Euphytica (2020) 216:191
https://doi.org/10.1007/s10681-020-02726-3

Mapping quantitative trait loci (QTL) for plant nitrogen isotope ratio ($\delta^{15}N$) in soybean

Sumandeep K. Bazzer · Jeffery D. Ray · James R. Smith · Felix B. Fritschi · Larry C. Purcell

Received: 15 July 2020 / Accepted: 28 October 2020 / Published online: 15 November 2020

Abstract Soybean (*Glycine max* (L.) Merr.) meets a large portion of its nitrogen (N) need via biological N_2 fixation, which is highly sensitive to drought stress. Nitrogen isotope ratios between ^{15}N and ^{14}N ($\delta^{15}N$) can be used as a metric for relative differences among soybean genotypes for N_2 fixation, as $\delta^{15}N$ is negatively associated with N_2 fixation. This study aimed to dissect the genetic basis of $\delta^{15}N$ using a mapping population of 196 F_6-derived recombinant inbred lines developed from a cross between PI 416997 and PI 567201D that was evaluated in multiple environments. There was a wide range of $\delta^{15}N$ in all environments and narrow-sense heritability for $\delta^{15}N$ was 35% when estimated across environments. Analysis of variance of $\delta^{15}N$ showed significant effects of genotype and environment, whereas the genotype × environment interaction was not significant ($P < 0.05$). Inclusive composite interval mapping for individual environments identified 10 additive QTLs on seven chromosomes with additive effects ranging from 0.02 to 0.13‰ and that individually explained phenotypic variations from 1.72 to 9.34%. In total, eight QTL × environment interactions were found, and several genomic regions were involved in QTL × QTL interactions that were not identified as additive QTLs. These identified QTLs were co-localized with genomic regions associated with N_2 fixation and other physiological traits identified in previous studies. A search for candidate genes resulted in detection of genes for nodulation and N-metabolism underlying many additive and epistatic QTLs. These identified regions may serve as potential targets for enhancing N_2 fixation in soybean.

Keywords Soybean · Nitrogen isotope ratio · Nitrogen fixation · Quantitative trait loci

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10681-020-02726-3) contains supplementary material, which is available to authorized users.

S. K. Bazzer · F. B. Fritschi
Division of Plant Sciences, University of Missouri, 1-13 Agriculture Building, Columbia, MO 65211, USA

J. D. Ray · J. R. Smith
Agricultural Research Service, Crop Genetics Research Unit, USDA, 141 Experiment Station Road, Stoneville, MS 38776, USA

L. C. Purcell (✉)
Department of Crop, Soil, and Environmental Sciences, University of Arkansas, Fayetteville, AR 72704, USA
e-mail: lpurcell@uark.edu

Introduction

Soybean (*Glycine max (L.)* Merr.) is one of the most important leguminous crops grown and consumed worldwide due to its high protein (~ 40%) and oil (~ 18–19%) concentrations. Soybean establishes

symbiotic associations with *Bradyrhizobium japonicum* (Strodtman and Emerich 2009) which reduces atmospheric N_2 to ammonia and provides N to the plant. This association decreases the requirement for N fertilizers for soybean and other leguminous crops and improves soil fertility (Giller 2001; Jensen and Hauggaard-Nielsen 2003). In soils with little available soil N, symbiotic N_2 fixation can meet up to 85–90% of the soybean N requirement (Mastrodomenico and Purcell 2012).

Symbiotic N_2 fixation is sensitive to various abiotic stresses including drought, flooding, soil salinity, soil acidity, mineral deficiency or toxicity, and low/high temperature (Ramaekers et al. 2013). Water deficit conditions negatively impact N_2 fixation in soybean by reducing nodulation and nitrogenase activity, which ultimately decreases soybean yield (Márquez-García et al. 2015; Serraj et al. 1999a). Also, it has been reported that N_2 fixation is more sensitive to water deficit than photosynthesis under both controlled and field conditions (Adams et al. 2016; Djekoun and Planchon 1991; Durand et al. 1987; Kuo and Boersma 1971; Sinclair et al. 1987). Proposed mechanisms for decreased N_2 fixation during water deficit conditions include carbon shortage, oxygen limitation, and feedback inhibition by products of N_2 fixation (Purcell 2009; Serraj et al. 1999b).

Methods for quantifying N_2 fixation include the N-difference method (Weaver and Danso 1994), acetylene reduction assay (ARA) (Hardy et al. 1968), ^{15}N enrichment (Fried and Broeshart 1975; Fried and Middleboe 1977), ^{15}N natural abundance (Shearer and Kohl 1986), and relative abundance of ureides (Unkovich et al. 2008). Each method has specific advantages over others, but easy, rapid, inexpensive, and quantitative methods for estimation of N_2 fixation under both controlled and field conditions are still needed.

Among the various methods for estimating N_2 fixation, ^{15}N natural abundance ($\delta^{15}N$) is frequently used to quantify the fraction of N derived from the atmosphere (NDFA) in large scale field experiments and to serve as an index of N_2 fixation (Andrews and Lea 2013; Barrie et al. 1995; Letolle 1980). This method compares the abundance of the ^{15}N isotope in plant tissue, the atmosphere, and the soil environment with respect to the ^{14}N isotope. The atmosphere has a lower concentration of ^{15}N compared to the soil due to the N transformations in soil. The difference in ^{15}N and ^{14}N concentration between soil and atmosphere is expressed in terms of parts per thousand (‰) and is referred to as the N isotope ratio ($\delta^{15}N$) (Peoples et al. 1989). N_2 fixation dilutes the ^{15}N in plants actively fixing N_2 as compared to plants that depend on mineral N as a N source (Doughton et al. 1995; Shearer and Kohl 1986). A low $\delta^{15}N$ value is a favorable trait for selection because it indicates greater dilution of ^{15}N by biological N_2 fixation.

The percentage of NDFA from $\delta^{15}N$ (Kohl and Shearer 1981) is calculated according to the equation below:

$$\%NDFA = \frac{\delta^{15}N_{ref} - \delta^{15}N_{samp}}{\delta^{15}N_{ref} - \delta^{15}N_0} * 100$$

where $\delta^{15}N_{ref}$ is the composition of a plant totally dependent on soil N (non-nodulating genotype), $\delta^{15}N_{samp}$ is the composition of the individual samples, and $\delta^{15}N_o$ (−1.30 for soybean, Bergersen et al. 1989) is the $\delta^{15}N$ from a plant totally dependent on N_2 fixation. The reference genotype in this equation reduces the error/noise caused by soil N variability in calculating % NDFA. However, the $\delta^{15}N$ of the reference genotype is often relatively uniform across a field (Peoples et al. 2002), indicating that in the absence of a reference crop, $\delta^{15}N$ can be used directly to estimate the amount of N fixed by genotypes via N_2 fixation (Steketee et al. 2019).

The difference among genotypes for N_2 fixation under normal and stress conditions may help identify genomic regions controlling N_2 fixation under water deficit conditions. Quantitative trait loci (QTLs) mapping is the molecular approach used to understand the genetic architecture of many physiological and agronomical traits. Recent advances in high throughput genotyping and phenotyping platforms have revolutionized the dissection of the genetic basis of quantitative traits like N_2 fixation and will accelerate development of soybean lines with enhanced N_2 fixation.

Several studies have mapped QTLs for N_2 fixation or related traits in soybean (www.soybase.org). Tanya et al. (2005) used a population of 136 F_2-derived recombinant inbred lines (RILs) to identify a total of nine QTLs for nodule number per plant, nodule fresh and dry weight per plant, and acetylene reduction activity (ARA). Nicolás et al. (2006) identified two genomic regions associated with nodule number and

nodule dry weight. Santos et al. (2013) studied the genetic control of nodule number and individual nodule weight and confirmed a QTL for nodule number identified previously by Nicolás et al. (2006). Hwang et al. (2014) were the first to map QTLs for nodule number, nodule size, and nodule weight in field experiments.

Dhanapal et al. (2015b) used association mapping on a diverse panel of 374 maturity group 4 accessions to identify QTLs for NDFA and N concentration. This analysis identified 17 and 19 SNPs significantly associated with NDFA and N concentration, respectively. Steketee et al. (2019) used association mapping for $\delta^{15}N$ using a panel of 211 diverse soybean accessions and found 23 and 26 SNPs associated with $\delta^{15}N$ and N concentration, respectively. To date, more than 70 QTLs for N_2 fixation or traits directly or indirectly linked with N_2 fixation have been mapped on all 20 chromosomes of soybean (www.soybase.org).

