INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
18–0531/201x/00–0–000–000
DOI: 10.17957/IJAB/15.0794
http://www.fspublishers.org

Full Length Article

Genetic Analysis of Nodule Traits in Soybean *Via* Wild Soybean Background Population (*Glycin soja*) and High Generation Recombination Inbred Lines [*Glycine max* (L.) Merr.]

Changyu Li[1], Huilin Chang[1], Jinhui Wang[1], Jianan Zou[1], Yan Shi[1], Yanjiao Zhang[1], Wei Wei[1], Zhongyu Wang[1], Jieqi Wang[1], Qingying Li[1], Jingyi Zhu[1], Lin Chen[1], Jianyi Li[1], Shuping Li[1], Xueying Liu[1], Hongwei Jiang[1,2], Zhenbang Hu[1], Chunyan Liu[1*], Zhengong Yin[1,3], Zhaoming Qi[1*], Yanli Zheng[1], Qingshan Chen[1*] and Dawei Xin[1*]

[1]*Key Laboratory of Soybean Biology of Chinese Ministry of Education, Key Laboratory of Soybean Biology and Breeding/ Genetics of Chinese Agriculture Ministry, College of Science, Northeast Agricultural University, Harbin, Heilongjiang Province, China*
[2]*Land Reclamation Research & Breeding Centre of Heilongjiang, Harbin, Heilongjiang Province, China*
[3]*Crop Breeding Institute, Heilongjiang Academy of Agricultural Sciences, Harbin, Heilongjiang, China*
[*]For correspondence: xdawei@163.com; qshchen@126.com; cyliucn@126.com; zhaomingqi1860@126.com

Abstract

Biological nitrogen fixation ability of soybean is depending on the nodule traits. To identify QTL that underlying nodule traits by the chromosome segment substitution line (CSSL) (Sunong14×ZYD00006, n=160) and the recombination inbred line (RIL) (Charleston×Dongnong594, n=150) in the field condition at the same time. 121 simple sequence repeat (SSR) markers in CSSL and 5,308 specific length amplified fragment sequencing (SLAF-seq) in RIL were used to construct genetic map. Nodule traits were evaluated in the field trial. QTL underlying nodule number, nodule size, and nodule dry weight were analyzed. There were nine QTLs were found in CSSLs and these QTLs were located on (LGs) A2, C1, D1b, I, K, and L, respectively. A total of 14 QTL were identified and mapped on six different LGs A1, A2, D1b, E, I, and K in RILs. Interestingly, both CSSLs and RILs had the same cluster of QTLs on LGs D1b (chromosome 2) and I (chromosome 20). Focus on these two loci, we compared the genomic substituted region of special lines in CSSL and the SLAF markers of RIL, then we narrow down the Confident region to ~320kb and ~330kb, respectively. And the QTL confident region located on LG K (chromosome 9) was narrow down to ~540kb. Eight, 28 and 41 genes were found in these three confident regions, respectively. Three candidate genes *Glyma.20g106800*, *Glyma.09g183800*, *Glyma.02g190100* could be used for further study. These results could supply strong candidate genes for soybean-rhizobium interaction. © 2018 Friends Science Publishers

Keywords: Soybean; QTL; Nodule number; Nodule size; Nodule dry weight; Genetic; RIL

Introduction

Nitrogen is the most important limiting factor for crop growth, usually provided by the use of chemical fertilizers, but it has a great cost to farmers and has a potential adverse impact on the environment (Zahran, 1999; Purcell, 2009; Islam *et al*., 2011). Soybean *Glycine max* (L.) Merr., can be obtained by biological nitrogen fixation and can be improved based on biological nitrogen fixation, inorganic-N or both (Coale *et al.*, 1985; Rondon *et al.*, 2007; Kumudini *et al.*, 2008; Santos *et al.*, 2013). Biological nitrogen fixation is mutually beneficial: although bacteria provide nitrogen sources for the growth of plants, plants provide carbon sources for bacteria survival (Santos *et al.*, 2013). The nitrogen fixation of the rhizobium of soybean could reduce the amount of chemical nitrogen fertilizer, but also keep the soybean yield and reduce production cost (Hussain *et al*., 2011). Therefore, it is necessary to clear the gene locus underlying symbiotic relationship between soybean and rhizobia.

At present, biological nitrogen fixation applied in agriculture has aroused wide public concern (Burias and Planchon, 1990; Sinclair *et al.*, 1991; King and Purcell, 2001; Tanya *et al.*, 2005; Nicolás *et al.*, 2006; Hwang *et al.*, 2014; Scheifele *et al.*, 2017). By the detection of identified QTLs support that the number of nodules, weight and quantity of the nodules directly affected the yield of the seeds (Greder *et al.*, 1986; Burias and Planchon, 1990). Only a few studies are based on quantitative trait loci (QTL) analysis of the nodules of the soybean with genetic mapping, and the use of markers on the linkage group is also limited (Santos *et al.*, 2013).

For example, only 45 simple sequence repeat (SSR) markers was tagged in QTLs underlying nodule number and nodules weight utilized a single strain of 160 F_2 population

To cite this paper: Li, C., H. Chang, J. Wang, J. Zou, Y. Shi, Y. Zhang, W. Wei, Z. Wang, J. Wang, Q. Li, J. Zhu, L. Chen, J. Li, S. Li, X. Liu, H. Jiang, Z. Hu, C. Liu, Z. Yin, Z. Qi, Y. Zheng, Q. Chen and D. Xin, 201x. Genetic analysis of nodule traits in soybean *via* wild soybean background population (*glycin soja*) and high generation recombination inbred lines [*glycine max* (L.) merr.]. *Int. J. Agric. Biol*. 00: 000-000

(Nicolás *et al.*, 2006). Tanya *et al*. (2005) used 85 SSR markers in the F_5 population to evaluate the correlation traits of 136 recombination inbred lines (RILs). Even though many traits of nodules QTL was identified, but the linkage group they used is not a complete genome covered and only based on the single marker analysis (Tanya *et al.*, 2005). Santos *et al.* (2013) used compound interval mapping to analyze the QTL of nodule number and weight of 157 $F_{2:7}$ population. Only 50 percent of the genome was covered by 97 SSR markers. Most of the QTL location of nodule trait was completed in the greenhouse condition, only few experiment was completed in the field (Hwang *et al*., 2014).

