Genetic Factors Associated With Nodulation and Nitrogen Derived From Atmosphere in a Middle American Common Bean Panel

Atena Oladzad1*, Abiezer González2, Raul Macchiavelli2, Consuelo Estevez de Jensen2, James Beaver2, Tim Porch3 and Phillip McClean1*  

1Department of Plant Sciences, North Dakota State University, Fargo, ND, United States, 2Department of Agroenvironmental Sciences, University of Puerto Rico, Mayagüez, Puerto Rico, 3USDA-ARS, Tropical Agriculture Research Station, Mayagüez, Puerto Rico

Among grain legume crops, common beans (Phaseolus vulgaris L.) are considered to have poor biological nitrogen (N₂) fixation (BNF) capabilities although variation in N₂ fixing capabilities exists within the species. The availability of genetic panel varying in BNF capacity and a large-scale single nucleotide polymorphism (SNP) data set for common bean provided an opportunity to discover genetic factors associated with N₂ fixation among genotypes in the Middle American gene pool. Using nodulation and percentage of N₂-derived from atmosphere (%NDFA) data collected from field trials, at least 11 genotypes with higher levels of BNF capacity were identified. Genome-wide association studies (GWASs) detected both major and minor effects that control these traits. A major nodulation interval at Pv06:28.0–28.27 Mbp was discovered. In this interval, the peak SNP was located within a small GTPase that positively regulates cellular polarity and growth of root hair tips. Located 20 kb upstream of this peak SNP is an auxin-responsive factor AUX/indole acetic auxin (IAA)-related gene involved in auxin transportation during root nodulation. For %NDFA, nitrate (NO₃⁻) transporters, NRT1:2 and NRT1.7 (Pv02:8.64), squamosa promoter binding transcriptome factor (Pv08:28.42), and multi-antimicrobial extrusion protein (MATE) efflux family protein (Pv06:10.91) were identified as candidate genes. Three additional QTLs were identified on chromosomes Pv03:5.24, Pv09:25.89, and Pv11: 32.89 Mbp. These key candidate genes from both traits were integrated with previous results on N₂ fixation to describe a BNF pathway.

Keywords: Phaseolus vulgaris, RAB GTPase, biological nitrogen fixation pathway, NRT1 genes, MATE efflux proteins, genome-wide association study, AUX/IAA family

INTRODUCTION

Common bean (Phaseolus vulgaris L.) is a major crop of smallholder farmers in Latin America, the Caribbean, and Eastern and Southern Africa. In these regions, common beans are often grown on soils deficient in nitrogen (N₂; Fenta et al., 2020) and among grain legumes, the species has a relatively low biological N₂ fixation (BNF) capacity (Mnasri et al., 2007).
The application of N\textsubscript{2} fertilizer can increase bean yields, but the input is often too expensive and/or unavailable to smallholder farmers.

Biological N\textsubscript{2} fixation is a process in which rhizobia microorganisms in the soil convert atmospheric N\textsubscript{2} to ammonia (NH\textsubscript{3}). This process generates a usable form of nitrogenous compounds for the plant in an economical and environmentally friendly manner through nodulation (Akter et al., 2018). Nodulation is a symbiotic interaction between rhizobia and the legume plant in which the plant provides carbohydrate to the rhizobial species, and the bacteria convert atmospheric N\textsubscript{2} into ammonium for use by the plant (Wang et al., 2018). Although common bean is considered a poor N fixer, genetic diversity for N\textsubscript{2} fixation exists within the species. This diversity enables the selection of germplasm with enhanced BNF capacity (Buttery et al., 1987; Barron et al., 1999; Farid et al., 2017; Kamfwa et al., 2019). Pereira et al. (1993) were able to significantly increase the nodule number of common bean lines after three cycles of recurrent selection in a controlled environment. The results demonstrated that nodule number is a heritable trait, and Bliss (1993) suggested that nodule number could be used as a criterion to estimate BNF capacity.

The identification of key genetic factors associated with nodulation and the physiological response in a diverse collection of germplasm may enable the development of molecular markers to select for enhanced levels of N fixation capacity. Previously, genetic mapping (Nodari et al., 1993; Heilig et al., 2017; Kamfwa et al., 2019) and transcriptional regulation studies of the BNF response (Blanco et al., 2009) provided valuable information regarding BNF in common bean. However, few genetic studies to understand the genetic regulation of BNF across a diverse panel of common bean germplasm have been conducted.

Selecting bean lines with greater BNF capacity in breeding nurseries with high N\textsubscript{2} fertilizer soil amendments negates the plant's need for effective BNF capacity. Although quantifying BNF itself is difficult, studying related traits, such as seed yield, photosynthesis, biomass partitioning, root morphology, plasticity, and nodulation, along with environmental effects, can provide a framework for the genetic improvement of BNF in common bean (Diaz et al., 2017; Fenta et al., 2020).

