THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Doctor of Philosophy (Agronomy)
DEGREE

Agronomy	Agronomy
FIELD	**DEPARTMENT**

TITLE: Locating SSR Markers Linking to QTLs Associating with N_2 fixation Components in Soybean

NAME: Miss Patcharin Tanya

THIS THESIS HAS BEEN ACCEPTED BY

Peerasak Srinives **THESIS ADVISOR**
(Professor Peerasak Srinives, Ph.D.)

COMMITTEE MEMBER
(Associate Professor Apichart Vanavichit, Ph.D.)

T. Toojinda **COMMITTEE MEMBER**
(Mr. Theerayut Toojinda, Ph.D.)

Sontichai Chan **DEPARTMENT HEAD**
(Assistant Professor Sontichai Chanprame, Ph.D.)

APPROVED BY THE GRADUATE SCHOOL ON 21 November 2015

Vinai Artkongharn **DEAN**
(Associate Professor Vinai Artkongharn, M.A.)

THESIS

LOCATING SSR MARKERS LINKING TO QTLs ASSOCIATING WITH N_2 FIXATION COMPONENTS IN SOYBEAN

PATCHARIN TANYA

A Thesis Submitted in Partial Fulfilment of

the Requirements for the Degree of

Doctor of Philosophy (Agronomy)

Graduate School, Kasetsart University

2005

ISBN 974-9843-91-6

Patcharin Tanya 2005: Locating SSR Markers Linking to QTLs Associating with N_2 Fixation Components in Soybean. Doctor of Philosophy (Agronomy), Major Field: Agronomy, Department of Agronomy. Thesis Advisor: Professor Peerasak Srinives, Ph.D. 86 pages.
ISBN 974-9843-91-6

The amount of N_2 fixation in soybean (*Glycine max*) relies mainly on interaction between the genotypes of soybean and the symbiotic bacteria *Bradyrhizobium japonicum.* Five N_2 fixation traits, viz. nodule number per plant, nodule fresh weight per plant, nodule dry weight per plant, plant dry weight, and acetylene reduction activity (ARA) were determined in 136 recombinant inbred lines (RILs) derived from a cross between SJ2 and Suwon157. A genetic linkage map of RILs was constructed using 78 simple sequence repeat (SSR) and 7 SSR labeled M13 (-21) primers. The map consisted of 20 linkage groups covering 1093.9 cM. Five QTLs located on linkage group (LG) O, J, I, D1b+W and A1 were found associating with nodule number. These QTL explained 19.78% of phenotypic variance. Suwon157 contributed the positive alleles on LG O, J, I and D1b+W and SJ2 contributed the positive alleles on LG A1. Three QTLs explained 23.75% of phenotypic variance of nodule fresh weight was mapped to LG O, J and A1. The favorable alleles of these QTLs were derived from Suwon157 for LG O and J and from SJ2 for LG A1. Two QTLs explained 16.86% of phenotypic variance of nodule dry weight was identified on LG O and K. Suwon157 contributed the positive allele at both QTL loci. Four QTLs were also associated with plant dry weight. These QTLs explained 18.61% of phenotypic variance were located on LG O, J and I in which Suwon157 contributed all of the positive allele. Two QTLs conditioning ARA were mapped on LG O and D1b+W, with the positive allele from Suwon157. These QTLs explained 11.26% of phenotypic variance. The coincidence of QTL associated with N_2 fixation traits on LG O and D1b+W indicated the important of these two genomic segments on N_2 fixations. Markers linked to these QTLs are useful tools for improving N_2 fixations in soybean breeding program.

Patcharin Tanya Peerasak Srinives 14/November/2005

Student's signature Thesis Advisor's signature

ACKNOWLEDGEMENTS

This dissertation could not be completed without the support of the Royal Golden Jubilee Ph.D. Program of the Thailand Research Fund (TRF). I would like to express my sincere thanks to my thesis advisor, Professor Dr. Peerasak Srinives for his kind support, advice, encouragement, and suggestion throughout the study. I also thank Associate Professor Dr. Apichart Vanavichit, Dr. Theerayut Toojinda, and Assistant Professor Dr. Julapark Chunwongse for giving me valuable suggestions, comments and kindness.

I would like to pay special thanks to Professor Dr. Suk-Ha Lee, Division of Plant Science, College of Agriculture and Life Science, Seoul National University, Korea who partially supported me the grant (code no. CG 3121) from Crop Functional Genomic of 21st Century Frontier Program of the Ministry of Science and Technology while I was in Korea. Mr. Bo-Keun Ha of the same lab who introduced to me the DNA technique, and all the lab members in Suwon and Seoul in Korea who gave me new experience of Korean style.

Thank to the Molecular Breeding Soybean Lab and National Instrumentation Center for Environmental Management at Seoul National University, the Republic of Korea for the service of ABI Prism ®377 sequencer (AB-PEC, Foster City, CA) and to their staff who gave me kind assistance.

Thank to Mr. Somsak Kotepong, Dr. Achara Nuntakij, and their staff at the Soil Microbiology Group of the Department of Agriculture, Ministry of Agriculture and Cooperatives, Thailand, who provided helps on *Bradyrhibium* and Leonard jar's method.

Thank to the staff of the Tropical Vegetable Research Center and the Asian Vegetable Research and Development Center (ARC-AVRDC), Kasetsart University, Kamphaeng Saen for the helps in field work.

Thank to the Project on Genetics and Breeding of Field Legumes for Thailand, funded to Kasetsart University by the TRF for partial funding of my research and the friendship and helps provided by its staff.

Finally, I thank my parents, Mr. Arb and Mrs. Somruay, and my older sister Miss Nuantip Tanya for their encouragement and kind supports.

Patcharin Tanya

September, 2005

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

LIST OF TABLES

LIST OF TABLES (Cont'd)

LIST OF FIGURES

LIST OF FIGURES (Cont'd)

LOCATING SSR MARKERS LINKING TO QTLs ASSOCIATING WITH N_2 FIXATION COMPONENTS IN SOYBEAN

INTRODUCTION

Soybean (*Glycine max* (L.) Merr.) is originated in China, Manchuria, and Korea. There was an evidence that soybean was domesticated around the eleventh century B.C. in the eastern half of northern China and extended to central and southern China, as well as peninsular Korea in the first century A.D. After 15^{th} to 16^{th} century, soybean was brought into many countries, including northern India, Indonesia, Japan, Malaysia, Myanmar, Nepal, the Philippines, Thailand and Vietnam (Hymowitz, 1990). Due to the expansion in feed industry, Thailand could not be self-sufficiency in soybean. In the crop year 2002-3, harvested area, production, and yield per ha were 0.175 Mha, 260 Kton and 1.49 t/ha, respectively (Anon, 2004). The factors responsible for low yield in Thailand are low seed quality, biotic and abiotic problems, and low production efficiency. Fertilizer application in soybean is limited due to high cost. However, like most legumes, soybean can fix N_2 from the air through symbiosis between soybean and the bacteria *Bradyrhizobium japonicum* associated in nodules. N_2 gas is transformed from inorganic to organic compounds by nitogenase enzyme (Mylona, *et al*., 1995). The advantages of N_2 fixation having are to decrease the use of N fertilizer, to enhance seed protein production, and to contribute N for plants grown in the same area. However, efficiency of N_2 fixation in a legume-rhizobium system is not easily measured. Five fixation components have been used as the indicator so far. They are nodule number per plant, nodule fresh weight, nodule dry weight, plant dry weight, and acetylene reduction activity (ARA). Each component cannot account for the major part of the efficiency and they should be considered together. Accumulation of the fixation components into the same soybean line can be made possible through marker-assisted selection. With this technique, soybean lines carrying all the components can be identified through the markers linking to them.

To develop such markers, this project used recombinant inbred lines derived from the cross between SJ2 and Suwon157 for constructing linkage map from 77 SSR markers and 7 SSR labeled M13 (-21) primers. Then the QTL controlling each fixation component can be assigned into the map for future uses.

OBJECTIVES

1. To determine genetic variation in major Thai and Korean soybean varieties. The emphasis will be placed on the variation of N_2 fixation components, viz. nodule number per plant, nodule fresh weight per plant, nodule dry weight per plant, plant dry weight, and acetylene reduction activity (ARA).

2. To determine the relationship between N_2 fixation components, and between the components and rhizobium strains.

3. To identify the SSR and SSR labeled M13 (-21) markers linking to N_2 fixation components in the RILs derived from a soybean cross from Thai and Korean varieties.

LITERATURE REVIEW

The population in the world is still increasing exponentially while the food production is continuously reduced due to overgrazing, deforestation, salinisation, erosion, etc. To increase crop production, inputs like chemical fertilizers are required, especially to produce high protein crops. In this instance legumes have an advantage over other crops in fixing its own nitrogen fertilizer. Like most legumes, soybean (*Glycine max* (L.) Merr.) can fix N_2 from the air through symbiosis between soybean and the bacteria *Bradyrhizobium japonicum* associated in nodules. N_2 fixation in leguminous crops helps reducing the use of N fertilizer, minimizing ground water pollution, increasing seed protein production and giving N fertilizer to the succeeding crops. Soybean can use N from atmosphere, soil and fertilizer (Haper, 1999). People *et al.* (1994) suggested two strategies to enhance N_2 fixation, viz. through crop and soil management, and through plant breeding and selection. In the first strategy, the crop that can fix more N_2 should be planted first in a cropping system, especially at early wet season to maximize N_2 fixation. Other cultural practices that can enhance N_2 fixation are

P-fertilization, irrigation, no or minimum tillage, and the use of rhizobium inoculation. In the second strategy, legume plants should be bred for more specific to rhizobium strain to improve plant yield and tolerance to environmental stress. Thus traits related to N_2 fixation should be directly or indirectly evaluated to support breeding and selection. A direct method is to grow soybean in N-free medium and observe the N_2 fixation ability of soybean genotypes. The alternative indirect measurement is to observe on plant characters affecting N_2 fixation potential, viz. nodule number per plant, plant dry weight, nodule fresh weight, and acetylene reduction activity (ARA) (King and Purcell, 2001; Pazdernik *et al*., 1996). Nitrogen fixation activity increases sharply after flowering stage and gradually decreases after green pod stage (Hardy, 1968). The best stage for determining ARA is at R_3 stage, since the sensitivity decreases rapidly after this stage (Attewell and Bliss, 1985)

Soybean (*Glycine max* (L.) Merr) and its uses

The legumes are classified into three subfamilies, *Mimosoideae, Caesalpinoidaeae,* and *Papilionoideae*. The genus *Glycine* belongs to Tribe *Phaseolae* and composes of two subgenera, *Glycine* and *Soja*. The wild perennial soybean belongs to subgenus *Glycine*, but cultivated and wild annual soybean belong to subgenus *Soja* (Palmer *et al*., 1996). The soybean plant is called a legume because it collects nitrogen from the air and releases it back into the soil. This is important for growing healthy crops and maintaining soil quality. When crushed and pressed, soybeans produce an oil that is

used in margarine, cooking oils, and many prepared foods. Soybean oil is also used to make paints, varnishes, cosmetics and plastics. Soybean oil is even used to make ink for printing newspapers and magazines. After the oil was removed in processing, the remaining flakes are processed into food products or protein meal for animal feed. Soybean meal is the most important protein source for livestock and poultry. Whole soybeans are used for bean sprouts, tofu, soy sauce and soymilk. Researchers are still looking for new uses of soybeans. Recent developments include soy diesel, building materials, candles, road dust suppressants, and preventive foods for osteoporosis, cancer and diabetes.

The soybean genome and genetic linkage map

The three types of genetic markers can be used in the genome. Morphological, protein-based, and DNA based markers usually used for genome mapping. Soybean mapping was first published in 1990 which 150 restriction fragment length polymorphism (RFLP) markers were located (Keim *et al*., 1990). Also in 1990, William *et al*. (1990), another marker based on the polymerase chain reaction (PCR), amplification of random DNA segments with single primers of arbitrary nucleotide sequence was random amplified polymorphic DNA (RAPD) markers were used to located in the genome. In 1992, Shoemaker *et al*. (1992) identified molecular probes that can detect polymorphism among soybean genotypes, can be used for analyzing genome structure and molecular pedigree analysis. The SSR or microsatellite method was primarily used as a highly polymorphic genetic marker in soybean (Akkaya *et al*., 1992) while Morgante (1994) demonstrated that SSR always have higher number of alleles per locus than RFLP markers. In 1995, Shoemaker and Specht (1995) used 7 pigmentation, 6 morphological, 7 isozyme, 8 RAPD, and 110 RFLP markers to map 60 $F_{2:3}$ progeny lines derived from crossing between Clark and Harosoy. In addition, Akkaya *et al*. (1995) mapped 40 SSR markers to compare with those reported by Shoemaker and Specht (1995) of researching in 60 F_2 plants from a cross between near isogenic lines of Clark and Harosoy. The RAPD marker was modified to increase the percentage of RAPD polymorphism by digesting the DNA template with restriction enzymes before amplification (Ferreira *et al*., 1997). Mansur *et al*. (1996) mapped agronomic traits using recombinant inbred lines (RILs). They stated that the RILs constitute a permanent population that can be evaluated over space and time, permitting reduction of experimental error and linkage error as compared to using F_2 or backcross populations. Keim *et al*. (1997) reported 840 marker map consisting of 165 RFLP, 25 RAPD, and 650 AFLP markers spreading over 28 linkage groups that represent 3441 cM distance. The AFLP marker was a useful approach for generating high density genetic map in soybean. Cregan *et al*. (1999)

reported 606 SSRs, 689 RFLP, 79 RAPD, 11 AFLP, 10 isozyme, and 26 classical loci that were mapped to one or more of 3 populations: the USDA/Iowa State *G. max* x *G. soja* F_2, the University of Utah Minsoy x Noir 1 recombinant inbred lines, and the University of Nebraska Clark x Harosoy F_2 population. Zhang *et al*. (2004) used 452 markers to map on 184 recombinant inbred lines (RILs) derived from the cross Kefeng No.1 and Nannong 1138-2. They finally covered 21 linkage groups with the distance of 3,595.9 cM. The map markers consisted of 189 RFLPs, 219 SSRs, 40 ESTs, 3 *R* gene, and 1 phenotypic marker. Song *et al*. (2004) demonstrated a soybean genetic map spanning 2,523.6 cM of Kosambi map distance across 20 linkage groups that contained 1,849 markers, including 1,015 SSRs, 709 RFLPs, 73 RAPDs, 24 classical traits, 6 AFLPs, 10 isozymes, and 12 others which mapped in one or more of five soybean mapping populations: Minsoy x Noir 1, Minsoy x Archer, Archer x Noir 1, Clark x Harosoy, and A81-356022 x PI468916.

How soybean respond to nitrogen fixation

Leguminous plants produce root nodules within which symbiotic bacteria capture atmospheric N_2 and convert it into nitrogen that can be used by the plant (Downie and Parniske, 2002). The beneficial plant-associated microbes can profoundly influence plant health by suppressing disease, enhancing nutrient uptake, fixing atmospheric nitrogen, and promoting plant growth (Smith and Goodman, 1999). Bacteria from the family Rhizobiaceae establish a symbiosis with leguminous plants to form nitrogen-fixing root nodules. The specific rhizobium of soybean is *Bradyrhizobium japonicum*, which is slow growing, rod-shaped, gram negative, aerobic, 0.5-0.9 μm x 1.2-3.0 μm in size, non-spore forming, motile by flagella, forming white and opaque colonies (Young, 1999).

Root infection by rhizobia is a multistep process that is initiated by preinfection events in the rhizosphere (Figure 1). Rhizobia respond by positive chemotoxis to plant root exudates as flavonoid and move toward localized sites on the legume roots (Ferguson and Mathesius, 2003). The rhizobia attach to the surface of young growing root hair, the primary target sites of infection, and then induce curling and branching. Consequently, the nodule is developed (Figure 2). Simultaneously, pericycle and outer (determinate) cortical cells are activated and divide to form the nodule primodium that contained large amounts of amyloplasts. Carbon flowed through the stele is required for starting with the initial steps of nodule development. In determinate nodules, cell expansion of a transient meristematic region determined nodule growth (Crespi and Gálvez, 2000).

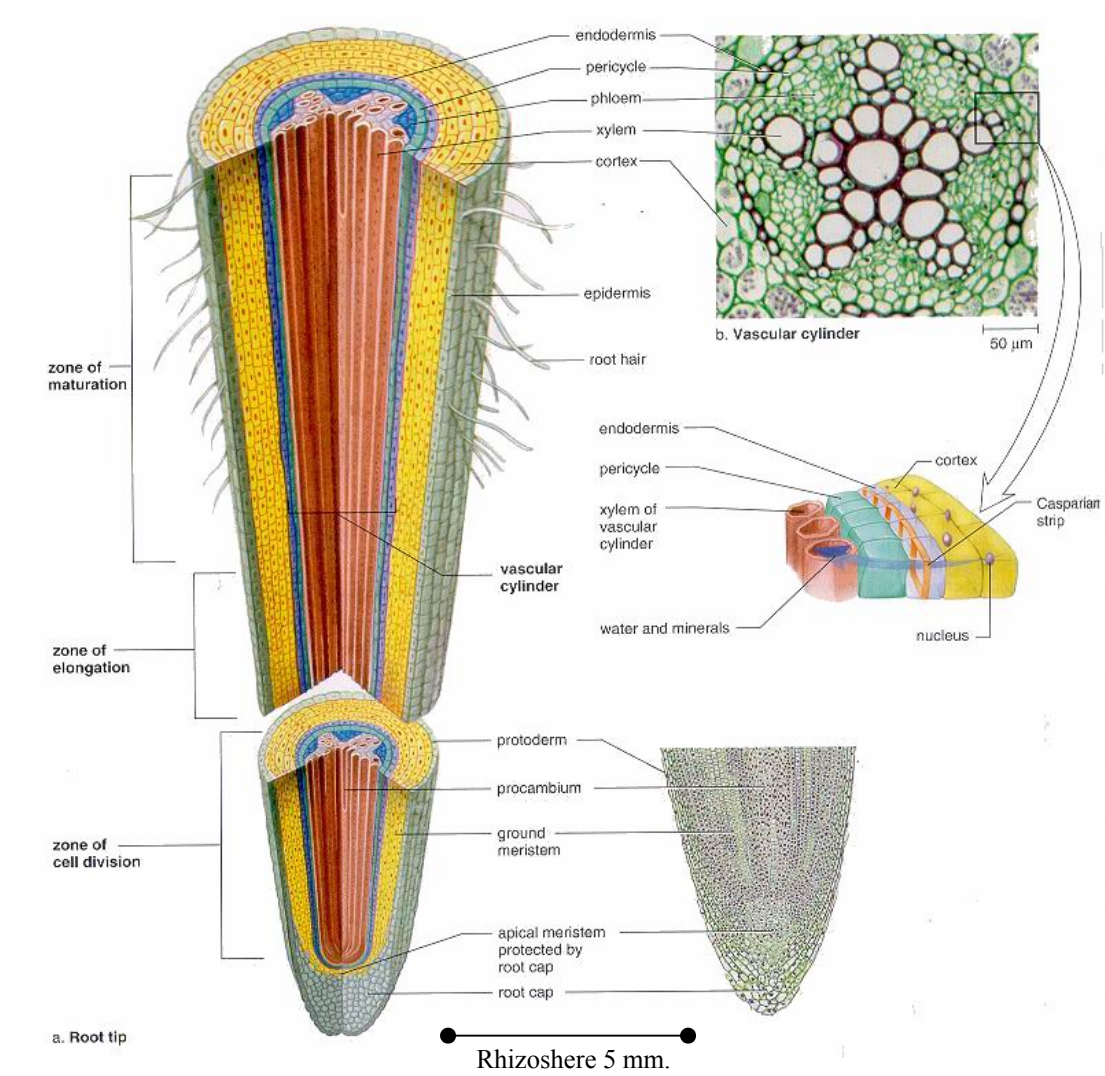


Figure 1 Structure of root and the corresponding rhizosphere: a = The root tip is divided into three zones and longitudinal section, and b = The vascular cylinder of eudicot root.

Source: Mader (2004)

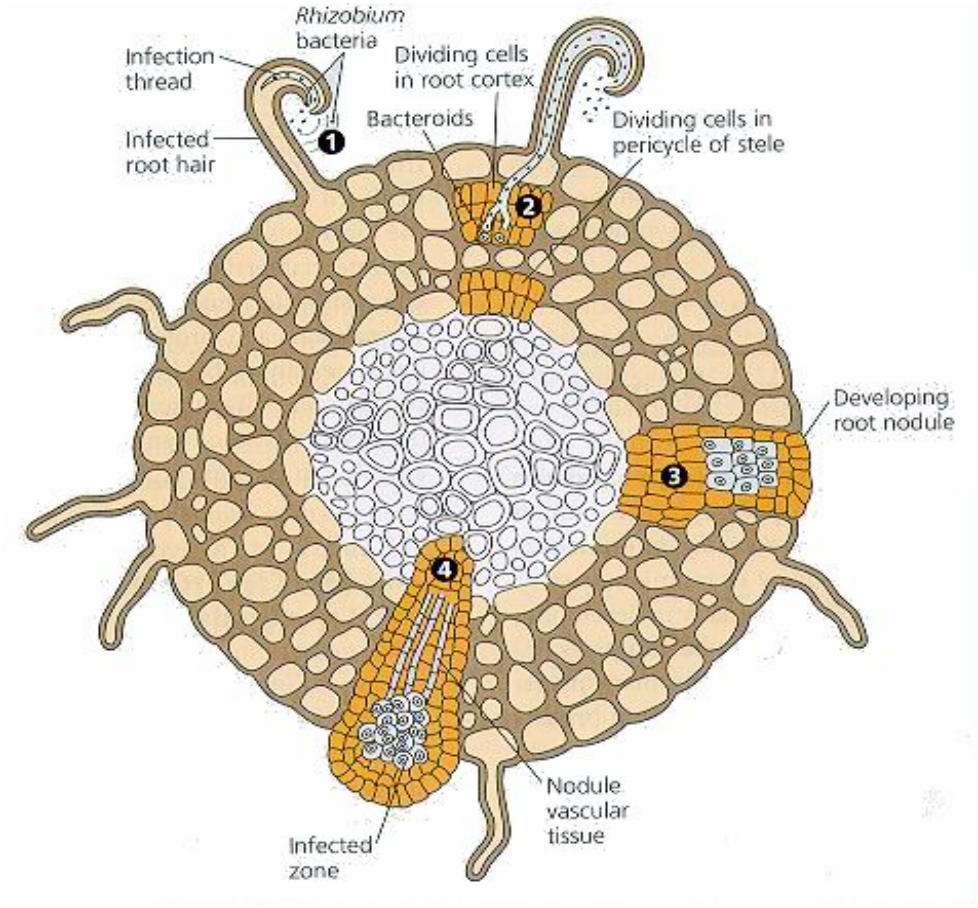


Figure 2 Nodule initiation and development: 1 = Roots emit chemical signals that attract rhizobium bacteria. The bacteria then emit signals that stimulate roots hairs to elongate and to form an infection thread by an invagination of plasma membrane, 2 = The bacteria penetrate the root cortex within the infection thread. Cells of the root cortex and the pericycle of the stele begin dividing, and vesicles containing the bacteria bud into the cortical cells from the branching infection thread, 3 = Growth continues in the affected regions of the cortex and pericycle, and these two masses of dividing cells fuse, forming nodule, 4 = The nodule continue to grow, and vascular tissue supplies nutrients to the nodule and carries nitrogeneous compounds from the nodule into stele for distribution to the rest of the plant

Source: Campbell and Reece (2002)

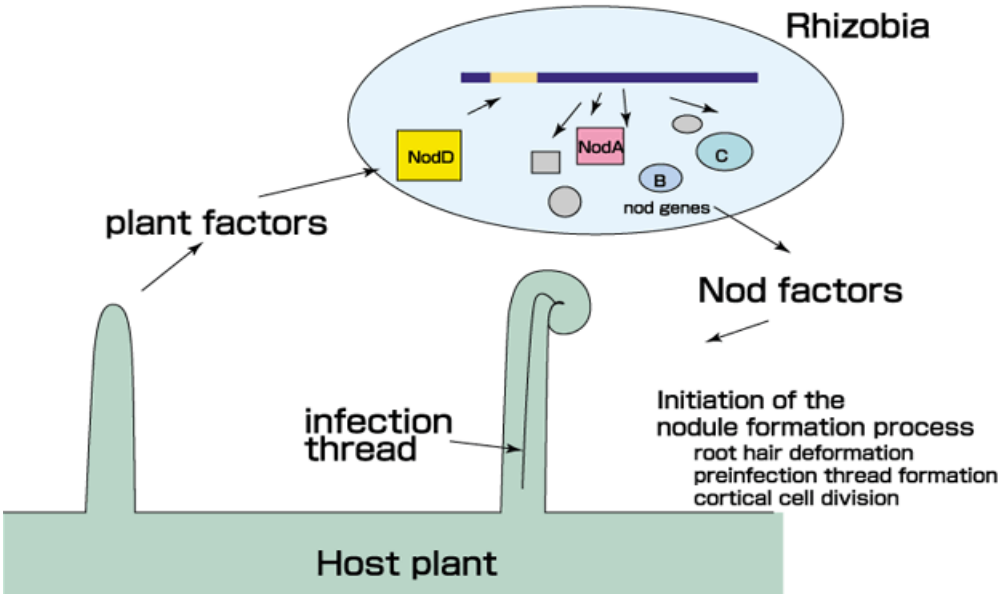


Figure 3 Diagram showing how rhizobium recognized an infecting position on soybean root.

