chromosomes 3 and 6 in the PI Panel were also identified but were not found to be a significant source of variation in elite material. A tolerance allele from Fiskeby III was independently mapped to the same locus on *Gm*05 in a biparental mapping population and was then fine-mapped to a 137-kb region containing 17 putative gene models. The results from this study provide an important starting point for identifying the causal polymorphism within this QTL, which will be valuable for understanding the molecular mechanism underlying the variation in IDC tolerance among elite soybean varieties.

Supplemental Information

Supplemental Table 1. Uniform Regional Trial yearly test information from 1989 to 2017.

Supplemental Table 2. Primers used in custom MassARRAY genotyping for fine mapping of the *Gm*05 IDC QTL

Conflict of Interest Disclosure

The authors declare that there is no conflict of interest.

ACKNOWLEDGMENTS

The authors are grateful to Adrian Stec for laboratory assistance. This work was supported in part by a United Soybean Board fellowship to R.M. and the United Soybean Board sponsored project #1920-172-0116-C.

REFERENCES

Broman, K.W., H. Wu, Ś. Sen, and G.A. Churchill. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889–890. doi:10.1093/bioinformatics/btg112

Burton, A.L., K.O. Burkey, T.E. Carter, J. Orf, and P.B. Cregan. 2016. Phenotypic variation and identification of quantitative trait loci for ozone tolerance in a Fiskeby III × Mandarin (Ottawa) soybean population. Theor. Appl. Genet. 129:1113–1125. doi:10.1007/s00122-016-2687-1

Butenhoff, K.J. 2015. QTL mapping and GWAS identify sources of iron deficiency chlorosis and canopy wilt tolerance in the Fiskeby III × Mandarin (Ottawa) soybean population. MSc diss, Univ. of Minnesota, Minneapolis-St. Paul.

Carter, T.E., H.R. Boerma, G.J. Lee, X. Zhou, M.R. Villagarcia, A. Cardinal, et al. 2006. On-farm QTL mapping of salt tolerance in the genetic base of North American soybean. 11th Biennial Conference on the Molecular and Cellular Biology of the Soybean, Lincoln, Nebraska. 5–8 Aug. 2006.

Charlson, D.V., T.B. Bailey, S.R. Cianzio, and R.C. Shoemaker. 2005. Molecular marker Satt481 is associated with iron-deficiency chlorosis resistance in a soybean breeding population. Crop Sci. 45:2394–2399. doi:10.2135/cropsci2004.0510

Chatterjee, A., S. Lovas, H. Rasmussena, and R.J. Goos. 2017. Foliar application of iron fertilizers to control iron deficiency chlorosis of soybean. Crop Forage Turfgrass Manag. 3:2017-05-0037. doi:10.2134/cftm2017.05.0037

Cianzio, S.R., and W.R. Fehr. 1982. Variation in the inheritance of resistance to iron deficiency chlorosis in soybeans. Crop Sci. 22:433–434. doi:10.2135/cropsci1982.0011183X002200020055x

Cianzio, S.R., W.R. Fehr, and I.C. Anderson. 1979. Genotypic evaluation for iron deficiency chlorosis in soybeans by visual scores and chlorophyll concentration. Crop Sci. 19:644–646. doi:10.2135/cropsci1979.0011183X001900050024x

De Mendiburu Delgado, F. 2009. Una herramienta de análisis estadístico para la investigación agrícola. MSc. dissertation, Univ. Nacional De Ingeniería, Lima, Peru.

Devaiah, B.N., A.S. Karthikeyan, and K.G. Raghothama. 2007. WRKY75 transcription factor is a modulator of phosphate acquisition and root development in *Arabidopsis*. Plant Physiol. 143:1789–1801. doi:10.1104/pp.106.093971

Do, T.D., T.D. Vuong, D. Dunn, S. Smothers, G. Patil, D.C. Yungbluth, et al. 2018. Mapping and confirmation of loci for slat tolerance in a novel soybean germplasm, Fiskeby III. Theor. Appl. Genet. 131:513–524. doi:10.1007/s00122-017-3015-0

Doerge, R.W., and G.A. Churchill. 1996. Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285–294.

Endelman, J.B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4:250–255. doi:10.3835/plantgenome2011.08.0024

Fan, J.B., K.L. Gunderson, M. Bibikova, J.M. Yeakley, J. Chen, E. Wickham Garcia, et al. 2006. Illumina universal bead arrays. Methods Enzymol. 410:57–73. doi:10.1016/S0076-6879(06)10003-8

Fehr, W.R. 1982. Control of iron-deficiency chlorosis in soybeans by plant breeding. J. Plant Nutr. 5:611–621. doi:10.1080/01904168209362989

Fehr, W.R., and C.E. Caviness. 1977. Stages of soybean development. Special Report 87. Iowa State Univ. http://lib.dr.iastate.edu/specialreports/87 (accessed 26 Aug. 2019).