Initially, the lack of a simple, rapid, and economical method to measure BNF capacity limited genetic improvement (Purcell, 2009). The discovery that BNF capacity in common bean is related to root nodule number and weight (Pereira et al., 1993), enabled the development of an effective screening procedure. BNF capacity can also be indirectly quantified by calculating percentage of N\textsubscript{2} derived from atmosphere (%NDFA). This procedure compares the amount of \textsuperscript{15}N (heavy N\textsubscript{2} isotope) for a genotype relative to the \textsuperscript{15}N value of a control genotype that does not have the capacity to nodulate and fix N\textsubscript{2}. This value, known as \begin{math}\delta^{15}N\end{math}, is obtained from the ratio of \textsuperscript{15}N: \textsuperscript{14}N. The abundance of \textsuperscript{15}N in different parts of a legume plant can be obtained by calculating the relative deviation from the \begin{math}\delta^{15}N\end{math} ratio. In a recent study, the significant genetic variation among recombinant inbred lines (RILs) derived from Middle American genotypes for %NDFA demonstrated that the genetic variation among RILs would be mostly additive (Farid et al., 2017). Furthermore, significant QTL for %NDFA was detected within Andean gene pool (Ramaekers et al., 2013). These significant and repeatable differences between lines in both common bean gene pools suggest additive genetic effects play important role for %NDFA. Previous studies in soybean (Glycine max L. Merr.) confirmed that the differences in \begin{math}\delta^{15}N\end{math} between nodulated roots from different lines or tissues are a robust means for quantifying %NDFA (Wanek and Arndt, 2002). Polania et al. (2016) noted that the natural abundance of \textsuperscript{15}N estimated using seed tissue would be an appropriate estimate for BNF capacity of common bean.

Given the importance of the nodule numbers and %NDFA in BNF estimation, genetic mapping these two traits may lead to a better understanding of the genetic factors affecting BNF capacity in common bean. In this study, we evaluated a panel specifically developed for abiotic stress tolerance. The aim of this study was to: (1) characterize common bean genotypes with high capacity for N\textsubscript{2}-fixation in tropical soil conditions and (2) discover the genomic regions and candidate genes associated with nodulation and %NDFA under tropical soil conditions.

**MATERIALS AND METHODS**

Bean Abiotic Stress Evaluation (BASE_120) field trials (Oladzad et al., 2019) were conducted at the Isabela Substation of the Agricultural Experiment Station of the University of Puerto Rico. The substation is located on the northwestern coastal plain of Puerto Rico at 18.468 N, −67.042 W at an altitude of 128 m above the sea level. The average annual minimum/maximum temperature at the Isabela Experimental Substation is 22.2/27.8°C, with an average annual rainfall of 1,630 mm. The trial was conducted in a field with Coto Clay soil, a very-fine, kaolinitic, isohyperthermic Typic Eutrustox. Soil subsamples were taken at random from different positions in the fields, where the BASE_120 trials were conducted. The subsamples were mixed to prepare a composite sample for each field. The soil analyses were conducted at AgSource Harris Laboratories in Lincoln, NE (Supplementary Table S1).

Field trials were planted in a randomized complete block design with five replications in four growing seasons; June and November 2015, June 2016, and June 2018. The experimental units were single 3-m rows with 0.76-m spacing between rows. Each BASE_120 trial included 118 common bean and two tepary bean (Phaseolus acutifolius L.) genotypes (Supplementary Table S2). Entries in the BASE_120 trials included elite bean breeding lines and cultivars from Zamorano University in Honduras, the International Center for Tropical Agriculture (CIAT) in Colombia, the University of Puerto Rico, USDA-ARS Tropical Agriculture Research Station in Puerto Rico, and Michigan State University. The BASE_120 genotypes were almost exclusively of race Mesoamerican origin, while including the Andean cultivars “Calima,” “Indeterminate Jamaica Red” (IJR), and “Quimbaya.” The seeding rate was 14 seeds per linear meter. The field trial received supplementary aerial irrigation to avoid drought stress.
Weeds were controlled manually. Standard agronomic practices for weed and insect control were used in the trials.

No fertilizer was applied to the BASE_120 trials in order to produce soil conditions that favor nodule formation and symbiotic N₂ fixation. In the June 2015 trial, seeds were inoculated with a liquid suspension (10⁷ rhizobium cells per milliliter) of two strains of *Rhizobium*: *R. etli* (CIAT 632) and *R. tropici* (CIAT 899). The use of inoculant containing a mixture of more than one strain of *Rhizobium* is a recommended agronomic practice in common bean (Liu et al., 2020a,b). The inoculant was applied to the seed pre-planting using a sticker, PREMAX (1 ml/kg of seed), and the same CIAT 632 and CIAT 899 inoculant (100 g/kg of seed) by stirring the inoculant and the seed of each genotype in a container. The inoculum was also applied directly to the seeds in the row at planting, and to the base of the seedlings 1 week after planting, using a backpack sprayer. The inoculations for the second, third, and fourth growing seasons were performed using a peat-based inoculant.

