Theor Appl Genet (2005) 110: 1003–1010
DOI 10.1007/s00122-004-1887-2

ORIGINAL PAPER

M. Y. Kim · K. Van · P. Lestari
J.-K. Moon · S.-H. Lee

SNP identification and SNAP marker development for a *GmNARK* gene controlling supernodulation in soybean

Received: 16 October 2004 / Accepted: 18 November 2004 / Published online: 25 February 2005

Abstract Supernodulation in soybean (*Glycine max* L. Merr.) is an important source of nitrogen supply to subterranean ecological systems. Single nucleotide-amplified polymorphism (SNAP) markers for supernodulation should allow rapid screening of the trait in early growth stages, without the need for inoculation and phenotyping. The gene *GmNARK* (*Glycine max* nodule autoregulation receptor kinase), controlling autoregulation of nodulation, was found to have a single nucleotide polymorphism (SNP) between the wild-type cultivar Sinpaldalkong 2 and its supernodulating mutant, SS2-2. Transversion of A to T at the 959-bp position of the *GmNARK* sequence results in a change of lysine (AAG) to a stop codon (TAG), thus terminating its translation in SS2-2. Based on the identified SNP in *GmNARK*, five primer pairs specific to each allele were designed using the WebSnaper program to develop a SNAP marker for supernodulation. One A-specific primer pair produced a band present in only Sinpaldalkong 2, while two T-specific pairs showed a band in only SS2-2. Both complementary PCRs, using each allele-specific primer pair were performed to genotype supernodulation against F_2 progeny of Sinpaldalkong 2 × SS2-2. Among 28 individuals with the normal phenotype, eight individuals having only the A-allele-specific band were homozygous and normal, while 20 individuals were found to be heterozygous at the SNP having both A and T bands. Twelve supernodulating individuals showed only the band specific to the T allele. This SNAP marker for supernodulation could easily be analyzed through simple PCR and agarose gel electrophoresis. Therefore, use of this SNAP marker might be faster, cheaper, and more reproducible than using other genotyping methods, such as a cleaved amplified polymorphic sequence marker, which demand of restriction enzymes.

Keywords Supernodulation · Autoregulation of nodulation · Single nucleotide-amplified polymorphism · Single nucleotide polymorphism · Marker-assisted selection · *Glycine max* nodule autoregulation receptor kinase · Functional marker

Communicated by S.J. Knapp

M. Y. Kim · K. Van · P. Lestari · S.-H. Lee (✉)
Department of Plant Science, Seoul National University,
San 56-1, Shillim-dong, Kwanak-gu, Seoul,
151-921, The Republic of Korea
E-mail: sukhalee@snu.ac.kr
Tel.: +82-2-8804545
Fax: +82-2-8732056

J.-K. Moon
National Institute of Crop Science,
Rural Development Administration, Suwon,
441-770, The Republic of Korea

Introduction

Nodulation in soybean (*Glycine max* L. Merr.) is an important source of nitrogen supply to subterranean ecological systems. Development of nodules on soybean roots is triggered by invasion of *Bradyrhizobium* bacteria and regulated by factors both internal and external to the hosts. The specific mechanism during the serial process, directed by autoregulation, is known to control nodulation systemically (Pierce and Bauer 1983; Carroll et al. 1985a).

Many soybean mutants with deficient autoregulation have been created by chemical mutagenesis (Carroll et al. 1985a, 1985b; Akao and Kouchi 1992) or fast neutron mutagenesis (Men et al. 2002). The mutagen ethylmethane sulfonate (EMS) has been commonly applied to produce several super/hypernodulating mutants derived from the cultivar Bragg, including a class of nitrate-tolerant symbiotic (nts) mutants (Carroll et al. 1985a, 1985b). Supernodulating soybean mutants produce nodule numbers ten times greater than wild types, whereas hypernodulating mutants have two to five times more nodules than wild types, exhibiting nodulation in between supernodulation and wild-type nodulation.