In the present study, a high-density genetic linkage map was constructed using 196 F_6-derived RILs developed from PI 416997 × PI 567201D. The parents of this population were originally chosen because they were extremes for the ratio between ^{13}C and ^{12}C (Bazzer et al. 2020a), which serves as a surrogate measure of water use efficiency (Farquhar and Richards 1984). Although the parents were not selected for $\delta^{15}N$ or N_2 fixation, their RILs segregated for $\delta^{15}N$. Therefore, the main objectives of our study were to identify additive QTLs for $\delta^{15}N$, epistatic QTLs, and QTL × environment interactions. Further characterization of genes underlying the QTLs identified in this study will help to understand the biological mechanisms regulating N_2 fixation in soybean and the genetic basis of N_2 fixation.

Materials and methods

Development of RIL population

The cross between PI 416997 and PI 567201D was made at Stoneville, MS in 2011. The F_1 generation was grown during the winter of 2011–2012 at the Tropical Agricultural Research Station at Isabela, Puerto Rico. The F_2 generation was grown in Stoneville in 2012, where over 200 individual F_2 plants were harvested without selection. Leaf tissue was harvested from each tagged F_2 plant for DNA extraction and genotyping of the population. The $F_{2:3}$ and $F_{4:5}$ generations were grown in Homestead, FL during the winters of 2012–2013 and 2013–2014, respectively, harvesting one random plant from each single-plant-derived row in each nursery. The $F_{3:4}$, $F_{5:6}$, and $F_{6:7}$ generations were grown in Stoneville, with the former two generations being advanced by single-plant descent in 2013 and 2014, respectively, and the latter generation being bulk harvested in 2015 to create bulked F_6-derived lines for phenotyping.

Field trials

A mapping population consisting of 196 F_6-derived RILs generated from a cross between PI 416997 × PI 567201D was used to identify the genomic regions associated with $\delta^{15}N$. The RIL population and parents were evaluated in four environments: at Stoneville, MS (33.42° N, 90.90° W) on a Bosket very fine sandy loam soil (fine-loamy, mixed, active, thermic, Mollic Hapludalfs) in 2016 and on a Dundee silty clay loam soil (Dundee fine-silty, mixed, active, thermic, Typic Endoaqualfs) in 2017, at the Milo J Shult Arkansas Agricultural Research Center, Fayetteville, AR (36.05° N, 94.15° W) on a Captina silt loam soil (fine-silty, siliceous, active, mesic Typic Fragiudult) in 2017, and at the Bradford Research Center near Columbia, MO (38.95° N, 92.33° W) on a Mexico silt loam soil (fine, smectitic, mesic Vertic Epiaqualf) in 2017. The combinations of locations and years were considered as individual environments and designated as ST16 (Stoneville in 2016), ST17 (Stoneville in 2017), FAY17 (Fayetteville in 2017), and CO17 (Columbia in 2017). Plantings occurred on 6 May 2016 at ST16 and 16 May 2017 at ST17 in one-row plots (0.66 m wide by 2.74 m long), 10 June 2017 at FAY17 in two rows plots (0.45 m wide by 6 m long), and 14 May 2017 at CO17 in single row plots (0.76 m wide by 3.05 m long). At each environment, the experimental design was a randomized complete block design with two replications. Experiments were irrigated as needed. Recommended practices were followed for insect and weed control.

Data collection

Shoot biomass of four random plants was sampled from each plot between beginning bloom (R1) and the

full bloom (R2) stages (Fehr and Caviness 1977) on 29 June 2016 at ST16, 21 June 2017 at ST17, and 21 July 2017 at FAY17 and CO17. Biomass samples were dried at 60°C and coarse ground with a Wiley Mill (Thomas Model 4 Wiley® Mill, Thomas Scientific, NJ USA). Subsamples were finely-ground to pass a 1 mm sieve, and then ground to a fine powder with a Geno Grinder (SPEX CertiPrep, Inc., NJ USA) as described by Bazzer et al. (2020a). About 3–5 mg of the powdered sample was weighed into tin capsules, for $\delta^{15}N$ isotope analysis which was conducted at the University of California-Davis Stable Isotope Facility (https://stableisotopefacility.ucdavis.edu/) using an elemental analyzer interfaced to a continuous flow isotope ratio mass spectrometer. Data from the stable isotope facility were expressed as $\delta^{15}N$ in per mil (‰) and determined according to the equation below:

$$\delta^{15}N = \frac{R_{sample}}{(R_{air} - 1)} * 1000$$

where R_{sample} and R_{air} are the isotope ratios ($^{15}N/^{14}N$) of the sample and air, respectively.

For more information refer to the Stable Isotope facility website, http://stableisotopefacility.ucdavis.edu/13cand15n.html.

Statistical analysis

Descriptive statistics of $\delta^{15}N$ for each environment and correlation coefficients between different environments for $\delta^{15}N$ were calculated with SAS version 9.4 (SAS, Institute 2013). The difference between parents for $\delta^{15}N$ in different environments was determined using a *t*-test. Analysis of variance (ANOVA) was conducted using SAS version 9.4 (SAS Institute 2013) with the PROC MIXED procedure ($\alpha = 0.05$). Genotype and environment were considered as fixed effects and replication nested within environment was considered as a random effect (Bondari 2003). Heritability (h^2, Holland et al. 2003) of $\delta^{15}N$ for each environment and averaged across environments was computed using the PROC VARCOMP procedure of SAS 9.4 based on the following formula: Across environments : $h^2 = \frac{\sigma_G^2}{\sigma_G^2 + \left(\frac{\sigma_{GE}^2}{e}\right) + \left(\frac{\sigma_e^2}{re}\right)}$ Within environments : $h^2 = \frac{\sigma_G^2}{\sigma_G^2 + \left(\frac{\sigma_e^2}{r}\right)}$ where σ_G^2, σ_{GE}^2, and σ_e^2 are the genotypic variance, genotypic × environment interaction variance, and residual error variance, respectively, e is the number of environments, and r is the number of replications within environment. This heritability should be considered a narrow sense estimate, as F_6-derived RILs have a minimal level of heterozygosity within lines. Hence, most of the genotypic variance is composed of additive variance, with negligible variance due to dominance effects and its interaction with additive effects. The best linear unbiased prediction (BLUP) values for each individual environment and across environments were calculated using a mixed model to reduce environmental variance. All factors were considered as random effects in the case of individual environments. For calculation of BLUP values averaged across environments, environment was considered a fixed effect and genotypes and replications were considered as random effects (Littell et al. 1996; Piepho et al. 2008). QTL analysis was conducted using BLUP values for individual environments and across environments. The $\delta^{15}N$ BLUP values of 196 RILs for all individual environment and across environments are provided in Supplementary File 1.

Selection of lines with extreme values for $\delta^{13}C$ and $\delta^{15}N$

As described previously (Bazzer et al. 2020a), this population was also evaluated for carbon isotope ratio ($\delta^{13}C$) (a proxy measurement for water use efficiency, WUE) as the parents were different in their $\delta^{13}C$ values based on phenotypic values and genomic estimated breeding values (GEBVs) (Dhanapal et al. 2015a; Kaler et al. 2017). The linear regression between $\delta^{13}C$ and $\delta^{15}N$ was performed using phenotypic values from each individual environment and averaged across environments. The biplots were divided into four quadrants using median values of $\delta^{13}C$ and $\delta^{15}N$ to select the lines that were extremes for both traits.

Genotyping-by-sequencing and construction of linkage map

Detailed information on genotyping and linkage map construction of this population were provided by Bazzer et al. (2020a), which are summarized below.

The 196 RILs, together with their two parents, were sequenced and genotyped using genotype-by-sequencing (GBS). DNA was isolated from lyophilized leaf tissue of the 196 RILs and both parents, and GBS library construction was conducted at LGC Genomics GmbH (Berlin, Germany). The sequence reads were mapped to the 'Williams 82' soybean reference genome (assembly 1). Markers with more than 15% missing data, markers that were heterozygous, or did not follow a 1:1 segregation ratio pattern (chi-square P-value ≤ 0.01) were removed, resulting in a total of 3234 polymorphic markers. Missing marker data were imputed using a LD-kNNi method which was implemented in TASSEL software (https://www.maizegenetics.net/tassel).

The filtered and imputed 3234 polymorphic SNPs were used initially for construction of the linkage map. The MAP functional module of IciMapping software v4.1 (Meng et al. 2015) was used for the genetic linkage map construction using 2468 out of 3234 polymorphic markers after dropping redundant and low-quality markers. Genetic linkage groups, marker order and distances between markers were determined as described previously (Bazzer et al. 2020a). Linkage groups were numbered as soybean chromosomes according to the genomic position of SNPs on the reference genome. Finally using 2,466 SNPs, a map with 20 linkage groups was constructed that corresponded to the 20 soybean chromosomes. The SNP markers information along with their position (in base pairs and cM) on specific chromosomes are provided in Supplementary File 3 reported by Bazzer et al. (2020a).