Two plants per experimental unit were carefully dug up about 0.75 m from the end of the second row approximately 45 days after planting during flowering. Root nodule numbers were evaluated by assigning values from 1 to 9 using the CIAT scale (Schoonhoven and Pastor-Corrales, 1987), where 1 indicates that the root has >80 nodules, 3 = 41–80 nodules, 5 = 21–40 nodules, 7 = 10–20 nodules, and 9 = less than 10 nodules per plant. Seed samples from three replications of the trials conducted at the Isabela Substation in 2016 and 2018 were used to estimate %NDFA. Approximately five seeds of each sample were dried at 70°C until a constant weight was achieved; then ground using a Wiley mini-mill (Thomas Scientific, Swedesboro, New Jersey, United States), and passed through a #40 mesh sieve, resulting in a fine powder. Samples containing 4.2 mg of the seed powder were packaged in 5 × 8 mm tin capsules (D1008, EA Consumables, Pennsauken, NJ, United States) and shipped to the University of California, Davis Stable Isotope Facility in 96-well plates. The ¹⁵N natural abundance method (Unkovich et al., 2008) was used to calculate %NDFA. The white bean R-99 was used as the non-N-fixing reference line. %NDFA was calculated as:

\[
%\text{NDFA} = \frac{\delta^{15}N_{\text{R99}} - \delta^{15}N_{\text{BASE,locus}}}{\delta^{15}N_{\text{R99}} - \beta} \times 100
\]

where \(\beta\) represents the \(\delta^{15}N\) of the bean line grown under N-free conditions and relies on BNF for all N requirements. In this study, \(\beta\) was assigned the value of 0 because the R99 reference line does not nodulate. It was assumed that there was no isotopic fractionation in seed harvested from the genotypes grown under N-free conditions. The entries in the BASE_120 trial and the R99 reference bean line trial have similar patterns of phenological development and seed size.

Statistical analyses were performed for nodule scores and %NDFA estimates using a GLIMMIX model from SAS/STAT14.1 (SAS Institute, Cary, NC, United States) to fit the model, estimate the main effects and interactions, and to compare the least squares (LS) means. Planting dates and bean genotypes were considered fixed effects, whereas replication within planting date was considered a random effect. A 95% probability level was used to establish statistical differences.

A genome-wide association study (GWAS) was performed using the phenotypic data for the BASE_120 genotypes, and their corresponding SNPs. The Base_120 was genotyped with SNPs \((n = 2,052,923)\) generated from genotype-by-sequencing (GBS) reads of 469 Middle American genotypes (Oladzad et al., 2019). Only SNPs with minor allele frequencies ≥5% were extracted from this HapMap for the GWAS analysis using the GEMMA software (Zhou and Stephens, 2012, 2014). The kinship matrix, obtained from the centered relatedness algorithm, was included in the GWAS model as a random effect, and the structure matrix, obtained from the principal component analysis (PCA) using the Pcom function in R3.5, was considered a fixed effect (Price et al., 2006). The smallest number of principle components that accounted for at least 25% of the cumulative variation was used in the mixed GWAS model. Those SNPs whose values of \(p\) were in the lower 0.01 and 0.05% tail of the bootstrap distribution were considered significant. Bootstrapping was performed in R with 1,000 resampling. This approach is sensitive to the fact that the more genetic factors affecting a phenotype, the corresponding values of \(p\) would be higher. Artificial thresholds work best for traits in which a few genetic factors are involved (such as disease resistance loci). These will typically pass the conservative Bonferroni test. But for other traits, we believe the best way is to define the cut-off value based on the trait and the specific population under study. That is why in the current study, 1,000 permutation tests and a \(p\) value bootstrap of two cutoff values were used. The amount of phenotypic variation explained by these significant SNPs was estimated by a likelihood-ratio-based (R², LR) analysis in R using the GenABEL package that included population structure and/or relatedness effects (Aulchenko et al., 2007; Sun et al., 2010). While GWAS results are indeed related to LD, it is important to note (and often little appreciated) that LD is a function of recombination. Genomic studies clearly show that recombination is not equal across the genome, especially as it relates to LD. For example, Moghadam et al. (2016) clearly demonstrated LD must be considered in a regional context, and that regional context is population specific. The regional and population differences mean that LD values averaged across chromosomes regions, across whole chromosomes, and across populations have little value. What is important is the degree of LD in the narrow region, where we find the genetic factors associated with a specific phenotype. The question is what is the solution with regard to selecting a window size for candidate gene selection? The only true solution is to calculate the Mb per cM ratio around each genetic signal discovered by the GWAS analysis and choosing a specific cM distance to define the physical boundaries for candidate gene selection. The challenge though is that we do not know what genetic map to use since each biparental parent will have a different recombination frequency in the region of interest. Therefore, choosing a window size (admittedly somewhat arbitrarily), searching for a gene(s) in that region, and, based on previous analysis of the function of the gene as outlined in referred literature, makes a case for its possibility as a
candidate gene. In the current study, candidate genes were identified within a conservative window size of ±50 kb interval of the significant SNPs. The possible effect of SNPs located inside a gene model or its promoter region was investigated using the SNPeff (Cingolani et al., 2012) database for common bean reference genome sequence version 2 (Oladzad et al., 2019). Finally, to check whether significant SNP markers near or within the major candidate genes associated with both traits, exhibit hitchhiking effects due to extensive LD, the pairwise LD in null model, using Tassel v5.0 (Bradbury et al., 2007), as well as in mixed model accounting for both kinship and population structure using R package LDcorSV (Mangin et al., 2012) were estimated.