In soybean, several genes derived from QTL underlying nodulation have been described. These genes including gene *Rj1*, *Rj2*, *Rj3* and *Rj4*, then *Rj1* is a recessive gene as spontaneous non-nodulating mutant (Williams and Lynch, 1954). *Rj2*, *Rj3* and *Rj4* have the function of control the host specificity (Caldwell, 1966; Vest, 1970; Vest and Caldwell, 1972; Htwe and Yamakawa, 2017). The non-nodulation ability of the gene *Rj5* and *Rj6* (Harper and Nickell, 1995) has been identified by chemical induced mutations and corresponds to the same sites of the *Rj1*. Chemical mutagenesis also produced supernodulating mutants (*Rj7* and *Rj8*) (Vuong *et al.*, 1996; Santos *et al*., 2013). Since then, there has been less research on the regions of genes that underlying biological nitrogen fixation (BNF) in soybeans. In soybean only *Rj2* and *Rj4* were cloned based on the genetic location (Yang *et al*., 2010; Tang *et al*., 2016). In other legume, such as pea and common bean (*Phaseolus vulgaris* L.) also were used to identify the QTL underlying nodulation ability (Nodari *et al.*, 1993; Tsai *et al.*, 1998; Bourion *et al.*, 2010). So that it is helpful to use QTL find the pivotal gene relate with symbiotic traits.

In soybean, there are still few genetic was found for underlying symbiosis establishment. In this study the QTL underlying nodulation relative characters would be detected via two high generation genetic populations in the field experiment. As the nodule phenotype between the two parents of RIL and chromosome segment substitution lines is different. We wish to use this two population in the same time to locate the QTL underlying the nodule trait. Then the overlap interval appeared in two populations was further compared by high density map, which constructed by the specific length amplified fragment (SLAF). Two populations used in the same time could detect more QTL information in the same time, this is helpful to identify the same location region in different genetic map. Although the plant grown in the field versus a greenhouse are likely quite large, some novel information could also be found. In the same time the wild soybean (*Glycin soja*) genetic background would helpful to find some novel locus which is not found in normal populations (*Glycine max* (L.) Merr.). This work would find valuable candidate genes underlying soybean nodulation character, this is important for molecular mechanism understanding of symbiosis.

Materials and Methods

Genetic Populations Used in this Study

The recombinant inbred line RIL population used in the study (C×D, n=150) was derived from two individual lines, 'Charleston' and 'Dongnong594' (Chen *et al.*, 2005). The high density genetic map was constructed by the specific-length amplified fragment sequencing (SLAF-seq) (Qi *et al.*, 2014). A wild soybean chromosome segment substitution lines (CSSLs) including 160 lines has been constructed (Xin *et al.*, 2016). In these lines, different chromosomal segments of wild soybean ZYD00006 (*Glycine soja* Sieb & Zucc) had been substituted into the genetic background of Suinong 14 (*Glycine max* (L.) Merr.).

Two populations and their parents were planted in Xiangyang crop breeding base of Northeast Agricultural University (45.75ºN, 126.53ºE) in 2017. Each line was planted in row with 5 m long, the row space is 0.65 m and the seed interval is 5 cm. The management method was the same as the general field (Zhang *et al*., 2018).

Trait Evaluation

Two months after soybean germination (V10-V13 stage), phenotype of nodule traits was evaluated in all RILs and CSSLs and their parental lines. The nodule traits including nodule size (small nodule number (SNN) (diameter≦2 mm), big nodule number (BNN) (diameter>2 mm), total nodule number (TNN) and total nodule dry weight (NDW). Five plants were randomly selected for each row of each plot of each line and the parents of populations were digged out from the field carefully to detect the nodule traits. The average phenotypic value was used for QTL identification. For the evaluation of nodule dry weight (NDW) was tested after treatment in an oven at 65ºC for 48 h.

QTL Analysis

The nodule traits mapped in the RIL and CSSL population were SNN, BNN, TNN and NDW. QTL mapping and estimation of their effects were generated by the method of composite interval mapping (CIM) using the software of WinQTL Cartographer (Wang *et al.*, 2005). The control marker number and window size were 5 and 10 c*M*, respectively. A walk speed of 0.5 c*M* and the forward regression method were selected. LOD score peaks greater than 2.5 indicated the existence of QTL for all four nodule traits reported in this study. The exact method following the known protocol (Brensha *et al.*, 2012).

Genes Identification in the Confident Interval of QTL

The identification of genes located in the QTL regions was performed as described method (Qi *et al.*, 2014; Xin *et al.*, 2016). The overlap region between the location of RIL and

CSSL was compared, and narrow down by the SLAF marker. Then the genes found in the overlap region were annotated by the reference genome of Willams 82 (www.phytozome.net/soybean).

Results

Phenotype Analysis

Means, standard errors and coefficients of variation of RILs and CSSLs were calculated for all phenotypic nodule presented in Table 1 and Table 2. There is significant difference in the nodule number and nodule dry weight between Charleston and Dongnong594, Suinong14 and ZYD00006, respectively. The phenotype of nodule number in RIL all lines are more than Charleston, then the phenotype of nodule dry weigh all lines are similar as Dongnong5694 (Table 1). The phenotype of nodule number in CSSLs all lines are more than the parents, then the phenotype of nodule dry weigh all lines are also more than the parents (Table 2). Frequency distributions of nodule trait were shown in Fig. 1 and Fig. 2.