QTL analysis

The BLUP values calculated for each individual environment and averaged across environments were used for QTL analysis. QTL mapping was performed using the QTL IciMapping v4.1 software (http://www.isbreeding.net/) through BIP and MET functional modules. A BIP module of inclusive composite interval mapping (ICIM) was used to detect the additive QTLs within and across environments. QTL × QTL interactions were identified by using the Inclusive Composite Interval Mapping of Epistatic QTL (ICIM-EPI) function. The LOD threshold to declare significant additive QTLs and interactions between QTLs was calculated using 1,000 iteration permutation tests with a genome-wide significance level of 0.05 (Doerge and Churchill 1996; Li et al. 2007). The mapping parameters were 1.0 cM walking speed with P-value inclusion threshold of 0.01 for ICIM-ADD and 5 cM walking speed with a P-value inclusion threshold of 0.0001 for ICIM-EPI. Multi-Environment Traits (MET) module was used for detection of QTL × environment interactions using $\delta^{15}N$ BLUP data from all four environments. The missing phenotypic values were calculated by using the 'mean replacement' method. The specific parameters for detecting QTL × environment interactions were 1.0 cM walking speed and a probability of 0.01 in stepwise regression. Finally, the position of SNPs on different chromosomes and the position of identified QTLs on the genetic map were drawn using Prism software (https://www.graphpad.com/).

Identification of putative candidate genes

The search for putative candidate genes related to nodulation and N-metabolism underlying the genomic regions associated with $\delta^{15}N$ identified in the present study was performed using the genome browser option (William 82 assembly 1) of Soybase (www.soybase.org). Genes between flanking markers and up to $\pm$ 1 MB outside of the confidence interval for flanking markers were considered as potential candidate genes. Additionally, the position of 54 soybean genes (28 nodulin + 24 regulatory genes) associated with nodulation or biological N_2 fixation (Schmutz et al. 2010) were compared to the genomic regions of the $\delta^{15}N$ QTLs identified in the present research.

Results

Phenotypic evaluation

The phenotypic values of parents and descriptive statistics of the RIL population are presented in Table 1. The parent PI 416997 had lower $\delta^{15}N$ values than PI 567201D in all environments, but parents were not significantly ($P < 0.05$) different in any single environment. However, ANOVA across environments did indicate that $\delta^{15}N$ for PI 416997 was significantly ($P < 0.01$) lower than PI 567201D (data not shown). There was wide segregation in the RIL population for $\delta^{15}N$ (Fig. 1; Table 1) as indicated by $\delta^{15}N$ ranges of

Table 1 Phenotypic variation for $\delta^{15}N$ (‰) in the parents (PI 416997 and PI 567201D) and RIL population grown in four environments (Env.): Stoneville in 2016 (ST16), Stoneville in 2017 (ST17), Fayetteville in 2017 (FAY17), and Columbia in 2017 (CO17)

Env.	Parent			RIL Population			
	PI 416997	PI 567201D	Mean ± SD	Max.–Min.[a]	Range	Skewness	Kurtosis
ST16	1.99	2.78	2.79 ± 0.70	5.08− 1.12	3.96	0.17	0.01
ST17	3.15	3.65	4.20 ± 0.84	6.28–2.20	4.08	0.19	− 0.11
FAY17	2.25	4.41	3.01 ± 0.68	5.00−1.57	3.43	0.14	− 0.35
CO17	1.07	3.12	2.02 ± 0.88	4.80− 0.25	4.55	0.45	0.06

[a] Indicated the maximum (Max.) and minimum (Min.) $\delta^{15}N$ values for different environments

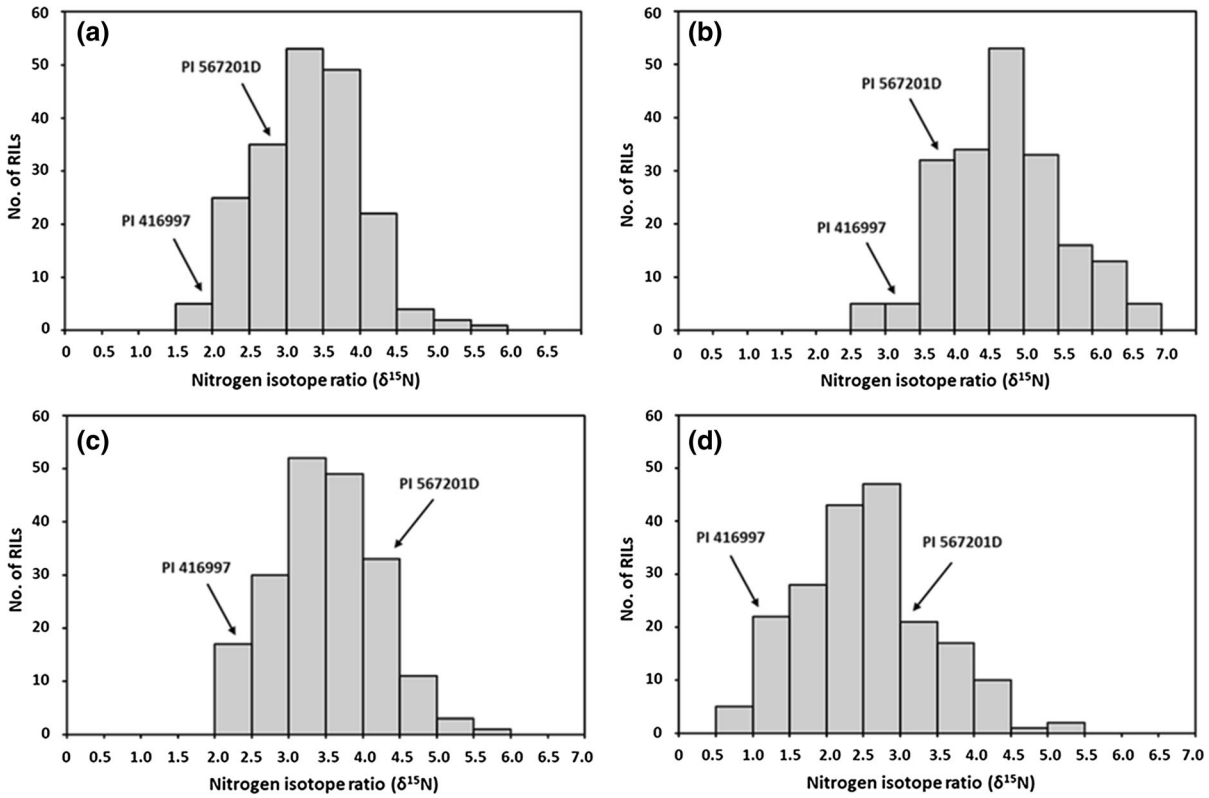

Fig. 1 Distribution of δ15N among recombinant inbred lines and parental genotypes at Stoneville, MS in 2016 (**a**), Stoneville, MS in 2017 (**b**), Fayetteville, AR in 2017 (**c**), and Columbia, MO in 2017 (**d**)

3.96‰ in ST16, 4.08‰ in ST17, 3.43‰ in FAY17, and 4.55‰ in CO17. The frequency distribution of $\delta^{15}N$ was normal in all environments except CO17, as indicated by Shapiro-Wilk test (data not shown, Shapiro and Wilk 1965) and absolute values of skewness and kurtosis (less than 1.0, Table 1) . Transgressive segregants exceeding both parents were observed, which indicates that favorable alleles for $\delta^{15}N$ were distributed between both parents. A significant positive correlation ($P < 0.05$) was found between ST16 and ST17 ($r = 0.15$), ST17 and

FAY17 ($r = 0.15$), and ST17 and CO17 ($r = 0.14$) (data not shown).

Analysis of variance (ANOVA) averaged across environments showed significant ($P < 0.05$) effects of genotype (G) and environment (E), whereas the interaction of genotype and environment (G × E) was not significant (Table 2). The narrow sense heritability of $\delta^{15}N$ averaged across environments was 35%. Estimates of narrow sense heritability for $\delta^{15}N$ within environments were 8% (ST16), 13% (ST17), 27% (FAY17), and 24% (CO17). Overall, low narrow sense heritability estimates across and within environments indicate that environmental effects play a major role in the expression of this trait and that phenotypic selection for this trait may only be successful using replicated trials of homogeneous lines across multiple environments.