RESULTS

The BNF capacity of a legume is associated with nodule-mediated nitrate (NO$_3$) production and general NO$_3$ uptake and mobilization. Integrating nodulation and %NDFA data with genotyping data in a genetic analysis enabled the identification of genomic intervals and candidate genes associated with BNF capacity and N metabolism under tropical soil conditions in common bean.

Phenotypic Analyses of %NDFA Estimates and Nodulation Scores

Significant differences between seasons and among entries were found for %NDFA (Table 1). Although an overall season × entry interaction for %NDFA was not observed, several entries in the BASE_120 trial had significant F tests for the season × entry LS means slice (Supplementary Table S3). LS means for %NDFA from the June 2018 planting were generally lower than the LS means from the June 2016 planting. The range of LS means was comparable to %NDFA estimates for common bean in previous studies based on seed tissue (Polania et al., 2016). The BASE_120 control variety “ICA Pijao” had among the highest estimated %NDFA for both growing seasons.

Based on results from previous studies, Polania et al. (2016) estimated that the average %NDFA for common bean was 39%. Several BASE_120 lines including “Beníquez,” FBN 1205-31, FBN 1210-48, “INTA Centro Sur,” “Paraisito,” PR0443-151, PR1418-15, and SB2-4. Among the 112 lines with data for both traits, the Pearson correlation between mean %NDFA and mean nodulation score was highly significant ($r = 0.60$) (Diaz et al., 2017) reported that %NDFA was correlated with seed yield and suggested that enhanced BNF may be beneficial under abiotic stress.

| TABLE 1 | Type III tests of fixed effects for %NDFA and nodulation scores from bean abiotic stress evaluation (BASE) 120 field trials conducted at Isabela, Puerto Rico from 2016 to 2018. |
| Effect | Num DF | Den DF | %NDFA | Nodulation score (1–9) |
| Season | 1 | 4 | 38.25 | 8 |
| Entry | 111 | 426 | 1.61 | 949 |
| Season × entry | 111 | 426 | 0.95 | 949 |

Association Mapping and Identification of Candidate Genes

After filtering for MAF ≥ 0.05, a total of 1,25,745 SNPs out of 2,11,764 were extracted from the original HapMap (Oladzad et al., 2019). Two principal components explained 29.1% of the genetic variation and were used in the mixed linear model analysis. For nodulation score, a block of SNPs in high LD located on Pv06 between 28.19 and 28.25 in the 0.01% cutoff threshold were associated $[\log_{10}(P) = 7.46]$ with the trait (Figure 1A). These SNPs explained 11% of the phenotypic variation. When SNPs in the 0.05% cutoff threshold were considered, the interval extended to 28.73 Mbp and cumulatively explained 13% of the phenotypic variation (Tables 2, 3). A total of 21 gene models were detected ±50 kb of this block. Four SNPs in complete LD at Pv06:28.198 Mbp; $[-\log_{10}(P) = 7.54]$ were located inside the gene model Phvul.006G18000, an ortholog of the Arabidopsis RAB GTPase. The auxin-responsive factor AUX/indole acetic auxin (IAA)-related gene model (Phvul.006G181200) was located 20 Kb upstream of RAB GTPase related gene model.
For %NDFA, a total of 25 smaller effect, significant intervals containing 53 SNPs were detected in the 0.05% cutoff threshold (Figure 1B). The peak SNPs among these intervals cumulatively explained 36% of the phenotypic variation in %NDFA (Tables 2, 4). The peak SNP \([-\log_{10}(P) > 5.31]\) was detected at P09:25.89 Mbp and was located \(\sim 189\) kb upstream of another annotated RAB GTPase gene. The second most significant SNP interval was 98 bp in length at P08:28.42 Mbp and was located \(\sim 2.5\) kb upstream region of gene model Phvul.008G157200, an ortholog of a squamosa promoter binding protein-like 8 (SPL8). The SNPEff results revealed that this SNP was placed in the promoter regions of SPL8 and potentially has a modifier effect. A SNP at P11:32.89 Mbp was detected 17 Mbp downstream of a small auxin up RNA (SAUR)-like auxin responsive protein family cluster. NO\(_3\) transporters, NRT1:7 and NRT1:2 (orthologs to gene models Phvul.002G067700 and Phvul.002G067214, respectively) flanked the P02:8.64 Mbp peak SNP within the heterochromatic region of P02. Finally, the P06:10.91 Mbp SNP interval is located \(\sim 40\) kb downstream of a MATE efflux protein (gene model; Phvul.006G028701) known to be involved in flavonoid metabolism (Marinova et al., 2007), which is also critical for nodule development (Ng et al., 2015). Despite the importance of these candidate genes potentially involved in N metabolism, due to the quantitative nature of these traits, all the gene models within \(\pm 50\) kb of the significant intervals were presented in the Supplementary Table S4. Furthermore, for more powerful trait discovery, single GWAS analyses were performed on phenotypic data from each season and year. The results and Manhattan plots are presented in the Supplementary Table S5 and Supplementary Figure S1. The results showed that the significant SNPs on P06 for nodulation which was detected from Ls-mean values within a LD block, was also identified in 3 seasons/years (November 2015, June 2018, and June 2015) and interestingly the significant SNP on P02; 14.33Mbp was identified in both nodulation and %NDFA in June 2016 and June 2018, respectively. This SNP was upstream of gene model Phvul.002G088100. Furthermore, the most significant SNPs on P09, P08, P06, and P11 that were detected from Ls-mean values for %NDFA, were also identified in the single GWAS analysis of June 2016 and June 2018. The pairwise LD values were reported as squared allele frequency correlation in the Supplementary Table S6 for both null model and mixed model accounting for both population structure and relatedness. Based on these values, the significant SNPs are not in high LD specifically when population structure and relatedness are accounted in the model. However, some degree of LD was observed between P08:28.42 and two other SNPs on P02:8.64 and P09:25.89Mbp in the null model and may very well be due to the gene by gene interactions within this pathway. As a final point, for each peak SNP, the allele associated with higher %NDFA and lower nodulation scores (higher number of nodules) were determined, and the mean of each trait of all possible haplotypes was calculated (Table 5). Selecting the genotypes for negative alleles for nodulation and positive alleles for %NDFA revealed that although both nodulation and %NDFA are indicators of BNF capacity, different genetic factors control each of these traits, and selection based on marker data should be independent for each trait (Supplementary Table S7).

