Title

Identification of novel candidate genes associated with the symbiotic compatibility of soybean with rhizobia under natural conditions

Masayoshi Teraishi[1], Kosuke Sakaguchi[1], Takahiro Tsuchimoto[1], Takanori Yoshikawa[2]

[1]Graduate School of Agriculture, Kyoto University, Kyoto, Japan

[2]National Institute of Genetics, Mishima, Japan

Correspondence:

Masayoshi Teraishi

Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

Email: teraishi.masayoshi.5e@kyoto-u.ac.jp

Abstract

A robust symbiotic relationship between soybean and rhizobia can enhance the yield and quality of soybeans by reducing nitrogen fertilizer inputs, thereby contributing to sustainable agriculture. The genetic interplay between soybean cultivars and the rhizobial species colonizing their roots under natural conditions remains underexplored. This study builds on the observation that the prevalence of rhizobial species associated with the soybean cultivars ‘Peking’ and ‘Tamahomare’ varies significantly. Herein, we performed a quantitative trait loci (QTL) analysis of the proportion of *Rhizobium* species present in the root nodules of these cultivars using recombinant inbred lines derived from a cross between ‘Peking’ and ‘Tamahomare.’ A major QTL was identified on chromosome 18, accounting for 42% of the phenotypic variation, and was subsequently localized to a 240 kb region. The RNA-seq analysis indicated that a single gene featuring nucleotide binding site–leucine-rich repeat domains exhibited markedly different expression levels in parent cultivars within the QTL region. As this locus is distinct from the chromosomal regions containing known nodule-related genes, such as *Rj* or *rj*, it likely represents a novel gene involved in symbiosis between rhizobia and soybeans. Further research of the function and role of this new gene has potential to improve soybean yield and contribute to sustainable agriculture under low nitrogen

fertilizer conditions.

KEYWORDS: Soybean, root nodule, quantitative trait loci (QTL) analysis, RNA-seq

1 INTRODUCTION

Soybean (*Glycine max*) is one of the most economically important crops worldwide. It is highly suitable for human and animal diets and is a source of 30% of the world's oil derived from processed crops. Soybean originated in northern China and is expanding worldwide, particularly in North and South America. The use of soybean in biofuel production is further increasing its economic impact. The rising economic importance of soybean has led to increased efforts over the past several years to improve soybean productivity.

Legume plants, including soybeans, exhibit the unique characteristic of symbiosis with rhizobia for nitrogen fixation (Doyle & Luckow, 2003). Nitrogen fixed by legumes is a major nitrogen source in soil (Schaedel et al., 2021). Legume crops occupy 12–15% of the arable land worldwide (Graham & Vance, 2003), and nitrogen fixation in legume crops provides 40 million tons of nitrogen annually to agricultural land (Galloway et al., 2008). The N-fixing capacity of legumes contributes to the nitrogen supply of both natural and arable ecosystems. The effective utilization of nitrogen-fixing capacity is necessary for developing sustainable and low-input agriculture.

Several genetic factors, called *Rj* or *rj* genes, which regulate nodulation traits, have been identified in soybeans. They were identified under inoculation conditions using specific

rhizobial strains. Genetic factors under uninoculated and natural conditions have not been investigated effectively.

Our previous research (Ramongolalaina et al., 2018), performed in a natural environment under non-inoculated conditions, showed that novel genes located on chromosome 18 in soybeans control the compatibility of rhizobial strains with host plants. No genes responsible for nodulation were identified in this region. Here, we narrowed down the quantitative trait loci (QTL) region and selected candidate genes using RNA-seq analysis.

2 MATERIALS AND METHODS

2.1 Plant materials

In this study, we used 113 F_6-derived RILs, derived from a cross between their parent cultivars, ‘Tamahomare’ and ‘Peking.’ The major nodulating rhizobial species in the root nodules of ‘Tamahomare’ was *B. japonicum*, whereas the dominant rhizobial species in the root nodules of ‘Peking’ was *B. elkanii*.