Source: Anon (2005)

For the plant genes acting in nodules, signals from rhizobial bacteria (call nod factors which appear to be lipo-oligosaccharides) are crucial for the symbiotic response of legume. This response leads to recognition of bacteria by root-hair cells, curling of root hairs, growth of infection threads, and finally the formation of root nodules (nodulation) (Cullimore and Dènariè, 2003) (Figure 3). The flavonoid activates the *nod D* gene (regulatory nod gene) of specific rhizobial strains. The *nod D* gene in turn switches on other bacterial *nod* gene, enabling the cell to synthesize and secrete specific chemicals (nod factors), which trigger nodule formation the host Nod factors appears to be determined by subset of nod gene present in all rhizobia (*nod ABC*). Strain-specific structure modifications of nod factor framework are encoded by the variable *nod EFGHLMN.* The specific association between host plant and the bacterium requires the chemical lectins, which recognize specific structural determinants on the bacterium. Lectins together with flavonoids contribute to the specificity of the host-symbiont interaction. Among the steps in nodulation triggered by nod factors are curling of root hairs, forming of infection threads, and initiating of cortical cell division for prenodule formation. Once the rhizobium moves through the infection thread to the cortex of the root it will change to the form of bacteroid. The nod factor then activates the level of Ca that alters the shape of root hair. The process causes depolarization at cell membrane of plant root, coupled with formation of cytokinin which activates mitosis in the cell to become a nodule. (Gustavo *et al*., 1991).

Eighty percent of nitrogen in the atmosphere is not used by plant in the gas (N_2) form. Nitrogen is normally changed from a gaseous form before being incorporated into organic compounds in living organisms. Certain of steps involving in the nitrogen cycle are shown in Figure 4. Ammonium is made available to plants by two types of soil bacteria, one that fixes atmospheric N_2 (nitrogen-fixing bacteria) and another that decomposes organic material (ammonifying bacteria). Plants absorb less ammonium than nitrate from the soil. The nitrate is produced from ammonium by nitrifying bacteria. Plants reduce nitrate back to ammonium before incorporating the nitrogen into organic compounds. Xylem transports nitrogen from roots to shoots in the form of nitrate, amino acid, and various other organic compounds. Denitrification bacteria change nitrate form to N_2 and release to the atmosphere again. Nitrogen is fixed into ammonium, a reduced form that is very toxic for the plant cell. Thus, the plant cell must rapidly assimilate the ammonium released by the bacteroid into amino acids catalyzed by nitrogenase enzyme (Mylona, 1995). This process is a highly complex system with dinitrogenase reductase and dinitrogenase as the main component. The following equation is a conversion of atmospheric nitrogen into ammonia:

$$N_2 + 8H^+ + 8e^- + 16\text{Mg-ATP} \longrightarrow 2NH_3 + H_2 + 16\text{Mg-ADP} + 16\text{Pi}$$

The complex system is highly conserved and present in the cytoplasm of bacteroid. Leghemoglobins are the most abundant host gene products in nitrogen fixing legume root nodules formed as a result of symbiotic association between plant and rhizobium. They are localized in the cytoplasm and function in providing an adequate supply of oxygen for bacteroid respiration while protecting the oxygen sensitive nitrogenase enzyme (Lee *et al*., 1983). The property of nitrogenase enzyme is changed N_2 from the air to ammonium ion or ammonia. The enzyme consists of 2 components, Mo-Fe protein and Fe protein. The first mode of action of nitrogenous begins from releasing an electron of NADH in the citrate cycle. The electron is transferred via soluble ferredoxin to dinitrogenase reductase. The electron from ferredoxin passes through Fe protein and catch Mg and ATP from photosynthesis. The Fe protein then sends electron to Mo-Fe protein and reduce N_2 into NH_3 (Peter *et al*., 1995) (Figure 5).

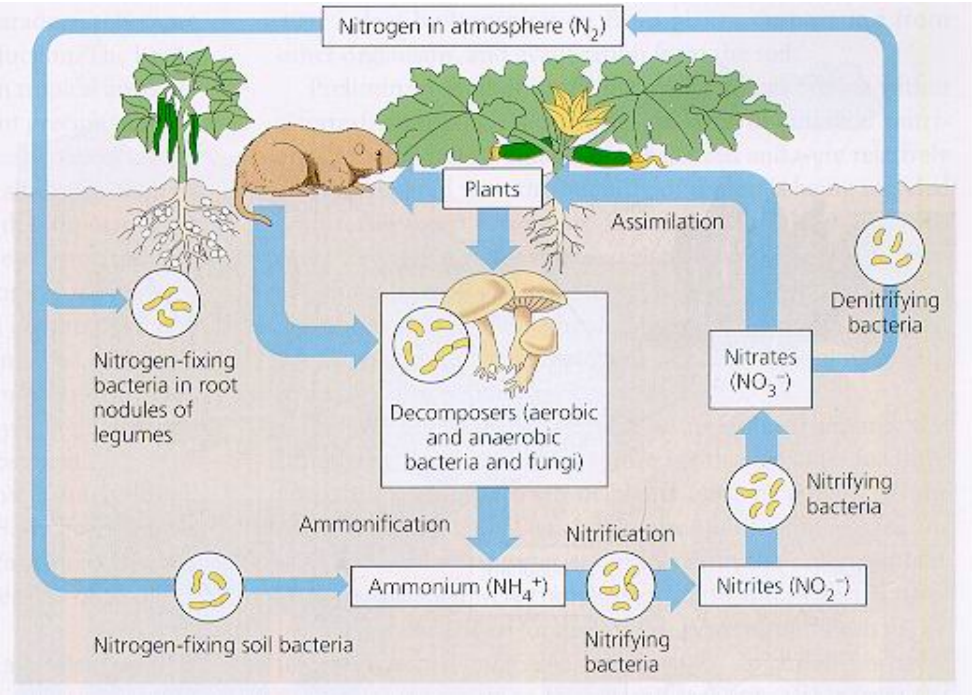


Figure 4 The nitrogen cycle. The nitrogen from the atmosphere can be fixed by N_2 fixing bacteria, assimilated into nitrates useful to plants, and finally returned via denitrification to the air

Source: Campbell and Reece (2002)

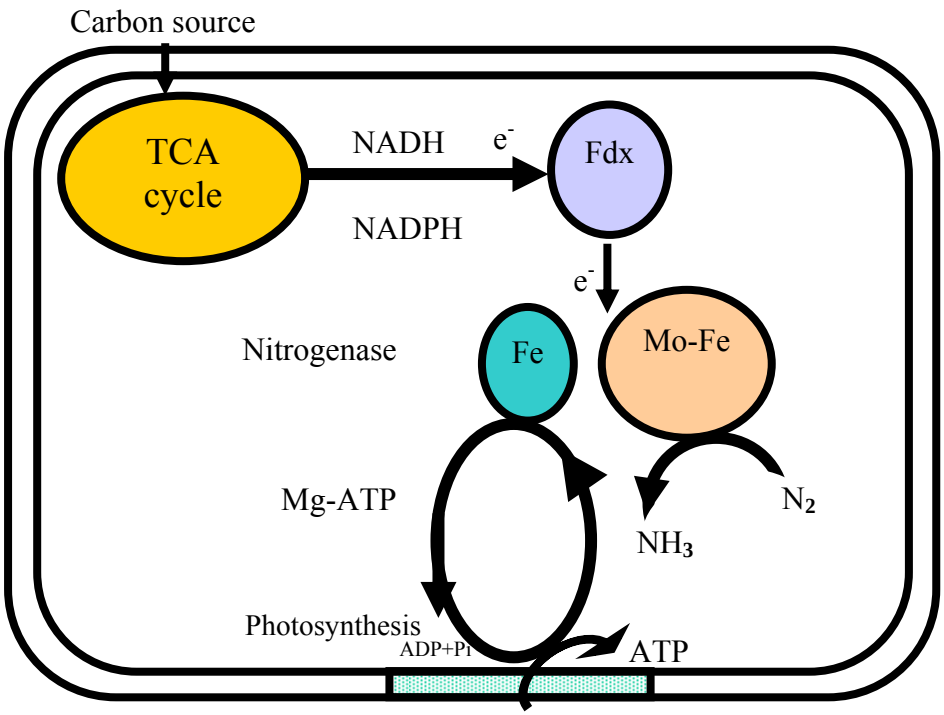


Figure 5 Reaction catalyzed by nitognenase in bacteroid of root nodules.

Inheritance of nodulation in soybean

Nodulation is a process of fixing nitrogen gas into root hairs of soybean by the symbiotic rhizobium bacteria (Gresshoff, 1990). A number of plant alleles controlling nodulation have been identified. The rj_1 allele was found conditioning restriction of nodulation in a broad range of rhizobium strains (Caldwell, 1966), whereas the dominant allele Rj_2 conditioned an ineffective nodulation in strain USDA7, 14 and 122 (Devine *et al*., 1991). The dominant allele Rj_3 conditioned an ineffective nodulation upon inoculated with USDA33 (Vest 1970). Rj_4 is a dominant allele conditioning an ineffective nodulation upon inoculated with USDA61 (Vest and Caldwell, 1972). The two recessive genes (rj_5 and rj_6) condition non-nodulation NN5 and that nod139 is allelic to NN5 (Pracht *et al*., 1993). Hypernodulation in soybean mutants is controlled by rj_7 (Vuong *et al*., 1996, Vuong and Harper, 2000), and rj_8 (Vuong *et al*., 1996). Whereas supernodulation was controlled by *nts* (Kolchinsky *et al*., 1997).

In previous studies, an integrated genetic linkage map of soybean has been constructed with RFLP, AFLP, RAPD, SSR, and isozymes. Rj_1 allele was found to locate on linkage group D1b+W, while the gene for ineffective bradyrhizobia nodulation (Rj_2) was on linkage group J (Cregan *et al*., 1999). The gene Rj_2, powdery mildew resistance (*Rmd*), and phytophthora root and stem rot resistance (Rps_2) were located in classical linkage group 19 in the BARC-4 crossed with Clark63 population (Devine *et al*., 1991). A study of the progenies derived from two isolines of the soybean cultivar 'Williams' showed that linkage group J was related to classical linkage group 19 with the genes Rj_2, *Rmd,* and Rps_2 (Polzin *et al*., 1994). A supernodulation was found in linkage group H using RFLP markers located on the linkage group (Kolchinsky *et al*., 1997).

DNA markers

Restriction fragment length polymorphism (RFLP)

RFLP is a technique that uses restriction enzymes to cut DNA into fragments in different lengths. Restriction endonucleases recognize specific 4, 5, 6, or 8 base pair (bp) nucleotide sequences and cut them accordingly. Restricted fragments are transferred to a filter for detecting polymorphism by specific probe using Southern blot analysis (Southern, 1976). The major strengths of RFLP marker are its codominant nature, no need prior information of DNA sequence, and all alleles are seen at a time. The disadvantages of this marker are its low throughput, laborious and expensive, and yet it requires high quality DNA.

Randomly Amplified Polymorphism DNA (RAPD)

RAPD is a molecular marker based on the differential PCR amplification of a sample of DNAs from short oligonucleotide sequences (Williams *et al*., 1990). This molecular marker is of dominant type and based on the PCR amplification of random locations in the genome. With this technique, approximately 10 nucleotide single stand is used to amplify the genomic DNA. The number of amplification products is directly related to the number and orientation of the sequences that are complementary to the primer in the genome. The advantage of RAPD marker is that the primers are commercially, it does not require prior knowledge of the target DNA sequence, yet it is a multilocus amplification, cheap, and easily to use. The disadvantage of this maker is its reproducibility and interpretation.

Amplified Fragment Length Polymorphism (AFLP)

AFLP is a molecular marker generated by a combination of restriction digestion and PCR amplification (Vos *et al*., 1995). The unique feature of the technique is the addition of adaptors of known sequence to the DNA fragments generated by digestion of whole genomic DNA. Pre-selective and selective primers are used to detect bands and separated by gel electrophoresis. The power of this procedure is that it can generate a large number of mapable loci with a single amplification. This facilitates saturating a region of the genome (Keim *et al*., 1997). The efficiency of this technique is its rapid generation, high reproductability, and high polymorphism. However this maker is expensive to generate as the bands are detected by silver strain, fluorescent dye and radioactive compounds.

Simple Sequence Repeat (SSR)

Simple sequence repeat (SSR), also known as microsatellite or short tandem repeat (STR) or simple sequence length polymorphisms (SSLPs), is a repeated nucleotide sequence of 2-7 base pair units. Repetitive sequences resulted from slippage in replication (Schlötterer and Tautz, 1992), and unequal crossover (Smith, 1976). This method has many advantages such as rapid, reliable (Diwan and Cregan, 1997), abundance (Lagercrantz *et al*., 1993), co-dominance (Akkaya *et al*., 1992), high heterozygosity (Powell *et al*., 1996), and high polymorphism (Akkaya *et al*., 1995). The SSR technique uses PCR to amplify DNA fragments by repeated cycles of DNA denaturing, annealing and extension

using DNA polymerase enzyme. The resulting DNA was separated by banding on gel (Akkaya *et al*., 1992).

Single Nucleotide Polymorphism (SNPs)

SNP is a single base variation between two otherwise identifical DNA sequences. Brookes (1999) reported that SNPs are single base pair positions in genomic DNA at which different sequence alternatives (alleles) exist in normal individuals in natural populations. SNPs is a powerful tool and the most frequent type of variation in the human genome (Wang *et al*., 1998). The technology allows greater number of tests to be run at a significantly lower cost than the other technologies. SNPs have been applied in genome sequencing, SNP identification and typing, screening for genetically-linked diseases, identification of genetic drug targets, screening individuals for potential drug side effects, gene cloning, screening potential tissue donors, screening cancer cells for genes conferring chemotherapy resistance, typing pathogenic and antibiotic resistant bacteria, forensics and pharmacogenomics. This technique has advantages over the other markers such as isozymes, RFLP, RAPD, or SSR, in that it expresses very large number of polymorphic loci distributing throughout the genome. The markers present within the coding regions, introns and regions that flank genes. Yet the technique is simple and unambiguous to assay with high levels of polymorphism in the population, stable Mendelian inheritance, and low level of spontaneous mutation (Brown, 1999).

Genetic mapping

Genetic mapping is the process in making-up parental crosses to quantify recombination frequencies, which is equivalent to measuring the degree of crossing over between the given two markers. More markers mean more resolution and better maps, as double-crossovers between markers become evident by employing three-factor crosses (also called three-point crosses). Fifty percent recombination indicates that two genes are genetically unlinked. They may be on the same chromosome but far apart, or on the different chromosomes. Genome mapping can be used to determine markers associating with interesting phenotype. There are 3 steps in conducting genetic mapping, viz. production of a mapping population, map construction, and QTL analysis.

Mapping population

Mapping populations are the population of backcross (BC), double haploids (DH), F_2 and recombinant inbred lines (RILs) used for genetic mapping (Liu, 1998). Backcross population is produced from crossing the first generation (F_1) with one parent or an individual genetically identical to one of the two parents. The genotypes in a backcrossing are AA, and Aa or Aa and aa. Double haploid (DH) is derived by doubling chromosomes of anthers from an F_1 plant with chochicine. The resulting genotypes are either AA or aa. The F_1 individuals can be randomly crossed among themselves to develop an F_2 population. The F_2 genotypes can be AA : Aa : aa at the expected ratio of 1 : 2 : 1. A population of RILs were developed by single-seed descent from individual plants of an F_2 population. The genotypes of the RILs are either AA : aa at the 1 : 1 ratio. The disadvantages of each population type are that RILs require long time to develop. DH is difficult to generate through anther culture. However, these 2 types of population can be replicated in the experiment. The population of F_2 can give large genetic information, while both F_2 and BC require relatively short time to develop.

Map construction

The data required for marker map construction concluding morphological markers, protein marker, and various types of DNA markers. Genotypic data can be used to construct map by a number of steps. In single locus analysis, each mapping population will give a specific segregation ratio at each locus. Information from these ratios is required to determine if the population is expressing a skewed segregation at any locus. In an F_2 population, a ratio of co-dominant marker is 1 : 2 : 1, while that of dominant marker is 3 : 1. In backcross and RIL populations the ratio of both codominant and dominant markers are 1 : 1. This ratio can be checked for goodness-of-fit by Chi square method. In two-locus analysis, the order and distance between two loci can be measured. In three locus analysis, the more accurate analysis allowing double crossing over can be obtained with the corresponding map distance. In this instance, one percent recombination equals to one map unit or centimorgan (cM). A map distance can be calculated with Morgan, Haldane, or Kosambi's map function. A popular software for analyzing map distance among markers for constructing map is MAPMAKER version 3.0 (Lincoln *et al*., 1993).

QTL analysis

Traditional quantitative genetic research defined a quantitative trait in terms of variances. The total phenotypic was first partitioned into genetic and environmental variances. The genetic variance could be further divided into additive, dominance and epistatic effects. From this information, it was possible to estimate heritability of the trait and predict response of the trait to selection. It was also possible to estimate the minimum number of genes controlling the trait. Mapping markers linked to QTLs identifies regions of the genome that may contain genes involved in the expression of the quantitative trait. The markers associated with a QTL each account for only a portion of the genetic variance. Likewise, each of these genes of known function may only account for a portion of the final phenotypic value. An important question is whether any known genes mapped are QTLs (Mohan *et al*., 1997). Various methods can be used to detect QTL such as ANOVA and regression, to analyze simple interval mapping, and composite interval mapping. Lee *et al*. (2001) conducted ANOVA method for detecting level for significant loci and locating QTL by marker regression. Koa *et al*. (1999) mentioned the single interval mapping method that used one marker interval at a time to construct a putative QTL, and the composite interval mapping method to combine between single interval mapping and multiple regression to detect the QTL.

MATERIALS AND METHODS

Genetic diversity of Thai and Korean soybean varieties

Plant materials

Thirteen Thai and 19 Korean soybean varieties (cultivars) were used in this study (Table 1). They were sown in the field of Seoul National University, Suwon during June to September 2000 for extracting DNA and observing verietal characteristics, viz. hypocotyl color, flower color, pubescence color and leaf shape.

DNA isolation and microsatellite primer

Total genomic DNA was isolated from 3 unfolded leaflets of a leaf of each variety and kept in lyophillizer at -50 °C. DNA was extracted using the modified protocols of Keim *et al*. (1988). The leaf was taken into 96-well plate containing a single steel ball. The tissue was pulverized by a reciprocal saw. Then 700ul CTAB was added into each well and taken to shaker-incubator at 225 rpm, 65 °C for 35 min, 600ul chloroform:isoamyl (24:1) was added into the plate, taken this into shaker-incubator at 24 °C for 10 min, then centrifuged in Eppendrof 5804 at 4,000 rpm for 30 minutes. Cold isopropanol were added at 420 µl into each well of new 96-well plate, take 200 µl supernatant to new plate and bring to refrigerator for 15 minutes. This processes was repeated 3 times. The plate was spun in Eppendrof 5804 at 4,000 rpm for 20 minutes, poured supernatant and then added at least 500 µl 70% ethanol to each well of plate, kept on shaker-incubator at 24 °C for 10-15 min, and centrifuged once more. The supernatant were pored and incubated at 37 °C for 15 min, then resuspend each well by 300 µl TE, shaked at 225 rpm for 1 day and kept in a refrigerator. DNA concentration was measured using F-4500 Fluorescence Spectrophotometer.

Eighteen of the fluorescent 5'-end forward primers were labeled with either 6-FAM (blue), NED (Yellow) and HEX (green) at by PE-ABI (Foster City, CA). DNA amplification were proceeded in genomic DNA at the concentration of 10ng/µl, 10x Buffer, 2.5mM of each nucleotide, 5 unit *Taq* DNA polymerase, 20mM $MgCl_2$ and 5pM Primer Mix with 32 cycles of 25 s of denaturation at 94ºC, 25 s of annealing at 46ºC, and 25 s of extension at 68ºC on PCT-100™ Thermal Controller (MJ research, Watertown, Mass). The PCR product was examined in 2 % agarose gel.

Combined 2 µl each of the PCR product of 6-FAM, NED, HEX which had different allele size in the same gel lane and then drew 4 µl of cocktail (consisted of 110 µl Formamide deionized, 55 µl loading buffer, 15 µl Genscan 500 (500XL)) and added to 1.5 µl of each combined PCR product. The PCR product was loaded and separated in ABI Prism® 377 (AB –PEC, Foster City, CA), analyzed for allele size by GeneScan® Analysis software, version 2.1.1(AB –PEC, Foster City, CA) and Genotyper® software, version 2.0 (AB –PEC, Foster City, CA)

Data analysis

Allelic polymorphic information content was calculated using the formula PIC_i (polymorphic information content) = $1-\sum p_{ij}^2$. Where p_{ij} is the frequency of the j allele for marker i (Anderson, 1992). The data were used to analyze genetic diversity from the scored bands of each SSR loci by giving 0 = absent and 1 = present for common band. The dendrograms were created based on Jaccard's similarity coefficients (Jaccard, 1908), and the similarity matrix was subjected to cluster analysis using UPGMA (unweighted pair-group method with arithmetric average) on NTSYS-pc version 2.0 (Rohlf, 1998).

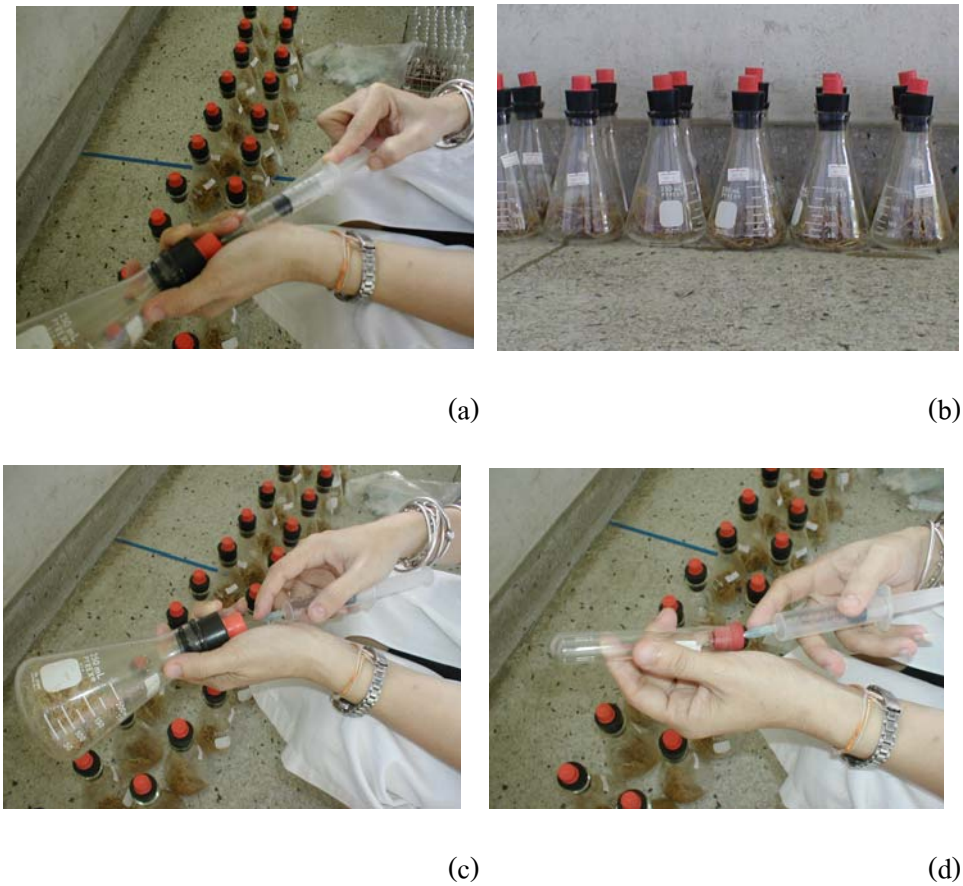
Screening of parental genotypes

Thirteen Thai and 21 Korean soybean cultivars (name list shown in Table 2) were inoculated with three *B. japonicum* strains (DASA01026, DASA01042, and DASA01054) in a factorial manner (i.e. all possible combinations). A control treatment (uninoculated) from each cultivar was also included in the experiment. The experimental design used was a completely randomized design (CRD) with 3 replicates.

The inoculation method was modified from that proposed by Somasegaran and Hoben (1985). Briefly, the soybean seeds from each genotype were surface-sterilized in 3% hydrogen peroxide in a beaker for 5 min, then the solution was drained off. The seeds were rinsed with sterile water for 4 - 5 times, poured more sterile water to submerge the seed and let stand for 4 h. The seeds were rinsed with sterile water 2 - 3 times more, transferred onto a plate laid with wet sterile cotton wool, and left overnight under room temperature until the radicles were 0.5 - 1 cm long.

Four germinating seeds from each cultivar were sown in each plastic cup filled with sterilized sand pre-inoculated with one of the three strains of *B. japonicum*. The cup was perforated at the bottom

and placed on a Leonard jar containing N-free medium. The germinating seed was covered with autoclaved gravel. Five to 10 days after planting, 2 seedlings were cut with sterilized scissors and finally 2 plants were left in each cup. The N-free medium was added weekly into the jar thereafter. At R_3 stage (pod size of around 0.5 cm in diameter), shoots of the 2 plants in each cup were collected in a paper bag and oven-dried at 70^{o} C for 2 days and weighed for plant dry weight. The roots were collected in a flask with an air-tight rubber cap and the air was subsequently replaced was with acetylene gas (C_2H_2). After one hour, ethylene gas (C_2H_4) from the flask was collected to a tube by piercing the rubber lid and drawing the gas with a syringe (Figure 6). The ethylene produced from reduction of the acetylene, as a result of rhizobial nitrogenase activity, was measured by a gas chromatograph to determine an acetylene reduction activity (ARA) value in µmole of C_2H_4 per plant per hour. The roots were then cleaned and determined for nodule number per plant, nodule fresh weight (g) per plant, and nodule dry weight (g) per plant.

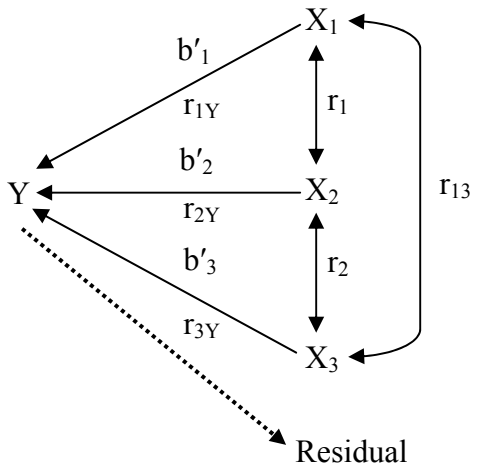


(a) (b)

(c) (d)

Figure 6 Steps in conducting ARA from roots kept in a flask with an air-tight rubber plug. (a) The air was replaced with acetylene gas, (b) let stand for one hour, (c) the ethylene from the flask was collected to a tube by piercing the rubber lid with a syringe, (d) the gas was ejected and kept in a test tube.

