Theoretical and Applied Genetics
https://doi.org/10.1007/s00122-019-03471-5

ORIGINAL ARTICLE

Whole-genome resequencing identifies quantitative trait loci associated with mycorrhizal colonization of soybean

Michelle L. Pawlowski[1] · Tri D. Vuong[2] · Babu Valliyodan[2] · Henry T. Nguyen[2] · Glen L. Hartman[1,3]

Received: 25 June 2019 / Accepted: 1 November 2019

Abstract

Key message **A whole-genome resequencing-derived SNP dataset identified six quantitative trait loci (QTL) significantly associated with colonization of soybean by an arbuscular mycorrhizal fungus (*Rhizophagus intraradices*). Candidate genes identified in these QTL regions include homologs to known nodulin protein families and other symbiosis-specific genes.**

Abstract Arbuscular mycorrhizal fungi (AMF) form associations with over 80% of all terrestrial plant species and assist their host plants by increasing their nutrient uptake, drought tolerance, and resilience against pathogens and pests. Genotypic variation of crop plants to AMF colonization has been identified in crops, including soybean; however, the genetics controlling levels of AMF colonization in soybean are unknown. The overall goal of our study was to identify genomic regions associated with mycorrhizal colonization in soybean using genome-wide association analysis. A diverse panel of 350 exotic soybean genotypes inoculated with *Rhizophagus intraradices* were microscopically evaluated for root colonization using a modified gridline intersect method. Root colonization differed significantly ($P < 0.001$) among genotypes and ranged from 11 to 70%. A whole-genome resequencing-derived SNP dataset identified six quantitative trait loci (QTL) significantly associated with *R. intraradices* colonization that explained 24% of the phenotypic variance. Candidate genes identified in these QTL regions include homologs to known nodulin protein families and other symbiosis-specific genes. The results showed there was a significant genetic component to the level of colonization by *R. intraradices* in soybean. This information may be useful in the development of AMF-sensitive soybean cultivars to enhance nutrient uptake, drought tolerance, and disease resistance in the crop.

Abbreviations

AMF	Arbuscular mycorrhizal fungi
CHR	Chromosome
GWAS	Genome-wide association study
LD	Linkage disequilibrium
PC	Principal component
PCA	Principal component analysis
QTL	Quantititatve trait loci
SNP	Single nucleotide polymorphisms
WGRS	Whole-genome resequencing

Communicated by Istvan Rajcan.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-019-03471-5) contains supplementary material, which is available to authorized users.

✉ Glen L. Hartman
Glen.Hartman@ars.usda.gov; ghartman@illinois.edu

[1] Department of Crop Science, University of Illinois, 1101 W. Peabody Drive, Urbana, IL 61801, USA

[2] Department of Plant Sciences, University of Missouri, Columbia, MO, USA

[3] USDA, Agricultural Research Services, University of Illinois, 1101 W. Peabody Dr., Urbana, IL, USA

Introduction

With an ever-growing human population and increasing impacts of climate change, abiotic and biotic limitations are putting a major stress on food security (Lobell et al. 2008). Proactive measures are necessary to introduce more ecologically friendly production practices and to develop breeding stock that are more robust to challenging growth conditions. Microbial communities including arbuscular mycorrhizal fungi (AMF) are known to provide a plethora of benefits to plants (Berendsen et al. 2012). AMF form associations with over 80% of all terrestrial plant species and assist them by

Published online: 09 November 2019

increasing their nutrient uptake, drought tolerance, and resilience against pathogens and pests (Richardson et al. 2011). These benefits have sparked interest in the private sector to produce AMF inoculants for commercial application (Berruti et al. 2016). Commercial inoculants are frequently composed of multiple AMF species, including *Rhizophagus intraradices* (N.C. Schenck & G.S. Sm.) C. Walker & A. Schüßler 2010 and other species in the Glomeraceae family. *Rhizophagus intraradices* (syn. *Glomus intraradices*) is a fast-growing fungus that produces abundant spores, which is ideal for large scale production. This species is found worldwide and in AMF inoculants because it provides benefits to a wide range of plant species (Berruti et al. 2016). In a recent report, *R. intraradices* was used in 61% of AMF inoculant studies either alone or as part of a mixture of AMF species (Berruti et al. 2016).

Numerous studies have reported the benefits of AMF to soybean [*Glycine max* (L.) Merr.]. Soybean plants colonized by AMF had increased uptake of all macronutrients (Khalil et al. 1994) and several micronutrients, including Mn, Zn, Cu, and Cd (Pacovsky 1986; Heggo et al. 1990). Transfer of nutrients has been observed to occur by AMF from soybean to maize (Hamel et al. 1991; Meng et al. 2015). Inoculation with AMF has shown promise as a method for phytoremediation and to protect soybean against heavy metal contaminated soils (Heggo et al. 1990) and to increase tolerance to drought (Porcel and Ruiz-Lozano 2004). Soybean plants inoculated with AMF had reduced disease severity to soilborne pathogens, including *Fusarium* spp., *Macrophomina phaseolina*, and *Rhizoctonia solani* (Zambolim and Schenck 1983), and root-knot nematode (Kellam and Schenck 1980).