The RIL population used in the present study was also evaluated for $\delta^{13}C$ because the parents, PI 416997 and PI 567201D, also differed in WUE (Bazzer et al. 2020a). Although no significant correlation was found between these traits, regression analysis was performed between $\delta^{15}N$ and $\delta^{13}C$ phenotypic values for within and across environments to identify extremes among RILs for different combinations of $\delta^{15}N$ and $\delta^{13}C$. The selected five RILs for high relative N_2 fixation and high WUE in individual environments and averaged across environments are listed in Table 3. The lines RIL.14, RIL.25, RIL.75, and RIL.135 were identified in at least two individual environments and across environments as having a favorable combination (low $\delta^{15}N$ and high $\delta^{13}C$). Similarly, lines with the unfavorable combination of high $\delta^{15}N$ and low $\delta^{13}C$ values were identified in individual environments and averaged across environments. The lines RIL.24, RIL.127, and RIL.161 were identified as having the unfavorable combination of high $\delta^{15}N$ and low $\delta^{13}C$ values in at least two individual environments, and RIL.112 had this unfavorable combination across environments in addition to individual environments.

Table 2 Analysis of variance (ANOVA) for $\delta^{15}N$ in the RIL population along with parents evaluated in four environments (ST16, ST17, FAY17, and CO17)

Effect	DF	*P* value	h^2 (%)
Genotypes (G)	195	< 0.0001	35
Environments (E)	3	0.016	
G × E	585	0.267	

h^2, Narrow sense heritability

QTL analysis

The linkage map was 3,836 cM with individual linkage groups varying between 116 and 409 cM, based on the construction using 2,466 SNP markers (data not shown, Bazzer et al. 2020a). QTL analysis conducted using $\delta^{15}N$ BLUP values from individual environments identified a total of 10 additive QTLs within environments (Fig. 2; Table 4), which were distributed on seven chromosomes (Gm01, Gm04, Gm07, Gm08, Gm10, Gm13, and Gm17). Of these QTLs, there were four QTLs in ST16, two QTLs in ST17, and four QTLs in FAY17. No QTLs were identified in CO17. The phenotypic variation explained by individual QTLs (denoted as R^2), their additive effect, and parent contributing favorable alleles are presented in Table 4. These QTLs individually accounted for 1.72 to 9.34% of the phenotypic variation and had additive effects ranging from 0.02 to 0.13‰. The QTL present on Gm04 at 49,247,258 bp detected in ST17 had a high R^2 value (9.34) as compared to other QTLs.

QTL analysis by using the BLUP values averaged across environment (AE) by the ICIM-ADD mapping method identified eight QTLs (Fig. 2; Table 4). These QTLs were present on Gm01 (2), Gm04 (1), Gm07 (1), Gm13 (1), Gm14 (1), and Gm15 (2) and had additive effects that ranged from 0.04 to 0.12‰ that explained individually 1.83 to 14.39% of the phenotypic variation (Table 4). The QTLs on Gm01, Gm04, Gm07, and Gm13 appeared in both individual environments and across environments. The QTLs on remaining chromosomes were detected only in single environments or only across environments. An allele decreasing $\delta^{15}N$ values was considered as the favorable allele, and the favorable allele for these QTLs was equally distributed between parents (PI 416997 and PI 567201D). When considering overlapping confidence intervals, there were 13 loci detected within and across environments (Fig. 2; Table 4).

Table 3 RILs with high $\delta^{13}C$ and low $\delta^{15}N$ phenotypic values or RILs with low $\delta^{13}C$ and high $\delta^{15}N$ phenotypic values in individual environments (ST16: Stoneville in 2016, ST17: Stoneville in 2017, FAY17: Fayetteville in 2017, and CO17: Columbia in 2017) and across environments (AE). Values in parentheses are the phenotypic values of $\delta^{13}C$ and $\delta^{15}N$ for that RIL

ST16	ST17	CO17	FAY17	AE
RILs with high $\delta^{13}C$ (> WUE and > N_2 fixation)				
RIL.14 (−28.85, 1.89)	RIL.25 (−27.70, 3.39)	RIL.14 (−27.54, 0.25)	RIL.25 (−26.56, 2.43)	**RIL.14 (−28.02, 2.02)**
RIL.25 (−28.38, 2.18)	RIL.75 (−28.02, 3.31)	RIL.42 (−27.27, 1.02)	RIL.69 (−26.93, 1.98)	**RIL.25 (−27.34, 2.55)**
RIL.132 (−29.02, 1.79)	RIL.120 (−28.02, 3.63)	RIL.75 (−27.17, 1.00)	RIL.85 (−26.94, 1.93)	**RIL.75 (−27.65, 2.49)**
RIL.193 (−28.75, 2.13)	RIL.135 (−27.99, 3.03)	RIL.85 (−27.44, 1.50)	RIL.135 (−26.69, 2.52)	**RIL.135 (−27.59, 2.47)**
RIL.209 (−28.87, 2.30)	RIL.204 (−28.10, 3.19)	RIL.135 (−27.18, 1.13)	RIL.151 (−26.92, 1.71)	RIL.204 (−27.72, 2.86)
RILs with low $\delta^{13}C$ and high $\delta^{15}N$ (< WUE and < N_2 fixation)				
RIL.12 (−29.96, 3.77)	RIL.24 (−30.14, 6.25)	RIL.15 (−28.66, 3.25)	RIL.92 (−29.06, 3.73)	**RIL.112 (−29.18, 3.87)**
RIL.24 (−30.19, 3.19)	RIL.50 (−29.97, 4.91)	RIL.112 (−28.89, 3.55)	RIL.127 (−29.29, 4.23)	RIL.154 (−29.17, 3.81)
RIL.86 (−30.17, 3.59)	RIL.52 (−29.77, 5.09)	RIL.127 (−28.61, 3.60)	RIL.129 (−28.99, 3.79)	RIL.177 (−29.12, 3.81)
RIL.112 (−30.14, 3.28)	RIL.110 (−29.86, 4.76)	RIL.195 (−28.55, 3.97)	RIL.161 (−29.06, 3.76)	RIL.203 (−29.03, 4.24)
RIL.114 (−30.02, 4.32)	RIL.161 (−29.86, 5.77)	RIL.205 (−28.51, 4.80)	RIL.213 (−29.03, 3.82)	RIL.205 (−29.10, 4.01)

Bold values indicate the RILs appeared within multiple individual environments as well as across environments

QTL × environment and QTL × QTL interactions analysis

The interactions between QTL × environment and QTL × QTL play important roles in the genetic control of quantitative traits (Rebetzke et al. 2007; Reif et al. 2011). Eight QTLs present on Gm01 (2), Gm04, Gm07, Gm08 (2), Gm10, and Gm13 showed significant QTL × environment interactions as identified by MET functionality (Table 5). This interaction explained phenotypic variation that ranged from 1.17 to 28.25% and with additive effects from 0.01 to 0.03‰ (Table 5). Phenotypic variation due to additive × environment effects (PVE (A × E)) was greater than additive effects (PVE (A)) and the LOD score of additive effects (LOD (A)) was less than the LOD score for additive × environments effects (LOD (A × E)) for most of these QTLs, indicating that these QTLs had strong interaction with environments. The QTL on Gm01 (3,032,794 bp) had a greater LOD score and PVE for additive effect than additive × environment effect (Table 5), indicating the stability of this QTL across environments.

QTL × QTL interactions were detected using the Epistatic QTL (ICIM-EPI) method of BIP functional module for $\delta^{15}N$ values from individual environments. Epistatic interactions between different genomic regions were detected in ST16 and FAY17 and across environments (AE) (Table 6). The phenotypic variation explained by these interactions ranged from 3.53 to 7.78%, with the LOD score of these interactions being greater or equal to 3.5. The QTLs involved in epistasis were not identified as additive QTLs. No epistasis was detected in ST17 or CO17.

Identification of putative candidate genes

Of 13 additive loci, five loci (loci 1, 3, 6, 9, and 12) fell in the genomic regions carrying published soybean nodulation genes reported by Schmutz et al. (2010) (Fig. 2) that are directly involved in nodulation through production of nodulin proteins, nodulation signaling proteins, and different regulatory proteins. For example, *Glyma.01g03470* (Locus 1), *Glyma.04g43090* (Locus 3), *Glyma.08g05370* (Locus 6), *Glyma.14g05690* (Locus 9), and *Glyma.17g08110* (Locus 12) genes (Schmutz et al. 2010) are involved directly in the process of nodulation. The nodulation genes, *Glyma.11g06740, Glyma.13g40400*, and *Glyma.15g05010* were in the genomic regions of epistatic QTLs present on Gm11, Gm13, and Gm15, respectively. The genes having function related to N metabolism or N_2 fixation present between flanking markers and up to ± 1 MB outside of the confidence interval of flanking markers underlying $\delta^{15}N$ loci are provided in Supplementary File 2.