Genetic studies and linkage mapping for these mutants revealed that the mutant locus *NTS-1* is located on linkage group (LG) H and contains separate alleles (Landau-Ellis et al. 1991; Kolchinsky et al. 1997). More recently, map-based cloning with markers pUTG-132 and UQC-IS1, which flank *NTS-1* on LG H, was performed to identify a gene conferring autoregulation of nodulation (Searle et al. 2003). Expression of this gene in shoots is responsible for control of nodule meristem proliferation, communicating distantly with nodules, and thus is called the *Glycine max* nodule autoregulation receptor-like protein kinase (*GmNARK*) gene. Nonsense and missense mutations comprising single-base changes on the gene were found to lead to supernodulation and hypernodulation phenotypes in the mutants, respectively (Searle et al. 2003).

Single nucleotide polymorphisms (SNPs), which can be defined as single-base changes or indels at a specific nucleotide position, have recently been developed into genetic markers. SNPs are the most abundant variations in the human genome, with the average frequency of one SNP per 1.9 kb (Sachidanandam et al. 2001). In soybean, a total of 280 SNPs were recently found in a 76-kb sequence from 25 different soybean genotypes (Zhu et al. 2003). Since SNPs are highly stable markers and often contribute directly to a phenotype (Kim et al. 2004), they can serve as a powerful tool for marker-assisted selection (MAS) and map-based cloning when combined with high-throughput genotyping systems.

SNP genotyping techniques are being rapidly improved. Various methods have been developed based on allele-specific hybridization, primer extension, oligonucleotide ligation, endonuclease cleavage, Invader, Taqman, or allelic-specific PCR (reviewed in Gupta et al. 2001; Gut 2001; Shi 2001; Syvänen 2001). Genotyping methods using fluorescence energy transfer have also been developed, requiring specialized detection equipment with a flow cytometric platform and modified fluorescence-labeled primers (Livak 1999; Cai et al. 2000; Chen et al. 2000; Taylor et al. 2001; Ye et al. 2001). These methods are more suitable to multiplexing with a large number of the markers for genetic mapping and diversity studies.

If a SNP marker linked to a certain trait is to be used in MAS, crop breeders should preferentially consider simple methods, basic requisites, low cost, and reproducibility in choosing a genotyping method to use. Among the many SNP genotyping methods, allele-specific PCR largely satisfies the above requirements. This technique is simply performed using allele-specific primers with each terminal 3′ nucleotide complementary to the SNP site. The products can be resolved on a standard agarose gel. Allele-specific PCR has not been widely used, because a single-base pair change at the 3′ end is often not sufficient to ensure reliable discrimination between the two SNP alleles. To increase specificity, a single nucleotide-amplified polymorphism (SNAP) marker was developed that uses a modified allele-specific primer with a mismatched base pair within four bases of the 3′ end in addition to the 3′-end base complementary to the SNP site (Drenkard et al. 2000; Hayashi et al. 2004).

We have already presented data showing that the mutant SS2-2 was isolated from M_2 families of cultivar Sinpaldalkong 2 mutagenized, using 30 mM EMS. SS2-2 showed no inhibition of nodule formation, even in the presence of exogenous nitrate (Lee et al. 1997, 1998). SS2-2 was designated as a hypernodulating mutant at first, based on phenotypic assay of root nodulation (Ha and Lee 2001).

In the present study, a SNP residing in the sequence of the *GmNARK* gene was identified between wild-type Sinpaldalkong 2 and its mutant counterpart, SS2-2. Type and site of mutation in SS2-2 could indicate whether it was a hypernodulating mutation or not. In addition, the identified SNP was developed into a SNAP marker to allow direct MAS for supernodulation at an early growth stage, without the need to inoculate and phenotype roots.

Materials and methods

Plant materials

DNA sequences from Sinpaldalkong 2 and its nodulating mutant variant, SS2-2, were compared to identify a SNP for the gene *GmNARK*. For development and evaluation of a SNAP marker for *GmNARK*, 40 F_2 progenies were obtained from the cross between Sinpaldalkong 2 and SS2-2. Also, six normal-nodulating genotypes, Danbaekkong, Taekwangkong, Pureunkong, Jinpumkong 2, PI 96188, and Jangyeobkong, and two super-/hypernodulating mutants, nts382 and nts1116, were used. Supernodulating mutant nts382 (Carroll et al. 1985b) and hypernodulating mutant nts1116 (Carroll et al. 1985a) were derived from Bragg. Bragg and the two mutants were obtained from Dr. P.M. Gresshoff of the University of Queensland, Australia.