The extent of mycorrhizal colonization and potential benefits to crop plants is likely dependent on the host genotypes. Plant genotypic variation to AMF colonization and/or their benefits was identified in maize (An et al. 2010), onion (Taylor et al. 2015), and wheat (Hetrick et al. 1996). Quantitative trait loci (QTL) have been found associated with mycorrhizal colonization or mycorrhizal responsiveness in multiple crop species. A study using several *Allium* species, that included offspring of 96 trihybrid genotypes of *A. cepa* × (*A. fistulosum* × *A. roylei*), found three significant QTL associated with mycorrhizal responsiveness in the trihybrid population (Galván et al. 2011). In maize, one QTL associated with mycorrhizal responsiveness was found in a population of 197 recombinant inbred lines (Kaeppler et al. 2000). In addition, 12 maize genotypes, measured for both mycorrhizal responsiveness and mycorrhizal colonization, showed a positive correlation between colonization and root response but no significant correlation between colonization and shoot response (Kaeppler et al. 2000). In winter wheat, six significant QTL were found associated with mycorrhizal colonization in a diversity panel of 94 winter wheat genotypes (Lehnert et al. 2017). Seven QTL were identified to be associated with susceptibility to colonization by both *R. intraradices* and *Funneliformis mosseae* in a diversity panel of 108 durum varieties (De Vita et al. 2018). However, in *Medicago truncatula*, transcript levels of *RiTubulin* (*R. irregularis* β-tubulin), and *MtPT4* (arbuscule specific phosphate transporter in *M. truncatula*), did not differ in 32 lines (Dreher et al. 2017).

Although these studies identified genomic regions associated with mycorrhizal colonization, very few were able to describe candidate genes in these regions that may play a role in the level of colonization. In both durum and winter wheat, the majority of genes identified were defense related (Lehnert et al. 2017; De Vita et al. 2018). There are many transcriptomic studies on mycorrhizal colonization showing a broader picture of host genes necessary for mycorrhizal colonization. It is known that hundreds of genes are differentially expressed during mycorrhizal symbiosis, and multiple 'core sets' of genes important for mycorrhizal symbiosis have been compiled using phylogenetic and transcriptomic studies. Through a phylogenetic study, a core set of 88 genes in *M. truncatula* was identified to be conserved in plant species able to form mycorrhizal symbiosis (Bravo et al. 2016). A comparative transcriptomic study found a conserved set of 156 genes in *Poncirus trifoliata* that were also differentially expressed in other hosts, including *Clonorchis sinensis*, *Lotus japonicus*, *M. truncatula*, *Solanum lycopersicum*, and *Oryza sativa* (An et al. 2018). Another comparative transcriptomic study compared differential expression in *Casuarina glauca, M. truncatula,* and *O. sativa* and revealed a set of 84 genes that were differentially expressed in all three species (Tromas et al. 2012).

The USDA Soybean Germplasm Collection consists of 18,480 cultivated soybean (*Glycine max*) and 1168 accessions of a wild relative (*Glycine soja*) (https://www.ars-grin.gov/). The entire collection has been genotyped using the SoySNP50K iSelect BeadChip, which contains over 50,000 single nucleotide polymorphisms (SNPs) (Song et al. 2015). Recent advances in next generation sequencing and new comprehensive statistical models and multiple tests have made genome-wide analysis more effective and efficient in evaluating genomic diversity and trait discovery (Zhou et al. 2015; Valliyodan et al. 2016), in identifying significant QTL associated with the traits of interest (Patil et al. 2016) and in predicting candidate genes (Cao et al. 2016; Yano et al. 2016).

A more extensive 3.7 M SNP dataset was created at the Soybean Genetics and Genomics Laboratory at the University of Missouri from a whole-genome resequencing (WGRS) project using a subset of 350 soybean lines, representing a broad range of soybean genotypes, including wild, landraces, and elite populations (http://www.soybase.org). This extensive 3.7 M SNP dataset and the diverse set of 350 soybean lines were first used in a genome-wide association

study (GWAS) of soybean salt tolerance (Do et al. 2019) which confirmed a known salt tolerance locus on chromosome (Chr) 3 and also successfully detected new QTL on Chrs 1, 8, and 18 with several candidate genes harbored in these genomic regions. Along with the 3.7 M SNP dataset, there is extensive genomic annotation and expression data accumulated which gives us better insight when conducting GWAS and identifying potential candidate genes (Grant et al. 2010).

The economic importance of soybean and wealth of available genomic information make it an ideal organism in which to study the genetic components involved in AMF symbiosis. Thus, the objectives of the present study were to determine if (1) soybean genotypes differ in levels of colonization by *R. intraradices* and (2) genomic regions involved in mycorrhizal colonization can be identified in soybean. A genetically diverse panel of 350 soybean accessions were inoculated with *R. intraradices*, and root colonization was used in a GWAS to identify potential QTL associated with mycorrhizal colonization.

Materials and methods

Plant materials and genotypic datasets

All plant materials were provided by the USDA ARS Soybean Germplasm Collection (Urbana, IL) (Table S1). A panel of 350 soybean plant introductions (PIs) served as a core set selected to represent a broad genetic diversity per the SoySNP50K iSelect BeadChip analysis (Drs. Cregan and Song, pers. comm.). The set included a wide range of maturity groups (MG), from 000 to X, and originated from 28 different countries (Table S1) (https://www.ars-grin.gov/npgs). The sequenced reference genome, soybean cultivar Williams 82 (Schmutz et al. 2010), was also included.

A marker dataset, generated from the WGRS project at the University of Missouri, contained over 4.7 M SNPs was accessed from the soybean database (http://www.soybase.org). For association analysis, the SNP dataset was further filtered based on the exclusion of the missing data (> 5%) and minor allele frequency (< 5%) to obtain approximate 3.7 M qualified SNPs, which were hereafter referred to the WGRS-derived 3.7 M SNP dataset.

Mycorrhizal inoculum preparation

Isolate *R. intraradices* AU212B (https://INVAM.wvu.edu) was used to assess soybean genotype variability to AMF fungal colonization. To increase *R. intraradices*, a modified sheared root inoculum method was used (Sylvia and Jarstfer 1992). Cone-tainers (SC10-158 ml, Stuewe and Sons, Inc., Tangent, OR) were filled with 100 ml of sterilized torpedo sand. A 2.5 cm (30 ml) layer of inoculum, original stock from INVAM, was spread evenly over the top and covered with another 2.5 cm of sand. Surface-disinfected soybean seeds of soybean cultivar Williams 82 were sown 2 cm deep into the sand. Plants were grown in a greenhouse held at 25 °C with a 16-h photoperiod and watered regularly. Watering was discontinued after 6 weeks plants were allowed to dry completely, and the dried above ground material was discarded. Roots, infested with AMF, and sand were removed from cone-tainers and collected in 1-gallon plastic bags. Roots were cut into 1 cm fragments with a scissors and mixed back into sand. Inoculum was stored at 4 °C for 1 week prior to planting.