Analysis of variance was performed in each trait using Statistical Analysis System (SAS) program (SAS 1999-2000) to determine the significance of factors affecting N_2 fixation ability, viz. soybean cultivar, *B. japonicum* strain, and interaction between cultivar and strain. Once the F-test is significant, mean difference among cultivars and among rhizobial strains were declared by DMRT (Duncan's multiple-range test) at $P \leq 0.05$. The degree of direct association between the significant N_2-fixation traits was determined from their correlation coefficient (r). While the indirect association was measured through path coefficient analysis of the ARA with the other fixation traits using the model proposed by Steel and Torrie (1980). Briefly, let Y = a dependent variable dictated by, say 3 independent variables X_1, X_2, and X_3. The relationship between them can be depicted as followed: -



In this relationship b'_1, b'_2, and b'_3 are direct effects of X_i as measured by standardized partial regression Y on X_1, X_2, and X_3; r_{12}, r_{13}, and r_{23} are correlation coefficients between X_1 and X_2, X_1 and X_3, and X_2 and X_3; r_{1Y}, r_{2Y}, and r_{3Y} represent total effect of X_i as measured by correlation coefficient between X_1 and Y, X_2 and Y, and X_3 and Y, respectively.

Then $b'_i = b_i(S_i)/S_Y$; where b_i is the regression coefficient of Y on X_i; S_i and S_y are standard deviation of the independent variable i (i = 1-3 in this case) and the dependent variable Y, respectively.

The degree of linear relationship between these variable can be measured from the coefficient of determination, $R^2 = r_{1Y}\, b'_1 + r_{2Y}\, b'_2 + r_{3Y}\, b'_3$. Thus the residual portion (unable to explain by the model) is obtained from $1 - R^2$.

The advantage of the path coefficient is that a set of simultaneous equations can be written directly from the diagram and a solution of the equations provides information on direct and indirect contribution of these causal factors (X_i) to the effect (Y). If the correlation coefficient (r_{iY}) is positive,

but the direct effect of X_i (b'_i) is negative or negligible, the indirect effects are likely should responsible for the correlation. In such a case, the indirect causal factors involved should be considered simultaneously.

Production of the mapping population

One hundred and thirty-six recombinant inbred soybean lines (RILs) derived from the cross between the Thai cultivar ´SJ2` and the Korean one ´Suwon157` were used in this study. The two varieties were chosen as parental lines in this study because of their polymorphism in N_2 fixation components (Tanya et al. 2005). The RILs were developed at Kamphaeng Saen Campus of Kasetsart University, Thailand by single seed descent method during December 2000 – December 2003. Finally, 136 F_6 lines were obtained and treated as the RILs (Figure 7).

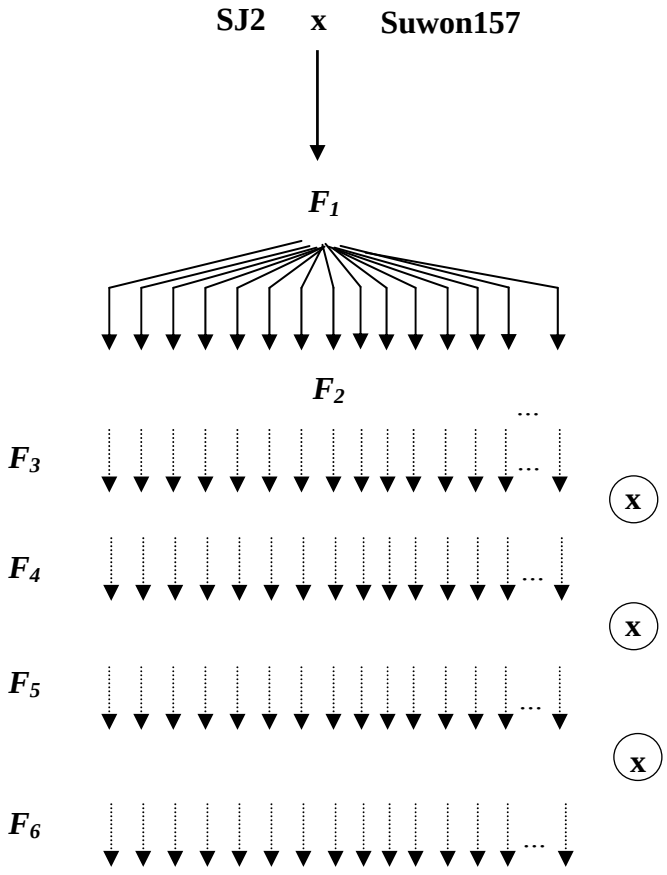


Figure 7 Diagram showing development of the RIL mapping population from the soybean cross SJ2 x Suwon157.

Phenotypic data

Preparing soybean seed and rhizobium inoculation

The 136 RILs and parental soybeans were inoculated and planted with *B. japonicum* strain DASA 01026 in a Completely Randomized Design (CRD) with 2 replications. The method modified from Somasegaran and Hoben (1985) was used to measure five phenotypic data, viz. nodule number per plant, nodule fresh weight per plant, nodule dry weight per plant, plant dry weight, and ARA. The phenotypic data were obtained from the experiments conducted at the laboratory of the Soil Microbiology Group, Soil Science Division, Department of Agriculture, Bangkok during March to November 2003.

The inoculation method was modified from that proposed by Somasegaran and Hoben (1985). Briefly, soybean seeds from each genotype were surface-sterilized in 3% hydrogen peroxide in a beaker for 5 min, then the solution was drained off. The seeds were rinsed with sterile water for 4 - 5 times, poured more sterile water to submerge the seed and let stand for 4 h. The seeds were rinsed with sterile water 2 - 3 times more, transferred onto a plate laid with wet sterile cotton wool, and left overnight under room temperature until the radicles were 0.5 - 1 cm long (Figure 8).

Four germinating seeds from each cultivar were sown in each plastic cup filled with sterilized sand pre-inoculated with *B. japonicum*. The cup was perforated at the bottom and placed on a Leonard jar containing N-free medium (Broughton and Dilworth, 1970). The germinating seed was covered with autoclaved gravel. Five to 10 days after planting, 2 seedlings were cut with sterilized scissors and finally 2 plants were left in each cup. The N-free medium was added weekly into the jar thereafter (Figure 9).

Traits related to N_2 fixation ability

Traits related to N_2 fixation components, viz. nodule number per plant, nodule fresh weight per plant (g), nodule dry weight per plant (g), plant dry weight (g), and ARA were recorded on 136 RILs following the same protocol as in screening of parental genotypes (p21).

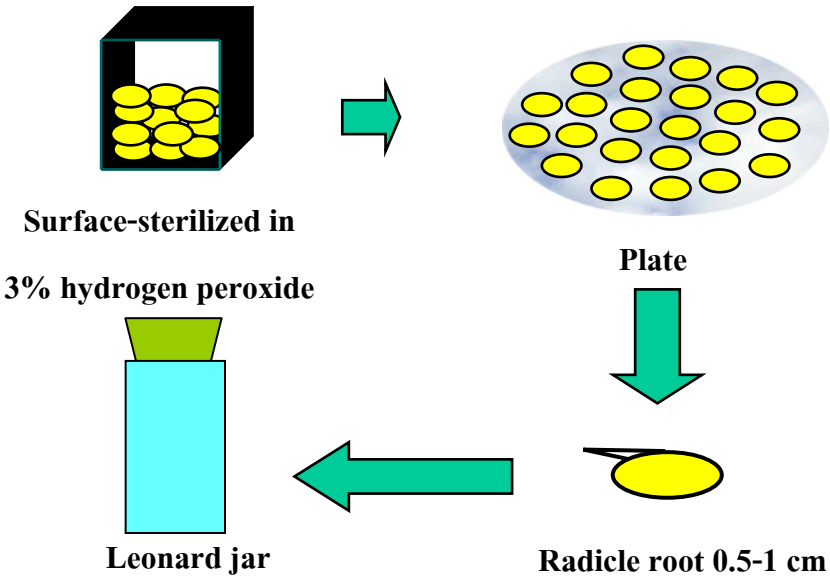


Figure 8 Surface sterilization of soybean seeds before planting in Leonard jar.

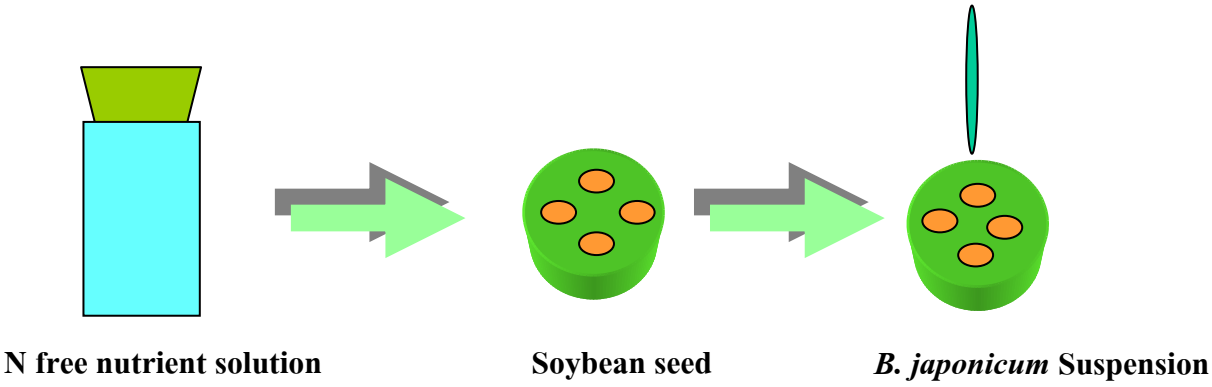


Figure 9 Planting of soybean seeds and inoculating of rhizobium into soybean seedlings.

DNA extraction

Genomic DNA was extracted from 3 folded leaflets of a leaf at R_3 stage (50 days after planting) (Figure 10-a) kept in lyophilizer (Figure 10-b). The used protocol was modified from Lambrides et al. (2000). Lyophilized leaflets were ground in a mortar and the powder was transferred into 1.5 ml tube, then added with 500 ul pre-warmed extraction buffer (1mM Tris HCl pH 8.0, 5mM EDTA, 5mM NaCl, 20% Sodium Dodecyl Sulfate (SDS), Sodium bisulfide) in each tube and incubated at 65^{o}C in a waterbath for 30 min. Each tube was added with 250 µl 5M KOAc and gently inverted for 15 min, moved into ice box for 1 h, then centrifuged at 13,000 rpm for 30 min. The supernatant was poured to a new tube for 750 µl and added with 95% ethanol to precipitate the DNA. The tubes were inverted to mix the solution well and kept in a refrigerator at –20^{o}C for 10 min, then the supernatant was discarded. Next, 250 µl of 70% cold ethanol was added to purify the DNA for 15 min. The tubes were centrifuged at 10,000 rpm for 5 min, then the supernatant was discarded. The process was repeated one more time and the resulting DNA was let dry under 37^{o}C. The DNA was resuspended in 50 µl TB buffer and measured for concentration using F-4500 Fluorescence Spectrophotometer (Figure 10-c).

Genotypic data

One hundred and ninety five SSR and 7 M13 (-21) SSR labeled primers were chosen for this study. The PCR reaction for SSR followed the suggestion by Diwan and Cregan (1997), while the reaction for SSR labeled M13 (-21) followed Schuelke (2000). Initially, the forward primers were labeled with fluorescent color tags (PE-ABI, Foster City, CA), viz. blue (6-FAM), yellow (NED), and green (HEX). The genomic DNA concentration was adjusted to 10 ng/µl before being amplified in a PCR reaction. The SSR reaction contained the genomic DNA (10 ng/µl), 10x Buffer (w/$MgCl_2$), 2.5 mM of each nucleotide, 2 unit *Taq* DNA polymerase, and 5 µM Primer Mix. Each PCR cycle consisted of 25 sec of denaturation at 94ºC, 25 sec of annealing at 46ºC, and 25 sec of extension at 68ºC in PCT-225^{TM} Thermal Controller (MJ research, Watertown, Mass) (Figure 10-d). The PCR process was repeated until 32 cycles. The reaction for SSR labeled M13 (-21) primers contained genomic DNA (10 ng/µl), 10x Buffer (w/$MgCl_2$), 2.5 mM of each nucleotide, 2 unit *Taq* DNA polymerase, 5 µM forward primer with M13 tail, 10 µM reverse primer, and fluorescent labeled with M13 (-21). Two PCR steps were run in this experiment. The first step required 30 sec of denaturation at 95ºC, 45 sec of annealing at 54ºC, and 45 sec of extension at 72ºC for 30 cycles. The second step required 30 sec of denaturation

at 95ºC, 45 sec of annealing at 53ºC, and 45 sec of extension at 72ºC for 10 cycles. The PCR product quality was checked by 3% agarose gel before mixing into a set of 6 primers. To do so, 2 µl each of 6-FAM, NED, HEX was mixed into the same well. Each color was labeled into 2 primers with different allele size and thus made the total solution of 12 µl per well. There were altogether 384 wells in each plate. The PCR mixture was taken into new well with 2 µl per well. A standard DNA marker of known size (0-500 bp) was prepared and added to the wells. Each standard composed of 2 µl mixed of a standard size cocktail consisted of 200 µl Formamide deionized, 100 µl loading buffer, 40 µl Genscan 500 (500XL) was added to each 2 µl mixed PCR product. The PCR product mixed was loaded into PCT-225™ Thermal Controller for denaturing at 95°C for 5 min and transferred to an icebox. The final mixture was loaded in ABI Prism® 377 sequencer (AB –PEC, Foster City, CA) (Figure 10-e). The GeneScan® Analysis software, version 2.1.1 (AB –PEC, Foster City, CA) and Genotyper® software, version 2.0 (AB –PEC, Foster City, CA) were used to detected the allele size. The laboratory work was conducted at the Molecular Breeding Soybean Lab and National Instrumentation Center for Environmental Management at Seoul National University, the Republic of Korea during March – September 2004.

Scoring genotypic data

The genotypic data were used to construct linkage map by scoring the bands of each SSR loci by A = same banding pattern as SJ2, B = same banding as Suwon157, and - = missing data.

Data analysis

Checking segregation ratio

The chi-square (χ^2) method was calculated between the observed and the expected band number for checking a 1: 1 segregation ratio of each marker in the RILs population, using the following formula.

$$\chi^2 = \frac{(O_1-E_1)^2}{E_1} + \frac{(O_2-E_2)^2}{E_2}$$

O_1 = observed number of SJ2 bands

E_1 = expected number of SJ2 bands

O_2 = observed number of Suwon157 bands

E_2 = expected number of Suwon157 bands

The goodness-of-fit can be checked with a table of χ^2 with the degrees of freedom, n. Where n is the number of band classes (= 2 in this case; i.e. SJ2 and Suwon157). If the segregation pattern is significantly different from 1 : 1, then the population may have different genetic configuration regarding the bands developed from these primers. If no difference was detected, then the population can be regarded as homogenous in genotypic segregation and can be pooled into one analysis.

Statistical analysis

The difference between RILs in each component trait was determined by an analysis of variance (ANOVA) using SAS program (SAS 1999-2000). The expected mean square (EMS) components (Table 1) were used to estimate the heritability of each trait from the formula $h^2 = \sigma_L^2/(\sigma_L^2 + \sigma^2/r)$. Where σ_L^2 and σ^2 are the variance components associated with RILs and error; r is number of replications (= 2 in this case).

Construction of linkage map and QTL analysis

The MAPMAKER program version 3.0 (Lincoln et al. 1993) was used to construct linkage maps from the genotypic data. Markers were assigned to linkage groups using LOD threshold of 3.0 and maximum distance of 50 cM. The marker order followed a new integrated genetic linkage map of soybean (Song *et al*., 2004). Map distance was estimated using Kosambi function. In each trait, a single factor analysis of variance (SF-ANOVA) was used to evaluate the association between the genotypic and phenotypic data. The significant markers from SF-ANOVA were assigned into each linkage group. Then, a multiple regression analysis was conducted by including all the significant markers on that linkage group in the model (SLG-Regr). They were assumed to detect unique QTLs on that linkage group. Then all the significant markers from the SLG-Regression analysis and unlinked single markers identified from SF-ANOVA were combined in a multiple linkage group regression model (MLG-Regr) to identify the markers linking to the trait. The regression analysis in this study was forward and stepwise selection of the significant markers at $P < 0.05$. The percent of phenotypic variation explained by the markers was estimated from the coefficient of determination (R^2) from MLG-Regr using SAS program (SAS 1999-2000). MQTL software was used to confirm the QTL by simple interval mapping

(SIM) and simplified composite interval mapping (sCIM). A single marker analysis using regression based software, STATGRAPHIC version 3.0.

(a)

(b)

(c)

(d)

e)

Figure 10 Steps in extracting DNA and detecting genotypic data of RIL population: (a) = three leaflets of optimum size for DNA extraction, (b) = lyophilizer machine for drying tissue before extracting DNA, (c) = F-4500 Fluorescence Spectrophotometer for checking DNA concentration, (d) = PCT-225™ Thermal Controller (MJ research, Watertown, Mass) for running PCR product, (e) = ABI Prism® 377 (AB –PEC, Foster City, CA) for separating marker alleles each loci.

Places and duration of studying

The RIL population was prepared at the Tropical Vegetable Research Center (TVRC) of Kasetsart University, Kamphaeng Saen Campus The phenotyping of N_2 fixation components were done at the laboratory of the Soil Microbiology Group, Soil Science Division, Department of Agriculture, Bangkok. The DNA genotyping by SSR markers was executed in Molecular Breeding Soybean Lab at Seoul National University in The Republic of Korea. The research was executed between the years 2000 to 2004.

RESULTS AND DISCUSSIONS

Morphological variation among Thai and Korean soybean varieties

Simple quantitative characters of 13 Thai and 19 Korean soybean varieties were shown in Table 1. Two Thai (CM 60, and KKU 35), and ten Korean (Dajang, Sunheuk, Jangkyung, Doremi, Keumkang, Danwon, Duyu, Milyang, and Songhak) varieties are green hypocotyl, the rests are purple ones. Two Thai (ST 1, and CM 1) and four Korean (Sowon, Dukyu, Jangkyung, and Songhak) varieties are white pubescence, other show brown pubescence on main stem. RM 1, KKU 50, Kumjung 2, and Milyang have rather dense pubescence. Kumjungol was the best in germination, while IT 182307 and IT 184222 were the least germination. The latter 2 lines have rhomboid-lanceolate leaf shape, while the others are either lanceolate or oval.

Table 1 Simple morphological traits of 13 Thai and 19 Korean soybean varieties

No.	Varieties name	HC	FC	PC	LS
1	SJ 1	Purple	Purple	Brown	Oval
2	SJ 2	Purple	Purple	Brown	Lanceolate
3	SJ 4	Purple	Purple	Brown	Oval
4	SJ 5	Purple	Purple	Brown	Oval
5	ST 1	Purple	Purple	White	Lanceolate
6	ST 2	Purple	Purple	Brown	Lanceolate
7	CM 1	Purple	Purple	White	Oval
8	CM 60	Green	White	Brown	Oval
9	NS 1	Purple	Purple	Brown	Oval
10	KUSL 20004	Purple	Purple	Brown	Oval
11	RM 1	Purple	Purple	Brown	Oval
12	KKU 35	Green	White	Brown	Oval
13	CKP	Purple	Purple	Brown	Oval
14	Kumjung -	Purple	Purple	Brown	Oval
15	Dajang	Green	White	Brown	Oval
16	Sowon	Purple	Purple	White	Lanceolate
17	Dukyu	Purple	Purple	White	Oval

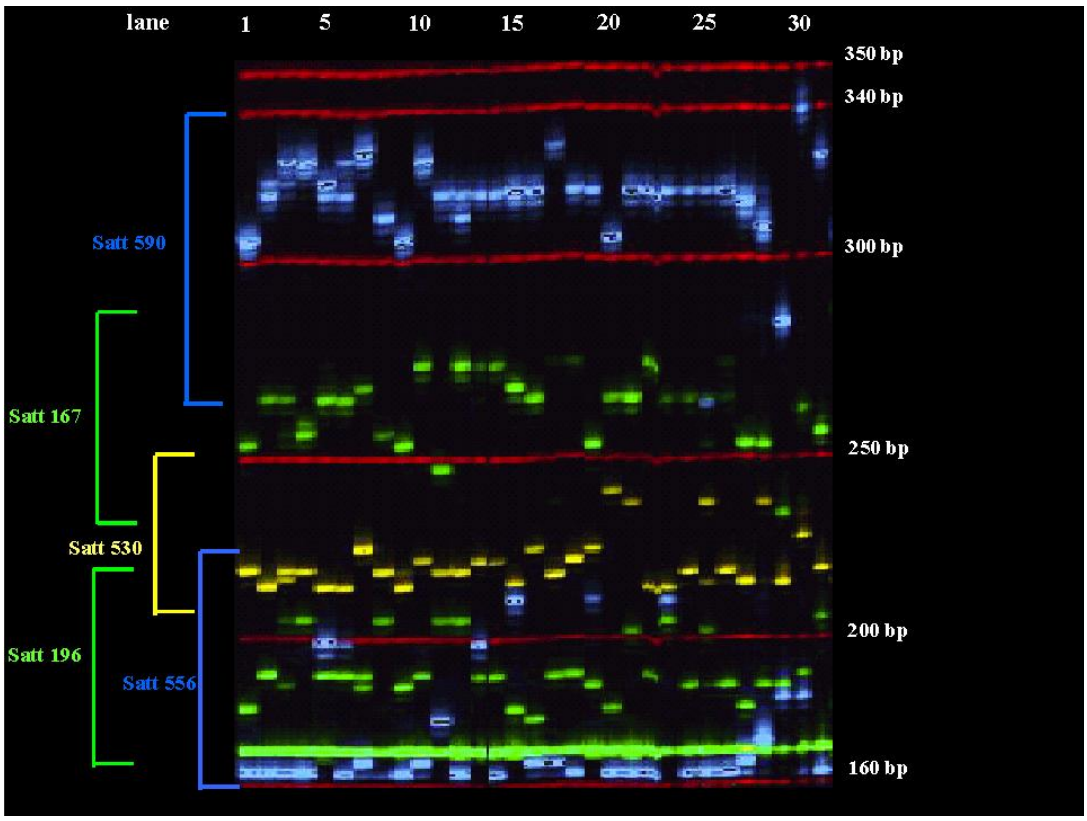
Table 1 (Cont'd)

No.	Varieties name	HC	FC	PC	LS
18	Sunheuk	Green	White	Brown	Oval
19	Jangkyung	Green	White	White	Oval
20	Doremi	Green	White	Brown	Lanceolate
21	Kumjungol	Purple	Purple	Brown	Oval
22	Keumkang	Green	White	Brown	Oval
23	Ilmi	Purple	Purple	Brown	Oval
24	IlpumKumjung	Purple	Purple	Brown	Oval
25	Danwon	Green	White	Brown	Oval
26	Duyu	Green	White	Brown	Oval
27	Milyang	Green	White	Brown	Oval
28	Suwon157	Purple	Purple	Brown	Lanceolate
29	Songhak	Green	White	White	Oval
30	IT 161471	Purple	Purple	Brown	Rhomboid – lanceolate
31	IT 182307	Purple	Purple	Brown	Rhomboid –lanceolate
32	IT 184222	Purple	Purple	Brown	Oval

HC = hypocotyl color; FC = flower color; PC = pubescence color, LS = leaf shape

Automated sizing of microsatellite DNA polymorphism in the soybean varieties

Nineteen microsatellite fluorescent labeled primers were used in DNA amplification summarized in Table 2. Its gel image from ABI Prism® 377 is shown in Figure 11. The total of 138 alleles were detected from 18 primers. The number of alleles per locus ranged from 3 to 12, with an average of 7.67. Three alleles was amplified by Satt271, 4 alleles by Satt187, 5 alleles by Satt038, 6 alleles by each of 4 loci (Satt143, Satt177, Satt196, and Satt253), 8 alleles by each of 5 loci (Satt141, Satt192, Satt294, Satt414, and Satt556), 9 alleles by each of 2 loci (Satt175 and Satt530), 10 alleles by Satt167, 11 alleles by each of 2 locus (Satt445 and Satt590) and 12 alleles by Satt545. Genetic diversity as measured by polymorphic information content (PIC) ranged from 0.56 (Satt271) to 0.83 (Satt545) with mean value of 0.70.



Figture 11 Gel image of 13 Thai and 18 Korean soybean varieties with fluorescent 5'-end forward primers taken from ABI Prism® 377 DNA sequencing

Table 2 Allele number, size range, and gene diversity of 20 fluorescent labeled SSR loci in 13 Thai and 19 Korean soybean varieties.

Locus	Fluorescent Label	Linkage Group	Cultivars genotypes		
			Allele size Range (bp)	Number of allels	Genetic diversity (PIC)
Satt141	6-FAM	D1b+W	148-201	8	0.70
Satt175	6-FAM	M	143-183	9	0.71
Satt192	6-FAM	H	234-264	8	0.67
Satt414	6-FAM	J	266-313	8	0.76
Satt545	6-FAM	A1	155-203	12	0.83
Satt556	6-FAM	B2	163-211	8	0.63
Satt590	6-FAM	M	263-340	11	0.70
Satt167	HEX	K	235-273	10	0.76
Satt187	HEX	A2	243-280	4	0.57
Satt196	HEX	K	178-205	6	0.73
Satt253	HEX	H	130-175	6	0.64
Satt038	NED	G	157-184	5	0.60
Satt143	NED	L	235-276	6	0.71
Satt177	NED	A2	105-131	6	0.74
Satt271	NED	D1b+W	112-121	3	0.56
Satt294	NED	C1	252-296	8	0.70
Satt445	NED	O	162-228	11	0.78
Satt530	NED	N	215-241	9	0.82
Mean				7.67	0.70

Clustering of Thai and Korean soybean varieties based on SSR polymorphism

The genetic similarity (GS) coefficients among the soybean varieties ranged from 0.025 to 0.944. The dendrogram prepared through cluster analysis is shown in Fig 12. The genotypes can be grouped into five clusters, clusterⅠwith 7 genotypes (SJ1, SJ4, SJ5, CKP, CM60, KKU50, RM1), cluster Ⅱ with 16 genotypes (SJ2, ST2, ST1, Dajang, Jangkyung, Ilmi, Danwon, Milyang, Sowon, Doremi, Dukyu, Keumkang, Duyu , Songhak, NS1, and IlpumKumjang), cluster Ⅲ with 6 genotypes (CM1,

KUSL20004, Kumjung2, Sunheuk, Kumjungol, and Suwon157), cluster Ⅳ with 2 genotypes (IT 161471 and IT 182307), and cluster V containing a solitary genotype, IT 184222.