Inoculation and phenotyping of root nodules

To evaluate root nodulation types from F_2 plants of SS2-2 × Sinpaldalkong 2, the plants were inoculated with *Bradyrhizobium japonicum* (USDA110). Roots of each plant were infected by evenly pouring an equal volume of bacteria suspension of about 10^8 cells per milliliter ($OD_{600} = 2.0$) into sterilized vermiculite. Root nodulation type was recorded as normal or supernodulating, based on the number of nodules on roots at least 2 weeks after inoculation. Supernodulation type was called if the number of nodules was tenfold more than normal.

Sequencing and SNP discovery

Genomic DNA was extracted from healthy leaves, according to the procedure described by Shure et al.

(1983), with slight modifications. DNA concentration was determined using a Hoechst dye-based protocol and a fluorescence spectrophotometer (Model F-4500, Hitachi, Ibaragi, Japan). The DNA solution was diluted to a working concentration with TE buffer (pH 8.0), and stored at −20°C until used.

To obtain PCR products corresponding to the *GmNARK* gene, primers were designed based on the *GmNARK* base sequence (GenBank acc. no. AY16665), available from the National Center for Biotechnology Information databases (http://www.ncbi.nlm.nih.gov/). Primer sequences were analyzed and selected using Gene Runner, version 3.05 (Hastings Software, Hastings on Hudson, N.Y., USA). Several primers were synthesized for PCR products to cover full sequences of *GmNARK*.

The amplification reaction consisted of 50 ng of genomic DNA, 3.2 pmol of each forward and reverse primer, 200 μM of each dNTP, 1.5 mM $MgCl_2$, 1 × reaction buffer [10 mM Tris-HCl (pH 8.3) 50 mM KCl], and 1 U of Ampli *Taq* Gold DNA polymerase (Applied Biosystems, Foster City, Calif., USA) in a total volume of 50 μl. The reaction mixture was denatured at 94°C for 4 min and subjected to 30 cycles of 94°C for 30 s, annealing at 50–65°C for 30 s, and 72°C for 1 min on a PTC-225 Peltier Thermal Cycler (MJ Research, Watertown, Mass., USA). PCR products were resolved by electrophoresis on 1.0% ethidium bromide-stained agarose gels. Only PCR products amplified as a single fragment were used as templates for further sequencing reactions.

After the PCR product was purified using a NucleoSpin Extract kit (MACHEREY-NAGEL, Easton City, Pa., USA), a sequencing reaction was performed with the BigDye terminator cycle sequencing kit (Applied Biosystems) and resolved on an ABI PRISM 3700 DNA sequencer. To identify a SNP in the *GmNARK* gene among different soybean genotypes, the sequences were aligned using SeqScape software, version 1.1 (Applied Biosystems).

SNP-specific primer design and SNAP-PCR

To obtain primers specific to the identified SNP, the segment of the *GmNARK* sequence containing the SNP site was entered into the Web-available SNAPER program (http://ausubellab.mgh.harvard.edu/). Five primer pairs corresponding to the SNP were tested using standard PCR amplification. PCR was performed in a total volume of 20 μl, which contained 10–100 ng of template DNA, 5 μM each of the forward and reverse primers, 100 μM of each dNTP, 1.5 mM $MgCl_2$, 1 × reaction buffer (10 mM Tris-HCl pH 8.3, 50 mM KCl), and 1 U of Ampli *Taq* Gold polymerase (Applied Biosystems). Template DNA was initially denatured at 94°C for 5 min, according to Drenkard et al. (2000), followed by 28 or 38 cycles for PCR amplification, using the following conditions: denaturation at 94°C for 30 s and annealing and extension at 64°C for 1 min, and then final extension at 72°C for 10 min on a PTC-225 Peltier Thermal Cycler (MJ Research). Amplified products were separated on a 1.5% agarose gel to estimate each allele in the SNP site as presence or absence of a band.