**DISCUSSION**

The contribution of BNF to common bean growth and development is key to its performance as a field crop, especially under field stress conditions. While it has been demonstrated...
previously that degree of nodulation and %NDFA are two key phenotypes related to BNF, the underlying genetic factors that control these traits were not studied before. In this study, a combination of detailed phenotypic data of a population with modern breeding alleles and in-depth genotyping of that population enabled the discovery of key genetic factors for nodulation and %NDFA. The individual genetic analysis for each trait was robust. These genetic analyses also enabled a description of the interactions of a genetic pathway that controls BNF.
The phenotypic analysis confirmed that the range of LS means for seed %NDFA in this common bean population was comparable to previous studies based on seed tissue (Polania et al., 2016). While nodulation and N\textsubscript{2} fixation in common beans are sensitive to high temperatures (Hungria and Franco, 1993) and inadequate soil moisture (Hungria and Vargas, 2000; Polania et al., 2016) two conditions that affected the variability in nodulation phenotypes/scores in this study, the consistency of the range of phenotypic when compared with previous studies suggests the data was representative of the modern germplasm and appropriate for performance and genetic analysis. This study identified bean lines capable of obtaining a significant portion of the N needed for growth and development from BNF. The BASE\_120 trials were conducted in soils at Isabela, Puerto Rico that have low levels N and P. Therefore, these results should be relevant to small-scale bean producers in the tropics, who often do not have access to fertilizer inputs. Moreover, Heilig et al. (2017) evaluated %NDFA in a RIL population derived from the cross “Puebla 152 x Zorro” (Mesoamerican) at Isabela, Puerto Rico during 2012 and 2013. This study was conducted under conditions similar to the BASE 120 field trials in our study; an unfertilized oxisol (Coto Clay). Significant differences were observed among RILs for %NDFA. Although narrow sense heritability was not estimated, the significant difference among RILs suggests that additive genetic effects were important in the expression of %NDFA in this “Puebla x 152 x Zorro” RIL population.

### TABLE 5 | Efficiency of the selected markers.

| Trait | SNP locus | Allele | Mean positive allele | Mean negative allele | p value |
|-------|-----------|--------|----------------------|----------------------|---------|
| Nodulation | S06\_28200687 | C/G | 6.84 | 7.37 | 3.46E-08 |
| % NDFA | S09\_25892359 | C/T | 38.03 | 49.78 | 4.83E-06 |
| | S09\_28424570 | C/G | 33.08 | 50.88 | 2.81E-05 |
| | S06\_10913425 | C/T | 31.15 | 51.81 | 6.62E-05 |
| | S02\_8646280 | G/T | 36.91 | 40.32 | 3.30E-04 |

Data of %NDFA are used as a proxy for “N\textsuperscript{15} natural abundance” to quantify the contribution of N\textsubscript{2} fixation in common bean to N\textsubscript{2} metabolism. The GWAS results of that data identified two candidate genes related to nodulation, SPL8, and SAUR-like auxin responsive protein, and two candidate genes were related to NO\textsubscript{3} transport, NRT1.2 and NRT1.7. The SPL gene family is a target gene for miR156 (Yu et al., 2015), which affects the regulation of nodule development in plants (Wang and Wang, 2015). Overexpression of this miRNA downregulates nodulation in legumes by silencing downstream SPL genes. In both common bean and soybean, it has been shown that SPL activates the transcription of miR172, which is highly accumulated in the mature nodules. miR172 binds APETALA2 (AP2), protein positively nodule senescence genes in common bean. In mature nodules, AP2-1 is silenced by
miR172 (Fíguez et al., 2015; Nova-Franco et al., 2015). In soybean, it has been demonstrated that AP2-2 controls the non-symbiotic hemoglobin expression (Hb1), which affects the level of nodulation by regulating nitrogenase activity (Yan et al., 2013). However, it is not clear which subfamily of SPLs in this regulatory network is associated with nodulation in common bean. SAURs were also detected as potential candidate genes. These are the largest families of early auxin response genes (Ren and Gray, 2015), and auxin families are associated with root nodule symbiosis.