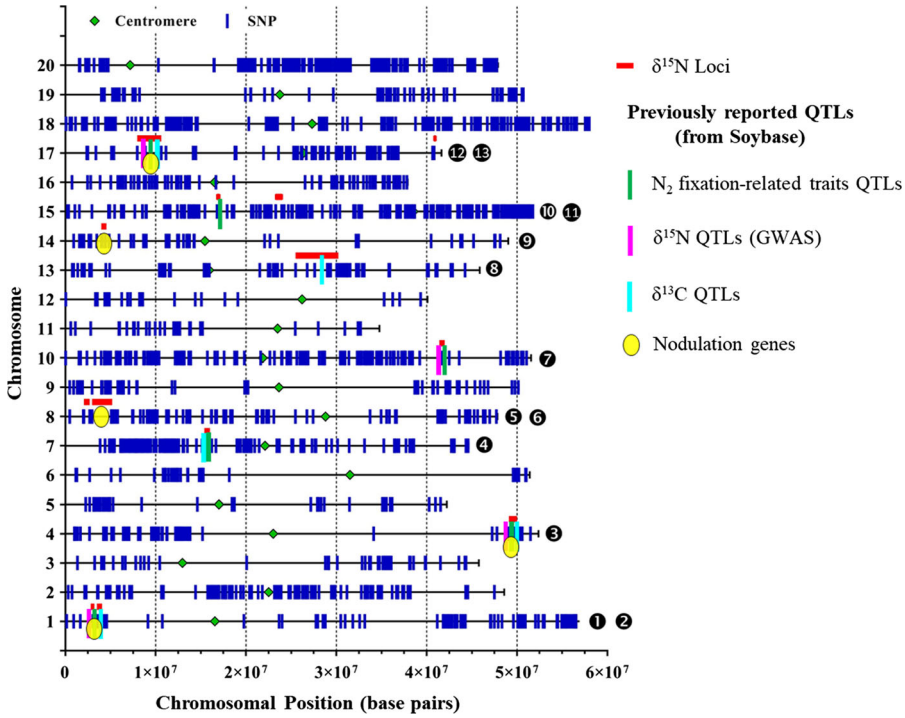

Fig. 2 Physical position of SNPs on soybean chromosomes and position of loci (horizontal red bars) associated with $\delta^{15}N$ identified by ICIM mapping for additive QTLs. The numbers in the black circles represent the loci number on a specific chromosome. Vertical colored bars (except blue) indicate the other QTLs found at the same positions in previous studies, and yellow circle indicate the nodulation genes

Discussion

A prerequisite for genetic improvement of N_2 fixation is adequate genetic variability for the trait, and understanding the genetic basis of this variability using a dense genetic map would be helpful for implementing the most appropriate strategies in a soybean breeding program. In this study, we investigated the variability in $\delta^{15}N$ as a proxy for biological N_2 fixation using a population of RILs. In previous studies, NDFA was used for estimation of N_2 fixation in soybean and other legumes (Dhanapal et al. 2015b; Heilig et al. 2017; Ramaekers et al. 2013). In the present study, $\delta^{15}N$ values were directly used as an estimate of N_2 fixation, as a non-nodulating/reference genotype was not planted with the experimental material in order to calculate NDFA values. Steketee et al. (2019) also used $\delta^{15}N$ values in mapping N_2 fixation in a GWAS panel for soybean. Our results found a significant difference ($P < 0.01$) between the parents when combined over environments, with PI 416997 having lower $\delta^{15}N$ values than PI 567201D. The low $\delta^{15}N$ value of PI 416997 indicates that the proportion of N from N_2 fixation was greater for PI 416997 compared to PI 567201D.

The RILs had a wide phenotypic range for $\delta^{15}N$ in all environments (ranged 3.43 to 4.55‰) (Table 1), but the specific range of $\delta^{15}N$ in soybean is not well defined (Dhanapal et al. 2015b; Ludidi et al. 2007). The presence of transgressive segregants indicates that selection of lines for both low and high $\delta^{15}N$ (along with low and high $\delta^{13}C$) values would be possible. Biplot analysis of $\delta^{15}N$ and $\delta^{13}C$ identified five RILs with favorable (low $\delta^{15}N$ and high $\delta^{13}C$) and unfavorable (high $\delta^{15}N$ and low $\delta^{13}C$) phenotypic combinations for $\delta^{13}C$ and $\delta^{15}N$ (Table 3). RIL.25 and RIL.135 were among the five top RILs (high WUE and high N_2 fixation) in three of the four environments as well as across environments. Comparison of lines with contrasting $\delta^{15}N$ and $\delta^{13}C$ signatures may be important in characterizing the physiological and interactions between of N_2 fixation and WUE.

Narrow sense heritability of $\delta^{15}N$ across environments was 35% and for individual environments ranged from 8 to 27% in this study. Steketee et al.

Table 4 Quantitative trait loci (QTLs) associated with $\delta^{15}N$ detected in individual environments (ST16, ST17, FAY17, and CO17) and across environment (AE) in the RIL population of PI 416997 and PI 567201D using BIP functional module of ICIM mapping

Locus[a]	Chrom.[b]	Env[c]	Nearest SNP position (bp)[d]	Flanking markers[e]	LOD[f]	R^{2g}	Add[h]	Favorable allele[i]
1	Gm01	ST16	3,284,926	*S_001_003_032_794– S_001_003_284_926*	10.56	3.14	− 0.08	PI567201D
		AE			18.36	7.93	− 0.09	
2	Gm01	ST16	3,955,325	*S_001_003_605_644– S_001_003_955_325*	6.08	1.72	0.06	PI416997
		AE	3,605,601	*S_001_003_382_064– S_001_003_605_601*	12.57	5.16	0.07	
3	Gm04	ST17	49,247,258	*S_004_049_095_809–S_004_049_247_258*	6.25	9.34	0.06	PI416997
		AE	49,618,681	*S_004_049_542_576– S_004_049_618_681*	4.87	1.83	0.04	
4	Gm07	FAY17	15,382,101	*S_007_015_382_101– S_007_015_546_393*	7.15	6.80	− 0.03	PI567201D
		AE	15,546,393	*S_007_015_546_393– S_007_015_805_651*	8.31	3.21	− 0.05	
5	Gm08	ST16	2,681,851	*S_008_002_049_802–S_008_002_681_851*	24.21	9.10	− 0.13	PI567201D
6	Gm08	ST16	2,960,542	*S_008_002_960_542– S_008_005_171_578*	18.44	6.26	0.11	PI416997
7	Gm10	FAY17	41,413,995	*S_010_041_413_938– S_010_041_413_995*	5.11	4.71	0.02	PI416997
8	Gm13	FAY17	27,584,266	*S_013_025_487_275– S_013_027_584_266*	5.98	6.88	0.03	PI416997
		AE	30,216,959	*S_013_030_039_666–S_013_030_216_959*	28.57	14.39	0.12	
9	Gm14	AE	4,448,348	*S_014_003_972_110–S_014_004_448_348*	4.37	1.63	− 0.04	PI567201D
10	Gm15	AE	17,072,416	*S_015_016_781_440–S_015_017_072_416*	6.17	2.35	− 0.05	PI567201D
11	Gm15	AE	23,341,546	*S_015_023_296_352– S_015_023_341_546*	12.63	5.22	0.07	PI416997
12	Gm17	ST17	7,961,686	*S_017_007_961_686–S_017_010_605_971*	4.29	7.64	− 0.06	PI567201D
13	Gm17	FAY17	40,829,268	*S_017_040_829_268–S_017_040_829_187*	3.40	3.02	− 0.02	PI567201D

[a]Closely spaced putative QTL falling within the same flanking markers were consider as one locus

[b]*Glycine max* chromosome on which putative QTL is present

[c]Environment in which a significant QTL was identified

[d]Nearest marker position in base pairs (bp) on specific chromosome

[e]Position of SNPs identified in the mapping analysis as flanking the putative QTL

[f]Log-likelihood at QTL peak position

[g]Phenotypic variation explained by putative QTL

[h]Additive effect of the QTL

[i]Allele that decreases $\delta^{15}N$ value

(2019) reported low broad sense heritability of $\delta^{15}N$ (H = 17%) in an association study conducted using a diverse panel of soybean accessions. Similarly, Dhanapal et al. (2015b) found that broad sense heritability of NDFA was low (H = 21%) in a GWAS panel. In previous research, several physiological traits linked with N_2 fixation were used for studying the genetic basis of N_2 fixation (Hwang et al. 2013; Ray et al. 2015; Santos et al. 2013; Vieira et al. 2006). Narrow sense heritability was 18% (Vieira et al. 2006) and 33% (Santos et al. 2013) for nodule number, and 27% for weight per nodule (Santos et al. 2013). Broad sense heritability ranged from 33% (Ray et al. 2015) to 73% (Hwang et al. 2013) for shoot ureides and 59% for shoot N concentration (Hwang et al. 2013). Therefore, heritability of traits related to N_2 fixation generally appear to be low to moderate, which is consistent with strong influence of environmental conditions on N_2 fixation (Mastrodomenico and Purcell 2012; Ramaekers et al. 2013; Serraj et al. 1999a; Sinclair et al. 1987).