Soybean seeds of each line, sterilized by soaking in 70% ethanol for 30 s and 2.5% sodium hypochlorite solution for 3 min, were sowed in a pot (10 cm x 20 cm x 15 cm) filled with soil derived from soybean planted field in a Kyoto University experimental farm and grown in an incubator set at 28□ with 18/6 h photoperiod. Plants were

removed from the soil at the fourth trifoliolate stage, a month after sowing. Twenty-four nodules per line were randomly harvested from roots.

2.2 DNA extraction

The nodules were washed to remove the soil, surface-sterilized with 70% ethanol for 1 min and 2.5% sodium hypochlorite solution for 3 min, and rinsed with sterilized water. Subsequently, the nodules were put into tubes one by one, submerged in 300 µl of TPS buffer containing 100 mM Tris-HCl, 1 M KCl, and 10.0 mM EDTA, and then crushed with multi-beads shocker (Yasui Kikai Co., Japan). After centrifugation at 10000 g for 1 min, the DNA in the supernatant was treated with isopropanol, washed twice with 70% ethanol, and dissolved in TE buffer.

2.3 PCR-based amplification and the classification of rhizobia species

The DNA sequence analysis of the ITS region between the 16S and 23S ribosomal RNA (rRNA) genes of nodule bacteria was used to classify the rhizobial species. In a previous report (Ramongolalaina et al., 2018), the classification of rhizobia species was based on restriction fragment length polymorphisms of the PCR products of the ITS region between the 16S and 23S rRNA genes. In this study, however, we adopted

Illumina MiSeq-based high-throughput analysis using the sequence information of a part of the ITS region for classifying nodule bacteria.

A conserved tRNA-Ala sequence within the ITS region of *Bradyrhizobium* spp. was used for designing the forward primer for the first round of PCR, whereas the reverse primer was designed using the ITS-R sequence at the end of the 23S rRNA gene (Akao, 2004). An eight-base index, i.e., the i5/i7 index, was added between the first primer and P5/P7 sequence (Figure S5). Primers used for amplicon sequencing are listed in Table S4. Twenty-four forward primers and 24 reverse primers were mixed to create 576 primer sets containing 10.0 μM concentration of each primer.

For the first PCR analysis, 0.1 μL of DNA template was mixed with 10 μL of EmeraldAmp MAX PCR Master Mix (TaKaRa Bío. Co., Japan) and 0.2 μL of a primer set containing 10.0 μM concentration of each primer, and the volume was made up to 20 μl with distilled water. The PCR cycle consisted of a pre-run at 96□ for 2 min, denaturation at 98□ for 20 s, annealing at 62□ for 15 s, and extension at 72□ for 1 min. The sequence was repeated for a total of 25 cycles and followed by a final post-run for 1 min at 72□. The PCR products were purified using the LaboPass Gel Extraction Kit (Hokkaido System Science Co., Japan) after electrophoresis and mixed with 576 PCR products in one tube, followed by amplicon sequencing on the MiSeq platform (2 × 300

bp) (Illumina, USA) conducted by Bioengineering Laboratory Co. Ltd. (Japan).

2.4 Estimation of rhizobial species

Primer and index sequences were trimmed from fastq files using Cutadapt (Martin, 2011), quality filtering was performed using fastp (Chen et al., 2018), and the processed fastq files were divided into RILs. The manifest file was imported into QIIME2 (Bolyen et al., 2019). Classification of nodule bacteria was performed using QIIME2 based on known *Bradyrhizobium* ITS sequences.

2.5 Construction of linkage map

A linkage map of RILs was constructed using the GRAS-Di technology developed by TOYOTA (Enoki & Takeuchi, 2018). The DNA samples of RILs were sent to Eurofins Genomics (Japan), which performed genotyping using GRAS-Di technology. Co-dominant DNA markers accrued by GRAS-Di were used for linkage map construction using MapMaker 3.0 (Lander et al., 1987).

The QTL analysis was performed using the composite interval mapping function of the R/qtl package and the Haley–Knott regression method. The LOD significance threshold for detecting QTLs was calculated by performing 1,000 iterations using the R/qtl

permutation test.