Nitrate transporters NRT1.2 and NRT1.7 are both members of the NRT1 NO$_3^−$ transporter subfamily system (later renamed the NPF family Lérant et al., 2014) that transfer a wide range of molecules including auxin and nitrates (Fan et al., 2009). The NPF family uses a proton gradient to transport NO$_3^−$ from the soil to different parts of the plant (Segonzac et al., 2007). The NRT1.2 mRNA was found in root hairs and in the epidermis of young and mature roots (Huang et al., 1999; Guo et al., 2002), while NRT1.7 mRNA was located in the phloem tissue of older leaves. This suggests a role of NRT1.7 in the remobilization of NO$_3^−$ from source to sink. This role was further confirmed by analyzing the N$^{15}$ content of the older and younger leaves in wild (NRT1.7) and mutant (nrt1.7) Arabidopsis (Fan et al., 2009). In wild type genotypes, N$^{15}$ was remobilized from older leaves to the younger leaves, while in the mutant; higher amount of NO$_3^−$ was accumulated in older leaves and was not transferred to the younger tissues.

Integrating the information from both nodulation score and %NDFA shows that the SNPs which mapped in or adjacent to candidate genes are associated with BNF in common bean. Figure 2 integrates our genetic results into a N$_2$ metabolism pathway based on previous research to show how the candidate genes interact to provide a high level of N$_2$ fixation in low fertile soil. For this genotype, rhizobacteria infection triggers the expression of the MATE efflux gene involves in flavonoid metabolism. Flavonoids prompt the transcription of genes related to the biosynthesis of the NFs in the soil which are perceived by the LRR domain of plasma membrane-localized receptor-like kinases (RLKs; Ferguson et al., 2010; Kamfwa et al., 2015). This complex would interact with a RAB family GTpase (Choudhury and Pandey, 2016), which alters the hair root cell polarity, resulting in nodule induction. The expression of auxin transporters from shoot to root also triggers flavonoids release, which would also lead to high auxin accumulation in the root resulting in greater lateral root development and consequently greater nodule number. Moreover, upon Rhizobia infection, the expression of SPL would increase miR172 levels until nodulation reaches it maximum accumulation in the mature plant. Higher levels of miR172 also improve root growth which subsequently increases the rhizobia infection and therefore a higher expression of nodulation regulatory-related genes. This leads to further nodule as the plant matures (Nova-Franco et al., 2015). The overall effect of this integrated genetic pathway is an increase in N fixation capacity. The ammonium (NH$_4^+$) produced in this process can be converted to NO$_3^−$ by nitrifying bacteria. The NO$_3^−$ will then be transported from root to shoot by NRT1.2, and later, remobilized by NRT1.7 from older leaves to the younger leaves. During plant maturation, NO$_3^−$ will be distributed to the seeds which increases %NDFA (Diaz et al., 2017). As a final point, it should not be rolled out that despite the statistical power of GWAS discovery of the candidate genes for complex traits, still there is potential for missing medium and low-effect associations specifically after filtering for allele frequency which is commonly employed in GWAS analysis. It has been suggested that sequencing the candidate genes specifically in individuals with extreme phenotypes of a quantitative traits may be able to detect other associations but the sample size at the extremes of a quantitative trait is

![Figure 2](www.frontiersin.org)
also debatable because it can be affected by small odds ratios (Manolio et al., 2009). However, using GWAS to introduce a breeder “friendly” marker was beyond the scope of this study, rather it aimed using GWAS as a genetic approach to understand the molecular pathways.

**DATA AVAILABILITY STATEMENT**

The original contributions presented in the study are included in the article/publication material, further inquiries can be directed to the corresponding authors.

**AUTHOR CONTRIBUTIONS**

AO, AG, RM, and CJ performed data analysis and JB, TP, and PM supervised the data analysis. JB and TP designed the phenotypic experiment and AO and PM designed the genomics approach for the association mapping. AO, AG, and RM discussed the results and the interpretation of the final data and JB, TP, and PM provided suggestions to improve it. All authors contributed to the article and approved the submitted version.

**FUNDING**

This research was supported by USAID Feed the Future Innovation Laboratory for Climate-Resilient Beans through grant USAID OAA-A-13-00077.

**ACKNOWLEDGMENTS**

We would like to thank Rian Lee, North Dakota State University, for his lab support.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.576078/full#supplementary-material

**REFERENCES**

Akter, Z., Pageni, B., Lupwayi, N. Z., and Balasubramanian, P. (2018). Biological nitrogen fixation by irrigated dry bean (Phaseolus vulgaris L.) genotypes. *Can. J. Plant Sci.* 98, 1159–1167. doi: 10.1037/cjs-2017-0301

Atzorn, R., Crozier, A., Wheeler, C., and Sandberg, G. (1988). Production of gibberellins and indole-3-acetic acid by rhizobium phaseoli in relation to nodulation of Phaseolus vulgaris roots. *Planta* 175, 532–538. doi: 10.1007/BF00393076

Aulchenko, Y. S., Ripke, S., Isaacs, A., and Van Duijn, C. M. (2007). GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 23, 1294–1296. doi: 10.1093/bioinformatics/btm108