Froechlich, D.M., and W.R. Fehr. 1981. Agronomic performance of soybeans with differing levels of iron deficiency chlorosis on calcareous soil. Crop Sci. 21:438–441. doi:10.2135/cropsci1981.0011183X002100030021x

Gizlice, Z., T.E. Carter, Jnr., and J.W. Burton. 1994. Genetic base for North American public soybean cultivars released between 1947 and 1988. Crop Sci. 34:1143–1151. doi:10.2135/cropsci1994.0011183X003400050001x

Goos, R.J., and B.E. Johnson. 2000. A comparison of three methods for reducing iron-deficiency chlorosis in soybean. Agron. J. 92:1135–1139. doi:10.2134/agronj2000.9261135x

Hansen, N.C., M.A. Schmitt, J.E. Andersen, and J.S. Strock. 2003. Iron deficiency of soybean in the Upper Midwest and associated soil properties. Agron. J. 95:1595–1601. doi:10.2134/agronj2003.1595

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martínez. 2010. Estimating and interpreting heritability for plant breeding: An update. Plant Breed. Rev. 22:9–112. doi:10.1002/9780470650202.ch2

Holmberg, S.A. 1973. Soybeans for cool temperate climates. Agric. Hort. Genet. 31:1–20.

Hyten, D.L., I.-Y. Choi, Q. Song, J.E. Specht, T.E. Carter, R.C. Shoemaker, et al. 2010. A high density integrated genetic linkage map of soybean and the development of a 1536 universal soy linkage panel for quantitative trait locus mapping. Crop Sci. 50:960–968. doi:10.2135/cropsci2009.06.0360

Inskeep, W.P., and P.R. Bloom. 1987. Soil chemical factors associated with soybean chlorosis in Calciaquolls of Western Minnesota. Agron. J. 79:779–786. doi:10.2134/agronj1987.00021962007900050005x

Jaganathan, D., K. Ramasamy, G. Sellamuthu, S. Jayabalan, and G. Venkataraman. 2018. CRISPR for crop improvement: An update review. Front. Plant Sci. 9:985. doi:10.3389/fpls.2018.00985

Jing, Y., and R. Lin. 2015. The VQ motif-containing protein family of plant-specific transcriptional regulators. Plant Physiol. 169:371–378. doi:10.1104/pp.15.00788

Kaiser, D.E., J.A. Lamb, P.R. Bloom, and J.A. Hernandez. 2014. Comparison of field management strategies for preventing iron deficiency chlorosis in soybean. Agron. J. 106:1963–1974. doi:10.2134/agronj13.0296

Kandel, H., T. Helms, S. Markell, M. Ostlie, B. Schatz, G. Endres, et al. 2017. North Dakota Soybean Variety Trial Results for 2017 and Selection Guide. North Dakota State Univ. www.ag.ndsu.edu/publications/crops/north-dakota-soybean-variety-trial-results-for-2017-and-selection-guide (accessed 26 Aug. 2019).

Karkosh, A.E., A.K. Walker, and J.J. Simmons. 1988. Seed treatment for control of iron-deficiency chlorosis of soybean. Crop Sci. 28:369–370. doi:10.2135/cropsci1988.0011183X002800020039x

Kisha, T.J., C.H. Sneller, and B.W. Diers. 1997. Relationship between genetic distance among parents and genetic variance in populations of soybean. Crop Sci. 37:1317–1325. doi:10.2135/cropsci1997.0011183X003700040048x

Le, S., J. Josse, and F. Husson. 2008. FactoMineR: An R package for multivariate analysis. J. Stat. Softw. 25:1–18. doi:10.18637/jss.v025.i01

Li, L., W. Gao, Q. Peng, B. Zhou, Q. Kong, Y. Ying, et al. 2018. Two soybean bHLH factors regulate response to iron deficiency. J. Integr. Plant Biol. 60:608–622. doi:10.1111/jipb.12651

Lin, S.F., D. Grant, S. Cianzio, and R. Shoemaker. 2000. Molecular characterization of iron deficiency chlorosis in soybean. J. Plant Nutr. 23:1929–1939. doi:10.1080/01904160009382154

Lin, S., S. Cianzio, and R. Shoemaker. 1997. Mapping genetic loci for iron deficiency chlorosis in soybean. Mol. Breed. 3:219–229. doi:10.1023/A:1009637320805

Minnesota Agricultural Experiment Station. 2017. 2017 Soybean field crop trials results. Univ. of Minnesota. www.maes.umn.edu/sites/maes.umn.edu/files/2017_soybean_final.pdf (accessed 26 Aug. 2019).