Experimental design and procedures

The experiment was arranged as a randomized block with two replicated blocks completed over time using different genotype randomizations for each block. Two soybean seeds of each soybean genotype were sown in sand 2.5 cm above a 30 ml layer of inoculum, consisting of approximately 60 spores, in cone-tainers as previously described. The spore rate was determined by removing five samples from the stock and counting the spores. Once this was determined, the stock was diluted by volume and mixed well before removing inoculum for cone-tainers.

Plants were thinned to one plant per cone-tainer after emergence. Six weeks after planting, above ground plant material was removed and roots were excised from the sand and washed several times in water to remove all sand particles. Roots were then placed in 50-ml centrifuge tubes filled with 70% ethanol and stored at 4 °C prior to evaluation. This experiment was conducted in a greenhouse at the University of Illinois held at 25 °C with a 16-h photoperiod. The plants were watered regularly and fertilized twice a week using a low-P formulation (150 ppm Peters Excel 15-5-15 Cal-Mag, ICL-SF, Dublin, OH).

Root colonization measurements

Excised roots were cleared with 20% bleach solution at room temperature for 16 h, rinsed three times with water, soaked in a 2% HCl solution for 30 min, and rinsed three times with water prior to being soaked in 0.05% chlorazol black E (Sigma Aldrich, St. Louis, MO) in 1:1:1 lactic acid, glycerol, and water solution for 48 h (Vierheilig et al. 2005). Roots were then destained by soaking in water for 2–3 days. To quantify root colonization, roughly one hundred 1 cm root segments were selected and placed in water in a 10 cm diameter Petri dish that was marked with gridlines to create 1 cm^2 boxes. Roots were observed using a stereoscope at ×20 magnification. A modified gridline intersect method was used to

quantify root colonization by dividing the number of root segments colonization by the total number of roots multiplied by 100 (Fig. 1) (Giovannetti and Mosse 1980). This was repeated twice for each sample.

Phenotypic data analysis

Statistical analysis of root colonization percentages was completed using JMP version 12 (SAS institute, Cary, NC, USA). A standard analysis of variance was performed, and means were separated using the JMP LSmeans Student's *t* procedure at $\alpha = 0.05$. To determine if there were any associations with other traits, mycorrhizal colonization means were compared to observations, including physiological and disease and pest resistance records that were obtained from the National Plant Germplasm System (http://www.npgsweb.ars-grin.gov), using either Pearson's correlation for numerical values or a *y* by *x* fit for categorical values.

Genome-wide association study

The WGRS-derived 3.7 M SNP dataset was employed to perform GWAS using the SNP Variation & Suites program (Golden Helix Inc., Bozeman, MT 59718, USA). Briefly, principal component analysis (PCA) was done for all filtered SNPs using a minor allele frequency > 0.05 and a call rate > 0.95 to select the correct number of principal components (PCs). BLUP genomic kinship matrix was calculated for population structure and relatedness in mixed linear models. Genome-wide associations between SNPs and root colonization percentage were identified using the efficient mixed-model association expedited and multi-locus mixed model (MLMM). False positives were controlled by multiple test correction with false discovery rate ≤ 0.05 and the threshold of $-\log10$ (*p* value) was used to identify significant associations calculated at a false discovery rate = 0.05.

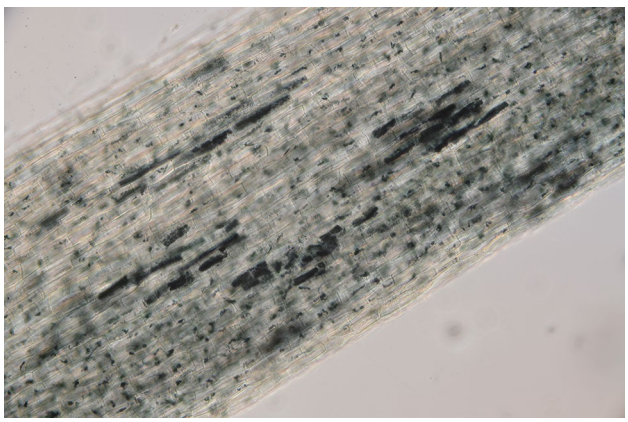

Fig. 1 Arbuscules of *Rhizophagus intraradices* in soybean roots stained with chlorazol black E at ×50 magnification

Manhattan plots for associated SNPs were visualized in GenomeBrowse v1.0 (Golden Helix, Inc, Bozeman, MT 59718, USA).

Linkage disequilibrium (LD) and candidate gene identification

Flanking SNPs located within ± 3 Mbp were used to assess LD decay in TASSEL5 (Bradbury et al. 2007). A sliding window size of 500 bp was used, and heterozygous alleles were treated as missing data. LD region cutoff was determined using a squared allele frequency correlation (r^2) of 0.2. Genes within the LD region were noted using the *G. max* genome assembly version Gmax2.0 (http://www.soybase.org). Several genes within each LD region were selected as candidate genes involved in AMF colonization. These putative AMF symbiosis-related genes were selected because they were (1) annotated in the most updated assembly version, (2) expressed in roots and/or nodules (Severin et al. 2010), and (3) identified in previous studies as being involved in or expressed during AMF colonization.