Barraza, A., Coss-Navarrete, E. L., Vizuete-De-Rueda, J. C., Martinez-Aguilar, K., Hernández-Chávez, J. L., Ordaz-Ortiz, J. J., et al. (2018). Down-regulation of PvtRXRxl increases nodule number and affects auxin, starch, and metabolic fingerprints in the common bean (*Phaseolus vulgaris* L.). *Plant Sci.* 274, 45–58. doi: 10.1016/plantsci.2018.05.006

Barron, J. E., Pasini, R., Davis, D., Stuthman, D., and Graham, P. (1999). Response to selection for seed yield and nitrogen (N2) fixation in common bean (*Phaseolus vulgaris* L.). *Field Crop Res.* 62, 119–128. doi: 10.1016/S0378-4290(99)00009-X

Blanco, F. A., Meschini, E. P., Zanetti, M. E., and Aguilar, O. M. (2009). A small GTPase of the Rab family is required for root hair formation and nitrogen source modulated pH fluctuations. *Plant Signal. Behav.* 4, 76–77. doi: 10.4161/psb.6.3.14523

Borg, S., Brandstrup, B., Jensen, T. J., and Poulsen, C. (1997). Identification of new protein species among 33 different small GTP-binding proteins encoded by cDNAs from Lotus japonicus, and expression of corresponding mRNAs in developing root nodules. *Plant J.* 11, 237–250. doi: 10.1046/j.1365-313X.1997.11020237.x

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Buttery, B., Park, S., and Findlay, W. (1987). Growth and yield of white bean (*Phaseolus vulgaris L.*) in response to nitrogen, phosphorus and potassium fertilizer and to inoculation with rhizobium. *Can. J. Plant Sci.* 67, 425–432. doi: 10.1111/j.1744-7909.2010.00899.x

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnPeff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80–92. doi: 10.1016/j.fly.19695

Coronado, C., Zuanazzi, J. S., Sallaud, C., Quirion, J.-C., Esnault, R., Husson, H.-P., et al. (1995). Alfalfa root flavonoid production is nitrogen regulated. *Plant Physiol.* 108, 533–542. doi: 10.1104/pp.108.2.533

Diaz, L. M., Ricuartre, J., Cajas, C., Galeano, C. H., Rao, L., Beebe, S., et al. (2017). Phenotypic evaluation and QTL analysis of yield and symbiotic nitrogen fixation in a common bean population grown with two levels of phosphorus supply. *Mol. Breed.* 37:76. doi: 10.1007/s11032-017-0673-1

Fan, S.-C., Lin, C.-S., Hsu, P.-K., Lin, S.-H., and Tsay, Y.-F. (2009). The Arabidopsis nitrate transporter NRT1.1, expressed in phloem, is responsible for source-to-sink remobilization of nitrate. *Plant Cell* 21, 2750–2761. doi: 10.1105/tpc.109.076703

Farid, M., Earl, H. J., Pauls, K. P., and Navabi, A. (2017). Response to selection for improved nitrogen fixation in common bean (*Phaseolus vulgaris L.*). *Euphytica* 213:99. doi: 10.1007/s10681-017-1885-5

Fenta, B. A., Beebe, S. E., and Kunert, K. J. (2020). Role of fixing nitrogen in common bean grown under water deficit conditions. *Food and Energy Security* 9:e183. doi: 10.1002/foes.183

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M. H., Lin, Y.-H., Reid, D. E., et al. (2016). Heterotrimeric G-protein complex (G protein-coupled receptor 57, G protein-coupled receptor 79) expressed in phloem, is responsible for source-to-sink remobilization of nitrate. *Plant Cell* 21, 2750–2761. doi: 10.1105/tpc.109.076703

Frid, M., Earl, H. J., Pauls, K. P., and Navabi, A. (2017). Response to selection for improved nitrogen fixation in common bean (*Phaseolus vulgaris L.*). *Euphytica* 213:99. doi: 10.1007/s10681-017-1885-5

Fenta, B. A., Beebe, S. E., and Kunert, K. J. (2020). Role of fixing nitrogen in common bean grown under water deficit conditions. *Food and Energy Security* 9:e183. doi: 10.1002/foes.183

Ferguson, B. J., Indrasumunar, A., Hayashi, S., Lin, M. H., Lin, Y.-H., Reid, D. E., et al. (2016). Molecular analysis of legume nodule development and autoregulation. *J. Integr. Plant Biol.* 52, 61–76. doi: 10.1111/j.1744-7909.2010.00899.x

Guo, F. Q., Wang, R., and Crawford, N. M. (2002). The Arabidopsis dual-affinity nitrate transporter gene ANRT1.1 (CHL1) is regulated by auxin in both shoots and roots. *J. Exp. Bot.* 53, 835–844. doi: 10.1039/jexbot/53/33.370.835

Heilig, H. A., Beaver, J. S., Wright, E. M., Song, Q., and Kelly, J. D. (2017). QTL analysis of symbiotic nitrogen fixation in a black bean population. *Crop Sci.* 57, 118–129. doi: 10.2135/cropsci2016.05.0348

Hirsch, A. M., Lum, M. R., and Downie, J. A. (2001). What makes the rhizobia-legume symbiosis so special? *Plant Physiol.* 127, 1484–1492. doi: 10.1104/pp.010866
nodulation in the Medicago truncatula cytokinin perception mutant crel. Plant Cell 27, 2210–2226. doi: 10.1105/tpc.15.00231

Nodari, R. O., Tsai, S. M., Guzman, P., Gilbertson, R. L., and Gepts, P. (1993). Toward an integrated linkage map of common bean.III. Mapping genetic factors controlling host-bacteria interactions. Genetics 134, 341–350.