2.6 RNA-seq analysis

For RNA-seq analysis, total RNA was extracted with RNAiso Plus (TaKaRa Bio Co., Japan) from soybean roots cultivated on soil for a month and purified with RNeasy Mini Spin Columns (Qiagen, Japan) following the manufacturer's protocols. Purified RNA was sent to the Bioengineering Lab. Co. Ltd. (Japan). The RNA-seq data were analyzed using the RaNA-Seq program (https://doi.org/10.1093/bioinformatics/btz854).

3 RESULTS

3.1 Analysis of QTL

Ramongolalaina et al. (2018) reported a major QTL associated with the compatibility of indigenous rhizobia on chromosome 18 using 93 recombinant inbred lines (F_{14}-derived) (RILs) derived from a cross between 'Peking' and 'Tamahomare.'

We redeveloped another set of RILs, consisting of 113 F_6-derived between 'Peking' and 'Tamahomare,' and subjected them to high-resolution linkage mapping using genotyping by random amplicon sequencing-direct (GRAS-Di) technology. The linkage map had a total length of 2,846.3 cm and contained 860 markers, with, and the average

distance between the markers was 3.3 cm (Table S1).

Rhizobial strains attached to the roots were divided into two groups, i.e., *Bradyrhizobium japonicum* and *Bradyrhizobium elkanii*, based on the amplicon sequence analysis of the partial internal transcribed spacer (ITS) region between tRNA-Ala and the end of 23S rRNA. The results of the QTL analysis in terms of the ratio of rhizobial strains showed that a single major QTL associated with the compatibility of indigenous rhizobia in soil was located between A08389 (52,593 kbp) and A35793 (57,317 kbp) on chromosome 18, responsible for 42% of the phenotypic variation (Table 1; Figure S1). The frequency distribution of RILs for the ratio of rhizobial strains between the two homozygous genotype groups in the QTL region showed a clear bias (Figure S2). The 'Peking' homozygous RILs showed tended toward a high frequency of *B. elkanii*, whereas the 'Tamahomare' homozygous RILs showed a high frequency of *B. japonicum*. The highest limit of detection (LOD) value was observed between A01978 (56,531 kbp) and A35793 (57,317 kbp) (Figure S3). A previous study (Ramongolalaina et al., 2018) detected QTLs in the same region near the simple sequence repeat marker Sat_064 (56,334 kbp) on chromosome 18.

3.2 Fine mapping of the QTL

The RILs harboring recombination within the QTL region were selected, and additional polymerase chain reaction (PCR)-based markers for detailed genotyping were designed (Table S2) to narrow down the QTL.

Finally, the location of genes determining compatibility with rhizobia was narrowed from 56,433 kbp to 56,675 kbp on chromosome 18 (Figure 1). There were 22 predicted gene models, named Glyma.18g280800-Glyma.18g283100, based on the genome assembly sequence data of Williams 82 (Glyma.Wm82.a2) (Table 2).

3.3 RNA-seq analysis and candidate genes

The RNA-seq analysis was performed using RNA derived from the roots of ‘Peking’ and ‘Tamahomare’ a month after germination to identify the potential candidate genes responsible for QTL. The results of the RNA-seq analysis revealed that only a single gene, Glyma.18g281700, was differentially expressed between the parents within the QTL region with padj<0.05 and fold change >2 (Figure S4; Table S3). The expression levels of Glyma.18g281700 were significantly higher in ‘Peking’ than in ‘Tamahomare.’ Glyma.18g281700 encodes NBS–ARC (apoptosis, R proteins, and CED-4) and LRR domains, often observed in disease resistance (R) genes.