Results

Evaluation of 350 diverse soybean accessions

Six weeks after sowing seeds, the plants were uprooted and at that time there were no genotypes that were in the pod stage and growth stage differences among the maturity groups were not noted. There were significant ($P < 0.0001$) differences among soybean genotypes for root colonization by *R. intraradices* (Table S1). Colonization percentages ranged from 12 to 69% with an average of 37% (Fig. 2). The top ten percent of genotypes had average colonization percentages ≥ 48.5%, while the bottom ten percent of genotypes had an average colonization of ≤ 23.5%. Colonization on cultivar Williams 82 averaged 33%. Even with limited replications, that are not uncommon in these large genotype experiments, the replicates were significantly ($P < 0.001$) correlated at $r = 0.53$. There were no correlations found for these genotypes between root colonization by *R. intraradices* and other physiological traits including maturity group or disease resistance. Although environmental conditions, like ideal photoperiod for each genotype, may have influenced the growth of the plants, the brevity of the experiment (6 weeks) did not allow plants to set seeds and no trend was observed between colonization by *R. intraradices* and plant growth stage.

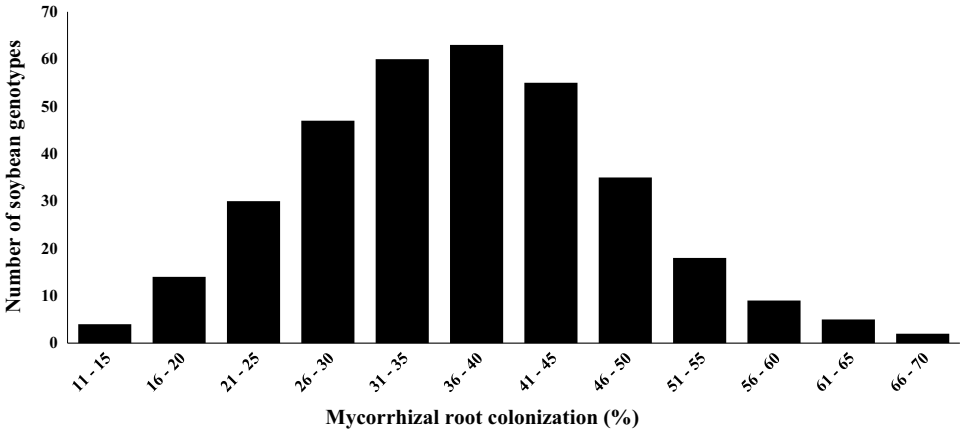

Fig. 2 Distribution of genotype means for root colonization by the mycorrhizal fungus, *Rhizophagus intraradices*, for 350 soybean genotypes. Means were averaged between two replicates and colonization percentages for each were based on arbuscule counts on 100 root segments using the grid line intersect method (Giovannetti and Mosse 1980)

Phylogeny, population stratification, and kinship

Phylogenetic relationships showed soybean accessions with close relationship were clearly clustered in several divergent subgroups (Fig. S1). The PCA showed dispersed genotypes among different components (Fig. S2). The PCA also showed that variation explained by eigenvalues of each PC rapidly reduced after the first four PCs (data not shown). The accumulative eigenvalues of the first four PCs accounted for 39.4% of the variance.

The cryptic relatedness among 350 soybean accessions in this study was evaluated by genomic kinship analysis, from which a kinship matrix was calculated using the WGRS-derived 3.7 M SNP dataset. The relatedness among soybean accessions was visualized in a heat map (Fig. S3), showing a similar clustering pattern of the accessions studied.

Table 1 Candidate genes associated with root colonization. Significant peak markers and corresponding candidate genes in soybean associated with root colonization by an arbuscular mycorrhizal fungus, *Rhizophagus intraradices*

Marker	Chr	Position	LOD	r^2	LD interval (Mbp)	Candidate genes	Gene description
S16_2603450	16	2603450	5.77	7.09	2.47–2.66	Glyma.16g026400	WRKY60 Transcription factor
						Glyma.16g026900	GH3 family protein
						Glyma.16g027200	Nodulin 25 family protein
						Glyma.16g027300	Serine–threonine protein kinase
						Glyma.06g272500, 274200	60S ribosomal protein L11
S6_46533686	6	46533686	5.29	4.47	46.14–46.76	Glyma.06g272600-274100, 275100-275300	Cysteine proteinase
						Glyma.06g275400	Chaperone protein dnaJ
						Glyma.06g276400	NBS-LRR
S9_49539462	9	49539462	4.67	5.11	49.22–49.56	Glyma.09g279100	Cytochrome P450
						Glyma.09g280400,280500	UDP-glucosyltransferase
S10_37487605	10	37487605	4.53	5.92	37.12–37.87	Glyma.10g139700	Germin-like protein 16
						Glyma.10g139800	Germin-like protein 17
						Glyma.10g141700	Germin-like protein 1
S14_43114203	14	43114203	4.44	4.54	42.66–43.18	Glyma.14g172200	VQ motif-containing protein
						Glyma.14g174300	Aquaporin NIP1-1
S8_17743016	8	17743016	4.36	5.98	17.24–18.06	Glyma.08g214100	VQ motif-containing protein
						Glyma.08g215700	Ethylene responsive factor
						Glyma.08g216600	Ethylene response factor
						Glyma.08g217400	Aquaporin NIP3
						Glyma.08g220800	AP2 transcription factor
						Glyma.08g221200	Kelch repeat-containing protein

Genome-wide association study

There were six significant peaks identified on Chrs 6, 8, 9, 10, 14, and 16 (Table 1 and Table S2). These markers individually explained between 4.54 and 7.09% of the phenotypic variation, and all six makers together explained 24% of the phenotypic variation. Using BLUPs as phenotypic data for association analysis, the WGRS-derived SNP dataset detected significant peaks at the same regions as the arithmetic mean phenotypic data, along with one more significant peak on Chr 7 (Table S3).