Results

SNP identification for *GmNARK*

Several primer pairs designed from the *GmNARK* gene sequence were used to identify SNPs between Sinpaldalkong 2 and SS2-2. A base change from A to T was discovered at the 959-bp position of the *GmNARK* sequence between both soybean genotypes (Fig. 1). This single-base mutation in SS2-2 changes a lysine codon (AAG) to a stop codon (TAG), terminating translation of the gene and thus producing a truncated protein. Searle et al. (2003) demonstrated that all supernodulating mutants had truncated proteins, whereas amino acid substitution was observed in hypernodulating mutants. The identified SNP confirmed that SS2-2 is supernodulating mutant, according to Searle et al. (2003).

Interestingly, the position of the mutation and the nucleotide change (A → T) in SS2-2 were found to be the same as those in nts246, another EMS-induced supernodulating mutant derived from Bragg (Fig. 2; Searle et al. 2003). The *GmNARK* sequences of Sinpaldalkong 2 and Bragg were mutually different where several SNPs were discovered (Fig. 2). Two SNPs, including an indel, and five single-base changes were identified in 5′UTR and intron regions, respectively. An additional six single-base changes were found in exon regions, leading to synonymous (no alteration in amino acid) changes of the *GmNARK* sequence. Therefore, no functional change in the *GmNARK* protein between Sinpaldalkong 2 and Bragg is expected, although two wild types carry several SNPs for *GmNARK*.

SNAP marker development for *GmNARK*

The identified SNP between Sinpaldalkong 2 and its supernodulating SS2-2 variant at the 959th base pair of *GmNARK* was used to develop a PCR-based SNP marker to facilitate MAS for supernodulation. Five primer pairs for the A/T SNP of *GmNARK* between Sinpaldalkong 2 and SS2-2 were comprising three pairs specific to the A allele and two pairs specific to the T allele (Table 1). The SNP-specific primers were designed so that the 3′-terminal nucleotide of a primer should be complementary to one allele of a SNP, and such that the primer should contain an artificial mismatch within 4 bp of the site of the SNP (Fig. 3). Optimal DNA concentration and the number of amplification cycles were initially determined to increase specificity and fidelity of

```
Sinpaldalkong 2  gcaaaatcag aga gacatga gaagctgtgt gtgctac acg ctattattgt ttattttctt
SS2-2            gcaaaatcag aga gacatga gaagctgtgt gtgctac acg ctattattgt ttattttctt

Sinpaldalkong 2  catatggctg cgcgtggcaa cgtgctcttc gttcactgac atggaatcgc ttctgaagct
SS2-2            catatggctg cgcgtggcaa cgtgctcttc gttcactgac atggaatcgc ttctgaagct

Sinpaldalkong 2  gaaggactcc atgaaaggag ataaagccaa agacgacgct ctccatgact ggaagttttt
SS2-2            gaaggactcc atgaaaggag ataaagccaa agacgacgct ctccatgact ggaagttttt

Sinpaldalkong 2  cccctcgctt tctgcacact gtttcttttc aggcgtaaaa tgcgaccgag aacttcgagt
SS2-2            cccctcgctt tctgcacact gtttcttttc aggcgtaaaa tgcgaccgag aacttcgagt

Sinpaldalkong 2  cgttgctatc aacgtctcgt ttgttcctct cttcggtcac cttccgccgg agatcggaca
SS2-2            cgttgctatc aacgtctcgt ttgttcctct cttcggtcac cttccgccgg agatcggaca

Sinpaldalkong 2  attggacaaa ctcgagaacc tcaccgtctc gcagaacaac ctcaccggcg tacttccc A a
SS2-2            attggacaaa ctcgagaacc tcaccgtctc gcagaacaac ctcaccggcg tacttccc T a
                                                                          *

Sinpaldalkong 2  ggagctcgcc gccctcactt ccctcaagca cctcaacatc tctcacaacg tcttctccgg
SS2-2            ggagctcgcc gccctcactt ccctcaagca cctcaacatc tctcacaacg tcttctccgg

Sinpaldalkong 2  ccatttcccc ggccaaatta tccttccgat gacgaaactg gaggtcctcg acgtctacga
SS2-2            ccatttcccc ggccaaatta tccttccgat gacgaaactg gaggtcctcg acgtctacga

Sinpaldalkong 2  caacaacttc accggaccgc ttcccgtaga gttggtgaaa ctggagaaat taaaatacct
SS2-2            caacaacttc accggaccgc ttcccgtaga gttggtgaaa ctggagaaat taaaatacct

Sinpaldalkong 2  gaagctcgac ggaaactatt tctccggcag cataccggag agtta ctcgg agtttaagag
SS2-2            gaagctcgac ggaaactatt tctccggcag cataccggag agtta ctcgg agtttaagag

Sinpaldalkong 2  cttgga gttt ttaagcttaa gcaccaatag cttatcgggg aagattccca agagtttgtc
SS2-2            cttgga gttt ttaagcttaa gcaccaatag cttatcgggg aagattccca agagtttgtc
```