QTL Mapping

In soybean C×D RIL population, a total of nine QTL for number of small nodule (SNN), number of big nodule (BNN), total of nodule number (TNN), nodule dry weight (NDW) were identified on six different linkage groups (LGs) (Table 3). The LOD scores of identified QTL ranging from 3.0 to 6.0 (Table 3). Two QTL were identified for SNN: one located on LG D1b (chromosome 2) (qSNN001) (Fig. 4) and another located on LG C1 (chromosome 4) (qSNN002) (Table 3 and Fig. 3). The trait BNN has only one QTL (qBNN001) on LG A2 (chromosome 8) (Table 3 and Fig. 3). Three QTLs underlying TNN were located on LG D1b (chromosome 2), LG C1 (chromosome 4) and LG L (chromosome 19). For the NDW, three QTLs were identified distributed in LG D1b (chromosome 2), LG K (chromosome 9) and LG I (chromosome 20) (Table 3 and Fig. 3). The QTLs underlying SNN, TNN and NDW are consistent located on the LG D1b in the same site, and the QTLs underlying SNN and TNN are consistent located on the LG C1 in the same site (Fig. 3).

In the chromosome segment substitution lines, there are 14 QTLs for SNN, BNN, TNN, NDW were identified on different LGs (Table 4). LOD scores ranging from 4 to 16.7 (Table 4). The QTL underlying SNN was identified on LG A2 (chromosome 8) (qSNN003) (Table 4 and Fig. 4). Three QTLs underlying BNN were found on LG D1b (chromosome 2), LG K (chromosome 9), and LG E (chromosome 15), respectively. Four QTLs underlying TNN were identified on the LG D1b (chromosome 2), LG A1 (chromosome 5), LG A2 (chromosome 8) and LG K (chromosome 9), respectively. There are six QTLs underlying NDW were found on LG D1b (chromosome 2), LG A1 (chromosome 5),

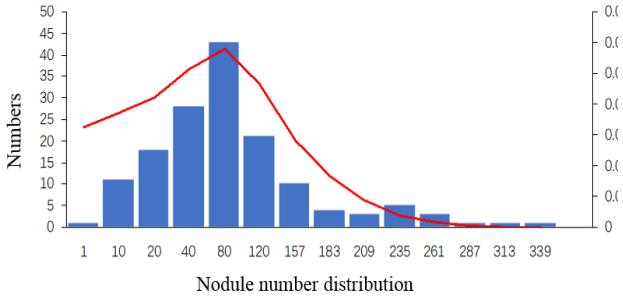

Fig. 1: Frequency distribution of phenotypic values of the nodule number in RIL

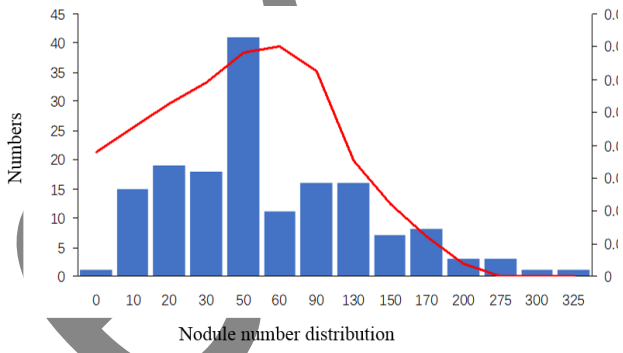

Fig. 2: Frequency distribution of phenotypic values of the nodule number in CSSL

LG A2 (chromosome 8), LG K (chromosome 9), LG E (chromosome 15) and LG I (chromosome 20), respectively. The QTLs underlying SNN and TNN are overlap located on the LG A2 in the similar region, and the QTLs underlying TNN and NDW are overlap located on the LG K in the similar region (Fig. 4).

Delimitation of QTL Regions

To narrow down the confidence region of QTL, the QTL location result of RIL and CSSL was compared. We found that the locus on D1b was identified in RIL and CSSL in the same time. The overlap region of the QTL located region between RIL and CSSL background is 74.8-76.8 c*M* (Fig. 5). Based on the SLAF marker of RIL, we narrow down the confident region to ~320kb. There are eight genes were found in this region by the prediction of Williams 82 genomic sequence (www.phytozome.net/soybean) (Table 5).

The QTL location on LG I was also found had overlap between RIL and CSSL. The overlap region of the QTL located region between RIL and CSSL background is 26.1-28.0 c*M* (Fig. 5). Based on the SLAF marker of RIL, we narrow down the confident region to ~330kb. There are 28 genes were found in this region by the prediction of Williams 82 genomic sequence (www.phytozome.net/soybean) (Table 6).

Table 1: Nodule traits in a soybean recombinant inbred line population between Charleston and Dong nong594 were evaluated

RILs (n=150)				Parents	
Trait	mean	StE	CV%	Charleston mean	Dong nong594 mean
Small Nodule Number(r≦2mm)	56.133	5.001	73.88	39	46
Big Nodule Number (r>2mm)	18.606	2.056	91.63	12	18
Total Nodule Number	74.74	5.609	91.92	51	64
Nodule Dry Weight (g)	0.166	0.011	87.80	0.098	0.163

Table 2: Nodule traits in a soybean chromosome segment substitution lines population between Suinong14 and ZYD00006 were evaluated

CSSLs (n=160)				Parents	
Trait	mean	StE	CV%	Sui nong14 mean	ZYD00006 mean
Small Nodule Number(r≦2mm)	45.162	3.745	95.29	26.25	21
Big Nodule Number (r>2mm)	17.906	2.484	57.54	13	1
Total Nodule Number	63.143	4.497	97.09	39.25	22
Nodule Dry Weight (g)	0.209	0.018	89.82	0.070	0.025

Table 3: QTL of nodule traits in soybean: number of small nodule (SNN), number of big nodule (BNN), total of nodule number (TNN), nodule dry weight (NDW) in the soybean C × D RIL population grown at Harbin in 2017