The automated sizing of allele using fluorescent-labeled taken from ABI Prism® 377 DNA sequencing was found to be rapid and reliable. Each gel can accommodate 5 or 6 primers depending on allelic range. The Genotyper® software was also practical in sizing and visualization. The cluster analysis of this set of primers revealed that SJ 5 and CKP are rather closely related (Fig. 12). The IT series (*Glycine soja*) in cluster IV and V were rather different from the other genotypes (*G. max*) and thus become distinct. Some Thai soybeans were genetically similar to Korean soybeans, showing that the Korean varieties may not be suitable as a source to increase genetic variation in Thai soybeans.

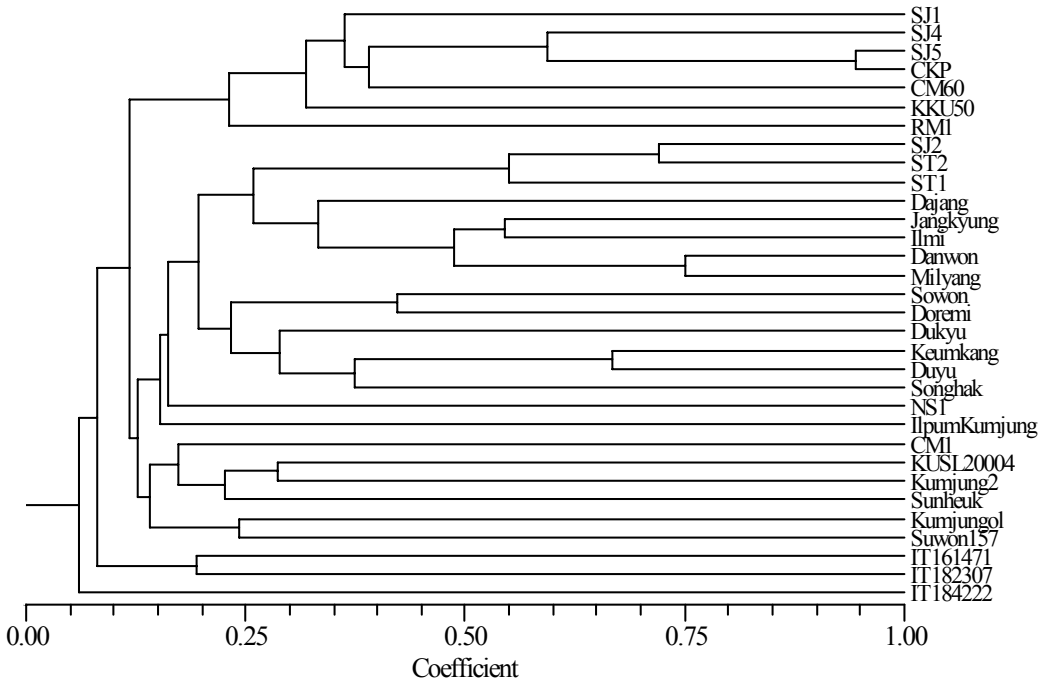


Figure 12 A dendrogram of 13 Thai and 19 Korean soybean varieties created by using UPGMA based on Jaccard's similarity coefficient computed from data matrix with 149 informative polymorphic DNA bands generated by 20 microsatellite primers.

Table 3 Mean squares from the analysis of variance of N_2 fixation components in 13 Thai and 21 Korean soybean varieties inoculated with 3 rhizobium strains. All observations were determined by Leonard jar method. Data from the control jars were not included in the analysis.

Source of Variation	df	Nodules number /plant	Nodule fresh wt./plant (g)	Nodule dry wt./plant (g)	Plant dry weight (g)	ARA (μmole C_2H_4/pl/hr)
Soybean varieties	33	450**	0.235**	0.013**	2.12**	27.2**
Thai vs Korean	*1*	136^{ns}	0.881**	0.076**	11.82**	97.9**
Rhizobium strains	2	806**	1.248**	0.061**	1.34**	165.1**
Varieties*Strains	66	116**	0.034**	0.003**	0.21^{ns}	7.4**
Thai variety x Strain	*24*	126**	0.056**	0.004**	0.32*	10.0**
Korean variety x Strain	*40*	81*	0.020^{ns}	0.002^{ns}	0.04^{ns}	5.3^{ns}
Thai vs Korean x Strain	*2*	699**	0.070*	0.004^{ns}	2.19**	15.7*
Error	190	50	0.022	0.002	0.19	4.0
Total	291					

*,** Significantly different at $P < 0.05$ and $P < 0.01$, respectively.

^{ns} non-significant ($P > 0.05$).

N_2 fixation components in Thai and Korean soybean varieties

Since the fixation data from the control jars were generally low with a number of zero values, they were not included in the analyses. Analysis of variance revealed significant difference among soybeans, rhizobia, and their interaction in all N_2 fixation traits in this study (Table 3). In general, Thai soybean varieties were not different from the Korean varieties in nodule number per plant (27.5 vs 26.6) but the mean of the formers was higher than that of the latter in nodule fresh weight (0.581 vs 0.482 g/plant), nodule dry weight (0.126 vs 0.096 g/plant), plant dry weight (1.913 vs 1.512 g/plant), and ARA (4.898 vs 3.845 μmole C_2H_4/pl/hr). The interaction between Thai soybeans and rhizobium strains was significant in all fixation traits, revealing that there was specificity between the soybean and the rhizobium. However, in Korean soybeans x strains, this specificity was significant only in nodule number. The highest nodule number in Thai varieties x strain was found in KKU35 inoculated with DASA01054 strain, and the lowest value was in CM60 inoculated with DASA01026. CM1 with

DASA01026 strain gave the highest nodule fresh and dry weight, while CM60 with DASA01042 gave the lowest value. The highest plant dry weight was found in CM1 with DASA01026, while the lowest value was in SJ2 with DASA01026. The highest ARA was obtained from ST1 with DASA01026 while the lowest was from SJ2 with DASA01054. The highest nodule number per plant in Korean varieties x strains was Jangkyung with DASA01026 and the lowest value was IT184222 with DASA01042.

The average values of the fixation components in each variety across 3 rhizobial strains were presented in Table 4. A rather high coefficient of variability (CV) was found associating with each trait. KKU35 gave the highest nodule number, which was not significantly different from Songhak, followed by CM1, Keumkang, Sunheuk, Jangkyung, and SJ4. While IT184222, Jangyup, Dajang, and IT161471 from Korea and CM60 from Thailand had the least nodule number. SJ4, ST1, RM1, Sunheuk, CM1, and SJ5 gave the highest nodule fresh weight, while IT184222, IT161471, CM60, SJ2, and Milyang were among the lowest. In nodule dry weight, SJ4, ST1, Danwon, Sunheuk, SJ5, CM1, and RM1 were the highest, while several varieties, including IT184222, CM60, SJ2, IT161471, and Milyang were the lowest. In plant dry weight, the Thai varieties, RM1, CM1, SJ4, ST1, KKU35, and SJ5 grew more vigorously than most Korean varieties. IT184222, IT161471, CM60, SJ2, and Milyang were low in plant dry weight. However, in ARA, the Thai and Korean varieties fell into both high and low groups. ST1, RM1, Sunheuk, and Songhak were in the high ARA group, while IT184222, SJ2, IT161471, Jangyup, and Sowon were in the low group. It should be noted that IT184222 from Korea gave the lowest value in all fixation traits observed.

When the N_2 fixation components were averaged across all varieties in each rhizobium strain, DASA01026 gave high nodule number, nodule fresh weight, and nodule dry weight. DASA01054 gave high nodule number, plant dry weight, and ARA, while DASA01042 gave only high plant dry weight (Table 5). This result is similar to that reported earlier (Somwang *et al*., 2002) who demonstrated that DASA01026 and DASA01054 exhibited ability to form more nodule number, nodule dry weight, and high nitogenase activity.

The N_2 fixation traits in each variety were not well correlated. For example, KKU35 was high in nodule number but low in the ARA assay. Aprison *et al*. (1954) demonstrated that soybean nodules with the diameter of less than 4 mm showed lower N_2 fixation ability than those with 5-6 mm in diameter. In mungbean, however, Tomooka *et al*. (1992) found that the genotypes with low ability to fix N_2 had larger nodules. Nodule size may not be correlated with fixation activity but rather with the quality of leghaemoglobin in the nodules. The active nodules should have high leghaemoglobin which gives red or

pink color inside the nodules while the less active ones are green, white, or brown (Damery and Alexander, 1969).Leghaemoglobin acts as does haemoglobin in blood as an oxygen-carrying pigment for nitrogen fixation.

The result showed shown significantly association between ARA and nodule number, nodule fresh weight, nodule dry weight, and plant dry weight, with the correlation coefficients (r) of 0.438, 0.738, 0.703, and 0.635, respectively (Table 6). This implies that an improvement of one fixation component will result in improvement of the others, including the ARA value. A more detail relationship between ARA and the other components can be demonstrated in a path coefficient relationship as shown in Figure 13. It was obvious that ARA value was the result of the other fixation traits, plus undefined factors designated by the residual. In this study, nodule fresh weight expressed a high positive direct effect ($b'_2 = 1.020$) and total effect on ARA ($r_2y = 0.738$). While the other fixation traits had small direct effect (b'_1) on ARA, they also showed small indirect effect through each others. For example, the indirect effect of nodule fresh weight through plant dry weight was the highest among all the effects, with the value of –0.208 ($b'_4 = -0.228$ and $r_{24} = 0.914$). Nodule fresh weight contributed both directly and indirectly to the ARA, while nodule number, nodule dry weight, and plant dry weight had a little direct effect. Similar to Pazdernik *et al*. (1996) who found a positively correlation between nodule fresh weight and ARA (r = 0.86**), and between nodule number and ARA (0.45*). DÖbereiner (1966) showed that nodule number is not always correlated with total N accumulated in shoots, while plant dry weight was a more reliable parameter. The total effect of all factors were moderate upon ARA because of the high indirect effect via nodule fresh weight. Thus it can be concluded that N_2 fixation activity in this soybean population can be assessed through nodule fresh weight alone. A breeding program for improving N_2 fixation can gear toward identifying the criteria for selection to improve nodule fresh weight. The criteria may be changed if the total N_2 fixed was used as the goal of N_2 fixation rather than the ARA (Giller, 2001). The ARA itself can be subdivided into component characters, for example, nodule number, nodule fresh weight, nodule dry weight, and plant dry weight. However, to measure these components will only give an approximate estimate of the reasons for the differences. Ultimately, the genes that affect these characters and ARA have specific roles to play in development of the plant. In this study, nodule fresh weight was an important component both direct and indirect effect upon ARA value, while the other components had a high total effect but low direct and indirect effect on ARA (Figure 13). Factors influencing ability of N_2 fixation and nodulation are rhizobium strains, temperature, pH, light intensity, and soil moisture. In this study, phenotyping of fixation components in the greenhouse faced with the fluctuation in temperature and sunlight. The best

condition for collecting the data were on the day that the soybean plants had received at least 2 hr of sunlight in the morning for nitrogenase activity (Attewell and Bliss, 1985). The best stage for determining nitrogen fixation activity increases sharply after flowering stage and gradually decreasing after green pod stage (Hardy, 1968). This suggestion is the same as Latimore *et al*. (1977) who repeated that N_2 fixation was declining from mid-pod to late pod fill stages.

Table 4 Average nodule number, nodule fresh weight, nodule dry weight, plant dry weight, and ARA per plant in 13 Thai and 21 Korean soybean varieties averaged across 3 rhizobium strains.

Number*	Varieties Names	Nodule number/plant	Nodule fresh wt/plant (g)	Nodule dry wt/plant (g)	Plant dry Weight (g)	ARA ($\mu moleC_2H_4$/pl/hr)
1	SJ 1	24.56^{e-j}	0.544^{d-h}	0.129^{c-f}	1.79^{b-g}	4.75^{c-g}
2	SJ 2	19.22^{h-j}	0.213^{l-n}	0.046^{j-k}	1.00^{j-l}	1.90^{i-j}
3	SJ 4	33.22^{b-d}	0.834^{a}	0.189^{a}	2.28^{a-b}	5.71^{b-c}
4	SJ 5	27.00^{d-h}	0.681^{a-e}	0.152^{a-d}	2.11^{a-d}	4.62^{c-g}
5	ST 1	24.56^{e-j}	0.780^{a-b}	0.181^{a-b}	2.28^{a-b}	9.06^{a}
6	ST 2	23.00^{f-j}	0.576^{c-g}	0.121^{c-g}	1.65^{d-i}	5.25^{b-e}
7	CM 1	36.33^{b-c}	0.693^{a-d}	0.146^{a-e}	2.55^{a}	3.16^{e-j}
8	CM 60	16.57^{j-k}	0.204^{m-n}	0.035^{k}	0.93^{l-k}	3.10^{e-j}
9	NS 1	29.44^{c-g}	0.582^{c-g}	0.137^{b-f}	2.07^{b-d}	4.97^{b-g}
10	KUSL 20004	25.67^{d-i}	0.550^{d-h}	0.107^{d-h}	1.63^{d-i}	5.05^{b-f}
11	RM 1	27.33^{d-h}	0.767^{a-b}	0.146^{a-e}	2.55^{a}	7.73^{a}
12	KKU 35	43.78^{a}	0.630^{b-f}	0.142^{b-e}	2.16^{a-c}	4.06^{c-i}
13	CKP 1	26.89^{d-h}	0.495^{f-j}	0.110^{d-h}	1.87^{b-g}	4.31^{c-h}
14	Kumjung 2	32.00^{b-e}	0.515^{e-i}	0.090^{f-j}	1.76^{c-g}	4.71^{c-g}
15	Dajang	18.33^{i-j}	0.538^{d-h}	0.102^{e-i}	1.82^{b-g}	3.47^{c-i}
16	Sowon	24.67^{e-j}	0.346^{j-m}	0.070^{h-k}	1.23^{h-l}	2.25^{h-j}
17	Dukyu	26.22^{d-i}	0.640^{b-f}	0.114^{d-h}	1.70^{c-h}	3.50^{c-i}
18	Sunheuk	34.00^{b-d}	0.737^{a-c}	0.160^{a-c}	1.92^{b-f}	7.93^{a}
19	Jangkyung	34.00^{b-d}	0.454^{g-k}	0.090^{f-j}	1.38^{g-k}	2.79^{f-j}
20	Doremi	20.78^{h-j}	0.363^{i-m}	0.074^{g-k}	1.21^{i-l}	3.05^{e-j}
21	Kumjungol	27.29^{d-h}	0.595^{c-g}	0.114^{d-h}	1.85^{b-g}	4.76^{c-g}
22	Keumkang	36.00^{b-c}	0.553^{d-h}	0.107^{d-h}	1.77^{c-g}	5.54^{b-d}
23	Ilmi	29.22^{c-g}	0.428^{g-k}	0.091^{f-j}	1.44^{f-j}	5.61^{b-d}
24	IlpumKumjang	22.00^{g-j}	0.368^{i-l}	0.070^{h-k}	1.44^{f-j}	3.65^{c-i}

Table 4 (Cont'd)

Number*	Varieties Names	Nodule number/plant	Nodule fresh wt/plant (g)	Nodule dry wt/plant (g)	Plant dry Weight (g)	ARA ($\mu mole C_2H_4$/pl/hr)
25	Danwon	31.89^{b-e}	0.635^{b-f}	0.164^{a-c}	2.04^{b-e}	3.85^{c-i}
26	Duyu	30.43^{c-f}	0.497^{f-j}	0.102^{e-i}	1.54^{e-i}	4.03^{c-i}
27	Milyang	31.88^{b-e}	0.343^{j-m}	0.060^{i-k}	1.02^{j-l}	2.61^{g-j}
28	Songhak	38.8^{a-b}	0.595^{c-g}	0.112^{d-h}	1.54^{e-i}	7.11^{a-b}
29	IT 161471	18.56^{i-j}	0.311^{k-n}	0.057^{i-k}	0.78^{l}	1.80^{i-j}
30	IT 184222	10.00^{k}	0.158^{n}	0.032^{k}	0.36^{m}	1.00^{j}
31	Danbaek	27.00^{d-h}	0.570^{c-h}	0.120^{c-g}	1.83^{b-g}	3.32^{d-i}
32	Taekwang	18.56^{i-j}	0.450^{g-k}	0.093^{f-i}	1.80^{b-g}	3.85^{c-i}
33	Jangyup	16.56^{j-k}	0.404^{h-k}	0.077^{g-k}	1.50^{f-i}	2.14^{h-j}
34	Suwon 157	30.78^{b-f}	0.631^{b-f}	0.127^{c-f}	1.83^{b-g}	3.78^{c-i}
	CV (%)	26.26	28.46	37.08	26.22	47.67

[a-n] Means followed by the same letters in each trait are not significantly different at $P < 0.05$ by DMRT.

* Varieties number 1 – 13 are Thai varieties, number 14-34 are Korean varieties.

Table 5 Average number of nodules , fresh nodule weight, dry nodule weight, dry plant weight, and ARA in 3 rhizobium strains averaged across 13 Thai and 21 Korean soybean varieties.

B. japonicum Strains	Nodule number/plant	Nodule fresh wt./plant (g)	Nodule dry wt./plant (g)	Plant dry weight (g)	ARA ($\mu mole\ C_2H_4$/pl/hr)
DASA01026 (T_1)	28.9^{a}	0.623^{a}	0.129^{a}	1.54^{b}	4.00^{b}
DASA01042 (T_2)	23.4^{b}	0.397^{c}	0.080^{c}	1.76^{a}	3.02^{c}
DASA01054 (T_3)	28.0^{a}	0.532^{b}	0.113^{b}	1.71^{a}	5.58^{a}

[a-c] Means followed by the same letter in each trait are not significantly different at $P < 0.05$ by DMRT.

Table 6 Correlation between N_2 fixation components across 13 Thai and 21 Korean soybean varieties and 3 rhizobium strains.

N_2 fixation components	Nodule fresh wt./plant (g)	Nodule dry wt./plant (g)	Plant dry weight (g)	ARA (μmole C_2H_4/pl/hr)
Nodule number/plant	0.606**	0.579**	0.557**	0.438*
Nodule fresh wt./plant (g)		0.964**	0.914**	0.738**
Nodule dry wt./plant (g)			0.899**	0.703**
Plant dry weight (g)				0.635**

*,** Significant at $P < 0.05$ and $P < 0.01$, respectively.

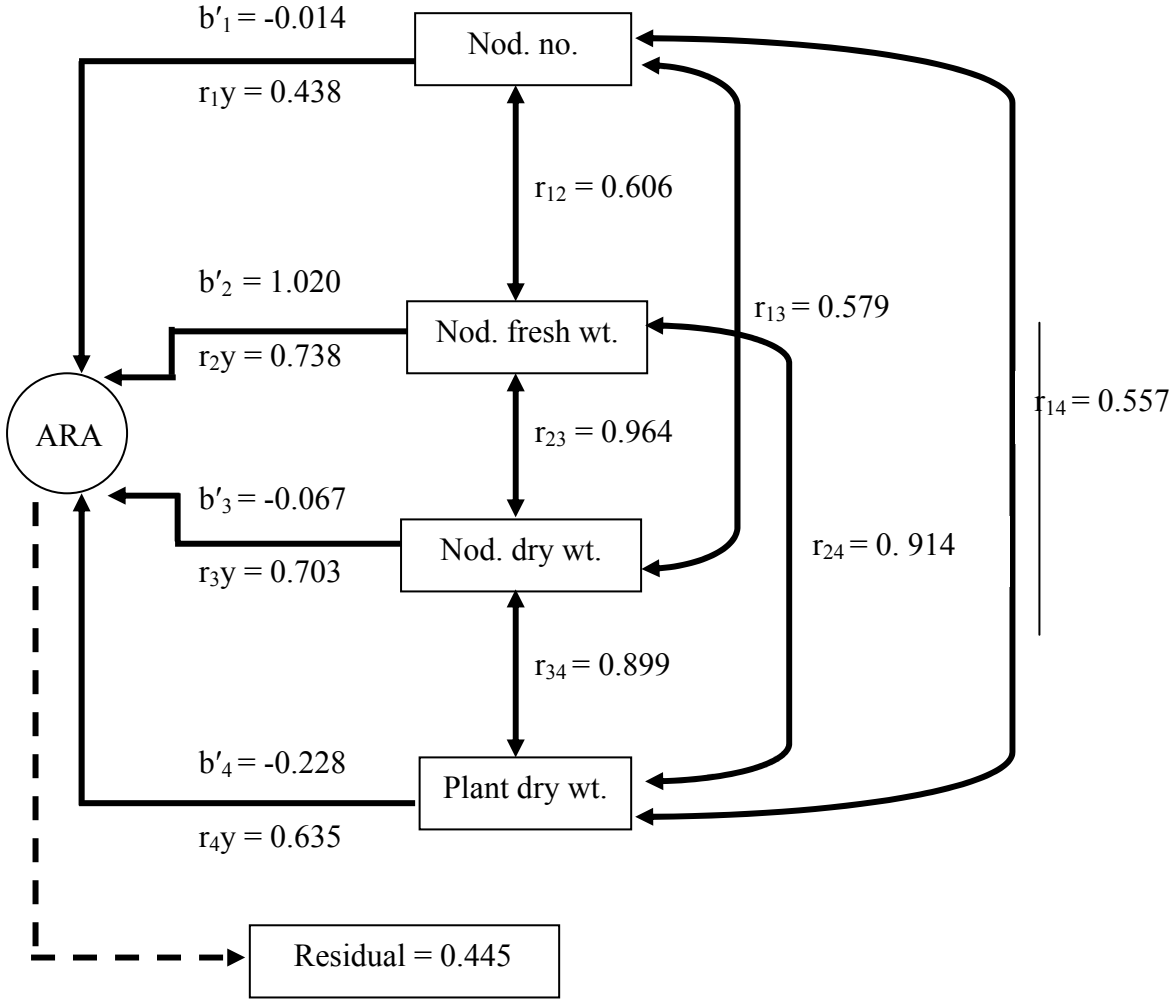


Figure 13 Path coefficient relationship between ARA and nodule number per plant, nodule fresh weight per plant, nodule dry weight per plant, and plant dry weight across 13 Thai and 21 Korean soybean varieties inoculated with 3 rhizobium strains.

Phenotypic variation of N_2 fixation components in the RIL population

N_2 fixation components were conducted on one hundred and thirty-six recombinant inbred lines (RILs) derived from the cross between SJ2 and Suwon157. They showed a quantitative distribution which nodule number per plant were 38 in SJ2 and 78 in Suwon157 (Figure 14). Nodule fresh weight per plant was 0.642 g in SJ2 and 1.368 g in Suwon157 (Figure 15). Nodule dry weight per plant was 0.157 g in SJ2 and 0.289 g in Suwon157 (Figure 16). Plant dry weight was 1.48 g in SJ2 and 2.36 g in Suwon157 (Figure 17). Acetylene reduction activity (ARA) was 8.1 μmole C_2H_4/p/hr in SJ2 and 26.3 μmole C_2H_4/p/hr in Suwon157 shown discrete classes allowing for Medalian analysis (Figure 18). The N_2 fixation components were significantly different among 136 RILs in Appendix Table 1 and Table 7. The N_2 fixation components among 136 RILs were significantly different (Table 8). The heritability estimates were moderate in acetylene reduction assay (ARA) to high in number of nodules per plant. This implied that these traits can be effectively improved through breeding and selection.

Table 7 Mean squares from the analysis of variance of N_2 fixation components in 136 RILs of SJ2 x Suwon157 inoculated with DASA01026 rhizobium strain.

Source of Variation	df	Nodule number/plant	Nodule fresh wt./plant (g)	Nodule dry wt./plant (g)	Plant dry weight (g)	ARA (μmole C_2H_4/pl/hr)	EMS
Between RILs	135	587**	0.339**	0.017**	0.74**	45.5**	$\sigma^2 + 2\sigma^2_R$
Error	136	73	0.056	0.005	0.15	15.0	σ^2
Total	271						
	h^2	0.78	0.72	0.55	0.66	0.50	

Table 8 Correlation coefficients of N_2 fixation components in 136 recombinant inbred lines derived from the soybean cross SJ2 and Suwon157.

N_2 fixation components	Nodule fresh wt./plant (g)	Nodule dry wt./plant (g)	Plant dry weight (g)	ARA (μmole C_2H_4/pl/hr)
Nodule number/plant	0.768**	0.638 **	0.697**	0.569**
Nodule fresh wt./plant (g)		0.867 **	0.906**	0.712**
Nodule dry wt./plant (g)			0.851**	0.655**
Plant dry weight (g)				0.686**

*,** Correlation coefficient was significant at $P < 0.05$ and $P < 0.01$, respectively.