4 DISCUSSION

Several genes regulating nodulation have been reported in soybeans. These are called *Rj*(s) and *rj*(s) genes, and *Rj2*, *Rj3*, *Rj4*, and *Rfg1* are the dominant alleles that restrict nodulation in specific strains (Hayashi et al., 2012), of which *Rj2* and *Rfg1* are allelic genes located on chromosome 16 and encode a member of the Toll/ interleukin-1 receptor–NBS–LRR resistance protein (Fan et al., 2017). The *Rj4* allele is located on chromosome 1 and encodes a thaumatin-like protein (Hayashi et al., 2014; Tang et al., 2014; Tang et al., 2016), and *Rj3* is restricted to nodulation by *B. elkanii* USDA33 (Vest 1970) but has not yet been identified. Other recessive genes that regulate nodulation induce nonnodulation (*rj1*, *rj5*, and *rj6*) (Pracht et al., 1993; Lee et al., 2011; Indrasumunar et al., 2011) and hypernodulation phenotypes (*rj7*) (Nishimura et al., 2002a; Nishimura et al., 2002b).

Ramongolalaina et al. (2018) demonstrated, for the first time, the chromosomal location of genes responsible for the proportion of nodulating bacterial strains in roots under natural non-inoculated conditions. In this study, we successfully narrowed down the location of the gene underlying the QTL to a 240 kb interval, wherein 22 predicted gene models were expected. RNA-seq revealed a single significantly differentially expressed gene between ‘Peking’ and ‘Tamahomare’ in the QTL interval.

Three consecutive genes, i.e., Glyma.18g281500, Glyma.18g281600, and Glyma.18g281700, have ubiquitin-like protease 1, NBS–ARC, and LRR domains, respectively (Jebanathirajah et al., 2002; Steele et al., 2019; Wei et al., 2023). These three tandemly arranged genes showed more than 90% sequence similarity. Glyma.18g281500 was located just outside the QTL region, and Glyma.18g281600 was within the QTL region, but these two genes did not exhibit a differential expression between the parents. Glyma.18g281700 alone differed substantially in its expression between the parents.

The genes *Rpp1* and *Rpp1b* that confer resistance to *Phakopsora pachyrhizi*, the causative agent of rust, are located within the same region on chromosome 18 (Pedley et al., 2019; Barros et al., 2023; Wei et al., 2023; Yamanaka et al., 2023). Pedley et al. (2019) reported that eight genes were homologous to the NBS–LRR family of disease R genes in four bacterial artificial chromosome contigs on chromosome 18, including the *Rpp1* locus. Three genes, namely Glyma.18g281500, 281600, and 281700, were considered he candidate genes for *Rpp1*, and Glyma.18g281600 was the most highly expressed candidate gene in PI200492. However, Barros et al. (2023) showed that the *Rpp1* gene in PI594756 was located immediately upstream of *Rpp1* in PI200492. Barros et al. (2023) indicated that genomic variations, such as presence/ absence and copy

number variations, cause the diversification of disease R genes.

In conclusion, a single major QTL located on chromosome 18 was associated with the symbiotic compatibility of nodules. No genes implicated in symbiosis between rhizobia and soybeans were observed on chromosome 18. A single candidate gene was identified with differential transcription levels between 'Peking' and 'Tamahomare.' This candidate gene bears the NBS−LRR domain and has been reported as a candidate gene for soybean rust resistance, *Rpp1*.

Supplemental data files

Supplemental table titles

Table S1: Genetic map summary generated by GRAS-Di

Table S2: List of primer sequences used for fine mapping of the QTL located on Chromosome 18

Table S3: Summary of differential expressed genes analysis within the QTL region.

Table S4: Primers used for amplicon sequence

Supplemental figure titles

Figure S1: QTL location on Chromosome 18

Figure S2: Distribution of the RILs for ratio of rhizobia strains used homozygous genotypes at the QTL region from the recombinant inbred lines from ‘Peking’ and ‘Tamahomare’

Figure S3: S3 Schematic view of position of molecular markers around the QTL region on Chromosome 18

Figure S4: Heatmap of the 22 candidate genes in ‘Peking’ and ‘Tamahomare’. Figure S5: Schematic representation of PCR amplification for ITS region.

ACKNOWLEDGMENTS

This work was supported by a Grant-in-Aid for Scientific Research (C) (General) (for MT) (JSPS KAKENHI Grant number 22K05571) and The Kyoto University Foundation (for MT).