Fig. 1 Sequences alignment of each PCR fragment derived from the *GmNARK* gene between cultivar Sinpaldalkong and SS2-2. The primer sequences are indicated by *italics*. The sites of single nucleotide polymorphisms (SNPs) were indicated as *asterisks*

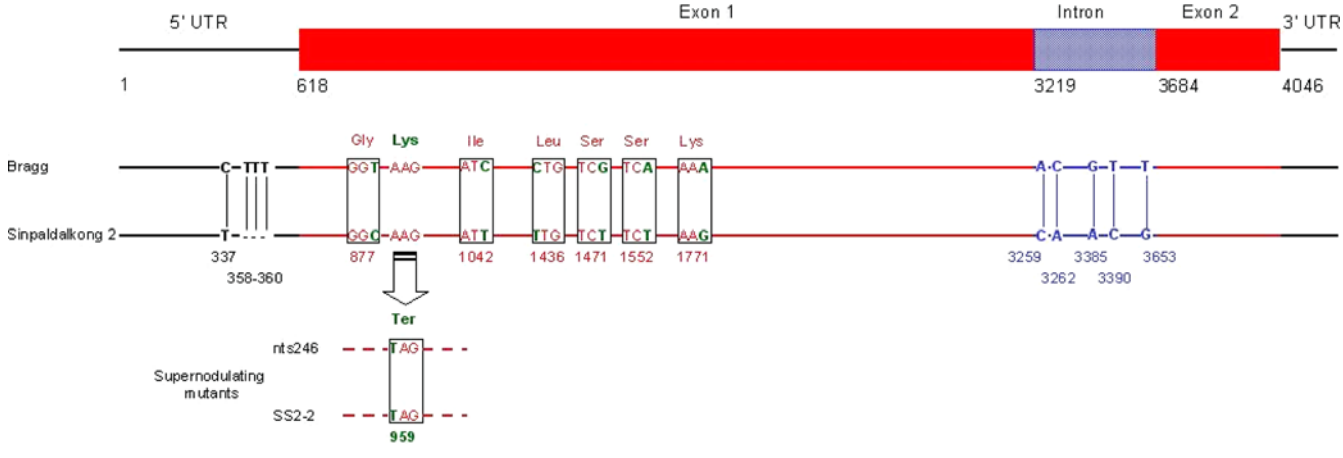

Fig. 2 Map of *GmNARK* genomic region. Exons and introns are indicated by *red-filled* and *blue-shaded boxes*. The second diagram represents SNP positions, replaced DNA, and amino acid sequences between cultivars Bragg and Sinpaldalkong 2. The A/T SNP at the 959-bp position of *GmNARK* in nts246 and SS2-2 is indicated at the *bottom*

SNAP-PCRs. One primer pair specific to the A allele was used to test different DNA concentrations and amplification cycles. Higher template-DNA concentrations and longer cycling numbers appeared to produce false positives (Fig. 4a). When less than 30 ng template DNA and 28 PCR cycles were used, the primer pair generated an amplification product for the A allele and no product for the T allele (Fig. 4a). Furthermore, the dNTPs were added into PCR mixture at a relatively low concentration of 100 μM, because low dNTP concentration minimizes mispriming at nontarget sites and reduces the likelihood of extending misincorporated nucleotides (Innis et al. 1988).