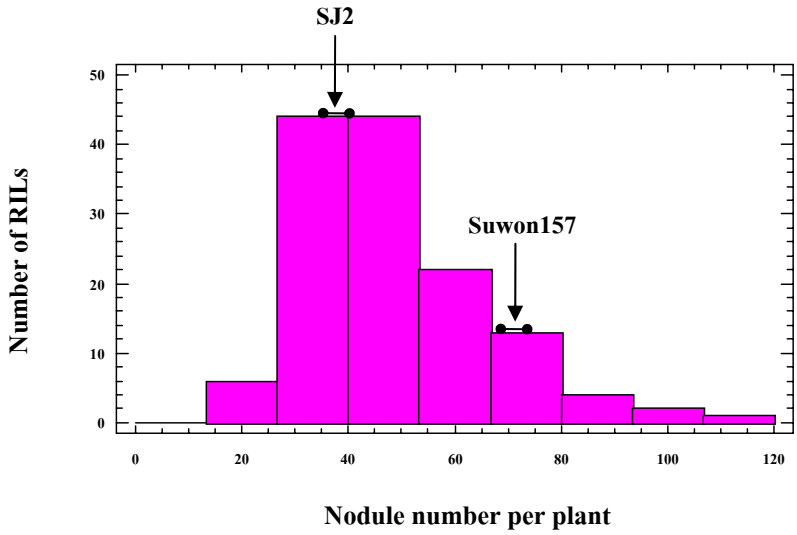


Figure 14 Frequency distribution of nodule number per plant of 136 RILs derived from crossing between Suwon157 (73 ± 1) and SJ2 (38 ± 3).

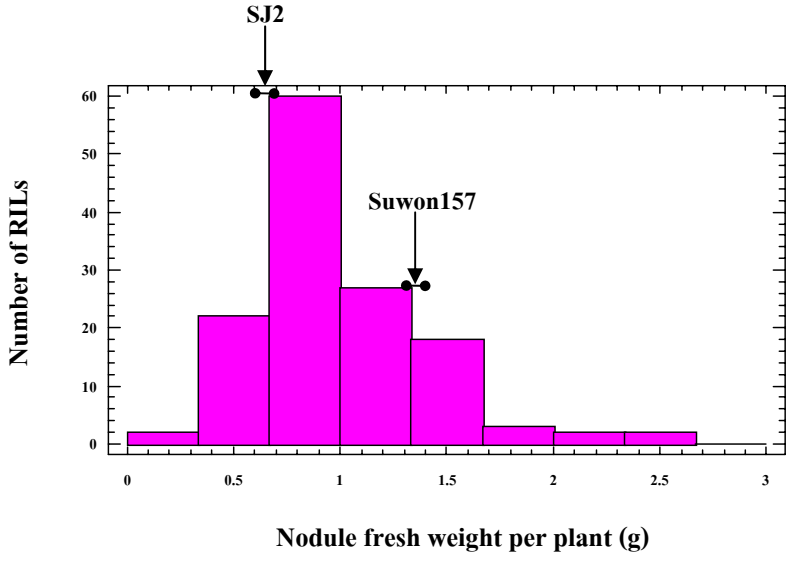


Figure 15 Frequency distribution of nodule fresh weight per plant (g) of 136 RILs derived from crossing between Suwon157 (1.368 ± 0.018) and SJ2 (0.642 ± 0.037).

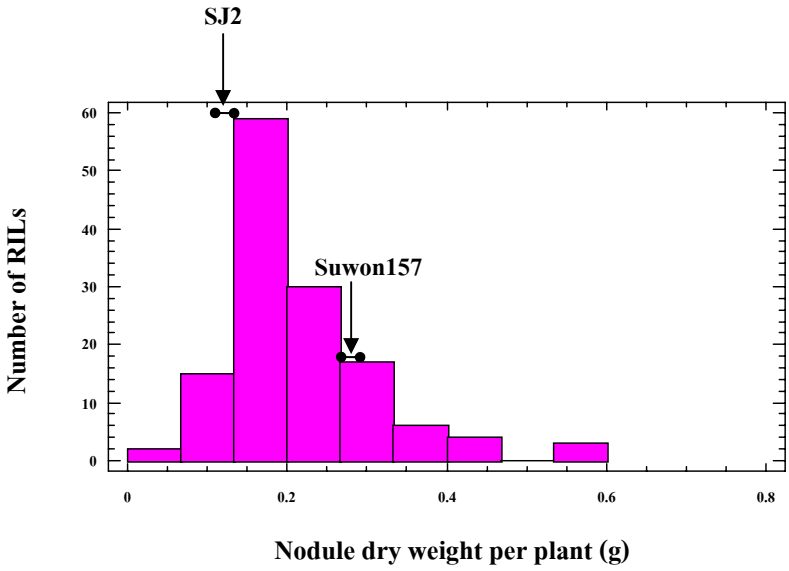


Figure 16 Frequency distribution of nodule dry weight per plant (g) of 136 RILs derived from crossing between Suwon157 (0.289 ± 0.004) and SJ2 (0.157 ± 0.008).

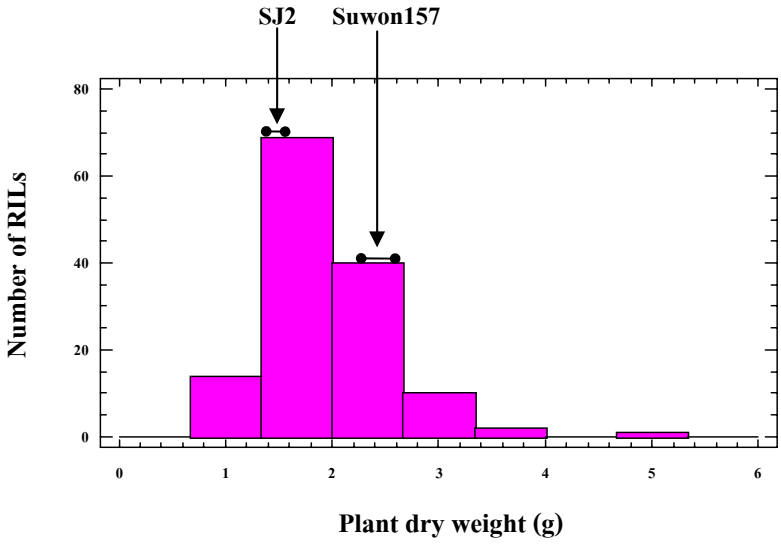


Figure 17 Frequency distribution of plant dry weight (g) of 136 RILs derived from crossing between Suwon157 (2.36 ± 0.25) and SJ2 (0.48 ± 0.08).

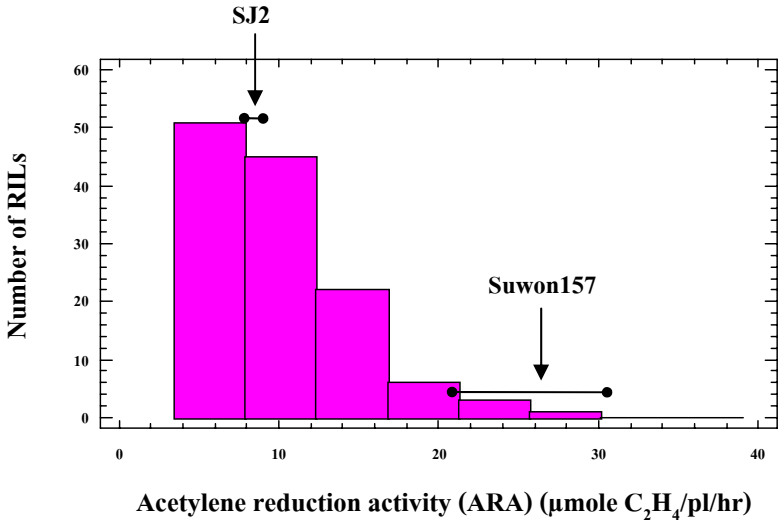


Figure 18 Frequency distribution of acetylene reduction activity (ARA) (µmole C_2H_4/pl/hr) of 136 RILs derived from crossing between Suwon157 (26.35 ± 6.30) and SJ2 (8.14 ± 0.03).

Construction of linkage map among the SSR markers

The cross between SJ2 and Suwon157 was confirmed in the F_1 plants using five SSR markers. Three of them, viz. Satt141, Satt288, and Satt596 identified 12 hybrid (heterozygous) plants (Figure 19). They were multiplied to obtain F_2 seeds from which the subsequent generations was derived by single seed descent method until F_6 and treated as the RIL mapping population. Two hundred and two markers of simple sequence repeat (SSR) and 7 SSR labeled M13 (-21) primers were screened for polymorphism between the parents. One hundred and thirty-one markers were found polymorphic between parents and thus used in amplification of the DNA from 136 RIL lines for collecting genotypic data (Figure 20). Analysis of variance (ANOVA) was used to detect the significant association between markers and RILs in each trait (Appendix table 2). The markers were discarded when there was non-significant effect in character. Then a chi-square (χ^2) test was used for checking the 1 : 1 segregation ratio of the RILs. Finally 123 markers were used in further study (Appendix Table 3).

One-hundred and twenty-four simple sequence repeat (SSR) and 7 SSR labeled M13 (-21) primers of the 195 SSR and 7 SSR labeled M13 (-21) primers produced informative polymorphic markers which could be incorporated into genetic linkage map. A map was constructed using 78

simple sequence repeat (SSR) and 7 SSR labeled M13 (-21) primers onto 20 linkage groups covering 1093.9 cM by MAPMAKER version 3 software at LOD = 3 and distance 50 cM (Figure 21). The smallest interval was found in chromosome H (0.0 cM) and the largest was identified on chromosome A2 (85.4 cM). There were 46 SSR markers remained unlinked.

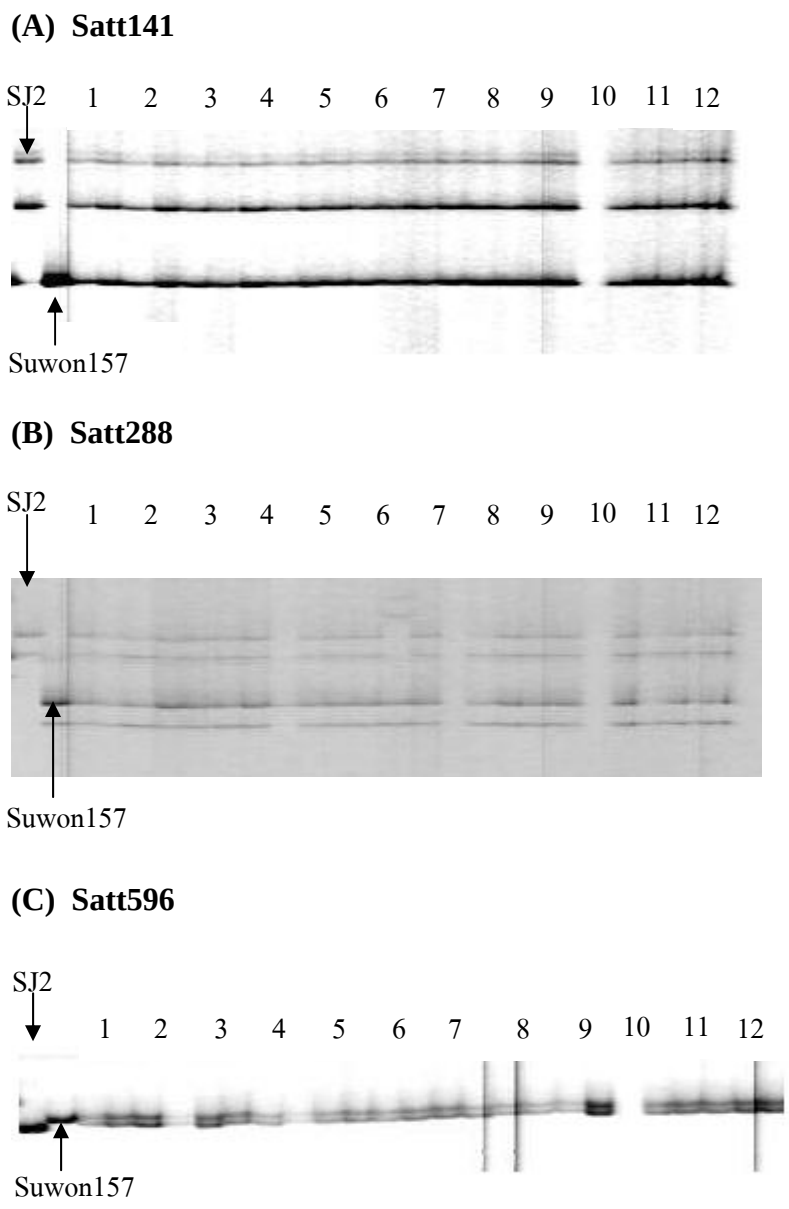


Figure 19 Satt141, Satt288, and Satt596 used to confirm the heterozygous (F_1) bands from crossing between SJ2 and Suwon157.

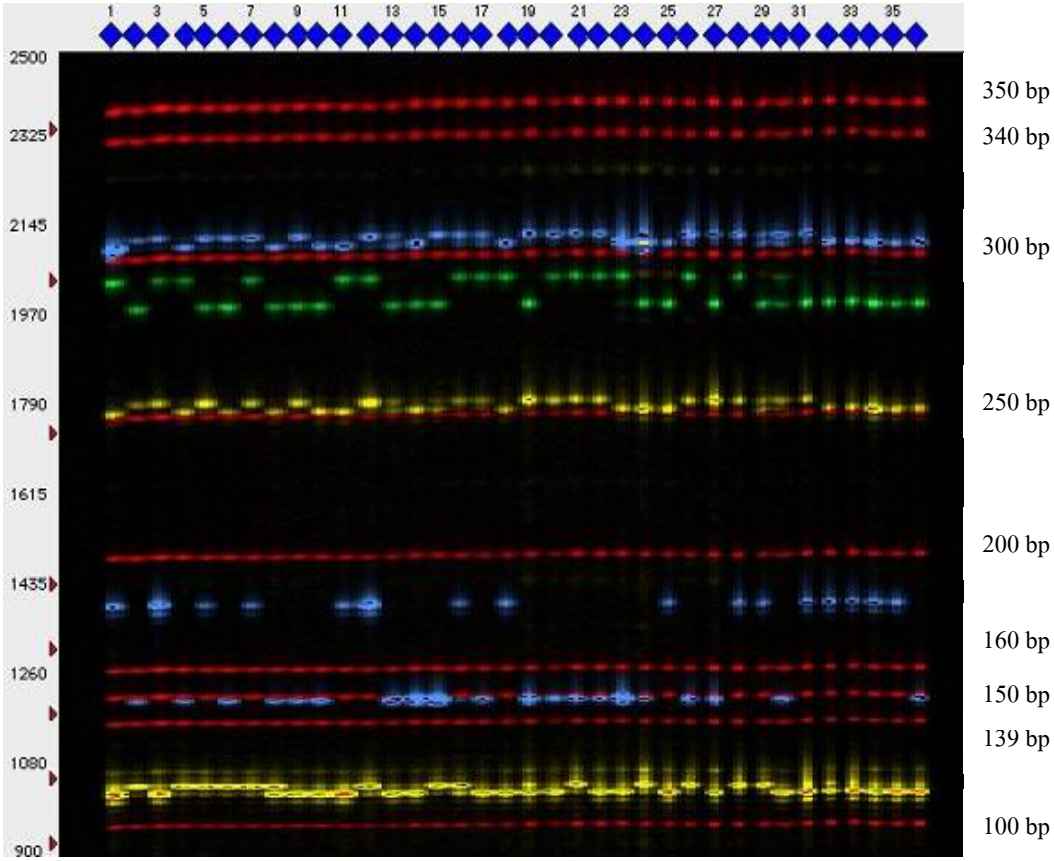


Figure 20 An example of five markers run in ABI 377 for detecting genotypic polymorphism between parents and 136 RIL. Two blue bands are Satt141 which gives the allele size of 183 bp in SJ2 and 148 bp in Suwon157 while Satt414 gives allele size of 303 bp in SJ2 and 306 bp in Suwon157. One green band is Satt388 which give the allele size of 293 bp in SJ2 and and 273 bp in Suwon157. Two yellow bands are Satt567 which gives the allele size of 111 bp in SJ2 and 114 bp in Suwon157 while Satt596 gives the alelle size of 251 bp in SJ2 and 254 bp in Suwon157. Well 1 is SJ2, well2 is Suwon157, and well 3 – 36 are RILs number 1 to 33.

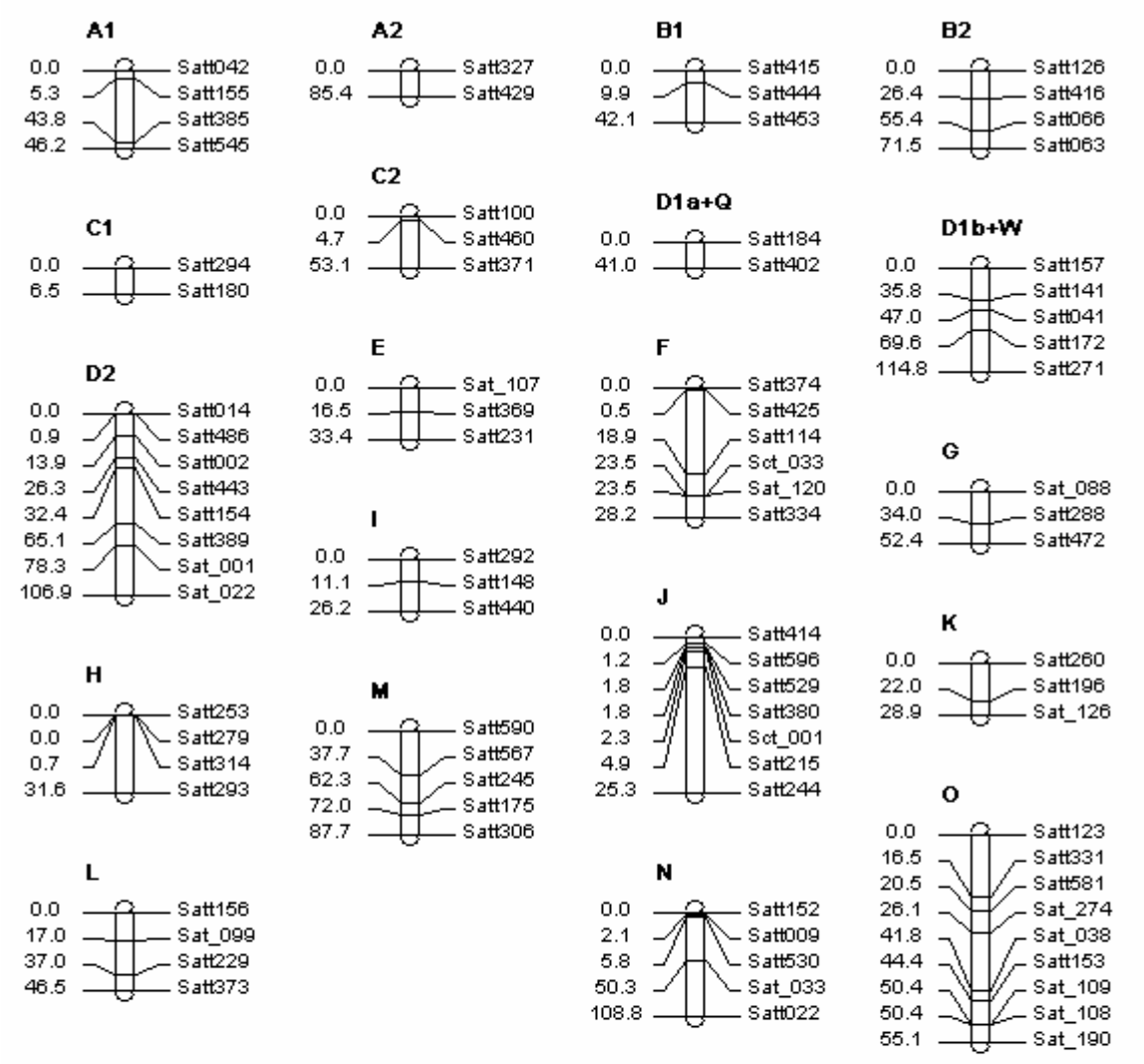


Figure 21 Linkage groups and map distance constructed from 136 RILs derived from crossing between SJ2 and Suwon157. The total map distance is 1093.9 in Kosambi cM unit.

QTL mapping of N_2 fixation components by ANOVA and regression analysis

For nodule number per plant, the SF-ANOVA showed that 19 markers associating with the QTLs at $P < 0.05$. The phenotypic variation due to each marker varied from 3.02 to 18.51 %. Alleles from Suwon157 that increase nodule number were located on linkage group (LG) D1b+W, I, J, L, and O, whereas SJ2 provided positive alleles (the allele that helped increasing number of nodules) on LG A1, C1, and H. Nine markers were assigned on LG O, two markers each on LG C1, D1b+W, and J, while the rest four markers on LG A1, H, I, and L were not linked with the other markers in this study. Satt385 (LG A1), Satt180 (LG C1), Satt157 (LG D1b+W), Satt314 (LG H), Satt440 (LG I), Satt529 (LG J), and Sat_038 (LG O) markers were significant in each linkage group when analyzed by SLG-Regression (Table 9). The MLG-Regression analysis showed that Satt157 (LG D1b+W), Satt440 (LG I), Satt529 (LG J), and Sat_038 (LG O) were linked to the QTLs conditioning nodule number per plant, with the positive alleles came from Suwon157 and the combined R^2 of 30.17 %. Whereas Satt385 (LG A1) had the positive allele derived from SJ2 with the R^2 of 2.98 %.

Sixteen markers were detected to link with the QTLs controlling nodule fresh weight per plant at $P < 0.05$, using SF-ANOVA. Each marker accounted for 3.48 to 19.65 % of the total variation of this trait. QTLs from LG J, L, and O increased nodule fresh weight by Suwon157 alleles, whereas SJ2 provided the positive alleles at the QTLs on LG A1. Satt545 (LG A1), Sct_001 (LG J), Sat_108 and Sat_274 (LG O) were detected to contribute to the QTLs by SLG-Regression analysis. Moreover, these same markers retained their significance after MLG-Regression analysis. Four markers, viz. Sct_001 (LG J), Sat_108 and Sat_274 (LG O) were found to associate with the positive alleles from Suwon157 with the combined R^2 of 29.67 %. Whereas Satt545 (LG A1) identified a positive allele from SJ2 with the R^2 of 4.24 % (Table 10).

Table 9 SSR markers linked to the QTLs controlling nodule number per plant in 136 RILs derived from the soybean cross SJ2 x Suwon157.

Markers	LG	SF-ANOVA[a]		Allelic means[d]		SLG-Regression[b]		MLG-Regression[c]	
		Nodule number/plant							
		P	R^2	Suwon157	SJ2	P	R^2	P	R^2
Satt385	A1	0.0186	4.10	46	53	0.0186	4.10	0.0373	2.98
Satt180	C1	0.0126	4.65	45	52	0.0148	4.51	-	-
Satt294	C1	0.0196	4.15	45	52	-	-	-	-
Satt041	D1b+W	0.0462	3.02	52	46	-	-	-	-
Satt157	D1b+W	0.0204	4.01	52	45	0.0210	4.09	0.0489	2.75
Satt314	H	0.0391	3.41	44	50	0.0391	3.41	-	-
Satt440	I	0.0075	5.49	53	45	0.0075	5.49	0.0279	3.54
Satt529	J	0.0341	3.57	52	45	0.0321	3.68	0.0030	6.84
Sct_001	J	0.0274	3.77	52	45	-	-	-	-
Satt388	L	0.0175	4.77	52	45	-	-	-	-
Sat_038	O	<0.0001	18.51	55	40	<0.0001	21.98	<0.0001	17.04
Sat_108	O	<0.0001	16.25	55	41	-	-	-	-
Sat_109	O	<0.0001	14.87	55	42	-	-	-	-
Sat_190	O	0.0008	8.15	53	43	-	-	-	-
Sat_274	O	0.0001	10.36	55	44	-	-	-	-
Satt123	O	0.0381	3.42	52	46	-	-	-	-
Satt153	O	<0.0001	15.64	55	41	-	-	-	-
Satt331	O	0.0008	8.38	54	44	-	-	-	-
Satt581	O	<0.0001	11.77	55	43	-	-	-	-
									33.15

[a] SF-ANOVA: single factor analysis of variance

[b] SLG-Regression: multiple regression with markers on each linkage group

[c] MLG-Regression: multiple regression with all significant markers from the SLG-Regression model

[d] SJ2: homozygous SJ2, Suwon157: homozygous Suwon157

P: Probability level of the marker

R^2: Coefficient of determination

Table 10 SSR markers linked to the QTLs controlling nodule fresh weight in 136 RILs derived from the soybean cross SJ2 x Suwon157.

Markers	LG	SF-ANOVA[a]		Allelic means[d]		SLG-Regression[b]		MLG-Regression[c]	
		Nodule fresh wt./plant (g)							
		P	R^2	Suwon157	SJ2	P	R^2	P	R^2
Satt385	A1	0.0171	4.20	0.920	1.089	-	-	-	-
Satt545	A1	0.0184	4.27	0.920	1.092	0.0184	4.27	0.0075	4.24
Satt380	J	0.0388	3.48	1.082	0.924	-	-	-	-
Satt414	J	0.0211	4.46	1.091	0.910	-	-	-	-
Satt529	J	0.0172	4.49	1.088	0.909	-	-	-	-
Sct_001	J	0.0090	5.26	1.092	0.900	0.0263	4.29	0.0296	2.69
Satt388	L	0.0219	4.45	1.068	0.893	-	-	-	-
Sat_038	O	<0.0001	16.39	1.137	0.802	-	-	-	-
Sat_108	O	<0.0001	13.43	1.137	0.831	0.0054	5.28	0.0018	6.15
Sat_109	O	<0.0001	12.74	1.139	0.843	-	-	-	-
Sat_190	O	0.0019	7.00	1.096	0.879	-	-	-	-
Sat_274	O	<0.0001	19.65	1.194	0.828	<0.0001	19.33	<0.0001	20.83
Satt123	O	0.0064	5.84	1.102	0.903	-	-	-	-
Satt153	O	<0.0001	14.73	1.117	0.816	-	-	-	-
Satt331	O	<0.0001	12.70	1.152	0.854	-	-	-	-
Satt581	O	<0.0001	17.84	1.185	0.833	-	-	-	-
									33.91

[a] SF-ANOVA: single factor analysis of variance

[b] SLG-Regression: multiple regression with markers on each linkage group

[c] MLG-Regression: multiple regression with all significant markers from the SLG-Regression model

[d] SJ2: homozygous SJ2, Suwon157: homozygous Suwon157

P: Probability level of the marker

R^2: Coefficient of determination

Twelve markers were identified to associate with nodule dry weight per plant by SF-ANOVA at $P < 0.05$. Each marker accounted for 1.67 to 13.65 % of the total variation of this trait. All markers indicated that all Suwon157 alleles increased nodule dry weight. There are two markers on LG J, one on LG K, and nine on LG O. SLG-Regression analysis retained Satt414 on LG J, Satt260 on LG K and two markers (Sat_108 and Sat_274) on LG O. Finally, with MLG-Regression analysis, only Satt260, Sat_108, and Sat_274 were found to identify positive QTL, with the combined R^2 of 28.74 % (Table 11).