	Trait	LG	QTL	Marker/interval	Position(cM)	LOD	R^2%	Additive
	SNN	D1b	qSNN001	Mark1027864-Mark983458	73.0-76.8	5.2	13.56	-7.120
		C1	qSNN002	Mark754603-Mark763760	30.3-31.4	3.4	8.49	-3.818
	BNN	A2	qBNN001	Mark1353907-Mark1344302	53.8-55.5	4.1	16.15	-2.412
	TNN	D1b	qTNN001	Mark1027864-Mark983458	73.0-76.8	4.3	9.84	-23.54
RILs		C1	qTNN002	Mark767306-Mark763760	30.3-31.4	6.0	14.09	-29.19
		L	qTNN003	Mark937111-Mark944473	42.3-47.4	5.0	9.91	31.24
	NDW	D1b	qNDW001	Mark1027864-Mark983458	73.0-76.8	4.0	9.27	-0.009
		K	qNDW002	Mark272164-Mark273071	41.8-44.2	3.2	6.26	-0.009
		I	qNDW003	Mark1168456-Mark1164942	25.7-28.0	3.0	7.05	0.008

Table 4: QTL of nodule traits in soybean: SNN, BNN, TNN, NDW in the soybean S × Z CSSL population grown at Harbin in 2017

	Trait	LG	QTL	Marker/interval	Position(cM)	LOD	R^2%	Additive
	SNN	A2	qSNN003	Block5149-Block5150	63.1-65.4	8.8	16.43	-6.681
	BNN	D1b	qBNN002	Block 961-Block 982	15.2-16.5	4.2	4.64	-2.535
		K	qBNN003	Satt260- BARC-065467-19490	80.1-82.3	5.8	0.93	-4.747
		E	qBNN004	Block 8562-Block8563	53.0-54.7	5.7	20.40	-7.811
	TNN	D1b	qTNN004	Block1185-Block1187	74.8-76.9	4	0.143	-3.47
		A1	qTNN005	Block3098-Block 3171	75.8-80.2	5.9	0.009	-0.318
CSSLs		A2	qTNN006	Block5148-Block5150	61.8-65.4	4.1	1.044	33.57
		K	qTNN007	Block 5479-Block 5481	76.2-77.8	5.5	1.816	-26.04
	NDW	D1b	qNDW004	Block1180-Block1184	62.5-64.5	15.4	0.003	-0.003
		A1	qNDW005	Block 2984-Block 2992	39.4-41.1	9.6	4.69	-0.002
		A2	qNDW006	Block 5303-Block 5307	119.6-128.86	13.6	0.08	0.0004
		K	qNDW007	Block 5479-Block 5481	75.9-77.8	6.2	0.70	0.0006
		E	qNDW008	Block 8267-Block 1841	12.8-13.7	13.9	0.26	-0.0015
		I	qNDW009	Block11244-Block11245	26.1-28.0	16.7	0.08	-0.0003

As there is overlap QTL located region in the LG K, although this locus was found in RIL. The overlap region of the QTL located in this region is 76.2-77.8 c*M* (Fig. 3). Based on the SLAF marker of RIL, we narrow down the confident region to ~540kb. There are 41 genes were found in this region by the prediction of Williams 82 genomic sequence (www.phytozome.net/soybean) (Table 7).

Discussion

Many reports have found that the QTL region of seed protein, oil, yield and nodule traits were co-localized (Hanson *et al.*, 1961; Thorne and Fehr, 1970; Wilcox and Cavins 1995; Helms and Orf, 1998; Chung *et al.*, 2003; Hwang *et al.*, 2014). In this study, using the SLAF map, a total of 23 QTL for three root nodule traits were detected and located on eight different chromosomes of the soybean genome. Clusters of QTLs that underlying different nodule traits were observed on several chromosomes which is in agreement with previous studies (Santos *et al.*, 2013; Hwang *et al.*, 2014). For example, a region on LG D1b (chromosome 2) was identified containing two QTLs

Table 5: The predicted genes in the chromosome 02 of the overlap region

Gene ID	Function annotation
Glyma.02g189700	Argonaute family protein
Glyma.02g189800	receptor kinase 3
Glyma.02g190000	Patatin-like phospholipase family protein
Glyma.02g190100	amino acid permease 3
Glyma.02g190300	Nucleic acid-binding, OB-fold-like protein
Glyma.02g190500	PIF1 helicase
Glyma.02g190600	WD-40 repeat family protein/small nuclear ribonucleoprotein Prp4p-related
Glyma.02g190700	UDP-glcnac-adolichol phosphate glcnac-1-p-transferase

Table 6: The predicted genes in the chromosome 02 of the overlap region

Gene ID	Function annotation
Glyma.20g106800	Protein phosphatase 2C family protein
Glyma.20g106900	translation initiation factor 3 (IF-3) family protein
Glyma.20g107100	Splicing factor, CC1-like
Glyma.20g107200	FAD/NAD(P)-binding oxidoreductase family protein
Glyma.20g107300	ubiquitin-conjugating enzyme 23
Glyma.20g107400	Cupredoxin superfamily protein
Glyma.20g107500	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
Glyma.20g107600	Arabidopsis phospholipase-like protein (PEARLI 4) family
Glyma.20g107700	GCIP-interacting family protein
Glyma.20g107800	hydroxypyruvate reductase
Glyma.20g107900	AP2/B3-like transcriptional factor family protein
Glyma.20g108000	AP2/B3-like transcriptional factor family protein
Glyma.20g108200	AP2/B3-like transcriptional factor family protein
Glyma.20g108300	AP2/B3-like transcriptional factor family protein
Glyma.20g108400	AP2/B3-like transcriptional factor family protein
Glyma.20g108500	NADH-ubiquinone oxidoreductase B18 subunit, putative
Glyma.20g108600	Homeodomain-like superfamily protein
Glyma.20g108800	Uncharacterised protein family (UPF0041)
Glyma.20g109000	Sec23/Sec24 protein transport family protein
Glyma.20g109100	Nucleotide-diphospho-sugar transferase family protein
Glyma.20g109300	UDP-glucosyltransferase 73B2
Glyma.20g109400	indeterminate(ID)-domain 4
Glyma.20g109600	20S proteasome alpha subunit E2
Glyma.20g109800	Tetratricopeptide repeat (TPR)-like superfamily protein
Glyma.20g109900	GHMP kinase family protein
Glyma.20g110000	TRICHOME BIREFRINGENCE-LIKE 36
Glyma.20g110100	RmlC-like cupins superfamily protein
Glyma.20g110200	Radical SAM superfamily protein