Using determined optimal DNA concentration and PCR cycles, the remaining four primer pairs were tested against Sinpaldalkong 2 and SS2-2. Two primer

Table 1 Sequences of primer pairs specific to the A/T single nucleotide polymorphism of the *GmNARK* gene used in single nucleotide-amplified polymorphism-PCR and sizes and phenotypes of the PCR products generated by each primer pair

Allele	Primer sequence	PCR product size	Phenotype
A specific	F: CAACCTCACCGGCGTACTTCCGA R: CCTCAGCGTCTTCAACTTCGACAAACTC	343	Polymorphic
	F: AACAACCTCACCGGCGTACTTCCAT R: CCTCAGCGTCTTCAACTTCGACAAACTC	345	Monomorphic
	F: GTGAGGGCGGCGAGCTCCCT R: AATCAGAGAGACATGAGAAGCTGTGTGTGCTA	374	Monomorphic
T specific	F: GAACAACCTCACCGGCGTACTTCCTT R: CCTCAGCGTCTTCAACTTCGACAAACTC	346	Polymorphic
	F: GTGAGGGCGGCGAGCTCCAA R: AATCAGAGAGACATGAGAAGCTGTGTGTGCTA	374	Polymorphic

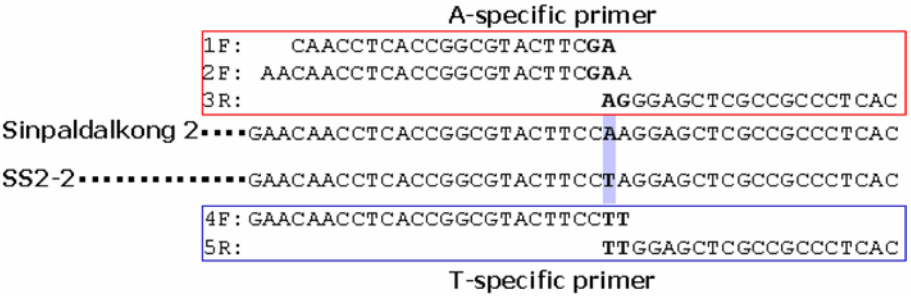

Fig. 3 Alignment between partial sequence of *GmNARK* containing the SNP and allele-specific primers. The extra base-mismatch were introduced into allele-specific primers. These produce a dramatic reduction in the PCR activity of the alternative allele but have a relatively minor effect on the amplification of the corresponding allele (Drenkard et al. 2000). *F* forward primer, *R* reverse primer