Identification of candidate genes

For all significant markers, 21 genes known to have important roles in mycorrhizal symbiosis were identified within the LD regions (Table 1 and Table S4). LD regions for the markers identified with the WGRS-derived dataset spanned 620 kb on Chr 6, 820 kb on Chr 8, 340 kb on Chr 9, 750 kb on Chr 10, 520 kb on Chr 14, and 190 kb on Chr 16. The LD region on Chr 16 included genes encoding for a GH3 auxin responsive promoter, a nodulin protein, and a serine–threonine protein kinase (Table 1). Candidate genes found on Chr 6 include genes encoding for a cysteine protease, a chaperone dnaj, and a nucleotide-binding site-leucine-rich repeat (NBS-LRR) protein. Candidate genes on Chr. 9 include genes encoding for a cytochrome P450, a WRKY DNA binding domain, and a UDP-glucosyltransferase. The LD region on Chr 10 includes three germin-like proteins. Candidate genes identified on Chr 14 include a valine-glutamine (VQ) motif-containing protein and an aquaporin transporter. The Chr 8 LD region includes a VQ motif-containing protein, two ethylene response factors, an aquaporin, and a kelch repeat-containing family protein (Table 1).

Discussion

We identified six significant GWAS loci within the soybean genome that are associated with the level of colonization by *R. intraradices*. These six regions explained 24% of the phenotypic variance with each peak SNP explaining from 4.5 to 7.1% of this variation. Similar studies have shown mycorrhizal associations in other crops. In durum wheat, individual QTL explained 7–16% of the phenotypic variance (De Vita et al. 2018), while in winter wheat, individual markers only explained 0.82–1.14% (Lehnert et al. 2017) of the phenotypic variance. Our study identified candidate genes involved in all aspects of mycorrhizal symbiosis ranging from pre-infection signaling to arbuscule senescence, but most candidate genes were related to host-microbe recognition and signaling pathways.

In our study, the 190 kb region located on Chr 16 explained 7.1% of the phenotypic variance. Within this region, a gene encoding for a GH3 (Gretchen Hagen) protein was found and is part of the GH3 protein family, a group of auxin responsive proteins first identified in soybean (Hagen et al. 1991). GH3 proteins conjugate free auxins and influence plant hormone signaling in Arabidopsis (Westfall et al. 2012). Recently, three GH3 genes were identified to be activated during mycorrhizal symbiosis in tomato (Liao et al. 2014). The expression GH3.25, one of the genes located on soybean Chr 16, was found highly upregulated in nodules, indicating it has a function in symbiosis (Severin et al. 2010; Singh et al. 2015). Other genes on Chr 16 within this LD region, included a gene encoding for a serine/threonine protein kinase (Ser/Thr), which activate during pre-infection signaling and, when phosphorylated, become fully activated as the symbiosis receptor kinase (Yoshida and Parniske 2005). A gene found in the casein family, a subgroup of Ser/Thr, was also identified through GWAS to be significantly associated with mycorrhizal colonization in winter wheat (Lehnert et al. 2017).

Several candidate genes were identified within the 620 kb region on Chr 6 including a large group of cysteine proteases which upregulate mycorrhizal symbiosis in many species such as *M. truncatula*, *Lotus japonicus*, and *Solanum lycopersicum* (Liu et al. 2003; Nair and Bhargava 2012; Handa et al. 2015). Cysteine proteases are specifically expressed in cells containing arbuscules and may be involved in degradation of arbuscules or in remodeling intracellular structures (Deguchi et al. 2007). Other candidate genes include those encoding for two 60S ribosomal proteins and a NBS-LRR resistance-like protein; both were highly expressed in nodules and may have a role in overall symbiosis (Severin et al. 2010). Ribosomal 60S proteins were found to be upregulated in mycorrhizal *L. japonicus* and *M. truncatula* (Gaude et al. 2012; Handa et al. 2015), indicating their potential role in overall symbiosis.

Many genes involved in symbiosis are also involved in defense against pathogens. It is not surprising that our study identified an overlapping region between the 340 kb LD region on Chr 9 and a QTL associated with SCN resistance (Zhang et al. 2017). Within this overlap is a gene encoding for cytochrome P450, which is known to be involved in both plant disease pathways and strigolactones biosynthesis, making it a potential candidate for both SCN resistance and mycorrhizal colonization (Nakamura and Asami 2014). Strigolactones are found in root-exudates and are associated with pre-infection signaling between AMF spores and plant roots (Lanfranco et al. 2017). A gene encoding for a UDP-glycosyltransferase (UGT) was also found within this LD region on Chr 9. UGTs are known to interact with phytohormones and metabolites during biotic and abiotic stress responses (Rehman et al. 2018). UGTs are overexpressed in

mycorrhizal rice, sunflower, and *L. japonicus* (Gutjahr et al. 2008; Handa et al. 2015; Vangelisti et al. 2018). The UGTs may play a role in the early stages of mycorrhizal infection (Vangelisti et al. 2018).

Two genes encoding for germin-like proteins (GLP) were found within the 750 kb region on Chr. 10. These genes are not expressed during abiotic stress and highly expressed in nodules, meaning they could be involved in symbiosis-specific processes (Severin et al. 2010; Li et al. 2016). Germins are well known in biotic stress responses by their oxalate activity (Yamahara et al. 1999). Therefore, GLPs are assumed to have some form of oxidative activity as well (Carter and Thornburg 2000). Mycorrhizal-specific GLPs may play a role in cell wall reorganization, but their function in mycorrhizal symbiosis remains unknown (Doll et al. 2003).