We would like to thank Editage (www.editage.jp) for English language editing.

AUTHOR CONTRIBUTIONS

MT conceived and designed the study. TY and TT developed the mapping population, constructed the genetic linkage map, and performed the fine mapping. TT and KS collected the phenotypic data. MT performed RNA-seq analysis and wrote the

manuscript. All the authors contributed to the manuscript and approved the submitted version.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The RNA-seq raw data used in this study were downloaded from the DRA database in the DDBJ under the BioProject accession number PRJDB17790.

ORCID

Masayoshi Teraishi: 0000-0002-6768-8052

Takanori Yoshikawa: 0000-0002-8631-4216

Kosuke Sakaguchi: 0009-0008-0943-7862

REFERENCES

Akao, S. (2004). Grouping of Bradyrhizobium USDA strains by sequence analysis of 168 rDNA and 168-238 rDNA internal transcribed spacer region. *Soil Science and Plant Nutrition*, *50*, 517–525.

Barros, L. G., Avelino, B. B., da Silva, D. C. G., Ferreira, E. G. C., Castanho, F. M., Ferreira, M. E., Lopes-Caitar, V. S., Marin, S. R. R., Arias, C. A. A., de ON Lopes, I., Abdelnoor, R. V., & Marcelino-Guimarães, F. C. (2023). Mapping of a soybean rust resistance in PI 594756 at the *Rpp1* locus. *Molecular Breeding*, *43*, 12.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., Cope, E. K., Da Silva, R., Diener, C., Dorrestein, P. C., Douglas, G. M., Durall, D. M., Duvallet, C., Edwardson, C. F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J. M., Gibbons, S. M., Gibson, D. L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G. A., Janssen, S., Jarmusch, A. K., Jiang, L., Kaehler, B. D., Kang, K. B., Keefe, C. R., Keim, P., Kelley, S. T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M. G. I., Lee, J., Ley, R., Liu, Y.-X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B. D., McDonald, D., McIver, L. J., Melnik, A. V., Metcalf, J. L., Morgan, S. C., Morton, J. T., Naimey, A. T., Navas-Molina, J. A., Nothias, L. F., Orchanian, S. B., Pearson, T., Peoples, S. L., Petras, D., Preuss, M. L., Pruesse, E., Rasmussen, L. B., Rivers, A., Robeson, M. S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S. J., Spear, J. R., Swafford, A. D., Thompson, L. R., Torres, P. J., Trinh, P., Tripathi, A., Turnbaugh, P. J., Ul-Hasan, S., van der Hooft, J. J. J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K. C., Williamson, C. H. D., Willis, A. D., Xu, Z. Z., Zaneveld, J. R., Zhang, Y., Zhu, Q., Knight, R., & Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*, 852–857.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultrafast all-in-one FASTQ preprocessor. *Bioinformatics* (Oxford, England), *34*, i884–i890.

Doyle, J. J., & Luckow, M. A. (2003). The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant Physiology*, *131*, 900–910.

Enoki, H., & Takeuchi, Y. (2018). New genotyping technology, GRAS-Di, using next

generation sequencer. Proceedings of the Plant and Animal Genome Conference XXVI. San Diego, CA.

Fan, Y., Liu, J., Lyu, S., Wang, Q., Yang, S., & Zhu, H. (2017). The Soybean Rfg1 Gene Restricts Nodulation by Sinorhizobium fredii USDA193. *Frontiers in Plant Science*, *8*, 1548-1548.

Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P., & Sutton, M. A. (2008). Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science*, *320*, 889–892.

Graham, P. H., & Vance, C. P. (2003). Legumes: importance and constraints to greater use. *Plant Physiology*, *131*, 872–877.

Hayashi, M., Saeki, Y., Haga, M., Harada, K., Kouchi, H., & Umehara, Y. (2012). Rj (rj) genes involved in nitrogen-fixing root nodule formation in soybean. *Breeding Science*, *61*, 544–553.