BLUP values of $\delta^{15}N$ were used in the QTL analysis as it increases the accuracy of detection of QTLs by reducing the impact of environment. QTL analysis for $\delta^{15}N$ by individual environment identified 10 QTLs present on seven chromosomes (Fig. 2; Table 4). No common QTLs were detected in two or more environments, but five QTLs were detected in specific environments that were also found across environments (Table 4). In addition to identified additive QTLs by individual environment analysis,

Table 5 QTLs showing QTL × environment interaction in four environments detected using MET functional module of ICIM mapping

Locus[a]	Chrom.[b]	Nearest SNP position[c]	LOD[d]	LOD (A)[e]	LOD (A × E)[f]	PVE[g]	PVE (A)[h]	PVE (A × E)[i]	Add[j]
1	Gm01	3,032,794	13.17	11.42	1.75	12.25	6.43	5.83	− 0.03
2	Gm01	3,955,325	6.18	2.11	4.07	5.75	1.21	4.53	0.01
3	Gm04	49,247,258	6.86	4.38	2.49	7.11	2.46	4.65	0.02
4	Gm07	15,382,101	7.25	0.90	6.35	1.40	0.50	0.90	− 0.01
5	Gm08	2,681,851	24.21	13.82	10.38	28.25	8.06	20.19	− 0.03
6	Gm08	2,960,542	18.44	7.87	10.57	19.68	4.45	15.23	0.03
7	Gm10	41,413,995	7.28	1.97	5.31	2.94	1.11	1.82	0.01
–[*]	Gm13	27,584,266	6.03	0.49	5.54	1.17	0.25	0.92	0.01

[*]New loci, not identified earlier in ICIM-ADD analysis

[a]Locus number assigned is the same as the locus number used for QTLs identified by BIP functionality of ICIM Mapping in Table 4

[b]*Glycine max* chromosome on which putative QTL is present

[c]Nearest marker position in base pairs (bp) on specific chromosome

[d]Total LOD score for QTL × environment interaction

[e]LOD score for additive effects

[f]LOD score for additive by environment effects

[g]Total phenotypic variance explained by QTL × environment interaction

[h]Phenotypic variance explained by additive effects

[i]Phenotypic variance explained by additive by environments effects

[j]Additive effect explained by the QTL

Table 6 Epistatic QTLs identified for $\delta^{15}N$ in the RIL population of PI 416997 × PI 567201D by the ICIM-EPI method of BIP functional module

Env.[a]	Chrom. 1[b]	Pos. 1[c]	Chrom. 2[d]	Pos. 2[e]	PVE[f]	Add × Add[g]
ST16	1	52,051,133–52,386,863	15	1,550,336–2,934,524	4.13	0.06
	5	41,542,487–36,175,591	6	12,165,433–12,361,723	7.78	0.08
	12	4,858,604–6,236,974	12	7,201,220–7,201,391	4.45	− 0.08
	13	35,909,641–40,053,812	16	33,339,297–33,694,437	6.08	− 0.07
FAY17	3	43,610,628–44,233,282	15	7,684,532–8,529,280	3.53	− 0.03
	5	41,542,487–36,175,591	2	4,726,469–5,730,965	4.44	0.03
AE	3	41,518,436–43,565,653	14	7,259,215–8,645,366	4.60	0.05
	6	14,794,344–18,148,587	11	585,930–1,088,794	6.17	− 0.07

[a]Environment in which a significant QTL × QTL interaction was identified

[b]*Glycine max* chromosome on which first QTL involved in epistasis present

[c]Position of flanking markers of first QTL in epistasis on specific chromosome

[d]*Glycine max* chromosome on which second QTL involved in epistasis present

[e]Position of flanking markers of second QTL in epistasis on specific chromosome

[f]Phenotypic variation explained by epistasis

[g]Additive by additive epistatic effect

one QTL on Gm14 and two QTLs on Gm15 were detected in QTL analysis across environments (Table 4). When considering QTLs detected in individual environments and across environments and overlapping confidence intervals, a total of 13 loci (Fig. 2; Table 4) were identified.

Of the eight QTLs that had significant QTL × environment interaction, six were also detected in individual environment analysis (Table 5). Although we did not find any QTLs common among individual environments, the detection of additive QTLs in joint-environment analysis by MET functionality increases the confidence of detection of these QTLs. The low phenotypic variation explained by the additive effect (PVE (A)) compared with the additive × environment effect (PVE (A × E)) indicates a large effect of environment for all QTLs except for the QTL on Gm01 (3,032,794 bp) (Table 5). The phenotypic variation explained by these additive QTLs is small ($R^2 < 10\%$), which indicates the complex nature of biological N_2 fixation (Santos et al. 2013).

In this study, eight epistatic interactions explained 41% of the phenotypic variation present on different chromosomes (Table 6). Further, these epistatic QTLs were not identified as additive QTLs (Table 4). An epistatic QTL present on Gm05 was detected in both ST16 and FAY17, but this QTL interacted with different epistatic QTLs in ST16 (Gm06) and FAY17 (Gm02). Also, an epistatic QTL present on Gm03 that was detected in both FAY17 and AE interacted with the QTLs present on Gm15 (FAY17) and Gm14 (AE) (Table 6). Our results indicate that additive QTLs, QTL × environment interactions, and epistasis were important factors influencing the variations in $\delta^{15}N$ in soybean.

The presence of QTLs associated with N_2 fixation and other N-related physiological traits in the genomic regions of identified $\delta^{15}N$ loci were screened in Soybase (www.soybase.org). Loci 3, 7, and 12 co-localized with previously identified QTLs for nodule-related traits such as nodule size (Hwang et al. 2013), nodule number (Shi et al. 2018), and nodule weight (Hwang et al. 2013; Shi et al. 2018). Locus 4 and Locus 10 coincided with ureide QTLs (Ray et al. 2015), and Locus 1 with a shoot N QTL (Dhanapal et al. 2015b) identified in a GWAS panel. Also, Loci 1, 7, and 12 co-localized with QTLs for $\delta^{15}N$ identified in an association study (Steketee et al. 2019). Among these loci, Loci 1, 3, and 12 were located in candidate genes involved in the nodulation process. Gene *N36 (Glyma.01g03470)*, underlying Locus 1, is an early nodulin gene involved in initiation of nodule development (Kouchi and Hata 1993). This gene also plays an important role in translocation of photosynthate into nodule tissue. *Glyma.04g43090*, underlying Locus 3, encodes for a nodulation signaling protein (NSP2), which is a Nod-factor activated transcriptional factor required for nodulation initiation (Murakami et al. 2006). Similarly, the *N315* gene (*Glyma.17g08110*, underlying Locus 12) is expressed at the time of nodule emergence and plays a unique role in nodule formation (Kouchi and Hata 1993).

The $\delta^{15}N$ QTLs at Loci 4 and 12 also overlapped with the genomic regions associated with $\delta^{13}C$ identified in the same population (Bazzer et al. 2020a). Additionally, Loci 1, 3, 8, and 12 coincided with $\delta^{13}C$ QTLs identified in GWAS mapping (Dhanapal et al. 2015a; Kaler et al. 2017). It is tempting to speculate that the greater $\delta^{13}C$ may indirectly lead to increased N_2 fixation under drought due to a greater supply of carbohydrates to nodules. Both water and carbohydrates are supplied to nodules through the phloem (Walsh et al. 1998), and allocation of greater amounts of carbon to nodules during drought is associated with prolonged N_2 fixation (Purcell et al. 1998; King and Purcell 2001). Likewise, among the 13 additive loci associated with $\delta^{15}N$ in the present research, five loci overlap with QTLs previously reported for $\delta^{13}C$ (Fig. 2). Although there was no significant correlation between $\delta^{15}N$ and $\delta^{13}C$ (Bazzer et al. 2020a) among RILs in the present research, both Loci 4 and 12 were associated with $\delta^{15}N$ and $\delta^{13}C$, but the favorable alleles for $\delta^{15}N$ were derived from PI 576201D and for $\delta^{13}C$ from PI 416997. Clearly, further research is needed to resolve the relationship between $\delta^{15}N$ and $\delta^{13}C$ in improving drought tolerance.