underlying TNN and NDW, respectively. In this locus, other QTLs of oil content, resistance protein, seed coat hardness, seed weigh, and sucrose concentration were also found by different populations (Liu *et al.*, 2011). Besides, it consistent with the linkage group of soybean nodule gene *rj6* (Indrasumunar *et al.*, 2010; Hwang *et al.*, 2014). Through gene annotation found one candidate gene *Glyma.02g190100* coded an Amino acid permease 3, had a special expression in root (https://phytozome.jgi.doe.gov/pz/portal.html#).

In particular, on LG K (chromosome 9) and I (chromosome 20) contain the same QTLs between the RIL and CSSL population. In previous research in soybean, Hwang *et al.* (2014) found a QTL on Gm20 associated with nodule dry weight close to the marker BARC-051035-10955. The marker BARC-061551-17265 was very approximately linked with BARC-051035-10955 (2 c*M* distance on consensus map), which was identified as a QTL for nodule dry weight trait. The QTL for nodule number trait on Gm20

Table 7: The predicted genes in the chromosome 02 of the overlap region

Gene ID	Function annotation
Glyma.09g179200	Protein phosphatase 2C family protein
Glyma.09g179300	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
Glyma.09g179500	poly(A) polymerase 1
Glyma.09g179600	aminopeptidase M1
Glyma.09g179800	Heavy metal transport/detoxification superfamily protein
Glyma.09g179900	Copper transport protein family
Glyma.09g180100	Octicosapeptide/Phox/Bem1p family protein
Glyma.09g180200	Uncharacterised conserved protein UCP031088, alpha/beta hydrolase
Glyma.09g180300	RING/U-box superfamily protein
Glyma.09g180400	Ribosomal protein S24/S35, mitochondrial
Glyma.09g180500	Tetratricopeptide repeat (TPR)-like superfamily protein
Glyma.09g180700	Putative membrane lipoprotein
Glyma.09g180800	Reticulon family protein
Glyma.09g181000	chlorophyllase 1
Glyma.09g181100	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
Glyma.09g181200	ATP binding microtubule motor family protein
Glyma.09g181300	Protein of unknown function (DUF789)
Glyma.09g181400	Protein kinase superfamily protein
Glyma.09g181500	Protein kinase superfamily protein
Glyma.09g181600	Protein kinase superfamily protein
Glyma.09g181700	Protein kinase superfamily protein
Glyma.09g181800	Protein kinase superfamily protein
Glyma.09g181900	Transducin/WD40 repeat-like superfamily protein
Glyma.09g182000	Protein kinase superfamily protein
Glyma.09g182100	Protein kinase superfamily protein
Glyma.09g182200	Protein kinase superfamily protein
Glyma.09g182400	Calmodulin binding protein-like
Glyma.09g182500	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
Glyma.09g182600	zinc finger protein-related
Glyma.09g182700	C2 calcium/lipid-binding and GRAM domain containing protein
Glyma.09g182900	ortholog of human splicing factor SC35
Glyma.09g183100	associated molecule with the SH3 domain of STAM 1
Glyma.09g183200	S-locus lectin protein kinase family protein
Glyma.09g183300	Ribosomal protein S7e family protein
Glyma.09g183400	myb domain protein 62
Glyma.09g183500	BR enhanced expression 3
Glyma.09g183600	Plant protein of unknown function (DUF639)
Glyma.09g183700	VQ motif-containing protein
Glyma.09g183800	C2 domain-containing protein
Glyma.09g183900	CRS1/ YhbY (CRM) domain-containing protein
Glyma.09g184000	alpha/beta-Hydrolases superfamily protein

was close to the QTL for nodule number trait by Santos *et al.* (2013). In this region, QTL was related to the high protein and low fat (Santos *et al.*, 2013), anti-soybean cystic nematodes and low beta-globulin (Panthee *et al.*, 2004). The results indicate that it may be that the genes with the marker chain have a multi-effect function. On the other hand, the functional genes of these traits may exist in clusters. Besides, on LG I (chromosome 20) found a candidate gene (*Glyma.20g106800*) coded a Protein Phosphatase 2C 64-Related, which had a high expression level in root and nodule (https://phytozome.jgi.doe.gov/pz/portal.html#). On LG K (chromosome 9), the gene (*Glyma.09g183800*) coded a Predicted Ca^{2+}-dependent phospholipid-binding protein was special high expressed in nodule. As Ca^{2+} signaling pathway play pivotal role in nodule organogenesis, so that this gene could a candidate gene for further analysis.