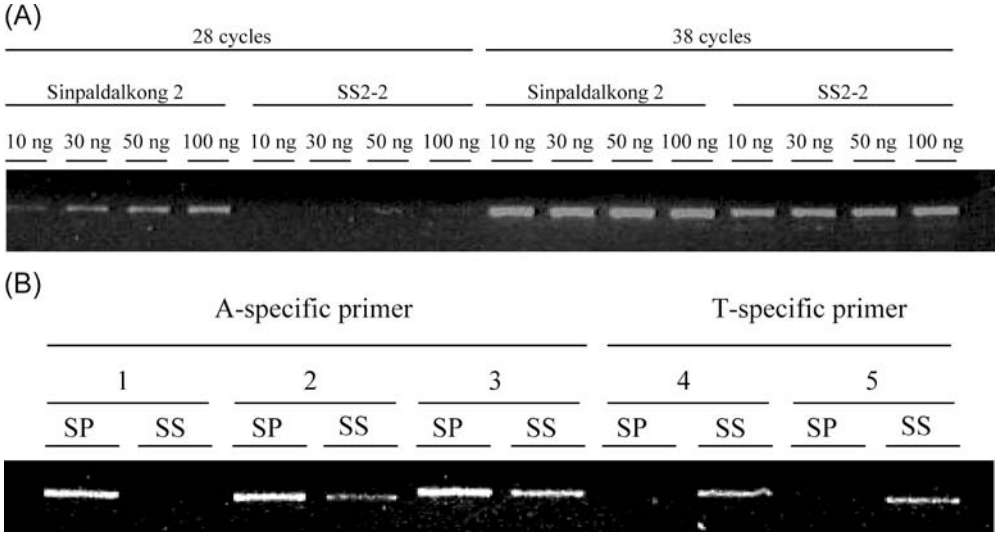

Fig. 4 **a** Analysis of specificity for representative single nucleotide-amplified polymorphism (SNAP) primer pairs, according to template DNA concentration and amplification cycles. The primer pairs specific to the A allele of *GmNARK* was used. **b** Specificity of five primer pairs designated for each allele of *GmNARK*. The numbers *1–5* represent five primer pairs, respectively. *SP* Sinpaldalkong 2, *SS* SS2-2

pairs specific to the A allele were monomorphic, producing PCR products for both alleles (Fig. 4b). T-allele-specific primer pairs designed for SNAP generated a PCR product in only SS2-2. The A- and T-specific primer pairs showing polymorphisms could be used in two complementary SNAP-PCRs to identify genotypes for supernodulation in breeding descendents, using SS2-2.

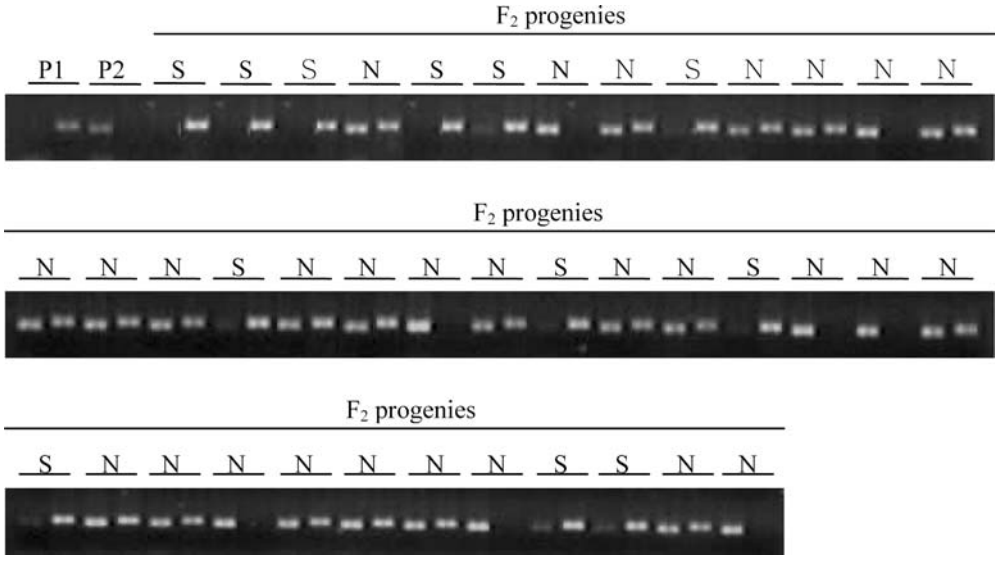

Fig. 5 Genotyping of the SNAP marker for supernodulation with the F_2 progeny of SS2-2 × Sinpaldalkong 2. Allele-specific primers for *GmNARK* were used to amplify 40 lines, and patterns of bands were scored as normal (*N*) and supernodulating (*S*). For each line, the *left lane* shows the amplification product obtained with the A-allele-specific SNAP primer and the *right lane* shows the amplification product obtained with the T-allele-specific primer. The lines showing both bands specific to each allele were heterozygous

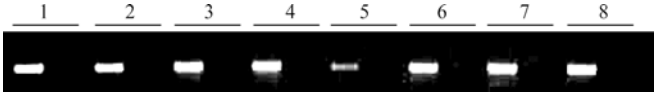

Fig. 6 Amplification patterns of the SNAP marker for supernodulation in eight soybean genotypes. For each genotype, the *left lane* shows the amplification product obtained with the A-allele-specific SNAP primer and the *right lane* shows the amplification product obtained with the T-allele-specific primer. *1* Danbaekkong, *2* Taekwangkong, *3* Pureunkong, *4* Jinpumkong 2, *5* PI 96188, *6* Jangyeobkong, *7* nts382, *8* nts1116