Similarly, epistatic QTLs (except QTLs present on Gm02) coincided with QTLs reported for nodule related traits (Hwang et al. 2013; Nicolás et al. 2006; Santos et al. 2013; Shi et al. 2018), ureide concentration (Ray et al. 2015), NDFA (Dhanapal et al. 2015b), and $\delta^{15}N$ (Stekettee et al. 2019). Epistatic QTLs on Gm11, Gm13, and Gm15 were in the genomic regions carrying nodulation genes *Glyma.11g06740, Glyma.13g40400*, and *Glyma.15g05010*, respectively (Schmutz et al. 2010). Epistatic QTLs present on Gm06, Gm11, Gm15, and Gm 17 coincided with

additive QTLs for $\delta^{13}C$ (Bazzer et al. 2020a). In addition, epistatic QTLs also co-localized with $\delta^{13}C$ QTLs found in various studies (Bazzer et al. 2020b; Dhanapal et al. 2015a; Kaler et al. 2017).

The co-localization of identified additive and epistatic $\delta^{15}N$ QTLs with N_2 fixation and WUE related traits supports the evidence that genetic links exist between these traits in soybean. To our knowledge, this is the first study on QTL analysis for $\delta^{15}N$ using a biparental population in soybean. Although many of the QTLs associated with $\delta^{15}N$ were environmentally specific, most of these QTLs co-localized with the QTLs associated with N_2 fixation and other N-related physiological traits (e.g., nodule size, nodule number, nodule weight, NDFA, ureide, and shoot N concentration) which increases the reliability of these $\delta^{15}N$ QTLs/loci for breeding purposes. Similarly, five of the 13 loci identified in this study as being associated with $\delta^{15}N$ are known locations for soybean nodulation genes. These QTLs/loci provide tools for identifying genotypes with favorable alleles for N_2 fixation that can be used in breeding programs to develop germplasm and understand the expression of N_2 fixation genes.

Conclusions

In the present study, a mapping population of 196 F_6-derived RILs was evaluated in multiple environments to understand the genetic basis of $\delta^{15}N$. A wide range of $\delta^{15}N$ values were observed in all environments and narrow sense heritability of $\delta^{15}N$ was low, indicating significant effects of environment on $\delta^{15}N$. Transgressive segregants for $\delta^{15}N$ were observed among the RILs, indicating that it is possible to create, from the specific parents used, extreme genotypes with lower and higher $\delta^{15}N$ values than observed in either parent. Both parents contributed to the higher and lower values observed. The extreme genotypes created in this population may be useful in future studies to better assess the physiological mechanisms of N_2 fixation. QTL analysis by environment identified 10 additive QTLs present on seven chromosomes that individually explained less than 10% of the observed phenotypic variation. Considering QTLs identified across environments, along with individual environments, there were 13 loci for $\delta^{15}N$ based on their overlapping confidence intervals. A lack of consistency of QTL detection was found as QTLs identified in an individual environment did not overlap with QTLs in any of the other environments. Co-localization of $\delta^{15}N$ QTLs with QTLs for important agronomic and physiological traits related to N_2 fixation, and the presence of reported nodulation genes associated with these QTLs, increases the likelihood that the newly identified regions are associated with N_2 fixation. Further studies are needed for fine mapping these QTLs to understand their expression and to determine how they interact with putative candidate genes.

Acknowledgements The authors gratefully acknowledge partial funding of this research from the United Soybean Board (project #2020-172-0116). Appreciation is also extended to Andy King, Marilynn Davies, and Arun Dhanapal for excellent technical support. Philip Handly, Hans Hinrichsen, Matt Kersh, Corey Wilkins, and Angelie Davis were instrumental in the success of the Stoneville trials (ST16 and ST17). Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. This research was supported in part by the U.S. Department of Agriculture, Agricultural Research Service. The USDA is an equal opportunity provider and employer.

References

Adams MA, Turnbull TL, Sprent JI, Buchmann N (2016) Legumes are different: leaf nitrogen, photosynthesis, and water use efficiency. P Natl Acad Sci 113:4098–4103

Andrews M, Lea PJ (2013) Our nitrogen "footprint": the need for increased crop nitrogen use efficiency. Ann Appl Biol 163:165–169

Barrie A, Debney S, Workman CT, Pullan C (1995) Recent developments in high productivity stable isotope analysis. In: International symposium on nuclear and related techniques in soil-plant studies for sustainable agricultural and environmental preservation. Vienna (Austria), 17–21 Oct 1994. IAEA TECDOC Vienna, Austria, pp 29–61

Bazzer SK, Kaler AS, Ray JD, Smith JR, Fritschi FB, Purcell LC (2020a) Identification of quantitative trait loci for carbon isotope ratio ($\delta^{13}C$) in a recombinant inbred population of soybean. Theor Appl Genet 133:2141–2155

Bazzer SK, Kaler AS, King CA, Ray JD, Hwang S, Purcell LC (2019) Mapping and confirmation of quantitative trait loci (QTLs) associated with carbon isotope ratio ($\delta^{13}C$) in soybean. Crop Sci 60:2479–2499

Bergersen FT, Brockwell J, Gault RR, Morthorpe L, Peoples MB, Turner GL (1989) Effects of available soil nitrogen and rates of inoculation on nitrogen fixation by irrigated soybeans and evaluation of delta 15N methods for measurement. Aust J Agric Res 40:763–780

Bondari K (2003) Statistical analysis of genotype × environment interaction in agricultural research. In: Paper SD15,

SESUG: The Proceedings of the South East SAS Users Group, St. Pete Beach, FL

Dhanapal AP, Ray JD, Singh SK, Hoyos-Villegas V, Smith JR, Purcell LC, King A, Cregan PB, Song Q, Fritschi FB (2015a) Genome-wide association study (GWAS) of carbon isotope ratio ($\delta^{13}C$) in diverse soybean [*Glycine max* (L.) Merr.] genotypes. Theor Appl Genet 128:73–91

Dhanapal AP, Ray JD, Singh SK, Hoyos-Villegas V, Smith JR, Purcell LC, King A, Fritschi FB (2015b) Genome-wide association analysis of diverse soybean genotypes reveals novel markers for nitrogen traits. Plant Genome 8:1–15

Djekoun A, Planchon C (1991) Water status effect on dinitrogen fixation and photosynthesis in soybean. Agron J 83:316–322

Doerge RW, Churchill GA (1996) Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285–294

Doughton JA, Saffigna PG, Vallis I, Mayer RJ (1995) Nitrogen fixation in chickpea. II. Comparison of 15N enrichment and 15N natural abundance methods for estimating nitrogen fixation. Aust J Agric Res 45:225–236

Durand JL, Sheehy JE, Minchin FR (1987) Nitrogenase activity, photosynthesis and nodule water potential in soybean plants experiencing water deprivation. J Exp Bot 38:311–321

Farquhar GD, Richards RA (1984) Isotopic composition of plant carbon correlates with water-use efficiency of wheat genotypes. Aust J Plant Physiol 11:539–552

Fehr WR, Caviness CE (1977) Stages of soybean development. Spec. Rep. 80. Iowa State Univ. Coop. Ext. Serv., Ames

Fried M, Broeshart H (1975) An independent measurement of the amount of nitrogen fixed by legume crops. Plant Soil 43:707–711

Fried M, Middelboe V (1977) Measurement of amount of nitrogen fixed by a legume crop. Plant Soil 47:713–715

Giller KE (2001) Nitrogen fixation in tropical cropping systems, 2 edn. CABI, New York

Hardy RWF, Holsten RD, Jackson EK, Burns RC (1968) The acetylene-ethylene assay for N_2 fixation: laboratory and field evaluation. Plant Physiol 43:1185–1207

Holland JB, Nyquist WE, Cervantes-Martinez CT (2003) Estimating and interpreting heritability for plant breeding: an update. Plant Breed Rev 22:9–112

Hwang S, King CA, Davies MK, Ray JD, Cregan PB, Purcell LC (2013) QTL analysis of shoot ureide and nitrogen concentrations in soybean [*Glycine max* (L.) Merr.]. Crop Sci 53:2421–2433