The SF-ANOVA identified 17 markers linking to QTLs for plant dry weight (Table 12). Each marker accounted for 3.44 to 14.75 %. All markers associating with Suwon157 alleles increased plant dry weight. Both SLG and MLG-Regression analyses resulted in four significant markers, viz. Satt440 on LG I, Sct_001 on LG J, Sat_038 and Sat_274 on LG O. Their combined R^2 effect was 28.06 %.

Ten markers were identified to associate with QTLs conditioning acetylene reduction activity (ARA) by SF-ANOVA at $P < 0.05$. Each marker accounted for 3.03 to 11.92 %. All markers of Suwon157 alleles increased ARA. Only two markers, Satt157 on LG D1b+W and Sat_274 on LG O identified QTLs with both SLG- and MLG-Regression analyses (Table 13). They contributed a rather low combined R^2 of 18.52 %.

Table 11 SSR Markers linked to the QTLs controlling nodule dry weight in 136 RILs derived from the soybean cross SJ2 x Suwon157.

Markers	LG	SF-ANOVA[a]		Allelic means[d]		SLG-Regression[b]		MLG-Regression[c]	
		Nodule dry wt./plant (g)							
		P	R^2	Suwon157	SJ2	P	R^2	P	R^2
Satt414	J	0.0254	4.20	0.236	0.198	0.0214	4.48	-	-
Sct_001	J	0.0409	3.25	0.236	0.202	-	-	-	-
Satt260	K	0.0279	3.64	0.234	0.199	0.0279	3.64	0.0446	2.60
Sat_038	O	<0.0001	10.72	0.244	0.183	-	-	-	-
Sat_108	O	<0.0001	10.97	0.247	0.186	0.0125	4.66	0.0005	8.38
Sat_109	O	0.0001	1.67	0.228	0.216	-	-	-	-
Sat_190	O	0.0034	6.21	0.240	0.194	-	-	-	-
Sat_274	O	<0.0001	13.27	0.255	0.187	0.0001	12.18	<0.0001	17.76
Satt123	O	0.0376	3.44	0.237	0.203	-	-	-	-
Satt153	O	0.0004	9.24	0.239	0.186	-	-	-	-
Satt331	O	0.0006	8.71	0.248	0.193	-	-	-	-
Satt581	O	<0.0001	13.65	0.256	0.187	-	-	-	-
									28.74

[a] SF-ANOVA: single factor analysis of variance

[b] SLG-Regression: multiple regression with markers on each linkage group

[c] MLG-Regression: multiple regression with all significant markers from the SLG-Regression model

[d] SJ2: homozygous SJ2, Suwon157: homozygous Suwon157

P: Probability level of the marker

R^2: Coefficient of determination

Table 12 SSR markers linked to the QTLs controlling plant dry weight in 136 RILs derived from the soybean cross SJ2 x Suwon157.

Markers	LG	SF-ANOVA[a]		Allelic means[d]		SLG-Regression[b]		MLG-Regression[c]	
		Plant dry wt./plant (g)							
		P	R^2	Suwon157	SJ2	P	R^2	P	R^2
Satt440	I	0.0293	3.68	2.07	1.84	0.0293	3.68	0.0069	4.80
Satt215	J	0.0205	4.43	2.09	1.83	-	-	-	-
Satt380	J	0.0058	6.12	2.11	1.80	-	-	-	-
Satt414	J	0.0019	7.98	2.13	1.77	-	-	-	-
Satt529	J	0.0016	7.73	2.12	1.77	-	-	-	-
Satt596	J	0.0053	6.39	2.11	1.79	-	-	-	-
Sct_001	J	0.0006	8.79	2.13	1.76	0.0023	8.30	0.0029	6.18
Satt388	L	0.0339	3.82	2.05	1.81	-	-	-	-
Sat_038	O	<0.0001	12.69	2.13	1.69	<0.0001	13.82	0.0320	2.91
Sat_108	O	<0.0001	11.76	2.14	1.72				
Sat_109	O	<0.0001	11.76	1.88	1.81	-	-	-	-
Sat_190	O	0.0026	6.55	2.09	1.78	-	-	-	-
Sat_274	O	<0.0001	14.75	2.20	1.73	0.0190	4.06	<0.0001	14.17
Satt123	O	0.0376	3.44	2.06	1.84	-	-	-	-
Satt153	O	<0.0001	11.49	2.11	1.71	-	-	-	-
Satt331	O	0.0005	8.99	2.15	1.78	-	-	-	-
Satt581	O	<0.0001	14.02	2.20	1.74	-	-	-	-
									28.06

[a] SF-ANOVA: single factor analysis of variance

[b] SLG-Regression: multiple regression with markers on each linkage group

[c] MLG-Regression: multiple regression with all significant markers from the SLG-Regression model

[d] SJ2: homozygous SJ2, Suwon157: homozygous Suwon157

P: Probability level of the marker

R^2: Coefficient of determination

Table 13 SSR markers linked to the QTLs controlling acetylene reduction activity (ARA) in 136 RILs derived from the soybean cross SJ2 x Suwon157.

Markers	LG	SF-ANOVA[a]		Allelic means[d]		SLG-Regression[b]		MLG-Regression[c]	
		ARA μmole C_2H_4/pl/hr							
		P	R^2	Suwon157	SJ2	P	R^2	P	R^2
Satt041	D1b+W	0.0324	3.50	10.58	8.68	-	-	-	-
Satt157	D1b+W	0.0147	4.46	10.61	8.49	0.0125	4.81	0.0021	6.18
Satt388	L	0.0239	4.32	10.50	8.48	-	-	-	-
Sat_038	O	0.0023	3.76	10.83	8.23	-	-	-	-
Sat_108	O	0.0436	3.03	10.25	8.77	-	-	-	-
Sat_109	O	0.0449	3.06	11.07	7.96	-	-	-	-
Sat_274	O	<0.0001	11.92	11.24	8.16	<0.0001	12.36	<0.0001	12.34
Satt153	O	0.0135	4.57	10.30	8.47	-	-	-	-
Satt331	O	0.0012	7.91	11.02	8.21	-	-	-	-
Satt581	O	0.0002	10.17	11.25	8.21	-	-	-	-
									18.52

[a] SF-ANOVA: single factor analysis of variance

[b] SLG-Regression: multiple regression with markers on each linkage group

[c] MLG-Regression: multiple regression with all significant markers from the SLG-Regression model

[d] SJ2: homozygous SJ2, Suwon157: homozygous Suwon157

P: Probability level of the marker

R^2: Coefficient of determination

Our experiment identified different markers from the earlier works, possibly due to different parents used in developing the mapping population and phenotyping. In this study, all major QTLs associated with the N_2 fixation components were found on LG O, while the minor ones were located on LG A1, D1b+W, I, J and K. Cregan *et al*. (1999) reported that *Rj_1* allele conditioning nodulation was located on LG D1b+W, while *Rj_2* allele conditioning ineffective nodulation was found on LG J in 3 mapping populations, viz. 59 F_2 plants from *G. max* x *G. soja*, 240 RILs from Minsoy x Noir, and 57 F_2 plants from Clark x Harosoy. The genes *Rj_2* for ineffective nodulation, *Rmd* for powdery mildew resistance and *Rps2* for Phytophthora root and stem rot resistance were located on the classical linkage

group 19 in the BARC-4 x Clark63 population (Devine *et al*., 1991). A study from the progenies derived from two isolines of the soybean cultivar 'Williams' showed that LG J was related to classical linkage group 19 with certain common loci such as *Rj_2, Rmd,* and *Rps_2* (Polzin *et al*., 1994). A gene for supernodulation (*nts*) was located with the RFLP anchor markers on LG H (Kolchinsky *et al*., 1997).

QTL mapping of N_2 fixation components by mQTL and STATGRAPHIC

Since LG O carries the maximum number of markers linking to the fixation components, the putative QTLs on this linkage group were determined by MAPMAKER version 3.0 construct to linkage group, and mQTL beta version 0.98 program locate markers linking to N_2 fixation components.

The interval of seven markers was predicted to come the QTL of nodule number per plant, with the value greater than simple main effect (SIM) 11.8. They were Satt331, Satt581, Sat_274, Sat_038, Satt153, Sat_109, and Sat_108. The highest value of SIM was detected in the QTLs between Sat_274 and Sat_038 on linkage group O. The distance was 10 cM between the QTL with Sat_274 and 5.7 cM with Sat_038. (Figure 22-blue line).

The interval of eight markers covered nodule fresh weight per plant, nodule dry weight per plant, and plant dry weight value on linkage group O, with the threshold value over simple main effect (SIM) of 11.9, 11.8, and 11.6, respectively. They were Satt123, Satt331, Satt581, Sat_274, Sat_038, Satt153, Sat_109, and Sat_108. The highest value of SIM indicating the QTLs was detected between Sat_274 and Sat_038 on linkage group O. The distance of the QTL from Sat_274 and Sat_038 were 5 and 10.7 cM, respectively (Figure 22-pink line, yellow line, and skyblue line).

The interval of three markers on linkage group O was covering the putative QTL of acetylene reduction activity (ARA) over simple main effect (SIM) of 11.4. They were Satt581, Sat_274, and Sat_038. The highest value of SIM was detected in the QTLs between Sat_274 and Sat_038. The distance from QTL to Sat_274 and Sat_038 were 5 and 10.7 cM, respectively (Figure 22- purple line).

All results confirmed by sCIM and LOD score confirmed that the QTLs of all N_2 fixation components were located between Sat_274 and Sat_038 (Figure 23 and 24).

Single and multiple regression (STATGRAPHIC version 3.0) was used to identify QTL linking to each trait by MAPMAKER version 3.0 and mQTL beta version 0.98 (Table 15). Nodule number per plant detected Sat_038 and Sat_108 (R^2 = 19.78) on linkage group O. Each marker can explain their contribution differently as measured by R^2. For example Sat_038 contributed 17.90%, while Sat_108 gave 15.62% by single regression analysis. Nodule fresh weight per plant was explained by Sat_274 and Sat_108 (combined R^2 = 23.75%) on linkage group O. Individually R^2 of Sat_274 was 19.05% and Sat_108 was 12.79% by single regression analysis. Nodule dry weight per plant was explained by Sat_274 and Sat_108 (combined R^2 = 16.86%) on linkage group O. Individually R^2 of Sat_274 was 12.62% and Sat_108 was 10.31 by % single regression analysis. Plant dry weight was explained by Sat_274 and Sat_108 (combined R^2 = 18.61%) on linkage group O. Individually, R^2 of Sat_274 was 14.12% and Sat_108 was 11.10% by single regression analysis. Acetylene reduction activity (ARA) was explained by Sat_274 (combined R^2 = 11.26%) on linkage group O. This confirmed that all these QTL located around marker Sat_274 to Sat_108.

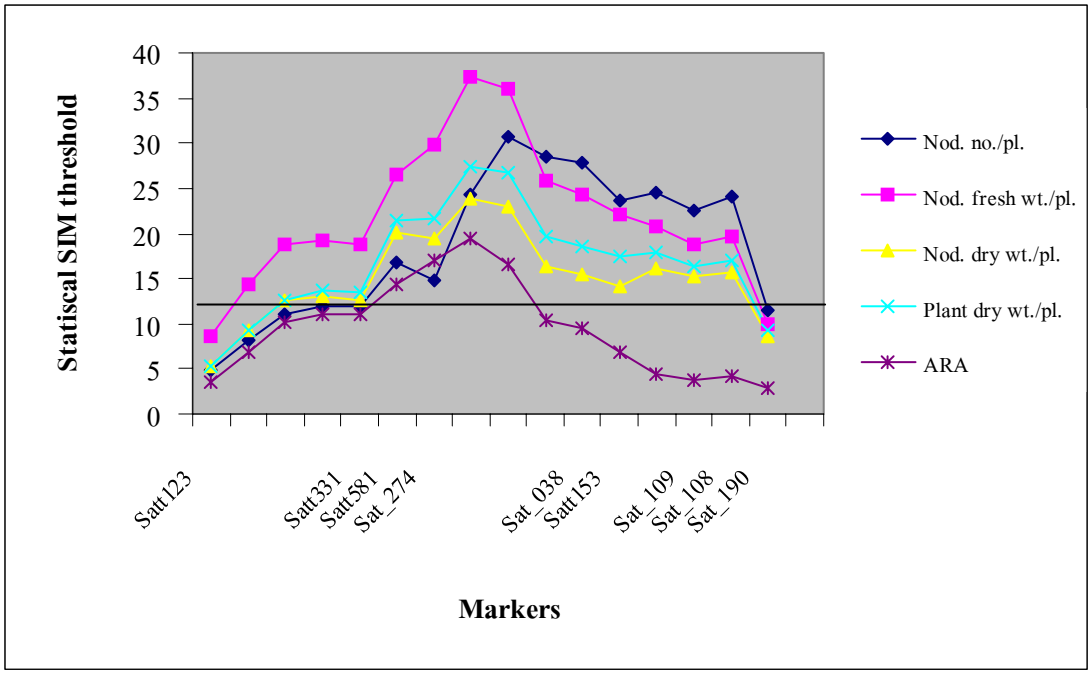


Figure 22 SIM test statistic from the QTL analysis of nodule number per plant, nodule fresh weight per plant, nodule dry weight per plant, plant dry weight, and acetylene reduction activity (ARA) on linkage group O. The X-axis shows ordering of the markers. The horizontal bar indicates the maximum significant thresholds (P = 0.05) (simple main effect = 12.6).

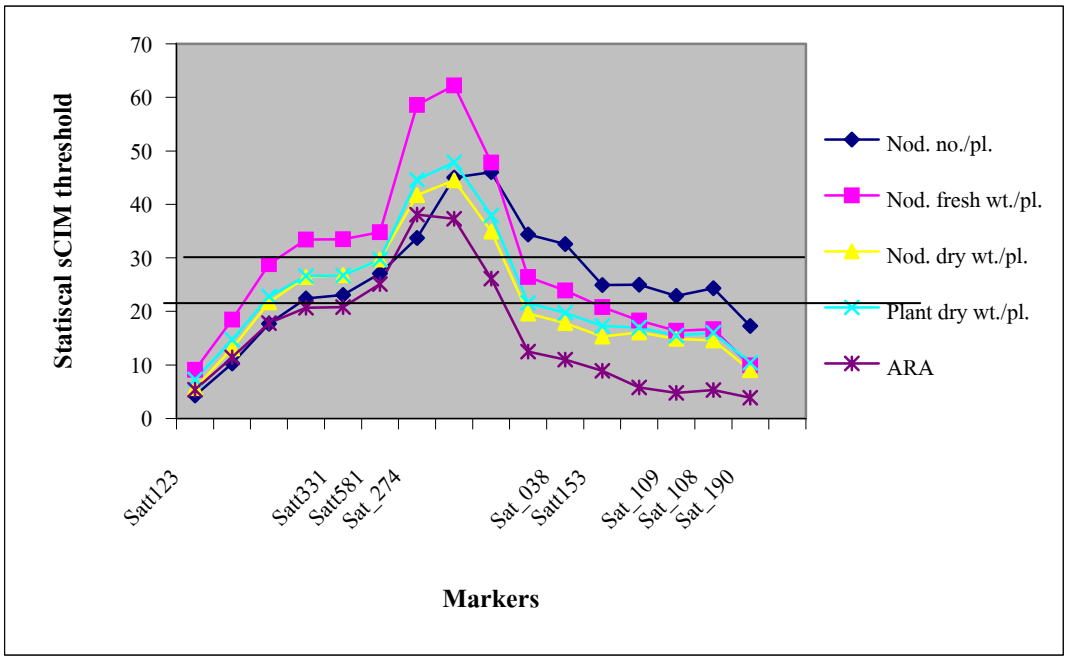


Figure 23 sCIM test statistic from the QTL analysis of nodule number per plant, nodule fresh weight per plant, nodule dry weight per plant, plant dry weight, and acetylene reduction activity (ARA) on linkage group O. The X-axis shows ordering of the markers. The horizontal bar indicates the maximum significant thresholds (P = 0.05) (composit main effect = 26.6).

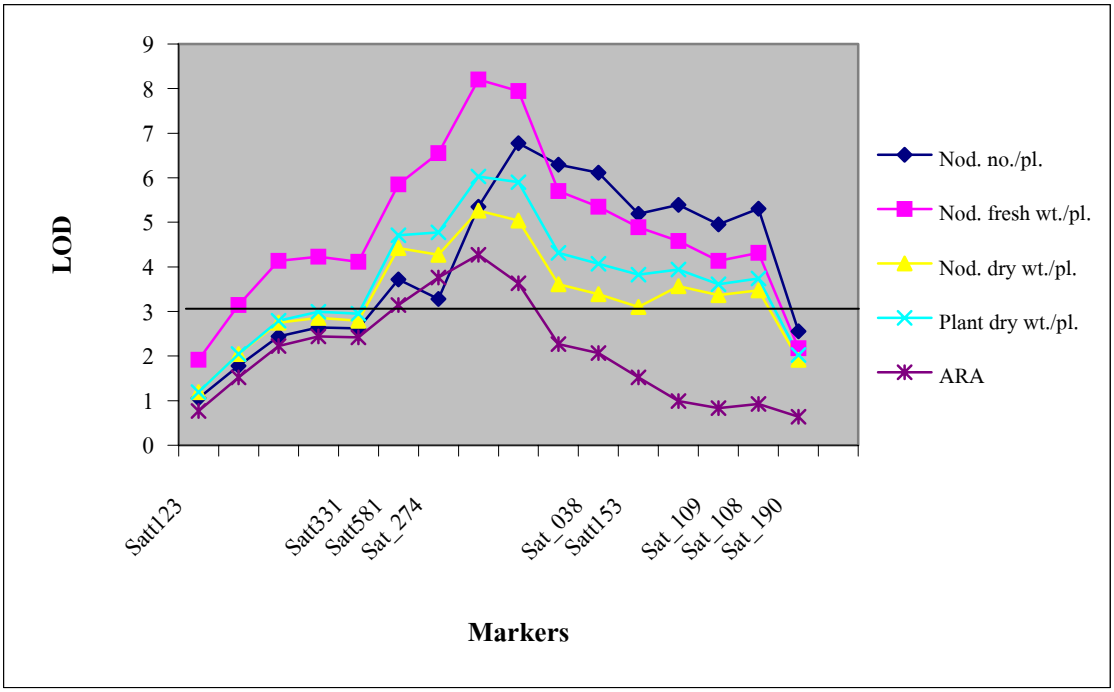


Figure 24 LOD score of the QTL analysis of nodule number per plant, nodule fresh weight per plant, nodule dry weight per plant, plant dry weight, and acetylene reduction activity (ARA) on linkage group O. The X-axis shows ordering of the markers.

Table 15 SSR markers on linkage group O linked to the QTLs controlling N_2 fixation components in 136 RILs derived from the soybean cross SJ2 x Suwon157.

N_2-fixation components	Markers	R^2	Mutilocus R^2
Nodule no./plant	Sat_038	17.90	19.78
	Sat_108	15.62	-
Nodule fresh wt/plant (g)	Sat_108	12.79	23.75
	Sat_274	19.05	-
Nodule dry wt/plant (g)	Sat_108	10.31	16.86
	Sat_274	12.62	-
Plant dry weight (g)	Sat_108	11.10	18.61
	Sat_274	14.12	-
ARA umole C_2H_4/pl/hr	Sat_274	11.26	11.26

Since this study employed only 202 SSR markers, the combined R^2 explaining the QTL in each fixation component was not high. The values varied from 18.52% in ARA to 33.91% in fresh nodule weight. More SSR markers should be used in combination with AFLP, RAPD, RFLP, SNP, etc., to fill up the gaps in each linkage group and to identify more exact markers correlating with each N_2 fixation component. The information will be useful in pyramiding all the desirable markers into one soybean genotype through marker–assisted selection so that the soybean line with the best N_2 fixation ability can be obtained in the future.

CONCLUSION

1. The genetic diversity of 14 Thai and 18 Korean soybean varieties was determined by SSR markers. Fluorescent labeling of alleles combined with automated sizing with internal size standard in each lane were observed. Gene diversity ranged from 0.56 to 0.83 with each of 18 SSR loci, giving the average alleles per locus of 7.67. The genetic similarity coefficient of these genotypes ranged from 0.025 to 0.944 and can be divided into 5 major groups.

2. All the N_2 fixation traits in this study were positively correlated and thus made it rather easy to simultaneously improve them. Generally, the Thai soybean cultivars were superior than Korean soybeans in all the traits, except in number of nodules per plant which was not statistically significant. Fresh nodule weight was the most important trait contributing to ARA and thus can be considered as the key trait for the improvement of N_2 fixation in this population.

3. The QTLs associating with N_2 fixation components were located between Sat_274 and Sat_038 or Sat_274 and Sat_108 on linkage group O.

LITERATURE CITED

Akkaya, M.S., A.A. Bhagwat, and P.B. Cregan. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. Genetics 132: 1131-1139.

Akkaya, M.S., R.C. Shoemaker, J.E. Specht, A.A. Bhagwat, and P. B. Cregan. 1995. Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci. 35: 1439-1445.

Anderson, J.A. 1992. Optimizing parental selection for genetic linkage maps. Genome 36:181-186.

Anon. 2004. Soybeans: Area production, yield, farm price and farm value, crop year 1993/94-2002/03. Thailand, Office of Agricultural Economics. Available Source: http://www.oae.go.th/statistic/yearbook/2002-03, December 3, 2004.

Anon. 2005. Interaction between leguminous plants and rhizobia mediated by nod factors. Available Source: http://www.glycoforum.gr.jp/science/word/: Saccharide/SA-A02E.htm, March 16, 2005.

Aprison M.H., R.H. Burris and W.E. Magee. 1954. Nitrogen fixing by excised soybean root nodules. J. Biol. Chem. 208, 28-39.

Attewell, J. and F.A. Bliss. 1985. Host plant characteristics of common bean lines selected using indirect measures of N_2 fixation. pp. 3-5. In: H.J. Evan, D.J. Bottamley and W.E. Newton, eds. Nitrogen Fixation Research Progress. Martinus Nijhoff Publishers, U.S.A.

Brookes, A.J. 1999. The essence of SNPs. Gene 234: 177-186.

Brown, S.M. 1999. Snapping up SNPs. Biotechniques 26(6): 1090-1094.

Broughton, W.J. and M.J. Dilworth. 1970. Control of leghaemoglobin synthesis in snake beans. Biochem. J. 125: 1075-1085.

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

Campbell, N.A. and J.B. Reece. 2002. Biology, 6th ed. Pearson, Inc., San Francisco, CA.

Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. VanToai, D.G. Lohnes, J. Chung and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci. 39: 1464-1490.

Crespi, M. and S. Gálvez. 2000. Molecular mechanisms in root nodule development. J. Plant Growth Regul. 19: 155-166.

Cullimore, J. and J. Dènariè. 2003. How legumes select their sweet talking symbionts. Science 302: 575-578.

Damery, J.T. and M. Alexander. 1969. Physiological differences between effective and ineffective strains of Rhizobium. Soil Sci. 108, 209-216.

Devine, T.E., T.C. Kilen, and J.J. O'Neill. 1991. Genetic linkage of phytophthora resistance gene Rps_2 and the nodulation response gene Rj_2 in soybean. Crop Sci. 31: 713-715.

Diwan, N. and P.B. Cregan. 1997. Automated sizing of fluorescent-labeled simple sequence repeat (SSR) markers to assay genetic variation in soybean. Theor. Appl. Genet. 95: 723-733.

Downie, J.A. and M. Parniske. 2002. Plant biology: fixation with regulation. Nature 420: 369-370.

DÖbereiner J. 1966. Evaluation of nitrogen fixation in legumes by the regression of total plant nitrogen with nodule weight. Nature 24: 153-166.

Ferguson. B.J. and U. Mathesius. 2003. Signaling interactions during nodule development. J. Plant Growth Regul. 22: 47-72.

Ferreira, A.R. and P. Keim. 1997. Genetic mapping of soybean (*Glycine max* (L.) Merr.) using random amplified polymorphic DNA (RAPD). Plant Mol. Biol. Report 15: 335-354.

Giller, K.E. 2001. Nitrogen Fixation in Tropical Cropping Systems, 2 nd ed, Biddles Ltd, Guildford and King's Lynn, London, UK.

Gustavo, C.A. and P.M. Gresshoff. 1991. Plant genetic control of nodulation. Annu. Rev. Microbiol. 45: 345-382.

Hardy, R.W.F., R.D. Holsten, E.K. Jackson and R.C. Burns. 1968. The acetylene-ethylene assay for N_2 fixation of the laboratory and field evaluation. Plant Physiol. 43: 1185-1207.

Haper, J.E. 1999. Nitrogen fixation - limitation and potential. pp. 235-243. *In*: R. Shibles, ed. Proceedings of the World Soybean Research Conference VI, 4-7 August 1999, Chicago, Illinois, USA.

Hymowitz, T. 1990. Soybeans: The Success Story, pp. 159-163. *In* J. Janick and J. Simon, eds. Advances in New Crops. Timber Press, Portland, Oregon, USA.

Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. Bull. Soc. Vaudoise Sci. Nat. 44: 223-270

Kao, C.H., A.B. Zeng and R.D. Teasdale. 1999. Multiple interval mapping for quantitative trait loci. Genetics 152: 1203-1216.