Hayashi, M., Shiro, S., Kanamori, H., Mori-Hosokawa, S., Sasaki-Yamagata, H., Sayama, T., Nishioka, M., Takahashi, M., Ishimoto, M., Katayose, Y., Kaga, A., Harada, K., Kouchi, H., Saeki, Y., & Umehara, Y. (2014). A thaumatin-like protein, Rj4, controls nodule symbiotic specificity in soybean. *Plant Cell Physiology*, *55*, 1679–1689.

Indrasumunar, A., Searle, I., Lin, M.-H., Kereszt, A., Men, A., Carroll, B. J., & Gresshoff, P. M. (2011). Nodulation factor receptor kinase 1α controls nodule organ number in soybean (Glycine max L. Merr). *The Plant Journal*, *65*, 39–50.

Jebanathirajah, J. A., Peri, S., & Pandey, A. (2002). Toll and interleukin-1 receptor (TIR) domain-containing proteins in plants: a genomic perspective. *Trends in Plant Science*, *7*, 388–391.

Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daly, M. J., Lincoln, S. E., & Newberg, L. A. (1987). MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, *1*, 174–181.

Lee, W. K., Jeong, N., Indrasumunar, A., Gresshoff, P. M., & Jeong, S. C. (2011). Glycine max non-nodulation locus rj1: a recombinogenic region encompassing a SNP in a lysine motif receptor-like kinase (GmNFR1α). *Theoretical and Applied Genetics*, *122*, 875–884.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, *17*, 10.

Nishimura, R., Hayashi, M., Wu, G.-J., Kouchi, H., Imaizumi-Anraku, H., Murakami, Y., Kawasaki, S., Akao, S., Ohmori, M., Nagasawa, M., Harada, K., &

Kawaguchi, M. (2002a). HAR1 mediates systemic regulation of symbiotic organ development. *Nature*, *420*, 426–429.

Nishimura, R., Ohmori, M., Fujita, H., & Kawaguchi, M. (2002b). A Lotus basic leucine zipper protein with a RING-finger motif negatively regulates the developmental program of nodulation. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 15206–15210.

Pedley, K. F., Pandey, A. K., Ruck, A., Lincoln, L. M., Whitham, S. A., & Graham, M. A. (2019). Rpp1 Encodes a ULP1-NBS-LRR Protein That Controls Immunity to Phakopsora pachyrhizi in Soybean. *Molecular Plant-Microbe Interactions*, *32*, 120–133.

Pracht, J. E., Nickell, C. D., & Harper, J. E. (1993). Genes Controlling Nodulation in Soybean: Rj5 and Rj6. *Crop Science*, *33*, 711–713.

Ramongolalaina, C., Teraishi, M., & Okumoto, Y. (2018). QTLs underlying the genetic interrelationship between efficient compatibility of Bradyrhizobium strains with soybean and genistein secretion by soybean roots. *PLoS ONE*, *13*, e0194671.

Schaedel, M., Hidrobo, G., & Grossman, J. (2021). From microns to meters: exploring advances in legume microbiome diversity for agroecosystem benefits. *Frontiers in Sustainable Food Systems*, *5*, 668195.

Steele, J. F. C., Hughes, R. K., & Banfield, M. J. (2019). Structural and biochemical studies of an NB-ARC domain from a plant NLR immune receptor. *PLoS ONE*, *14*, e0221226.

Tang, F., Yang, S., Liu, J., Gao, M., & Zhu, H. (2014). Fine mapping of the *Rj4* locus, a gene controlling nodulation specificity in soybean. *Molecular Breeding*, *33*, 691–700.

Tang, F., Yang, S., Liu, J., & Zhu, H. (2016). *Rj4*, a Gene controlling nodulation specificity in soybeans, encodes a Thaumatin-like protein but not the one previously reported. *Plant Physiology*, *170*, 26–32.

Vest, G. (1970). *Rj3* A gene conditioning ineffective nodulation in soybean. *Crop Science*, *10*, 34–35.