Hwang S, Ray JD, Cregan PB, King A, Davies MK, Purcell LC (2014) Genetics and mapping of quantitative traits for nodule number, weight, and size in soybean (*Glycine max* L. [Merr.]). Euphytica 195:419–434

Heilig JA, Beaver JS, Wright EM, Song Q, Kelly JD (2017) QTL Analysis of symbiotic nitrogen fixation in a black bean population. Crop Sci 57:118–129

Jensen E, Hauggaard-Nielsen H (2003) How can increased use of biological N_2 fixation in agriculture benefit the environment? Plant Soil 252:177–186

Kaler AS, Dhanapal AP, Ray JD, King CA, Fritschi FB, Purcell LC (2017) Genome-wide association mapping of carbon isotope and oxygen isotope ratios in diverse soybean genotypes. Crop Sci 57:1–16

King CA, Purcell LC (2001) Water and carbon allocation to soybean nodules differing in size and the importance to nitrogen fixation response to water deficit. Crop Sci 41:1099–1107

Kohl DH, Shearer GB (1981) The use of soils lightly enriched in ^{15}N to screen for N_2-fixing activity. Plant Soil 60:487–489

Kouchi H, Hata S (1993) Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. Mol Gen Genet 238:106–119

Kuo T, Boersma L (1971) Soil water suction and root temperature effects on nitrogen fixation in soybeans. Agron J 63:901–904

Letolle R (1980) Nitrogen-15 in the natural environment. In: Fritz P, Fontes JC (eds) Handbook of environmental isotope geochemistry. Elsevier, Amsterdam, pp 407–433

Li HH, Ye GY, Wang JK (2007) A modified algorithm for the improvement of composite interval mapping. Genetics 175:361–374

Littell RC, Milliken GA, Stroup WW, Wolfinger RD (1996) SAS system for mixed models. SAS Inst. Inc., Cary

Ludidi NN, Pellny TK, Kiddle G, Dutilleul C, Groten K, Van Heerden PD, Dut S, Powers SJ, Römer P, Foyer CH (2007) Genetic variation in pea (*Pisum sativum* L.) demonstrates the importance of root but not shoot C/N ratios in the control of plant morphology and reveals a unique relationship between shoot length and nodulation intensity. Plant Cell Environ 10:1256–1268

Márquez-García B, Shaw D, Cooper JW, Karpinska B, Quain MD, Makgopa EM, Kunert K, Foyer CH (2015) Redox markers for drought-induced nodule senescence, a process occurring after drought-induced senescence of the lowest leaves in soybean (*Glycine max*). Ann Bot 116(4):497–510

Mastrodomenico AT, Purcell LC (2012) Soybean nitrogen fixation and nitrogen remobilization during reproductive development. Crop Sci 52:1281–1289

Meng L, Li H, Zhang L, Wang J (2015) QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J 3:269–283

Murakami Y, Miwa H, Imaizumi-Anraku H, Kouchi H, Downie JA, Kawaguchi M, Kawasaki S (2006) Positional cloning identifies *Lotus japonicus* NSP2, a putative transcription factor of the GRAS family, required for NIN and ENOD40 gene expression in nodule initiation. DNA Res 13:255–265

Nicolás MF, Hungria M, Arias CAA (2006) Identification of quantitative trait loci controlling nodulation and shoot mass in progenies from two Brazilian soybean cultivars. Field Crops Res 95:355–366

Peoples MB, Faizah AW, Rerkasem B, Herridge DF (1989) Methods for evaluating nitrogen fixation by nodulated legumes in the field. ACIAR Monograph No 11:76

Peoples MB, Boddy RM, Herridge DF (2002) Quantification of nitrogen fixation. In: Leigh GJ (ed) Nitrogen fixation at the millennium. Elsevier, Amsterdam, pp 357–389

Piepho HP, Möhring J, Melchinger AE, Büchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209–228

Purcell LC (2009) Physiological responses of N2 fixation to drought and selecting genotypes for improved N2 fixation. In: Krishnan HB and Emerich DW (eds) Nitrogen fixation

in crop production. Agron Monogr 52 ASA CSSA SSSA, Madison, WI, pp 211–238

Purcell LC, de Silva M, King CA, Kim WH (1998) Biomass accumulation and allocation in soybean associated with differences in tolerance of nitrogen fixation to water deficits. Plant Soil 196:101–113

Ramaekers L, Galeano CH, Garzon N, Vanderleyden J, Blair MW (2013) Identifying quantitative trait loci for symbiotic nitrogen fixation capacity and related traits in common bean. Mol Breed 31:163–180

Ray JD, Dhanapal AP, Singh SK, Hoyos-Villegas V, Smith JR, Purcell LC, King CA, Boykin D, Cregan PB, Song Q, Fritschi FB (2015) Genome-wide association study of ureide concentration in diverse maturity group IV soybean [*Glycine max* (L.) Merr.] accessions. G3 5(11):2391–2403

Rebetzke GJ, Ellis MH, Bonnett DG, Richards RA (2007) Molecular mapping of genes for coleoptile growth in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 114:1173–1183

Reif JC, Maurer HP, Korzun V, Ebmeyer E, Miedaner T, Wurschum T (2011) Mapping QTL with main and epistatic effects underlying grain yield and heading time in soft winter wheat. Theor Appl Genet 123:283–292

Santos MA, Geraldi IO, Garcia AAF, Bortolatto N, Schiavon A, Hungria M (2013) Mapping of QTLs associated with biological nitrogen fixation traits in soybean. Hereditas 150:17–25

SAS Institute (2013) The SAS system for Windows. In: Release 9.4. SAS Inst., Cary

Schmutz J, Cannon S, Schlueter J, Ma J, Mitros T, Nelson W, Hyten D, Song Q, Thelen J, Cheng J, Xu D, Hellsten Uffe, May G, Yu Y, Sakurai T, Umezawa T, Bhattacharyya M, Sandhu D, Valliyodan B, Jackson S (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178–183

Serraj R, Sinclair TR, Purcell LC (1999a) Symbiotic N_2 fixation response to drought. J Exp Bot 50:143–155

Serraj R, Vadez V, Denison RF, Sinclair TR (1999b) Involvement of ureides in nitrogen fixation inhibition in soybean. Plant Physiol 119:289–296

Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52:591–611

Shearer G, Kohl D (1986) N_2 fixation in field settings: estimations based on natural ^{15}N abundance. Funct Plant Biol 13:699–756

Shi X, Yan L, Yang CY, Yan WW, Moseley DO, Wang T, Liu BQ, Di R, Chen PY, Zhang MC (2018) Identification of a major quantitative trait locus underlying salt tolerance in 'Jidou 12' soybean cultivar. BMC Res Notes 11:95

Sinclair TR, Muchow RC, Bennett JM, Hammond LC (1987) Relative sensitivity of nitrogen and biomass accumulation to drought in field-grown soybean. Agron J 79:986–991

Steketee CJ, Sinclair TR, Riar M, Schapaugh WT, Li Z (2019) Unraveling the genetic architecture for carbon and nitrogen related traits and leaf hydraulic conductance in soybean using genome-wide association analyses. BMC Genom 20:811

Strodtman KN, Emerich DW (2009) Nodule metabolism. In: Emerich DW, Krishnan HB (eds) Nitrogen fixation in crop production. Crop Science Society of America, Madison, pp 95–124

Tanya P, Srinives P, Toojinda T, Vanavichit A, Lee S-H (2005) Identification of SSR markers associated with N_2 fixation components in soybean [*Glycine max* (L.) Merr.]. Korean J Genet 27:351–359

Unkovich M, Herridge D, Peoples M, Cadisch G, Boddey R, Giller K, Alves B, Chalk P (2008) Measuring plant-associated nitrogen fixation in agricultural systems. ACIAR Monograph No. 136, Canberra, Australia, pp 85–124

Vieira A, de Oliveira A, Soares T, Schuster I, Piovesan N, Martinez C, de Barros E, Moreira M (2006) Use of the QTL approach to the study of soybean trait relationships in two populations of recombinant inbred lines at the F7 and F8 generations. Braz J Plant Physiol 18:281–290

Walsh KB, Canny MJ, Layzell DB (1989) Vascular transport and soybean nodule function. II: a role for phloem supply in product export. Plant Cell Environ 12:713–723

Weaver RW, Danso SKA (1994) Dinitrogen fixation. In: Weaver RW, Angle S, Bottomley P, Bezdicek D, Tabatabai A, Wollum A (eds) Methods of soil analysis. Part 2, Microbiological and biochemical properties. Soil Science Society of America, Madison, pp 1019–1045