We found one gene encoding for an aquaporin nodulin26 intrinsic protein (NIP) within the 520 kb LD region on Chr 14. Genes that encode nodulin proteins were identified as nodule-specific, but studies have shown some nodulins proteins are also expressed during mycorrhizal symbiosis (Wyss et al. 1990). In *L. japonicus,* an aquaporin gene, JtNIP1, was found to be expressed in cells containing arbuscules (Giovannetti et al. 2012). Along with a gene encoding for a NIP, a gene encoding for a valine-glutamine (VQ) motif-containing protein was identified as a possible candidate. VQ motif-containing proteins are a class of plant-specific proteins that regulate several plant development processes, including responses to abiotic and biotic stresses (Jing and Lin 2015). This gene was also considered the candidate gene within a locus identified to be associated with resistance to *Sclerotinia sclerotiorum* (Moellers et al. 2017). There are no previous studies on the function of VQ motif-containing proteins in mycorrhizal symbiosis, but their role in defense signaling makes them a viable candidate for involvement mycorrhizal symbiosis.

Another gene encoding for a NIP (explained above) occurred within the 820 kb LD region on Chr 8 along with several genes encoding for ethylene response factors (ERF) and another locus associated with resistance to *S. sclerotiorum* (Moellers et al. 2017). ERFs are a part of the ethylene response pathway and are known for their role in disease resistance to necrotrophic pathogens (Müller and Munné-Bosch 2015). Mycorrhizal symbiosis is known to manipulate host defense pathways and upregulate the ethylene pathway (Pozo et al. 2010). ERFs are highly expressed in mycorrhizal plants compared to control plants (Pozo et al. 2015). One gene, *Glyma.08g216600*, is shown to be highly expressed in nodules and may function in other symbiotic processes such as AMF symbiosis (Severin et al. 2010).

The results of our study showed there was a significant genetic component in soybean that dictates the level of colonization by *R. intraradices*. Potential candidate genes that influence the level of colonization are involved in other microbial signaling, including defense related genes and genes involved in nodulation. A better understanding of the function of these genes could provide the basis for developing enhanced AMF-sensitive soybean cultivars. Increasing AMF sensitivity in soybean may result in increases in nutrient uptake, drought tolerance, and disease resistance. This environmentally friendly approach to improving soybean production may help reduce the overuse of fertilizers and pesticides and promote more holistic crop production systems.

Acknowledgements We thank Amanda Bardeau and Theresa Herman for their help with the greenhouse experiments, T. Herman for assisting with revising the manuscript, and Hao-Xun Chang for his advice on LD analysis and candidate gene selection. We thank Dr. Joseph Morton at the International Culture Collection of Vesicular Mycorrhizal Fungi for his advice on AMF species selection and for providing the isolate used in this study. We thank the United Soybean Board and the USDA-ARS for funding, as well as the University of Illinois Department of Crop Sciences for support via the James B. Sinclair and Fraley-Borlaug Fellowships.

Author Contribution statement MPL, TDV, and GLH planned, and designed experiments. MPL completed root colonization analysis, LD analysis, and identified candidate genes. TDV performed genome-wide association study, phylogenetic and kinship analyses. BV and HTN selected and sequenced the 350 soybean accessions, MPL, TDV, and GLH wrote the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethic standards The author state that the experiments comply with the current laws of the country in which they were performed (USA).

References

An G-H, Kobayashi S, Enoki H, Sonobe K, Muraki M, Karasawa T, Ezawa T (2010) How does arbuscular mycorrhizal colonization vary with host plant genotype? An example based on maize (*Zea mays*) germplasms. Plant Soil 327:441–453

An J, Sun M, van Velzen R, Ji C, Zheng Z, Limpens E, Bisseling T, Deng X, Xiao S, Pan Z (2018) Comparative transcriptome analysis of *Poncirus trifoliata* identifies a core set of genes involved in arbuscular mycorrhizal symbiosis. J Exp Bot 69:5255–5264

Berendsen RL, Pieterse CMJ, Bakker PAHM (2012) The rhizosphere microbiome and plant health. Trends Plant Sci 17:478–486

Berruti A, Lumini E, Balestrini R, Bianciotto V (2016) Arbuscular mycorrhizal fungi as natural biofertilizers: let's benefit from past successes. Front Microbiol 6:1559

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635

Bravo A, York T, Pumplin N, Mueller LA, Harrison MJ (2016) Genes conserved for arbuscular mycorrhizal symbiosis identified through phylogenomics. Nat Plants 2:15208

Cao K, Zhou Z, Wang Q, Guo J, Zhao P, Zhu G, Fang W, Chen C, Wang X, Wang X (2016) Genome-wide association study of 12 agronomic traits in peach. Nat Commun 7:13246

Carter C, Thornburg RW (2000) Tobacco nectarin I purification and characterization as a germin-like, manganese superoxide dismutase implicated in the defense of floral reproductive tissues. J Biol Chem 275:36726–36733

De Vita P, Avio L, Sbrana C, Laidò G, Marone D, Mastrangelo AM, Cattivelli L, Giovannetti M (2018) Genetic markers associated to arbuscular mycorrhizal colonization in durum wheat. Sci Rep 8:10612

Deguchi Y, Banba M, Shimoda Y, Chechetka SA, Suzuri R, Okusako Y, Ooki Y, Toyokura K, Suzuki A, Uchiumi T (2007) Transcriptome profiling of *Lotus japonicus* roots during arbuscular mycorrhiza development and comparison with that of nodulation. DNA Res 14:117–133

Do TD, Vuong TD, Dunn D, Clubb M, Valliyodan B, Patil G, Chen P, Xu D, Nguyen HT, Shannon JG (2019) Identification of new loci for salt tolerance in soybean by high-resolution genome-wide association mapping. BMC Genom 20:318

Doll J, Hause B, Demchenko K, Pawlowski K, Krajinski F (2003) A member of the germin-like protein family is a highly conserved mycorrhiza-specific induced gene. Plant Cell Physiol 44:1208–1214

Dreher D, Yadav H, Zander S, Hause B (2017) Is there genetic variation in mycorrhization of *Medicago truncatula*? PeerJ 5:e3713

Galván GA, Kuyper TW, Burger K, Keizer LP, Hoekstra RF, Kik C, Scholten OE (2011) Genetic analysis of the interaction between *Allium* species and arbuscular mycorrhizal fungi. Theor Appl Genet 122:947–960