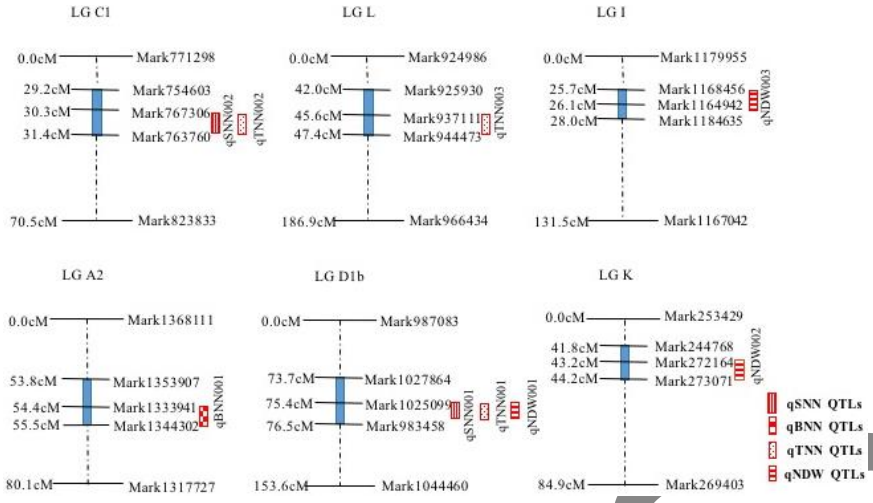

Fig. 3: Locations of DNA markers and the QTLs that underlying SNN, BNN, TNN, and NDW in RIL. Markers were assigned to LGs on the basis of soybean genetic map (Qi *et al.,*2014)

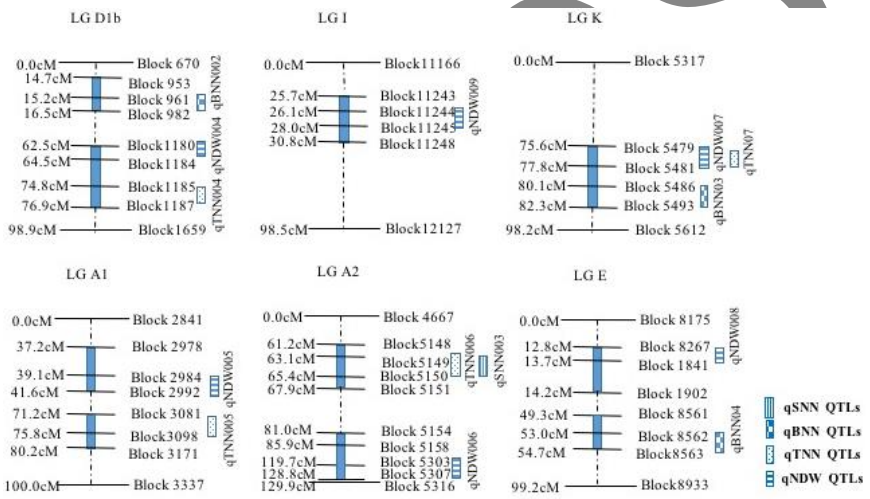

Fig. 4: Locations of DNA markers and the QTLs that underlying SNN, BNN, TNN, and NDW in CSSL

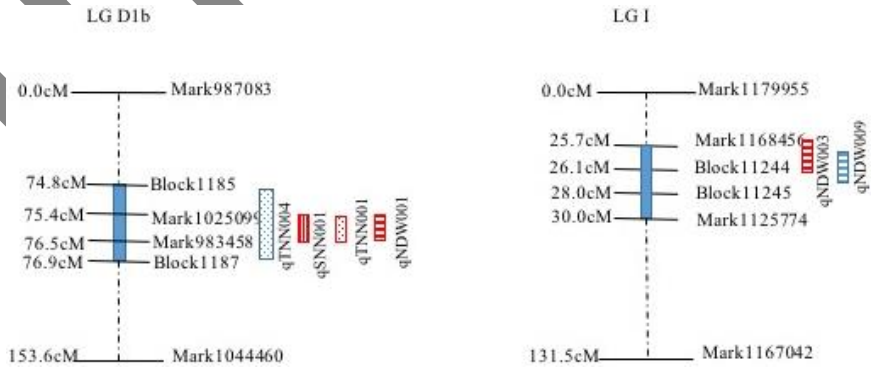

Fig. 5: Overlap regions of LG D1b and LG I between RIL and CSSL

Mapping QTL could supply powerful support for gene identification and marker-assisted selection (MAS) (Price, 2006). Many successful examples had been developed by using QTL to modify crop agronomy traits, and the exact genes also be identified for the disease resistance studying and application (Su *et al.,* 2002; Beaver and Osorno, 2009;

Blair *et al*., 2009; Sabouri, 2009; Swarbrick *et al*., 2009). Then the QTL application in root nodule traits is no so widespread, although they have the potential to increase crop stress tolerance. The pivotal first step in using these traits is to evaluate their phenotypic variation and detecting related QTL. In this study, QTL analysis was performed for nodule size, nodule number, nodule dry weight three traits in RIL (n=150) and CSSL (n=160) population, and overlaps sites were found on chromosome D1b, K, and I. This can improve the accuracy and effectiveness of locating, However, the clear background of CSSL is needed for the gene fine mapping. The CSSL used in this study would be backcrossed with the female parent Suinong 14 to get a clearer background. Then we could have more confident QTL location result.

The new QTLs we found in this study has several reasons. The different genetic background of material was the mainly reason and may be unique loci associated with nodule traits. Also, the limited genetic mapping information to identify QTL used in the earlier reports. In the current research used 3900 and 5300 markers which may have precluded discovery of QTLs in past research. Additionally, the environment may influence the expression of nodule traits, plants grown in fields and greenhouse have different effects on nodule traits.

As individual nodule weight and total nodule weight are closely associated with N_2 fixation (Tanya *et al*., 2005; Hwang *et al*., 2014), and ours identified QTL information on nodule traits in soybean from field experiments using a dense and complete linkage map. We wish that this result of nodule traits would be helpful in selecting genotypes with increased capacity for N_2 fixation.

Conclusion

In this study found 23 QTLs in total, three overlap regions was identified between the location result of RIL and CSSL. Three confident region were delimitated into ~320 kb, ~330 kb, and ~540 kb on chromosome 2, 9 and 20, respectively. Furthermore, three candidate genes, *Glyma.20g106800*, *Glyma.09g183800*, *Glyma.02g190100* were predicted involved in the nodule traits regulation. Future work could be conducted for these two genes.