Wei, W., Wu, X., Garcia, A., McCoppin, N., Viana, J. P. G., Murad, P. S., Walker, D. R., Hartman, G. L., Domier, L. L., Hudson, M. E., & Clough, S. J. (2023). An NBS-LRR protein in the *Rpp1* locus negates the dominance of *Rpp1*-mediated resistance against *Phakopsora pachyrhizi* in soybean. *The Plant Journal*, *113*, 915–933.

Yamanaka, N., Aoyagi, L. N., Hossain, M. M., Aoyagi, M. B. F., Muraki, Y. (2023). Genetic mapping of seven kinds of locus for resistance to Asian soybean rust. *Plants* (Basel), *12*, 2263.

FIGURE LEGEND

FIGURE 1 Fine mapping of the quantitative trait locus (QTL) on chromosome 18.

Table 1 QTLs detected for nodule number ratios of *B. japonicum* and *B. elkanii*

QTL name	Trait	Chromosome	Peak(cM)[1]	LOD[2]	Variance(%)
qBJ_18	Ratio of *B. japonicum*	18	77.6	15.13	42.33
qBE_18	Ratio of *B. elkanii*	18	77.6	15.39	42.92

(1) QTL position from the first marker.
(2) Maximum likelihood LOD score for the individual QTL.
(3) Phenotypic variation explained by each QTL.

Table 2 Candidate genes within the narrowed QTL region.

marker	Gene Name (Wm82.a4.v1)	Gene Name (Wm82.a2.v1)	Annotations
Mk1-4			
	Glyma18g51591	Glyma.18G280800	ROTUNDIFOLIA like 21
	Glyma18g51600	Glyma.18G280900	allene-oxide cyclase
	Glyma18g51626	Glyma.18G281000	uncharacterized protein
	Glyma18g51643	Glyma.18G281100	DNA-directed RNA polymerase I and III subunit Rpc19
	Glyma18g51660	Glyma.18G281200	60S ribosomal protein L30-like
	Glyma18g51670	Glyma.18G281300	Potential peptide N-acetyl transferase
	Glyma18g51680	Glyma.18G281400	ethylene-responsive transcription factor
Mk1-5			
		Glyma.18G281451	MEI2-like protein 5
	Glyma18g51715	Glyma.18G281500	Disease resistance protein
	Glyma18g51741	Glyma.18G281600	NB-ARC domain disease resistance protein
A01978			
	Glyma18g51765	Glyma.18G281700	NB-ARC domain disease resistance
	Glyma18g51790	Glyma.18G281800	myb/SANT-like DNA-binding domain protein
	Glyma18g51810	Glyma.18G282000	ADP-ribosylation factor GTPase-activating
	Glyma18g51820	Glyma.18G282100	Protein kinase superfamily protein
	Glyma18g51830	Glyma.18G282200	serine carboxypeptidase-like
	Glyma18g51840	Glyma.18G282300	C-terminal domain phosphatase-like
	Glyma18g51850	Glyma.18G282400	protein bicaudal C homolog 1-B-like
	Glyma18g51871	Glyma.18G282500	E3 ubiquitin protein ligase DRIP2-like
	Glyma18g51880	Glyma.18G282600	Disease resistance-responsive

Table 1 QTLs detected for nodule number ratios of *B. japonicum* and *B. elkanii*

QTL name	Trait	Chromosome	Peak（cM）(1)	LOD(2)	Variance（%）
qBJ_18	Ratio of *B. japonicum*	18	77.6	15.13	42.33
qBE_18	Ratio of *B. elkanii*	18	77.6	15.39	42.92

(1) QTL position from the first marker.
(2) Maximum likelihood LOD score for the individual QTL.
(3) Phenotypic variation explained by each QTL.