Gaude N, Bortfeld S, Duensing N, Lohse M, Krajinski F (2012) Arbuscule-containing and non-colonized cortical cells of mycorrhizal roots undergo extensive and specific reprogramming during arbuscular mycorrhizal development. Plant J 69:510–528

Giovannetti M, Mosse B (1980) An evaluation of techniques for measuring vesicular arbuscular mycorrhizal infection in roots. New Phytol 84:489–500

Giovannetti M, Balestrini R, Volpe V, Guether M, Straub D, Costa A, Ludewig U, Bonfante P (2012) Two putative-aquaporin genes are differentially expressed during arbuscular mycorrhizal symbiosis in *Lotus japonicus*. BMC Plant Biol 12:186

Grant D, Nelson RT, Cannon SB, Shoemaker RC (2010) SoyBase, the USDA-ARS soybean genetics and genomics database. Nucleic Acids Res 38:D843–D846

Gutjahr C, Banba M, Croset V, An K, Miyao A, An G, Hirochika H, Imaizumi-Anraku H, Paszkowski U (2008) Arbuscular mycorrhiza-specific signaling in rice transcends the common symbiosis signaling pathway. Plant Cell 20:2989–3005

Hagen G, Martin G, Li Y, Guilfoyle TJ (1991) Auxin-induced expression of the soybean GH3 promoter in transgenic tobacco plants. Plant Mol Biol 17:567–579

Hamel C, Barrantes-Cartin U, Furlan V, Smith D (1991) Endomycorrhizal fungi in nitrogen transfer from soybean to maize. Plant Soil 138:33–40

Handa Y, Nishide H, Takeda N, Suzuki Y, Kawaguchi M, Saito K (2015) RNA-seq transcriptional profiling of an arbuscular mycorrhiza provides insights into regulated and coordinated gene expression in *Lotus japonicus* and *Rhizophagus irregularis*. Plant Cell Physiol 56:1490–1511

Heggo A, Angle J, Chaney R (1990) Effects of vesicular-arbuscular mycorrhizal fungi on heavy metal uptake by soybeans. Soil Biol Biochem 22:865–869

Hetrick B, Wilson G, Todd T (1996) Mycorrhizal response in wheat cultivars: relationship to phosphorus. Can J Bot 74:19–25

Jing Y, Lin R (2015) The VQ motif-containing protein family of plant-specific transcriptional regulators. Plant Physiol 169:371–378

Kaeppler SM, Parke JL, Mueller SM, Senior L, Stuber C, Tracy WF (2000) Variation among maize inbred lines and detection of quantitative trait loci for growth at low phosphorus and responsiveness to arbuscular mycorrhizal fungi. Crop Sci 40:358–364

Kellam M, Schenck N (1980) Interactions between a vesicular-arbuscular mycorrhizal fungus and root-knot nematode on soybean. Phytopathology 70:293–296

Khalil S, Loynachan TE, Tabatabai MA (1994) Mycorrhizal dependency and nutrient uptake by improved and unimproved corn and soybean cultivars. Agron J 86:949–958

Lanfranco L, Fiorilli V, Venice F, Bonfante P (2017) Strigolactones cross the kingdoms: plants, fungi, and bacteria in the arbuscular mycorrhizal symbiosis. J Exp Bot 69:2175–2188

Lehnert H, Serfling A, Enders M, Friedt W, Ordon F (2017) Genetics of mycorrhizal symbiosis in winter wheat (*Triticum aestivum*). New Phytol 215:779–791

Li Y, Zhang D, Li W, Mallano AI, Zhang Y, Wang T, Lu M, Qin Z, Li W (2016) Expression study of soybean germin-like gene family reveals a role of *GLP7* gene in various abiotic stress tolerances. Can J Plant Sci 96:296–304

Liao D, Chen X, Chen A, Wang H, Liu J, Liu J, Gu M, Sun S, Xu G (2014) The characterization of six auxin-induced tomato GH3 genes uncovers a member, SlGH3.4, strongly responsive to arbuscular mycorrhizal symbiosis. Plant Cell Physiol 56:674–687

Liu J, Blaylock LA, Endre G, Cho J, Town CD, VandenBosch KA, Harrison MJ (2003) Transcript profiling coupled with spatial expression analyses reveals genes involved in distinct developmental stages of an arbuscular mycorrhizal symbiosis. Plant Cell 15:2106–2123

Lobell DB, Burke MB, Tebaldi C, Mastrandrea MD, Falcon WP, Naylor RL (2008) Prioritizing climate change adaptation needs for food security in 2030. Science 319:607–610

Meng LB, Zhang AY, Wang F, Han XG, Wang DJ, Li SM (2015) Arbuscular mycorrhizal fungi and rhizobium facilitate nitrogen uptake and transfer in soybean/maize intercropping system. Front Plant Sci 6:339

Moellers TC, Singh A, Zhang J, Brungardt J, Kabbage M, Mueller DS, Grau CR, Ranjan A, Smith DL, Chowda-Reddy R (2017) Main and epistatic loci studies in soybean for Sclerotinia sclerotiorum resistance reveal multiple modes of resistance in multi-environments. Sci Rep 7:3554

Müller M, Munné-Bosch S (2015) Ethylene response factors: a key regulatory hub in hormone and stress signaling. Plant Physiol 169:32–41

Nair A, Bhargava S (2012) Reduced mycorrhizal colonization (rmc) tomato mutant lacks expression of SymRK signaling pathway genes. Plant Signal Behav 7:1578–1583

Nakamura H, Asami T (2014) Target sites for chemical regulation of strigolactone signaling. Front Plant Sci 5:623

Pacovsky R (1986) Micronutrient uptake and distribution in mycorrhizal or phosphorus-fertilized soybeans. Plant Soil 95:379–388