Acknowledgements

Financial support was received from National Natural Science Foundation of China (31400074, 31271747, 31471516, 31401465, 31501332), National Key R&D Program of China (2016YFD0100500, 2016YFD0100300, 2016YFD0100201), Natural Science Foundation of Heilongjiang Province of China (Grant number ZD201213), 863 Program (Grant number 2013AA102602), Heilongjiang Postdoctoral Science Foundation (LBH-Q16014), Harbin Science Technology Project (2013RFQXJ005; 2014RFXXJ012), The Ministry of Science and Technology of People's Republic of China Project (2017YFE0111000); EUCLEG (727312); Academic Backbon' Project of Northeast Agricultural University (15XG02). The promote project for young talents of the colleges in Heilongjiang province UNPYSCT-2016008

References

Beaver, J.S. and J.M. Osorno, 2009. Achievements and limitations of contemporary common bean breeding using conventional and molecular approaches. *Euphytica*, 168: 145–175

Blair, M.W., T.A. Sandoval, G.V. Caldas, S.E. Beebe and M.I. Paez, 2009. Quantiative trait locus analysis of seed phosphorus and seed phytate content in a recombinant inbred line population of common bean. *Crop Sci.*, 49: 237–246

Bourion, V., S.M.H. Rizvi, S. Fournier, H. de Larambergue, F. Galmiche, P. Marget and J. Burstin, 2010. Genetic dissection of nitrogen nutrition in pea through a QTL approach of root, nodule, and shoot variability. *Theor. Appl. Genet.*, 121: 71–86

Brensha, W.B., S.K. Kantartzi, K. Meksem, I.V.R.L. Grier, A. Barakat, D.A. Lightfoot and M.A. Kassem, 2012. Genetic analysis of root and shoot traits in the 'Essex' by 'Forrest' recombinant inbred line (RIL) population of soybean [*Glycine max* (L.) Merr.]. *Plant Physiol. Biochem.*, 2012, 1: 1–9

Burias, N. and C. Planchon, 1990. Increasing soybean productivity through selection for nitrogen fixation. *Agron. J.*, 82: 1031–1034

Caldwell, B.E., 1966. Inheritance of a strain-specific ineffective nodulation in soybean. *Crop Sci.*, 6: 427–428

Chen, Q.S., C.Z. Zhang and C.Y. Liu, 2005. Construction and analysis of soybean genetic map using recombinant inbred line of Charleston× Dongnong 594. *Sci. Agric. Sin.*, 38: 1312–1316

Chung, J., H.L. Babka, G.L. Graef, P.E. Staswick, D.J. Lee, P.B. Cregan, R.C. Shoemaker and J.E. Specht, 2003. The seed protein, oil, and yield QTL on soybean linkage group I. *Crop Sci.*, 43: 1053–1067

Coale, F.J., J.J. Meisinger and W.J. Wiebold, 1985. Effects of plant breeding and selection on yields and nitrogen fixation in soybeans under two soil nitrogen regimes. *Plant Soil*, 86: 357–367

Greder, R.R., J.H. Orf and J.W. Lambert, 1986. Heritabilities and associations of nodule mass and recovery of *Bradyrhizobium japonicum* serogroup USDA 110 in soybean. *Crop Sci*., 26: 33–37

Htwe, A.Z. and T. Yamakawa, 2017. Incompatible Nodulation of *Bradyrhizobium elkanii* Strains BLY3-8 and BLY6-1 with *Rj3* Gene-Harboring Soybean Cultivars. *Amer. J. Plant Sci.*, 8: 178–179

Harper, J.E. and C.D. Nickell, 1995. Genetic analysis of non-nodulating soybean mutants in a hypernodulating background. *Soybean Genet. Newslett*., 22: 185–190

Hanson, W.D., R.C. Leffel and R.W. Howell, 1961. Genetic analysis of energy production in the soybean. *Crop Sci.*, 1: 121–126

Helms, T.C. and J.H. Orf, 1998. Protein, oil and yield in soybean lines selected for increased protein. *Crop Sci.*, 38: 707–711

Hussain, K., M. Islam, M.T. Siddique, R. Hayat and S. Mohsan, 2011. Soybean growth and nitrogen fixation as affected by sulfur fertilization and inoculation under rainfed conditions in Pakistan. *Int. J. Agric. Biol*., 13: 951–955

Hwang, S., J.D. Ray, P.B. Cregan, C.A. King, M.K. Davies and L.C. Purcell, 2014. Genetics and mapping of quantitative traits for nodule number, weight, and size in soybean (*Glycine max* L.[Merr.]). *Euphytia*, 195: 419–434

Indrasumunar, A., A. Kereszt, I. Searle, M. Miyagi, D. Li, C.D. Nguyen, A. Men, B.J. Carroll and P.M. Gresshoff, 2010. Inactivation of duplicated Nod factor receptor 5 (NFR5) genes in recessive loss-of-function non-nodulation mutants of allotetraploid soybean (*Glycine max* L. Merr.). *Plant Cell Physiol.*, 51: 201–214

Islam, M., S. Mohsan, S. Ali, R. Khalid, F.U. Hassan, A. Mahmood and A. Subhani, 2011. Growth, nitrogen fixation and nutrient uptake by chickpea (Cicer arietinum) in response to phosphorus and sulfur application under rainfed conditions in Pakistan. *Int. J. Agric. Biol.*, 13: 725–730

King, C.A. and L.C. Purcell, 2001. Soybean nodule size and relationship to nitrogen fixation response to water deficit. *Crop Sci.*, 41: 1099–1107

Kumudini, S., J. Omielan and D.J. Hume, 2008. Soybean genetic improvement in yield and the effect of late-season shading and nitrogen source and supply. *Agron. J.*, 100: 400–405 (278–290)

Liu, W.X., M.Y. Kim, K. Van, Y.H. Lee, H.L. Li, X.H. Liu and S.H. Lee, 2011. QTL identification of yield-related traits and their association with flowering and maturity in Soybean. *J. Crop Sci. Biotech.*, 14: 65–70