Novoa-Franco, B., Íñiguez, L. P., Valdés-López, O., Alvarado-Afiantranger, X., Leija, A., Fuentes, S. I., et al. (2015). The micro-RNA172c-APETALA2-1 node as a key regulator of the common bean-rhizobium eli nitrogen fixation symbiosis. Plant Physiol. 168, 273–291. doi:10.1007/pl00012858

Polania, J., Poschenrieder, C., Rao, L., and Beebe, S. (2016). Estimation of genetic variation for symbiotic nitrogen fixation of common bean under drought stress using 15N natural abundance in grain. Eur. J. Agron. 79, 66–73. doi:10.1016/j.eja.2016.05.014

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38, 904–909. doi:10.1038/ng1847

Purcell, L. C. (2009). Physiological responses of N2 fixation to drought and selecting genotypes for improved N2 fixation. Nitrogen fixation in crop production. 52, 211–238. doi:10.2134/agronmonog52.c8

Ramaekers, L., Galeano, C. H., Garzon, N., Vanderleyden, J., and Matthew, W. B. (2013). Identifying quantitative trait loci for symbiotic nitrogen fixation capacity and related traits in common bean. Mol. Breed. 31, 163–180. doi:10.1007/s11032-012-9780-1

Ren, H., and Gray, W. M. (2015). SAUR proteins as effectors of hormonal and environmental signals in plant growth. Mol. Plant 8, 1153–1164. doi:10.1016/j.molp.2015.05.003

Schoonhoven, A. and Pastor-Corrales, M. A. (eds.) (1987). Standard system for the evaluation of bean germplasm. Cali, CO: The International Center for Tropical Agriculture (CIAT), 53.

Segonzac, C., Boyer, J. -C., Ipotesi, E., Szponarski, W., Tillard, P., Touraine, B., et al. (2007). Nitrate efflux at the root plasma membrane: identification of an Arabidopsis excretion transporter. Plant Cell 19, 3760–3777. doi:10.1105/tpc.108.0648173

Sun, G., Zhu, C., Kramer, M., Yang, S., Song, W., Piepho, H., et al. (2010). Variation explained in mixed-model association mapping. Heredity 105, 333–340. doi:10.1038/hdy.2010.11

Unkovich, M., Herridge, D., Peoples, M., Cadisch, G., Boddey, B., Giller, K., et al. (2008). Measuring plant-associated nitrogen fixation in agricultural systems. Canberra, ACT, Australia: Australian Centre for International Agricultural Research (ACIAR).

Van Noorden, G. E., Verbeek, R., Ding, Q. D., Jin, J., Green, A., Ng, J. L. P., et al. (2016). Molecular signals controlling the inhibition of nodulation by nitrate in Medicago truncatula. Int. J. Mol. Sci. 17:1060. doi:10.3390/ijms17071060

Wanek, W., and Arndt, S. K. (2002). Difference in S15N signatures between nodulated roots and shoots of soybean is indicative of the contribution of symbiotic N2 fixation to plant N. J. Exp. Bot. 53, 1109–1118. doi:10.1093/jxb/erj119

Yang, W. (2004). Lipid signaling. Curr. Opin. Plant Biol. 7, 329–336. doi:10.1016/j.pbi.2004.03.012

Wang, Q., Liu, J., and Zhu, H. (2018). Genetic and molecular mechanisms underlying symbiotic specificity in legume-rhizobium interactions. Front. Plant Sci. 9:213. doi:10.3389/fpls.2018.00213

Wang, H., and Wang, H. (2015). The miR156/SPL module, a regulatory hub and versatile toolbox, gears up crops for enhanced agronomic traits. Mol. Plant 8, 677–688. doi:10.1016/j.molp.2015.01.008
Yan, Z., Hossain, M. S., Wang, J., Valdés-López, O., Liang, Y., Libault, M., et al. (2013). miR172 regulates soybean nodulation. *Mol. Plant-Microbe Interact.* 26, 1371–1377. doi: 10.1094/MPMI-04-13-0111-R

Yu, N., Niu, Q. W., Ng, K. H., and Chua, N. H. (2015). The role of miR156/SPL s modules in Arabidopsis lateral root development. *Plant J.* 83, 673–685. doi: 10.1111/tpj.12919

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–824. doi: 10.1038/ng.2310

Zhou, X., and Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* 11, 407–409. doi: 10.1038/nmeth.2848

Zou, H., Zhang, N.-N., Pan, Q., Zhang, J.-H., Chen, J., and Wei, G.-H. (2019). Hydrogen sulfide promotes nodulation and nitrogen fixation in soybean-rhizobia symbiotic system. *Mol. Plant-Microbe Interact.* 32, 972–985. doi: 10.1094/MPMI-01-19-0003-R

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Oladzad, González, Macchiavelli, de Jensen, Beaver, Porch and McClean. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.