Table 2　Candidate genes within the narrowed QTL region.

marker	Gene Name (Wm82.a4.v1)	Gene Name (Wm82.a2.v1)	Annotations
Mk1-4			
	Glyma18g51591	Glyma.18G280800	ROTUNDIFOLIA like 21
	Glyma18g51600	Glyma.18G280900	allene-oxide cyclase
	Glyma18g51626	Glyma.18G281000	uncharacterized protein
	Glyma18g51643	Glyma.18G281100	DNA-directed RNA polymerase I and III subunit Rpc19
	Glyma18g51660	Glyma.18G281200	60S ribosomal protein L30-like
	Glyma18g51670	Glyma.18G281300	Potential peptide N-acetyl transferase
	Glyma18g51680	Glyma.18G281400	ethylene-responsive transcription factor
Mk1-5			
		Glyma.18G281451	MEI2-like protein 5
	Glyma18g51715	Glyma.18G281500	Disease resistance protein
	Glyma18g51741	Glyma.18G281600	NB-ARC domain disease resistance protein
A01978			
	Glyma18g51765	Glyma.18G281700	NB-ARC domain disease resistance
	Glyma18g51790	Glyma.18G281800	myb/SANT-like DNA-binding domain protein
	Glyma18g51810	Glyma.18G282000	ADP-ribosylation factor GTPase-activating
	Glyma18g51820	Glyma.18G282100	Protein kinase superfamily protein
	Glyma18g51830	Glyma.18G282200	serine carboxypeptidase-like
	Glyma18g51840	Glyma.18G282300	C-terminal domain phosphatase-like
	Glyma18g51850	Glyma.18G282400	protein bicaudal C homolog 1-B-like
	Glyma18g51871	Glyma.18G282500	E3 ubiquitin protein ligase DRIP2-like
	Glyma18g51880	Glyma.18G282600	Disease resistance-responsive
	Glyma18g51890	Glyma.18G282700	histone-lysine N-methyltransferase SUVR5-like isoform X2
	Glyma18g51900	Glyma.18G282800	Ran binding protein
		Glyma.18G282900	Unknown protein
	Glyma18g51911	Glyma.18G283000	RNA recognition motif
	Glyma18g51920	Glyma.18G283100	Eukaryotic aspartyl protease family
Mk2-1			

Figure1 Fine mapping of the QTL locus on Chr.18

Marker name	A04404*	Mk1-1	Mk1-2	Mk1-3	Mk1-4	Mk1-5	A01978*	Mk2-1	Mk2-2	Mk2-3	Mk2-4	A35793*	*B. japonicum* (%)	*B.elkanii* (%)
Position	56,118kb	56,218kb	56,296kb	56,399kb	56,433kb	56,493kb	56,531kb	56,674kb	56,852kb	56,982kb	57,189kb	57,317kb		
Peking	A	A	A	A	A	A	A	A	A	A	A	A	27.8	72.2
Tamahomare	B	B	B	B	B	B	B	B	B	B	B	B	75.0	25.0
PT109	B	B	B	B	B	B	B	B	B	B	B	A	58.3	41.6
PT72	B	B	B	B	B	B	B	B	B	B	B	A	83.3	8.3
PT84	B	B	B	B	B	B	B	B	B	B	B	A	62.5	33.3
PT166	B	B	B	B	B	B	B	B	B	B	A	A	83.3	16.7
PT32	B	B	B	B	B	B	B	B	A	A	A	A	50.0	45.8
PT90	B	B	B	B	B	B	B	B	A	A	A	A	50.0	50.0
PT94	B	B	B	B	B	B	B	A	A	A	A	A	79.1	20.8
PT77	A	A	A	A	A	A	A	A	A	A	A	B	33.3	66.7
PT158	A	A	A	A	A	A	A	A	B	B	B	B	15.4	84.6
PT36	A	A	A	A	A	B	B	B	B	B	B	B	87.5	12.5
PT10	A	B	B	B	B	B	B	B	B	B	B	B	54.2	37.5
PT37	A	B	B	B	B	B	B	B	B	B	B	B	75.0	25.0
PT197	A	B	B	B	B	B	B	B	B	B	B	B	95.8	4.2

*: GrasDi marker

A and B means ‘Peking’ and ‘Tamahomare’ genotypes, respectively.

PT number indicates line number of progenies crossed between ‘Peking’ and ‘Tamahomare’.