Mamidi, S., S. Chikara, R.J. Goos, D.L. Hyten, D. Annam, S.M. Moghaddam, et al. 2011. Genome-wide association analysis identifies candidate genes associated with iron deficiency chlorosis in soybean. Plant Genome 4:154–164. doi:10.3835/plantgenome2011.04.0011

Mamidi, S., R.K. Lee, J.R. Goos, and P.E. McClean. 2014. Genome-wide association studies identifies seven major regions responsible for iron deficiency chlorosis in soybean (*Glycine max*). PLoS One 9:E107469. doi:10.1371/journal.pone.0107469

Niebur, W.S., and W.R. Fehr. 1981. Agronomic evaluation of soybean genotypes resistant to iron deficiency chlorosis. Crop Sci. 21:551–554. doi:10.2135/cropsci1981.0011183X002100040019x

O'Rourke, J.A., R.T. Nelson, D. Grant, J. Schmutz, J. Grimwood, S. Cannon, et al. 2009. Integrating microarray analysis and the soybean genome to understand the soybeans iron deficiency response. BMC Genomics 10:376. doi:10.1186/1471-2164-10-376

Peiffer, G.A., K.E. King, A.J. Severin, G.D. May, S.R. Cianzio, S.F. Lin, et al. 2012. Identification of candidate genes underlying an iron efficiency quantitative trait locus in soybean. Plant Physiol. 158:1745–1754. doi:10.1104/pp.111.189860

Ricachenevsky, F.K., R.A. Sperotto, P.K. Menguer, and J.P. Fett. 2010. Identification of Fe-excess-induced genes in rice shoots reveals a WRKY transcription factor responsive to Fe, drought and senescence. Mol. Biol. Rep. 37:3735–3745. doi:10.1007/s11033-010-0027-0

Rogers, E.E., X. Wu, G. Stacey, and H.T. Nguyen. 2009. Two MATE proteins play a role in iron efficiency in soybean. J. Plant Physiol. 166:1453–1459. doi:10.1016/j.jplph.2009.02.009

Severin, A.J., J.L. Woody, Y.T. Bolon, B. Joseph, B.W. Diers, A.D. Farmer, et al. 2010. RNA-Seq atlas of *Glycine max*: A guide to the soybean transcriptome. BMC Plant Biol. 10:160. doi:10.1186/1471-2229-10-160

Schmutz, J., S.B. Cannon, J. Schlueter, J. Ma, T. Mitros, W. Nelson, et al. 2010. Genome sequence of the palaeopolyploid soybean. Nature 463:178–183. doi:10.1038/nature08670

Song, Q., D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, et al. 2013. Development and evaluation of SoySNP50K, a high-density genotyping array for soybean. PLoS One 8:E54985. doi:10.1371/journal.pone.0054985

Song, Q., D.L. Hyten, G. Jia, C.V. Quigley, E.W. Fickus, R.L. Nelson, et al. 2015. Fingerprinting soybean germplasm and its utility in genomic research. G3 (Bethesda) 5:1999–2006. doi:10.1534/g3.115.019000

Song, Q., J. Jenkins, G. Jia, D.L. Hyten, V. Pantalone, S.A. Jackson, et al. 2016. Construction of high resolution genetic linkage maps to improve the soybean genome sequence assembly Glyma1.01. BMC Genomics 17:33. doi:10.1186/s12864-015-2344-0

Spiller, S., and N. Terry. 1980. Limiting factors in photosynthesis II: Iron stress diminishes photochemical capacity by reducing the number of photosynthetic units. Plant Physiol. 65:121–125. doi:10.1104/pp.65.1.121

Tuinstra, M.R., G. Ejeta, and P.B. Goldsbrough. 1997. Heterogeneous inbred family (HIF) analysis: A method for developing near-isogenic lines that differ at quantitative trait loci. Theor. Appl. Genet. 95:1005–1011. doi:10.1007/s001220050654

Weiss, M.G. 1943. Inheritance and physiology of efficiency in iron utilization in soybeans. Genetics 28:253–268.

Wang, J., P.E. McClean, R. Lee, R.J. Goos, and T. Helms. 2008. Association mapping of iron deficiency chlorosis loci in soybean (*Glycine max* L. Merr.) advanced breeding lines. Theor. Appl. Genet. 116:777–787. doi:10.1007/s00122-008-0710-x

Warnes, G., G. Gorjanc, F. Leisch, and M. Man. 2019. Genetics: Population genetics. R package version 1.3.8.1.2. R Foundation for Statistical Computing. https://CRAN.R-project.org/package=genetics (accessed 26 Aug. 2019).