Keim, P., B.W. Diers, T.C. Olson, R.C.Shoemaker. 1990. RFLP mapping in soybean: association between maker loci and variation in quantitative traits. Genetics 126: 735-742.

Keim, P., J.M. Schupp, S.E. Travis, K. Clayton, T. Zhu, L. Shi, A. Ferreira and D.M. Webb. 1997. A high-density soybean genetic map based on AFLP markers. Crop Sci. 37: 537-543.

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.

Kolchinsky, A., Landau-Ellis and P.M. Gresshoff. 1997. Map order and linkage distance of molecular markers close to the supernodulation (nts-1) in soybean. Mol. Gen. Genet. 254: 29-36.

Lagercrantz, U., H. Ellegran, and L. Anderson. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res. 21:1111-1115.

Lambrides, C.J., R.J. Lawn, I.D. Godwin, J. Manners and B.C. Imrie. 2000. Two genetic linkage maps of mungbean using RFLP and RAPD markers. Aust. J. Agric. Res. 51: 415-425.

Latimore, M., J.J. Giddens and D.A. Ashley. 1977. Effect of ammonium and nitrate nitrogen upon photosynthate supply and nitrogen fixation by soybeans. Crop Sci. 17: 399-404.

Lee, W.S., G.G. Brown and D.P.S. Verma. 1983. Chromosomal arrangement of leghaemoglobin genes in soybean. Nucleic Acids Res. 5541-5552.

Lee, S.H., K.Y. Park, H.S. Lee, E.H. Park and H.R. Boerma. 2001. Genetic mapping of QTLs conditioning soybean sprout yield and quality. Theor. Appl. Genet. 103: 702-709.

Lincoln, S.L., M.J. Dary and E.S. Lander. 1993. Constructing genetic linkage maps with MAPMAKER/EXP version 3.0, a tutorial and reference manual. Available Source: http://linkage.rockefeller.edu/soft /mapmaker/index.html, December 3, 2005.

Liu, B.H. 1998. Statistical Genomics: Linkage, Mapping and QTL analysis. CRC Press, Florida, USA.

Mader, S.S. 2004. Biology. 8th ed. McGraw-Hill, Inc., New York.

Mansur, L.M., J.H. Orf, K.Chase, T. Jarvik, P.B. Cregan and K.G. Lark. 1996. Genetic mapping of agronomic traits recombinant inbred lines of soybean. Crop Sci. 36: 1327-1336.

Mohan, M., S. Nair, A. Bhagwat, T.G. Krishna, M. Yano, C.R. Bhatia and T. Sasaki. 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding 3(87): 87-103.

Morgante, M. 1994. Genetic mapping and variability of seven soybean simple sequence repeat loci. Genome 37:763-769.

Mylona, P., K. Pawlowski and T. Bisselling. 1995. Symbiotic nitrogen fixation. The Plant Cell 7: 869-885.

Palmer, R.G., T. Hymowitz and R.L. Nelson. 1996. Germplasm diversity within soybean. *In* D.P.S. Verma and R.C. Shoemaker, eds. Soybean Genetic, Molecular Biology and Biotechnology. Biddles Ltd., Guildford, UK.

Pazdernik, D.L., P.H. Graham, C.P. Vance and J.H. Orf. 1996. Host genetic variation in the early nodulation and dinitrogen fixation of soybean. Crop Sci. 36: 1102-1107.

Peoples, M.B., D.F. Herridge, B. Rerkasem, A. Bhromsiri, C. Samet and J. Brockwell. 1997. Improving nitrogen fixation by soybean. pp. 270-276. *In*: B. Napompeth, ed., Proceedings of the World Soybean Research Conference V,21-27 February 1994, Chiang Mai, Thailand.

Peter, J.W., K. Fisher and D.R. Dean. 1995. Nitrogenase structure and function: A biochemical-genetic perspective. Annu. Rev. Microbiol. 49:335-366.

Polzin, K.M., D.G. Lohnes, C.D. Nickell, and R.C. Shoemaker. 1994. Integration of *Rps_2, Rmd*, and *Rj_2* into linkage group J of the soybean molecular map. J. Hered. 85(4): 300-303.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey, and A. Rafalski. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. 2:225-238.

Pracht, J.E., C.D. Nickell and J.E. Haper. 1993. Genes controlling nodulation in soybean: *Rj_5* and Rj_6. Crop Sci. 33: 711-713.

Rohlf, F. J. 1998. Numerical taxonomy and multivariate analysis system, NTSYS-pc program manual version 2.0. Applied Biostatistics Inc., 3 Heritage Lane, Setauket, NY 11733.

SAS institute. 1999-2000. SAS/STAT User's Guides, version 8.01st ed., SAS Institute, Cary, Nort Carolina, USA.

Schlötterer, C. and D. Tautz. 1992 Slippage synthesis of simple sequence DNA. Nuleic Acids Res. 20:211-215.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biotech. 18: 233-234.

Shoemaker, R.C., R.D. Guffy, L.L. Lorenzen and J.E. Specht. 1992. Molecular genetic mapping of soybean: map utilization. Crop Sci. 32: 1091-1098.

Smith, K.P. and R.M. Goodman. 1999. Host variation for interactions with beneficial plant-associated microbes. Ann. Rev. Phytopath. 37: 473-491.

Somasegaran, P. and H.J. Hoben. 1985 Methods in Legume-Rhizobium Technology. USAID, Hawaii. 365 p.

Somwang T, A. Yothasiri, C. Hongprayoon, A. Nuntagij, S. Kotepong and P. Srinives. 2002. Heritability of nodulation and N_2-fixation efficiency in soybean (*Glycine max* (L.) Merill). Korean J. Breed. 34 (4): 331-336.

Song, Q.J., L.F. Marek, R.C. Shoemaker, K.G. Kark, V.C. Concibido, X. Delannay, J.E. Specht and P.B. Cregan. 2004. A new Integrated genetic map of the soybean. Theor. Appl. Genet. 109: 122-128.

Southern, E.M. 1976. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

Smith, G.P. 1976. Evolution of repeated DNA sequences by unequal crossover. Science 191:528-535.

Steel, R.G.D. and J.H. Torrie. 1980. Principles and Procedures of Statistics: A Biometrical Approach, 2nd ed., McGraw-Hill, Inc., USA.

Tanya, P., P. Srinives, T. Toojinda, A. Vanavichit, A. Nuntakij, S. Kotepong and S. H. Lee. 2005. Evaluation of N_2-fixation traits in Thai and Korean soybean cultivars. (submitted to SciAsia)

Tomooka, N., C. Lairungreang, C. Thavarasook and T. Murakami. 1992. The difference of nitrogen fixation ability between cultivar and wild type of mungbean (*Vigna radiata*) and blackgram (*Vigna mungo*) In: *Nitrogen Uptake in Legume-Rhizobium Symbiosis in Thailand –Focus on Mungbean*, (Edited by Murakami T., S. Siripin, P. Wadisirisuk, N. Boonkerd, T. Yoneyama, T. Yokoyama and H. Imai), February 1992, pp. 63-68. Hewlett-Packard Company.

Vest, G. 1970. *Rj_3*-A gene conditioning ineffective nodulation in soybean. Crop Sci. 10: 34-35.

Vest, G. and B.E. Caldwell 1972. Rj_4-A gene conditioning ineffective nodulation in soybean. Crop Sci. 12: 692-693.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T.V.D. Lee, M. Hornes, A. Frijters, J. Pot, J. Pelemen, M. Kuiper and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprint. Nucleic Acids Rese. 23 (21): 4407-4414.

Vuong, T.D. and J.E. Harper. 2000. Inheritance and allelism analysis of hypernodulation genes in the NOD3-7 and NOD2-4 soybean mutants. Crop Sci. 40: 700-703.

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

Wang, D.G., J.B. Fan, C.J Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Huddon, R. Lipshutz, M. Chee and E. S. Lander. 1998. Large-scale identification, mapping, and genotyping of single nucleotide polymorphisms in the human genome. Science 208: 1077-1082.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18(22): 6531-6535.

Young, J.P.W. 1996. Phylogeny and taxonomy of Rhizobia. Plant and Soil 186: 45-52.

Zhang, W.K., Y.J. Wang, G.Z. Luo, J.S. Zhang, C.Y. He, X.L. Wu, J.Y. Gai and S.Y. Chen. 2004. QTL mapping of ten agronomic traits on the soybean (*Glycine max* (L.) Merr.) genetic map and their association with EST markers. Theor. Appl. Genet. 108: 1131-1139.

Appendix

Appendix Table 1 N_2 fixation components in SJ2, Suwon157 and their 136 RIL population inoculated with DASA 01026 rhizobium strain.

Parent/ RILs lines	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry weight/plant (g)		ARA (μmole C_2H_4/pl/hr)	
	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2
SJ2	40	36	0.668	0.615	0.162	0.151	1.53	1.42	8.2	8.1
Suwon157	72	74	1.380	1.355	0.286	0.292	2.53	2.18	30.8	21.9
1	59	79	1.361	1.701	0.258	0.323	2.05	2.21	8.8	11.8
2	40	48	0.377	1.091	0.067	0.247	1.07	2.22	4.0	9.5
3	38	39	0.788	1.130	0.149	0.248	1.62	2.07	8.7	11.4
4	34	48	0.673	0.639	0.126	0.117	1.90	0.96	3.0	2.5
5	54	57	1.147	0.968	0.268	0.240	2.06	1.83	13.9	8.1
6	35	24	0.697	0.778	0.150	0.161	1.44	1.45	5.5	5.9
7	46	32	0.755	0.435	0.174	0.098	1.82	1.26	9.0	5.1
8	36	41	0.666	0.482	0.150	0.119	1.43	1.31	4.7	4.6
9	57	50	1.046	1.069	0.245	0.276	2.19	2.53	0.9	6.2
10	28	35	0.928	1.139	0.237	0.234	2.32	2.18	7.6	16.4
11	73	55	0.887	0.515	0.160	0.112	1.80	1.32	12.2	5.5
12	38	32	0.735	0.965	0.157	0.220	1.84	2.16	5.7	9.0
13	22	35	0.854	0.696	0.217	0.150	1.81	1.71	4.8	6.7
14	17	28	0.282	0.697	0.058	0.133	0.90	1.58	3.1	4.9
15	38	31	0.942	0.697	0.221	0.171	2.09	1.89	5.2	10.5
16	60	41	1.815	1.397	0.404	0.329	3.69	2.90	15.2	13.9
17	22	22	0.313	0.938	0.600	0.195	1.97	1.19	2.7	5.9
18	53	49	1.203	1.158	0.216	0.205	2.06	1.96	11.8	11.0
19	41	59	0.657	0.600	0.144	0.139	1.64	1.35	3.9	4.4

Appendix Table 1 (cont'd)

Parent/ RILs lines	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry weight/plant (g)		ARA (µmole C_2H_4/pl/hr)	
	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2
20	32	22	0.771	0.623	0.157	0.126	1.36	1.25	3.3	1.9
21	39	36	1.005	1.158	0.210	0.240	2.00	2.24	7.8	8.8
22	41	39	0.683	0.562	0.144	0.128	1.59	1.40	10.6	7.2
23	46	47	0.809	0.555	0.182	0.138	1.67	1.68	7.5	2.9
24	46	46	0.880	0.879	0.222	0.183	1.73	1.83	7.2	8.9
25	52	46	1.554	0.807	0.306	0.168	2.48	1.43	9.2	5.8
26	45	46	1.003	0.974	0.192	0.224	1.84	1.93	9.4	12.0
27	47	30	1.241	0.740	0.244	0.147	2.19	1.67	9.9	4.3
28	40	37	1.315	0.852	0.275	0.178	2.16	1.90	10.5	4.3
29	21	18	0.245	0.551	0.057	0.122	0.52	1.31	2.2	4.1
30	28	49	1.072	1.619	0.203	0.335	1.88	2.71	11.7	23.4
31	52	35	0.555	0.905	0.143	0.228	1.55	1.90	7.0	11.2
32	60	79	1.464	1.334	0.307	0.255	2.16	2.44	6.3	11.0
33	50	18	0.934	1.005	0.194	0.206	2.20	2.04	4.9	8.5
34	41	58	0.941	1.065	0.203	0.238	1.85	2.29	3.6	2.5
35	33	32	0.778	0.704	0.152	0.156	1.27	1.49	7.3	7.4
36	75	56	1.622	0.901	0.335	0.186	3.00	1.35	12.2	22.7
37	21	22	0.738	0.611	0.166	0.141	1.48	1.36	6.8	3.6
38	51	51	1.732	1.554	0.378	0.326	3.53	2.89	17.6	5.6
39	23	37	0.688	0.727	0.480	0.144	1.79	1.56	2.8	4.6
40	33	24	0.303	0.267	0.060	0.059	0.84	0.77	3.3	2.4
41	35	21	0.291	0.398	0.066	0.072	0.95	0.90	6.6	6.6

Appendix Table 1 (cont'd)

Parent/ RILs lines	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry weight/plant (g)		ARA (µmole C_2H_4/pl/hr)	
	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2
42	63	47	0.912	0.797	0.191	0.146	1.94	1.65	9.7	9.7
43	37	30	0.778	0.665	0.165	0.154	1.51	1.53	7.8	4.8
44	23	34	0.431	0.802	0.081	0.165	0.90	1.52	.	.
45	35	56	0.950	0.823	0.223	0.173	1.85	1.87	13.4	7.2
46	27	28	0.613	0.815	0.127	0.183	1.80	1.84	8.1	8.1
47	26	17	0.536	0.949	0.120	0.202	1.29	1.91	7.8	10.6
48	77	65	1.409	1.306	0.294	0.265	2.25	2.03	12.6	12.6
49	57	61	0.646	1.357	0.134	0.226	1.84	2.71	2.7	11.0
50	32	41	0.643	0.662	0.148	0.184	1.50	1.92	3.6	5.3
51	97	98	2.355	2.554	0.512	0.589	4.78	4.91	28.0	30.8
52	84	72	2.607	2.039	0.483	0.406	3.24	3.07	17.8	13.8
53	75	83	1.258	1.835	0.264	0.375	2.54	3.02	9.6	14.6
54	49	43	0.695	0.964	0.134	0.194	1.70	1.18	6.3	5.5
55	57	57	1.255	1.375	0.236	0.269	2.56	2.38	7.2	16.0
56	49	42	0.919	0.919	0.156	0.173	1.57	1.37	8.9	8.5
57	35	30	1.351	0.975	0.294	0.205	2.46	1.93	18.0	13.6
58	39	24	0.725	0.217	0.154	0.053	1.43	0.70	8.2	3.0
59	43	48	0.996	1.210	0.205	0.240	1.67	2.10	11.5	11.8
60	42	22	0.950	0.390	0.202	0.087	1.94	1.03	6.3	4.6
61	19	34	0.438	0.573	0.098	0.116	0.99	1.21	4.9	4.9
62	43	56	0.747	1.041	0.135	0.241	1.32	2.19	7.0	7.0
63	35	39	0.446	0.545	0.108	0.117	1.03	1.05	3.5	4.1

Appendix Table 1 (cont'd)

Parent/ RILs lines	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry weight/plant (g)		ARA (µmole C_2H_4/pl/hr)	
	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2
64	31	42	0.425	0.536	0.081	0.123	1.29	1.56	5.4	5.9
65	42	41	1.024	0.699	0.201	0.158	1.57	0.44	14.2	4.0
66	72	52	1.771	0.781	0.391	0.164	3.18	1.51	14.8	14.8
67	60	55	1.261	1.224	0.336	0.261	2.46	2.37	3.9	11.5
68	44	32	0.948	0.901	0.216	0.199	1.80	2.01	7.9	6.7
69	69	53	1.742	1.470	0.370	0.338	2.60	2.47	19.1	9.0
70	29	46	0.879	0.836	0.198	0.204	1.62	1.84	11.6	10.7
71	48	46	0.754	0.716	0.156	0.180	1.51	1.62	5.2	5.3
72	34	33	0.680	0.976	0.138	0.197	1.69	2.03	8.7	8.5
73	39	45	0.762	0.993	0.171	0.214	1.50	1.88	4.9	7.5
74	29	32	0.404	0.779	0.081	0.150	0.93	1.42	5.5	4.7
75	26	46	0.771	1.592	0.163	0.329	1.26	2.64	6.6	14.1
76	94	73	1.877	1.830	0.442	0.430	4.32	3.29	7.6	13.0
77	42	34	0.624	0.554	0.132	0.119	1.47	1.35	3.6	6.1
78	56	62	1.792	1.433	0.348	0.313	2.76	2.74	24.5	17.3
79	38	43	0.899	0.584	0.199	0.122	2.00	1.33	10.7	5.7
80	49	37	0.961	0.597	0.228	0.112	2.04	1.18	14.0	10.6
81	82	74	1.623	1.600	0.338	0.241	2.19	1.93	9.0	6.8
82	37	53	0.486	0.637	0.104	0.161	0.90	1.74	3.7	7.9
83	50	29	1.171	0.712	0.232	0.128	2.01	1.28	10.1	26.2
84	32	45	1.186	1.488	0.272	0.305	2.43	2.53	9.8	6.2
85	87	93	1.344	1.700	0.262	0.344	1.75	2.08	14.9	14.9

Appendix Table 1 (cont'd)

Parent/ RILs lines	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry weight/plant (g)		ARA (µmole C_2H_4/pl/hr)	
	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2
86	29	44	0.618	1.384	0.134	0.356	1.28	2.66	7.9	3.9
87	55	44	0.851	0.739	0.178	0.158	1.88	1.42	5.8	5.9
88	56	53	1.064	0.883	0.219	0.179	1.82	1.90	4.5	8.5
89	56	50	1.023	0.889	0.196	0.196	1.81	1.73	4.7	16.5
90	38	37	1.153	1.060	0.255	0.242	2.17	1.93	14.6	15.0
91	106	93	1.756	1.848	0.415	0.412	3.33	3.30	23.6	23.6
92	63	77	1.553	1.574	0.319	0.318	2.39	2.38	12.4	15.4
93	74	83	2.534	2.546	0.543	0.539	3.61	3.98	15.0	29.0
94	74	57	1.608	1.414	0.335	0.352	2.82	2.76	30.5	12.1
95	45	35	0.829	0.595	0.187	0.149	1.84	2.19	4.6	8.0
96	48	67	1.422	1.296	0.324	0.337	2.99	2.89	4.9	9.2
97	61	47	1.063	0.880	0.240	0.201	1.86	1.86	13.2	5.5
98	49	55	0.981	0.820	0.221	0.171	2.05	1.71	7.9	7.8
99	41	40	0.774	0.798	0.175	0.197	1.64	2.08	12.6	12.3
100	37	40	0.821	0.860	0.168	0.173	1.72	1.66	15.0	15.9
101	50	51	0.999	0.873	0.204	0.168	2.01	1.61	7.0	9.8
102	45	32	0.676	0.563	0.146	0.131	1.37	1.58	4.8	5.3
103	73	84	1.173	1.621	0.206	0.309	2.03	2.64	4.6	17.6
104	33	56	1.266	0.930	0.262	0.186	2.53	1.66	6.4	6.4
105	44	40	0.411	0.508	0.116	0.101	1.25	1.00	4.2	8.8
106	31	34	0.720	0.676	0.153	0.200	1.49	1.53	13.4	7.0
107	43	40	0.797	0.866	0.187	0.940	1.76	1.81	14.2	13.9

Appendix Table 1 (cont'd)

Parent/ RILs lines	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry weight/plant (g)		ARA (µmole C_2H_4/pl/hr)	
	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2
108	45	59	0.946	0.949	0.181	0.192	2.20	2.06	7.1	12.7
109	40	42	0.463	0.617	0.099	0.147	1.15	1.52	5.0	9.1
110	30	46	0.796	0.986	0.169	0.193	1.55	1.85	13.5	13.8
111	76	59	0.840	0.854	0.190	0.200	1.75	1.93	13.6	13.4
112	49	47	1.034	0.746	0.221	0.152	2.39	1.83	2.7	15.9
113	65	71	1.804	1.907	0.383	0.387	3.18	3.10	16.4	11.9
114	43	58	1.005	1.414	0.216	0.302	2.11	2.79	6.8	14.3
115	39	46	0.717	0.880	0.137	0.186	1.57	1.87	14.2	13.7
116	45	37	0.590	0.525	0.122	0.114	1.47	1.49	1.9	2.9
117	18	41	0.303	0.355	0.066	0.066	0.99	1.10	1.0	1.4
118	39	46	0.627	0.948	0.162	0.221	2.24	2.37	11.2	17.5
119	73	58	1.256	1.280	0.280	0.260	2.53	2.38	12.5	13.1
120	55	39	1.025	1.126	0.191	0.230	2.18	2.08	12.1	11.1
121	65	80	0.870	0.690	0.221	0.183	2.06	1.47	2.3	8.3
122	42	46	0.574	0.800	0.121	0.150	1.37	1.65	10.5	8.1
123	55	64	0.851	1.159	0.165	0.232	1.76	2.31	13.5	7.5
124	44	46	1.095	1.098	0.209	0.222	2.04	2.07	18.6	15.4
125	48	45	1.155	0.812	0.228	0.164	1.82	1.35	9.6	10.2
126	108	117	2.085	2.393	0.383	0.419	3.28	3.29	19.2	11.7
127	62	50	0.919	0.784	0.214	0.175	2.03	1.37	14.0	0.8
128	67	88	1.206	1.173	0.265	0.242	2.11	2.24	8.0	3.9
129	46	65	0.474	1.410	0.103	0.283	1.02	2.01	5.3	14.1

Appendix Table 1 (cont'd)

Parent/ RILs lines	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry weight/plant (g)		ARA (µmole C_2H_4/pl/hr)	
	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2	Rep. 1	Rep. 2
130	81	85	0.859	1.875	0.207	0.351	1.71	3.44	14.3	18.3
131	59	67	0.733	0.844	0.166	0.229	1.84	2.59	9.6	18.2
132	55	37	1.191	0.589	0.222	0.107	1.95	0.97	13.8	4.0
133	41	48	1.252	0.983	0.291	0.224	2.38	1.88	13.8	14.8
134	70	50	1.031	1.185	0.232	0.262	1.81	2.20	11.3	12.4
135	72	91	1.209	1.505	0.280	0.294	1.96	2.16	10.9	22.9
136	65	51	0.191	1.205	0.223	0.228	2.22	2.19	10.4	11.6

Appendix Table 2 Association between SSR markers and N_2 fixation components evaluated by single-factor analysis (SF-ANOVA).