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

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

Panthee, D.R., P. Kwanyuen, C.E. Sams, D.R. West, A.M. Saxton and V.R. Pantalone, 2004 Quantitative trait loci for B-conglycinin (7S) and glycinin (11S) fractions of soybean storage protein. *J. Amer. Oil Chem. Soc.*, 81: 1005–1012

Price, A.H., 2006. Believe it or not, QTLs are accurate. *Trend Plant Sci*., 11: 213–216

Purcell, L.C., 2009. Physiological responses of N_2 fixation to drought and selecting genotypes for improved N_2 fixation. *Nitrogen Fixation Crop Prod*., 52: 211–238

Qi, Z., L. Huang, R. Zhu, D. Xin, C. Liu, X. Han, H. Jiang, W. Hong, G. Hu, H. Zheng and Q. Chen, 2014 A high-density genetic map for soybean based on specific length amplified fragment sequencing. *PLoS One*, 9: e104871

Rondon, M.A., J. Lehmann, J. Ramírez and M. Hurtado, 2007 Biological nitrogen fixation by common beans (*Phaseolus vulgaris* L.) increases with bio-char additions. *Biol. Fertil. Soil*, 43: 699–708

Sabouri, H., 2009 QTL detection of rice grain quality traits by microsatellite markers using an indica rice (*Oryza sativa* L.) combination. *J. Genet.*, 88: 81–85

Santos, M.A., I.O. Geraldi, A.A.F. Garcia, N. Bortolatto, A. Schiavon and M. Hungria, 2013 Mapping of QTLs associated with biological nitrogen fixation traits in soybean. *Hereditas*, 150: 17–25

Scheifele, M., A. Hobi, F. Buegger, A. Gattinger, R. Schulin, T. Boller and P. Mäder, 2017 Impact of pyrochar and hydrochar on soybean (*Glycine max* L.) root nodulation and biological nitrogen fixation. *J. Plant Nutr. Soil Sci.*, 180: 199–211

Sinclair, T.R., A.R. Soffes, K. Hinson, S.L. Albrecht and P.L. Pfahler, 1991. Genotypic variation in soybean nodule number and weight. *Crop Sci.*, 31: 301–304

Su, C.C., X.N. Cheng, H.Q. Zhai and J.M. Wan, 2002. Detection and analysis of QTL for resistance to the brown planthopper, *Nilaparvata lugens* (Stal), in rice (*Oryzasativa* L.), using backcross inbred lines. *Acta Genet. Sin.*, 29: 332–338

Swarbrick, P.J., J.D. Scholes, M.C. Press and J. Slate, 2009. A major QTL for resistance of rice to the parasitic plant *Striga hermonthica* is not dependent on genetic background. *Pest Manage. Sci.*, 65: 528–532

Tang, F., S. Yang, J. Liu and H. Zhu, 2016. Rj4, a gene controlling nodulation specificity in soybeans, encodes a thaumatin-like protein but not the one previously reported. *Plant Physiol.*, 170: 26–32

Tanya, P., P. Srinives, T. Toojinda, A. Vanavichit and S. Lee, 2005. Identification of SSR Markers Associated with N_2-Fixation Components in Soybean [*Glycine max* (L.) Merr.]. *Kor. J. Genet.*, 27: 351–359

Thorne, J.C. and W.R. Fehr, 1970. Incorporation of high-protein, exotic germplasm into soybean populations by 2 and 3 way crosses. *Crop Sci.*, 10: 652–655

Tsai, S.M., R.O. Nodari, D.H. Moon, L.E.A. Camargo, R. Vencovsky and P. Gepts, 1998. QTL mapping for nodule number and common bacterial blight in *Phaseolus vulgaris* L. *Plant Soil*, 204: 135–145

Vest, G., 1970. *Rj3* a gene controlling ineffective nodulation in soybean. *Crop Sci.*, 10: 34–35

Vest, G. and B.E. Caldwell, 1972. *Rj4* a gene controlling ineffective nodulation in soybean. *Crop Sci.*, 12: 692–693

Vuong, T.D., C.D. Nickell and J.E. Harper. 1996. Genetic and allelism analyses of hypernodulation soybean mutants from two genetic backgrounds. *Crop Sci.*, 36: 1153–1158

Xin, D., Z. Qi, H. Jiang, Z. Hu, R. Zhu, J. Hu, H. Han, G. Hu, C. Liu and Q. Chen, 2016. QTL location and epistatic effect analysis of 100-seed weight using wild soybean (*Glycine soja* Sieb. & Zucc.) chromosome segment substitution lines. *PLoS One*, 11: e0149380

Wang, S., C.J. Basten and Z.B. Zeng, 2005. Windows QTL Cartographer 2.5. Department of Statistics, NCSU, Raleigh, NC (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm)

Wilcox, J.R. and J.F. Cavins, 1995. Backcrossing high seed protein to a soybean cultivar. *Crop Sci.*, 35: 1036–1041

Williams, L.F. and D.L. Lynch, 1954. Inheritance of a non-nodulating character in the soybean. *Agron. J.*, 46: 28–29

Yang, S., F. Tang, M. Gao, H.B. Krishnan and H. Zhu, 2010. R gene-controlled host specificity in the legume–rhizobia symbiosis. *Proc. Natl. Acad. Sci. USA*, 107: 18735–18740

Zhang, D.Y., L.Q. Wu, H.D. Qi, X.R. Mao, Y. Shi, Z.Y. Wang, H.W. Jiang, Q.C. Gong, X.X. Wu, D.W. Xin, Q.S. Chen and Z.M. Qi, 2018. SNP-SNP Interaction Analysis on Soybean Seed Size in Multiple Years. *Int. J. Agric. Biol*., 20: 803–810

Zahran, H.H., 1999. Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol. Rev.*, 63: 968–989

(Received 07 April 2018; Accepted 04 June 2018)