Patil G, Do T, Vuong TD, Valliyodan B, Lee J-D, Chaudhary J, Shannon JG, Nguyen HT (2016) Genomic-assisted haplotype analysis and the development of high-throughput SNP markers for salinity tolerance in soybean. Sci Rep 6:19199

Porcel R, Ruiz-Lozano JM (2004) Arbuscular mycorrhizal influence on leaf water potential, solute accumulation, and oxidative stress in soybean plants subjected to drought stress. J Exp Bot 55:1743–1750

Pozo MJ, Jung SC, López-Ráez JA, Azcón-Aguilar C (2010) Impact of arbuscular mycorrhizal symbiosis on plant response to biotic

stress: the role of plant defence mechanisms. Arbuscular Mycorrhizas: Physiology and Function. Springer, Berlin, pp 193–207
Pozo MJ, López-Ráez JA, Azcón-Aguilar C, García-Garrido JM (2015) Phytohormones as integrators of environmental signals in the regulation of mycorrhizal symbioses. New Phytol 205:1431–1436
Rehman HM, Nawaz MA, Shah ZH, Ludwig-Müller J, Chung G, Ahmad MQ, Yang SH, Lee SI (2018) Comparative genomic and transcriptomic analyses of Family-1 UDP glycosyltransferase in three *Brassica* species and *Arabidopsis* indicates stress-responsive regulation. Sci Rep 8:1875
Richardson AE, Lynch JP, Ryan PR, Delhaize E, Smith FA, Smith SE, Harvey PR, Ryan MH, Veneklaas EJ, Lambers H (2011) Plant and microbial strategies to improve the phosphorus efficiency of agriculture. Plant Soil 349:121–156
Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J (2010) Genome sequence of the palaeopolyploid soybean. Nature 463:178
Severin AJ, Woody JL, Bolon Y-T, Joseph B, Diers BW, Farmer AD, Muehlbauer GJ, Nelson RT, Grant D, Specht JE (2010) RNA-Seq Atlas of *Glycine max*: a guide to the soybean transcriptome. BMC Plant Biol 10:160
Singh VK, Jain M, Garg R (2015) Genome-wide analysis and expression profiling suggest diverse roles of GH3 genes during development and abiotic stress responses in legumes. Front Plant Sci 5:789
Song Q, Hyten DL, Jia G, Quigley CV, Fickus EW, Nelson RL, Cregan PB (2015) Fingerprinting soybean germplasm and its utility in genomic research. G3: Genes Genom Genet 5:1999–2006
Sylvia DM, Jarstfer AG (1992) Sheared root inocula of vesicular arbuscular mycorrhizal fungi. Appl Environ Microbiol 58:229–232
Taylor A, Pereira N, Thomas B, Pink DA, Jones JE, Bending GD (2015) Growth and nutritional responses to arbuscular mycorrhizal fungi are dependent on onion genotype and fungal species. Biol Fertil Soils 51:801–813
Tromas A, Parizot B, Diagne N, Champion A, Hocher V, Cissoko M, Crabos A, Prodjinoto H, Lahouze B, Bogusz D (2012) Heart of endosymbioses: transcriptomics reveals a conserved genetic program among arbuscular mycorrhizal, actinorhizal and legume-rhizobial symbioses. PLoS ONE 7:e44742
Valliyodan B, Qiu D, Patil G, Zeng P, Huang J, Dai L, Chen C, Li Y, Joshi T, Song L (2016) Landscape of genomic diversity and trait discovery in soybean. Sci Rep 6:23598. https://doi.org/10.1038/srep23598
Vangelisti A, Natali L, Bernardi R, Sbrana C, Turrini A, Hassani-Pak K, Hughes D, Cavallini A, Giovannetti M, Giordani T (2018) Transcriptome changes induced by arbuscular mycorrhizal fungi in sunflower (*Helianthus annuus* L) roots. Sci Rep 8:4
Vierheilig H, Schweiger P, Brundrett M (2005) An overview of methods for the detection and observation of arbuscular mycorrhizal fungi in roots. Physiol Plant 125:393–404
Westfall CS, Zubieta C, Herrmann J, Kapp U, Nanao MH, Jez JM (2012) Structural basis for prereceptor modulation of plant hormones by GH3 proteins. Science 336:1708–1711
Wyss P, Mellor RB, Wiemken A (1990) Vesicular-arbuscular mycorrhizas of wild-type soybean and non-nodulating mutants with *Glomus mosseae* contain symbiosis-specific polypeptides (mycorrhizins), immunologically cross-reactive with nodulins. Planta 182:22–26
Yamahara T, Shiono T, Suzuki T, Tanaka K, Takio S, Sato K, Yamazaki S, Satoh T (1999) Isolation of a germin-like protein with manganese superoxide dismutase activity from cells of a moss, *Barbula unguiculata*. J Biol Chem 274:33274–33278
Yano K, Yamamoto E, Aya K, Takeuchi H, Lo P-c, Hu L, Yamasaki M, Yoshida S, Kitano H, Hirano K (2016) Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. Nat Genet 48:927
Yoshida S, Parniske M (2005) Regulation of plant symbiosis receptor kinase through serine and threonine phosphorylation. J Biol Chem 280:9203–9209
Zambolim L, Schenck NC (1983) Reduction of the effect of pathogenic, root-infecting fungi on soybean by the mycorrhizal fungus, *Glomus mosseae*. Phytopathology 73:1402–1405
Zhang H, Kjemtrup-Lovelace S, Li C, Luo Y, Chen LP, Song B-H (2017) Comparative RNA-seq analysis uncovers a complex regulatory network for soybean cyst nematode resistance in wild soybean (*Glycine soja*). Sci Rep 7:9699
Zhou Z, Jiang Y, Wang Z, Gou Z, Lyu J, Li W, Yu Y, Shu L, Zhao Y, Ma Y (2015) Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. Nat Biotechnol 33:408