SSR Marker	LG	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry Weight/plant (g)		ARA (µmole C_2H_4/pl/hr)	
		P	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*	R^2
Sat 001	D2	0.4708	0.4009	0.8871	0.0156	0.9082	0.0103	0.8181	0.0408	0.3624	0.6434
Sat_022	D2	0.8684	0.0215	0.454	0.4388	0.5582	0.2685	0.6064	0.208	0.6572	0.1556
Sat_033	N	0.7805	0.0604	0.6356	0.1746	0.4616	0.4209	0.7336	0.0901	0.8848	0.0163
Sat_038	O	<.0001	18.5092	<.0001	16.386	<.0001	10.7215	<.0001	12.6874	0.0023	6.7592
Sat_039	F	0.5007	0.3469	0.767	0.0672	0.3847	0.5773	0.8535	0.0261	0.6401	0.1686
Sat_088	G	0.4637	0.4014	0.6853	0.1229	0.2577	0.9549	0.4911	0.3545	0.6341	0.1709
Sat_099	L	0.8941	0.0136	0.7119	0.1044	0.5328	0.2977	0.2498	1.0097	0.3635	0.6356
Sat_105	I	0.3638	0.6158	0.2886	0.8402	0.1459	1.5713	0.0085	5.0532	0.1492	1.5575
Sat_107	E	0.5054	0.3419	0.4203	0.5002	0.3241	0.748	0.1014	2.0513	0.6597	0.1508
Sat_108	O	<.0001	16.2472	<.0001	13.4312	<.0001	10.9731	<.0001	11.7594	0.0436	3.0262
Sat_109	O	<.0001	14.8695	<.0001	12.7436	0.0001	10.6741	<.0001	11.7586	0.0449	3.0585
Sat_118	H	0.0753	2.3424	0.9112	0.0093	0.9999	0.0000	0.9264	0.0064	0.9019	0.0115
Sat_120	F	0.795	0.0533	0.7249	0.0978	0.6452	0.1674	0.7965	0.0526	0.5551	0.2749
Sat_126	K	0.0387	3.1986	0.3651	0.6219	0.4403	0.4519	0.2255	1.1111	0.7343	0.0882
Sat_127	H	0.4185	0.5003	0.5212	0.3148	0.5778	0.2371	0.369	0.6166	0.5562	0.267
Sat_190	O	0.0008	8.1494	0.0019	7.0018	0.0034	6.2113	0.0026	6.5534	0.0941	2.0934
Sat_274	O	0.0001	10.3627	<.0001	19.6529	<.0001	13.2672	<.0001	14.7511	<.0001	11.9238
Sat_307	O	0.0003	9.6425	0.0013	7.7727	0.008	5.3256	0.0042	6.1755	0.1321	1.7628
Sat_318	O	0.2944	0.8200	0.5352	0.2877	0.7691	0.0645	0.8554	0.0249	0.1478	1.5689
Sat_324	G	0.6183	0.1901	0.1502	1.5738	0.0605	2.6648	0.0426	3.1013	0.6974	0.1167
Satt002	D2	0.9205	0.0078	0.767	0.0688	0.8188	0.0411	0.2075	1.2384	0.5205	0.3259
Satt009	N	0.4045	0.5311	0.2257	1.1185	0.1539	1.5461	0.2706	0.9256	0.0583	2.7308
Satt014	D2	0.8996	0.0131	0.5258	0.3307	0.4351	0.5001	0.1441	1.7407	0.6691	0.1515

Appendix Table 2 (cont'd)

SSR Marker	LG	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry Weight/plant (g)		ARA (μmole C_2H_4/pl/hr)	
		P	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*	R^2
Satt022	N	0.4334	0.4917	0.3448	0.7144	0.408	0.5484	0.3987	0.5705	0.202	1.3094
Satt038	G	0.8077	0.0461	0.1911	1.3214	0.0375	3.311	0.3353	0.7199	0.2464	1.0483
Satt041	D1b+W	0.0462	3.0213	0.0829	2.2948	0.3744	0.6075	0.2217	1.1463	0.0324	3.4999
Satt042	A1	0.4758	0.3802	0.8469	0.0279	0.7852	0.0556	0.5671	0.245	0.9996	0.0000
Satt063	B2	0.9711	0.001	0.4439	0.4551	0.7952	0.0524	0.6631	0.1476	0.8987	0.0127
Satt066	B2	0.1664	1.4677	0.1418	1.6531	0.8282	0.0364	0.2368	1.0749	0.0973	2.1162
Satt100	C2	0.1379	1.7116	0.807	0.0468	0.9632	0.0017	0.7143	0.105	0.5463	0.2873
Satt114	F	0.3605	0.6695	0.5602	0.2722	0.4551	0.4471	0.4191	0.523	0.8816	0.0179
Satt117	E	0.8932	0.0137	0.4097	0.5155	0.5001	0.3452	0.4783	0.3816	0.9561	0.0023
Satt123	O	0.0381	3.4234	0.0064	5.8441	0.0376	3.4389	0.0376	3.4386	0.059	2.8675
Satt126	B2	0.2163	1.2205	0.2016	1.3015	0.3524	0.6923	0.1455	1.6873	0.358	0.6818
Satt135	D2	0.6748	0.1424	0.1842	1.4178	0.1268	1.8703	0.0765	2.5082	0.5829	0.2458
Satt137	K	0.596	0.222	0.94	0.0045	0.5857	0.2346	0.9556	0.0025	0.7761	0.0644
Satt138	G	0.4146	0.5011	0.1586	1.4888	0.1944	1.263	0.2687	0.9191	0.6728	0.1356
Satt 141	D1b+W	0.494	0.3605	0.3172	0.7694	0.5477	0.2788	0.3327	0.722	0.1007	2.0751
Satt143	L	0.0561	2.8666	0.1794	1.4257	0.3621	0.6596	0.2143	1.2211	0.3534	0.6893
Satt146	F	0.0349	3.4584	0.7142	0.106	0.7337	0.0915	0.7738	0.0653	0.2031	1.2826
Satt147	D1a+Q	0.2763	0.911	0.245	1.0382	0.4074	0.5286	0.3063	0.8049	0.0768	2.4072
Satt148	I	0.0612	2.6306	0.5534	0.2668	0.9124	0.0092	0.2496	1.003	0.2251	1.1214
Satt152	N	0.2377	1.0963	0.1734	1.4545	0.1109	1.9887	0.2443	1.0659	0.1335	1.7779
Satt153	O	<.0001	15.6405	<.0001	14.7344	0.0004	9.2412	<.0001	11.4849	0.0135	4.5732
Satt154	D2	0.8982	0.0124	0.8761	0.0185	0.6301	0.1762	0.3843	0.574	0.4734	0.3932
Satt155	A1	0.7806	0.0628	0.7524	0.0805	0.3828	0.6148	0.5534	0.284	0.9149	0.0093

Appendix Table 2 (cont'd)

SSR Marker	LG	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry Weight/plant (g)		ARA (μmole C_2H_4/pl/hr)	
		P	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*	R^2
Satt156	L	0.5307	0.3127	0.9404	0.0045	0.9382	0.0048	0.6156	0.2007	0.166	1.5295
Satt157	D1b+W	0.0204	4.0077	0.0829	2.261	0.1561	1.5177	0.1027	2.0048	0.0147	4.4583
Satt160	F	0.9963	0	0.8547	0.0259	0.5951	0.2178	0.8046	0.0473	0.9473	0.0034
Satt167	K	0.6676	0.1401	0.6503	0.1562	0.4439	0.4449	0.8195	0.0396	0.7451	0.0809
Satt172	D1b+W	0.1841	1.385	0.8974	0.0131	0.3258	0.7604	0.9176	0.0085	0.1836	1.3989
Satt175	M	0.3786	0.6014	0.1699	1.455	0.556	0.2695	0.1738	1.429	0.1837	1.3765
Satt180	C1	0.0126	4.6544	0.6955	0.1173	0.6702	0.1389	0.6117	0.1973	0.7313	0.091
Satt184	D1a+Q	0.8102	0.0442	0.9731	0.0009	0.5437	0.2821	0.8082	0.0451	0.449	0.4416
Satt196	K	0.0642	2.5513	0.4056	0.5207	0.1008	2.0123	0.119	1.8179	0.2228	1.1239
Satt215	J	0.324	0.8175	0.1581	1.6674	0.3576	0.7116	0.0205	4.4318	0.9651	0.0016
Satt216	D1b+W	0.2682	0.9494	0.7056	0.111	0.7554	0.0755	0.4549	0.4335	0.184	1.375
Satt229	L	0.8581	0.0249	0.7391	0.0863	0.7544	0.0761	0.9994	0	0.6493	0.162
Satt231	E	0.2908	0.8996	0.9157	0.0091	0.3065	0.8432	0.7946	0.0549	0.1217	1.9359
Satt232	L	0.4453	0.4744	0.6599	0.158	0.5207	0.3361	0.514	0.347	0.9376	0.005
Satt243	O	0.5959	0.2153	0.2761	0.9047	0.4678	0.4031	0.4757	0.389	0.9044	0.0111
Satt244	J	0.5442	0.2814	0.6988	0.1146	0.1364	1.685	0.2847	0.8733	0.789	0.0553
Satt245	M	0.9053	0.0112	0.5224	0.3229	0.9215	0.0077	0.7823	0.0604	0.7532	0.0781
Satt249	J	0.2838	0.8764	0.3765	0.5975	0.9889	0.0001	0.7406	0.0839	0.7802	0.0601
Satt251	B1	0.8469	84.69	0.469	0.395	0.1477	1.5695	0.3887	0.5592	0.5745	0.2394
Satt253	H	0.215	1.1626	0.4217	0.4897	0.5088	0.3313	0.3747	0.5975	0.6619	0.1465
Satt260	K	0.2338	1.0804	0.2572	0.9788	0.0279	3.6355	0.1728	1.4143	0.3022	0.8124
Satt268	E	0.4843	0.3742	0.1915	1.2986	0.3367	0.7047	0.3406	0.6936	0.5833	0.2321
Satt271	D1b+W	0.1762	1.5317	0.4028	0.589	0.7609	0.0781	0.6943	0.1302	0.2978	0.9107

Appendix Table 2 (cont'd)

SSR Marker	LG	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry Weight/plant (g)		ARA (μmole C_2H_4/pl/hr)	
		P	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*	R^2
Satt279	H	0.1831	1.3288	0.3314	0.7096	0.4159	0.4982	0.312	0.7685	0.7401	0.0837
Satt285	J	0.8669	0.0224	0.8689	0.0217	0.31	0.8178	0.3326	0.7451	0.7185	0.1043
Satt288	G	0.7561	0.0769	0.4551	0.4436	0.9045	0.0115	0.5824	0.2406	0.8056	0.0487
Satt292	I	0.0838	2.2643	0.2834	0.8777	0.4016	0.5376	0.294	0.8403	0.1548	1.5508
Satt293	H	0.3985	0.549	0.7898	0.0549	0.8495	0.0278	0.4706	0.4012	0.3359	0.7127
Satt294	C1	0.0196	4.1474	0.2009	1.2647	0.8374	0.0328	0.1625	1.5067	0.8761	0.0191
Satt306	M	0.3251	0.8069	0.5694	0.2705	0.8096	0.0486	0.4927	0.393	0.7274	0.1025
Satt314	H	0.0391	3.4136	0.4205	0.5283	0.3381	0.7463	0.4202	0.5289	0.6859	0.1345
Satt327	A2	0.0832	2.3269	0.9107	0.0099	0.9131	0.0093	0.7451	0.0829	0.2156	1.2048
Satt331	O	0.0008	8.378	<.0001	12.6977	0.0006	8.7089	0.0005	8.9942	0.0012	7.9147
Satt334	F	0.4267	0.5062	0.537	0.3056	0.4639	0.4299	0.4781	0.4033	0.7884	0.0579
Satt342	D1a+Q	0.7363	0.0855	0.5859	0.2238	0.0957	2.0731	0.2016	1.2234	0.8968	0.0128
Satt353	H	0.9081	0.0103	0.8683	0.0212	0.816	0.0418	0.8647	0.0224	0.9616	0.0018
Satt354	I	0.3678	0.6391	0.3404	0.7159	0.0511	2.9619	0.0104	5.0555	0.2484	1.056
Satt369	E	0.3766	0.6309	0.1882	1.3924	0.7329	0.0943	0.1104	2.0428	0.9815	0.0004
Satt371	C2	0.2575	1.0007	0.5007	0.3551	0.9584	0.0021	0.7365	0.0887	0.6618	0.1512
Satt373	L	0.7423	0.0847	0.6548	0.1567	0.6859	0.1282	0.7215	0.0996	0.9664	0.0014
Satt374	F	0.0721	2.43	0.0989	2.0491	0.0719	2.4336	0.0737	2.4034	0.6464	0.1611
Satt380	J	0.0888	2.3751	0.0388	3.4803	0.148	1.7219	0.0058	6.1194	0.3219	0.8108
Satt385	A1	0.0186	4.0969	0.0171	4.2005	0.2043	1.2089	0.1698	1.4124	0.3667	0.6178
Satt388	L	0.0175	4.7723	0.0219	4.4491	0.1512	1.768	0.0339	3.821	0.0239	4.3225
Satt389	D2	0.1222	1.8712	0.1426	1.6851	0.373	0.6253	0.6853	0.1298	0.2683	0.9715
Satt390	B1	0.9304	0.0061	0.7183	0.1036	0.5724	0.2536	0.5468	0.2889	0.6268	0.1882

Appendix Table 2 (cont'd)

SSR Marker	LG	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry Weight/plant (g)		ARA (μmole C_2H_4/pl/hr)	
		P	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*	R^2
Satt398	L	0.0491	3.0612	0.1672	1.5206	0.1063	2.0732	0.1886	1.3786	0.1939	1.3576
Satt402	D1a+Q	0.664	0.1456	0.3767	0.6015	0.2712	0.9306	0.4418	0.4558	0.6817	0.1308
Satt405	J	0.392	0.5599	0.4824	0.3774	0.9666	0.0013	0.6119	0.1971	0.9835	0.0003
Satt409	A2	0.3219	0.7492	0.6952	0.1176	0.3023	0.8119	0.4733	0.3933	0.2814	0.8919
Satt411	E	0.9551	0.0025	0.6658	0.1439	0.7167	0.1016	0.7272	0.0939	0.6633	0.1463
Satt414	J	0.0689	2.7997	0.0211	4.4613	0.0254	4.2006	0.0019	7.9816	0.0665	2.8741
Satt415	B1	0.6041	0.2012	0.3549	0.6391	0.8214	0.0381	0.6887	0.1202	0.7317	0.0887
Satt416	B2	0.5029	0.3433	0.4297	0.4767	0.4034	0.5335	0.2049	1.2237	0.2662	0.9502
Satt417	K	0.5632	0.2577	0.6662	0.1436	0.446	0.4475	0.8847	0.0163	0.6745	0.1361
Satt425	F	0.2507	1.0216	0.2328	1.1018	0.203	1.2532	0.1984	1.2789	0.2681	0.9574
Satt426	B1	0.6404	0.1671	0.8923	0.0141	0.1552	1.5363	0.5572	0.2637	0.8826	0.0168
Satt429	A2	0.7665	0.0702	0.5464	0.2895	0.3476	0.7003	0.9009	0.0123	0.6734	0.1426
Satt434	H	0.0173	4.1576	0.9404	0.0042	0.73	0.0892	0.8279	0.0354	0.3112	0.771
Satt440	I	0.0075	5.4915	0.1318	1.7797	0.3307	0.745	0.0293	3.6835	0.2969	0.8632
Satt442	H	0.2324	1.095	0.6269	0.1823	0.4338	0.4719	0.5947	0.2184	0.1685	1.4647
Satt443	D2	0.3964	0.5539	0.4134	0.5152	0.8994	0.0123	0.501	0.349	0.5251	0.3138
Satt444	B1	0.1459	1.6447	0.1868	1.3575	0.2933	0.8626	0.268	0.9576	0.0532	2.9109
Satt453	B1	0.3372	0.7088	0.0669	2.559	0.5046	0.3432	0.1364	1.6981	0.3152	0.782
Satt460	C2	0.0724	2.4423	0.945	0.0037	0.9268	0.0065	0.8156	0.0417	0.8187	0.0406
Satt472	G	0.6081	0.2077	0.3556	0.6724	0.7854	0.0586	0.2751	0.9374	0.8719	0.0205
Satt486	D2	0.3212	0.8199	0.2213	1.2439	0.1678	1.5796	0.0688	2.7332	0.2012	1.3693
Satt496	I	0.1463	1.5917	0.1707	1.4169	0.0191	4.092	0.0127	4.6126	0.2364	1.0684
Satt500		0.3035	0.8205	0.2644	0.9645	0.2018	1.26	0.157	1.5466	0.1689	1.4732

Appendix Table 2 (cont'd)

SSR Marker	LG	Nodule no./plant		Nodule fresh weight/plant (g)		Nodule dry weight/plant (g)		Plant dry Weight/plant (g)		ARA (μmole C_2H_4/pl/hr)	
		P	R^2	*P*	R^2	*P*	R^2	*P*	R^2	*P*	R^2
Satt510	F	0.4442	0.4475	0.5524	0.2702	0.9228	0.0072	0.3947	0.5536	0.4475	0.4445
Satt516	F	0.6091	0.1971	0.402	0.5286	0.5736	0.2387	0.6591	0.1468	0.9506	0.0029
Satt529	J	0.0341	3.5694	0.0172	4.4882	0.0762	2.5132	0.0016	7.7294	0.2394	1.1237
Satt530	N	0.4214	0.4867	0.2481	1.0015	0.1814	1.3392	0.3144	0.761	0.1663	1.4467
Satt534	B2	0.3426	0.6774	0.0672	2.4971	0.4694	0.3942	0.3244	0.7302	0.0692	2.4804
Satt541	H	0.537	0.2984	0.5326	0.3049	0.6762	0.1368	0.3575	0.6617	0.5225	0.3228
Satt545	A1	0.0938	2.1783	0.0184	4.2686	0.1995	1.2827	0.1533	1.5865	0.1898	1.3497
Satt546	D1b+W	0.0012	7.6791	0.4975	0.352	0.7019	0.1122	0.2539	0.9924	0.0069	5.4867
Satt551	M	0.6627	0.1446	0.8019	0.0479	0.6372	0.169	0.8172	0.0406	0.9283	0.0062
Satt554	F	0.2025	1.2861	0.9314	0.0059	0.7551	0.0775	0.9558	0.0024	0.7532	0.0793
Satt567	M	0.2111	1.2896	0.734	0.0958	0.8347	0.0361	0.8968	0.014	0.5484	0.301
Satt581	O	<.0001	11.7743	<.0001	17.8439	<.0001	13.6537	<.0001	14.0225	0.0002	10.1691
Satt590	M	0.2029	1.3045	0.3791	0.6245	0.304	0.852	0.5091	0.3523	0.2143	1.2512
Satt596	J	0.0983	2.2986	0.0529	3.1384	0.0759	2.6456	0.0053	6.3904	0.2759	1.0053
Sct_001	J	0.0274	3.7747	0.009	5.2554	0.0409	3.2504	0.0006	8.7852	0.2759	0.9415
Sct_033	F	0.6959	0.1179	0.6225	0.1869	0.3794	0.5949	0.6028	0.2089	0.6533	0.1557

Appendix Table 3 Contingency Chi-square test for marker segregation againsts the 1 : 1 Mendelian ratio of 131 markers in 136 derived from the soybean cross SJ2 x Suwon157.

Marker	A	B	Total	AB	E	$(O_A-E_A)^2/E_A$	$(O_B-E_B)^2/E_B$	Total
Sat 001	59	73	132	4	66.0	0.74	0.74	1.48
Sat_022	55	75	130	6	65.0	1.54	1.54	3.08
Sat_033	66	65	131	5	65.5	0.00	0.00	0.01
Sat_038	76	60	136	0	68.0	0.94	0.94	1.88
Sat_039	69	64	133	3	66.5	0.09	0.09	0.19
Sat_088	74	62	136	0	68.0	0.53	0.53	1.06
Sat_099	57	76	133	3	66.5	1.36	1.36	2.71
Sat_105	71	65	136	0	68.0	0.13	0.13	0.26
Sat_107	59	73	132	4	66.0	0.74	0.74	1.48
Sat_108	69	67	136	0	68.0	0.01	0.01	0.03
Sat_109	67	66	133	3	66.5	0.00	0.00	0.01
Sat_118	66	70	136	0	68.0	0.06	0.06	0.12
Sat_120	71	58	129	7	64.5	0.66	0.66	1.31
Sat_126	58	76	134	2	67.0	1.21	1.21	2.42
Sat_127	68	65	133	3	66.5	0.03	0.03	0.07
Sat_190	69	67	136	0	68.0	0.01	0.01	0.03
Sat_274	60	76	136	0	68.0	0.94	0.94	1.88
Sat_307	82	49	131	5	65.5	4.16	4.16	8.31*
Sat_318	63	73	136	0	68.0	0.37	0.37	0.74
Sat_324	51	82	133	3	66.5	3.61	3.61	7.23
Satt002	65	65	130	6	65.0	0.00	0.00	0.00
Satt009	77	56	133	3	66.5	1.66	1.66	3.32
Satt014	56	68	124	12	62.0	0.58	0.58	1.16
Satt022	56	71	127	9	63.5	0.89	0.89	1.77

Appendix Table 3 (cont'd)

Marker	A	B	Total	AB	E	$(O_A-E_A)^2/E_A$	$(O_B-E_B)^2/E_B$	Total
Satt038	56	75	131	5	65.5	1.38	1.38	2.76
Satt041	56	76	132	4	66.0	1.52	1.52	3.03
Satt042	76	60	136	0	68.0	0.94	0.94	1.88
Satt063	65	66	131	5	65.5	0.00	0.00	0.01
Satt066	62	70	132	4	66.0	0.24	0.24	0.48
Satt100	63	67	130	6	65.0	0.06	0.06	0.12
Satt114	62	65	127	9	63.5	0.04	0.04	0.07
Satt117	60	74	134	2	67.0	0.73	0.73	1.46
Satt123	61	65	126	10	63.0	0.06	0.06	0.13
Satt126	58	69	127	9	63.5	0.48	0.48	0.95
Satt135	68	58	126	10	63.0	0.40	0.40	0.79
Satt137	58	71	129	7	64.5	0.66	0.66	1.31
Satt138	69	66	135	1	67.5	0.03	0.03	0.07
Satt141	61	71	132	4	66.0	0.38	0.38	0.76
Satt143	63	65	128	8	64.0	0.02	0.02	0.03
Satt146	68	61	129	7	64.5	0.19	0.19	0.38
Satt147	60	72	132	4	66.0	0.55	0.55	1.09
Satt148	63	71	134	2	67.0	0.24	0.24	0.48
Satt152	69	60	129	7	64.5	0.31	0.31	0.63
Satt153	72	62	134	2	67.0	0.37	0.37	0.75
Satt154	57	77	134	2	67.0	1.49	1.49	2.99
Satt155	77	49	126	10	63.0	3.11	3.11	6.22*
Satt156	56	72	128	8	64.0	1.00	1.00	2.00
Satt157	64	70	134	2	67.0	0.13	0.13	0.27
Satt160	64	68	132	4	66.0	0.06	0.06	0.12
Satt167	58	76	134	2	67.0	1.21	1.21	2.42

Appendix Table 3 (cont'd)

Marker	A	B	Total	AB	E	$(O_A-E_A)^2/E_A$	$(O_B-E_B)^2/E_B$	Total
Satt172	63	66	129	7	64.5	0.03	0.03	0.07
Satt175	74	57	131	5	65.5	1.10	1.10	2.21
Satt180	72	61	133	3	66.5	0.45	0.45	0.91
Satt184	63	70	133	3	66.5	0.18	0.18	0.37
Satt196	63	72	135	1	67.5	0.30	0.30	0.60
Satt215	56	65	121	15	60.5	0.33	0.33	0.67
Satt216	57	74	131	5	65.5	1.10	1.10	2.21
Satt229	60	71	131	5	65.5	0.46	0.46	0.92
Satt231	61	65	126	10	63.0	0.06	0.06	0.13
Satt232	47	78	125	11	62.5	3.84	3.84	7.69
Satt243	69	64	133	3	66.5	0.09	0.09	0.19
Satt244	74	59	133	3	66.5	0.85	0.85	1.69
Satt245	65	64	129	7	64.5	0.00	0.00	0.01
Satt249	76	57	133	3	66.5	1.36	1.36	2.71
Satt251	63	72	135	1	67.5	0.30	0.30	0.60
Satt253	64	70	134	2	67.0	0.13	0.13	0.27
Satt260	73	60	133	3	66.5	0.64	0.64	1.27
Satt268	68	65	133	3	66.5	0.03	0.03	0.07
Satt271	61	60	121	15	60.5	0.00	0.00	0.01
Satt279	65	70	135	1	67.5	0.09	0.09	0.19
Satt285	65	63	128	8	64.0	0.02	0.02	0.03
Satt288	68	60	128	8	64.0	0.25	0.25	0.50
Satt292	71	62	133	3	66.5	0.30	0.30	0.61
Satt293	62	70	132	4	66.0	0.24	0.24	0.48
Satt294	74	57	131	5	65.5	1.10	1.10	2.21
Satt306	63	59	122	14	61.0	0.07	0.07	0.13

Appendix Table 3 (cont'd)

Marker	A	B	Total	AB	E	$(O_A-E_A)^2/E_A$	$(O_B-E_B)^2/E_B$	Total
Satt314	52	73	125	11	62.5	1.76	1.76	3.53
Satt327	61	69	130	6	65.0	0.25	0.25	0.49
Satt331	61	70	131	5	65.5	0.31	0.31	0.62
Satt334	62	65	127	9	63.5	0.04	0.04	0.07
Satt342	76	59	135	1	67.5	1.07	1.07	2.14
Satt353	74	58	132	4	66.0	0.97	0.97	1.94
Satt354	70	59	129	7	64.5	0.47	0.47	0.94
Satt369	72	54	126	10	63.0	1.29	1.29	2.57
Satt371	63	67	130	6	65.0	0.06	0.06	0.12
Satt373	61	69	130	6	65.0	0.25	0.25	0.49
Satt374	58	76	134	2	67.0	1.21	1.21	2.42
Satt380	63	60	123	13	61.5	0.04	0.04	0.07
Satt385	77	58	135	1	67.5	1.34	1.34	2.67
Satt388	56	62	118	18	59.0	0.15	0.15	0.31
Satt389	62	67	129	7	64.5	0.10	0.10	0.19
Satt390	38	90	128	8	64.0	10.56	10.56	21.13*
Satt398	64	63	127	9	63.5	0.00	0.00	0.01
Satt402	69	63	132	4	66.0	0.14	0.14	0.27
Satt405	72	61	133	3	66.5	0.45	0.45	0.91
Satt409	71	62	133	3	66.5	0.30	0.30	0.61
Satt411	61	71	132	4	66.0	0.38	0.38	0.76
Satt414	61	58	119	17	59.5	0.04	0.04	0.08
Satt415	47	89	136	0	68.0	6.49	6.49	12.97*
Satt416	42	91	133	3	66.5	9.03	9.03	18.05*
Satt417	59	73	132	4	66.0	0.74	0.74	1.48
Satt425	59	72	131	5	65.5	0.65	0.65	1.29

Appendix Table 3 (cont'd)

Marker	A	B	Total	AB	E	$(O_A-E_A)^2/E_A$	$(O_B-E_B)^2/E_B$	Total
Satt426	63	70	133	3	66.5	0.18	0.18	0.37
Satt429	62	66	128	8	64.0	0.06	0.06	0.13
Satt434	23	113	136	0	68.0	29.78	29.78	59.56*
Satt440	59	70	129	7	64.5	0.47	0.47	0.94
Satt442	64	68	132	4	66.0	0.06	0.06	0.12
Satt443	60	72	132	4	66.0	0.55	0.55	1.09
Satt444	63	67	130	6	65.0	0.06	0.06	0.12
Satt453	65	67	132	4	66.0	0.02	0.02	0.03
Satt460	68	65	133	3	66.5	0.03	0.03	0.07
Satt472	63	66	129	7	64.5	0.03	0.03	0.07
Satt486	54	68	122	14	61.0	0.80	0.80	1.61
Satt496	74	60	134	2	67.0	0.73	0.73	1.46
Satt500	58	73	131	5	65.5	0.86	0.86	1.72
Satt510	59	74	133	3	66.5	0.85	0.85	1.69
Satt516	69	66	135	1	67.5	0.03	0.03	0.07
Satt529	64	62	126	10	63.0	0.02	0.02	0.03
Satt530	75	60	135	1	67.5	0.83	0.83	1.67
Satt534	58	77	135	1	67.5	1.34	1.34	2.67
Satt541	63	67	130	6	65.0	0.06	0.06	0.12
Satt545	72	58	130	6	65.0	0.75	0.75	1.51
Satt546	18	115	133	3	66.5	35.37	35.37	70.74*
Satt551	75	59	134	2	67.0	0.96	0.96	1.91
Satt554	72	59	131	5	65.5	0.65	0.65	1.29
Satt567	68	55	123	13	61.5	0.69	0.69	1.37
Satt581	60	72	132	4	66.0	0.55	0.55	1.09
Satt590	58	68	126	10	63.0	0.40	0.40	0.79

Appendix Table 3 (cont'd)

Marker	A	B	Total	AB	E	$(O_A-E_A)^2/E_A$	$(O_B-E_B)^2/E_B$	Total
Satt596	62	57	119	17	59.5	0.11	0.11	0.21
Sct_001	65	64	129	7	64.5	0.00	0.00	0.01
Sct_033	68	64	132	4	66.0	0.06	0.06	0.12

* Marker segregation significantly derived from the 1 : 1 Mendelian ratio at $P < 0.05$.