SNAP marker evaluation

To assess the feasibility of using the SNAP marker in MAS for supernodulation, we compared the genotype data with phenotypic data for nodulation in the 40 F_2 progenies of SS2-2 × Sinpaldalkong 2 (Fig. 5). In this population, 28 seedlings were scored as normal and 12 as supernodulating. Among 28 normal individuals, eight individuals had only the band specific to the A allele, indicating that they were homozygous. The remaining individuals with normal phenotype were found to be heterozygous for the A/T SNP, because bands for both bases were present. Twelve supernodulating individuals showed only a band specific to the T allele. The allelic segregation ratio for *GmNARK* was found to fit to the expected ratio of 1:2:1 for a single recessive gene in the F_2 population. Thus, the A/T SNAP marker co-segregated completely with the supernodulation trait, and the alternative nucleotides at the SNP site should represent normal and supernodulating phenotypes, respectively.

The SNAP marker for *GmNARK* was also evaluated using eight soybean genotypes, including six normal, one supernodulating, and one hypernodulating. All genotypes represented produced a band identical to that of Sinpaldalkong 2 (Fig. 6). The supernodulating mutant, nts382, has a SNP at a different site on the *GmNARK* gene than does SS2-2 (Searle et al. 2003). Although SS2-2 and nts382 have the same supernodulating character, they could be discriminated by using the SNAP marker corresponding to the different SNP for *GmNARK*.

Discussion

Genetic mapping in a population of SS2-2 × nts382 revealed that the mutation in SS2-2 was in the same position as *NTS-1* (Ha and Lee 2001). The locus was 18.9 cM away from Satt353 on LG H, and near the RFLP marker pUTG-132, which is tightly linked to the *nts* gene (Landau-Ellis et al. 1991; Kolchinsky et al. 1997). Nodulation of SS2-2 might be thus controlled by the same nts382 supernodulation locus (Ha and Lee 2001). Recently, Searle et al. (2003) isolated the *GmNARK* gene, which is involved in autoregulation of nodule formation, based on the physical map of the *NTS-1* locus. Since both mutant loci of SS2-2 and nts382 were located at the same position on LG H, SNPs within *GmNARK* were surveyed between Sinpaldalkong 2 and SS2-2.

A single-base change from A to T induced by EMS was detected at the 959th base pair in an exon region of the *GmNARK* sequence (Fig. 1). The SNP affected a codon at 5′ upstream of exon in *GmNARK* (Fig. 2) and resulted in the loss of most of the leucine-rich repeats, transmembrane, and kinase domains. This truncated

protein could not therefore deliver extracellular signals into the intracellular space (Searle et al. 2003). In contrast, the hypernodulating mutant resulting from a missense mutation carries an incomplete and weak signal, and autoregulation of nodulation would be partial (Searle et al. 2003).

SS2-2 was originally designated as a hypernodulating mutant in terms of the absolute number of nodules, because it had three to five times more nodules than its wild type, Sinpaldalkong 2 (Lee et al. 1998). However, SS2-2 could be also considered as a supernodulating mutant, based on the occurrence of nonsense mutation in *GmNARK*.

The alkylating agent, EMS, produces O^6-ethylguanine, pairing not with C, but with T (Ashburner 1990; Horsfall et al. 1990). Mutations induced by EMS are thus largely characterized by C/G or T/A transitions (Pastink et al. 1991; Vidal et al. 1995; Greene et al. 2003). Among six different mutants, however, the A/T transversion was discovered in two supernodulating mutants, nts246 and En6500 (Searle et al. 2003). The EMS mutation mechanism in these mutants, including SS2-2, does not appear to be involved in direct mispairing of alkylated bases, but rather in an error-prone repair pathway. If the DNA were heavily damaged by EMS, a number of proteins involved in DNA repair, recombination, and replication would be synthesized. Some of these proteins allow DNA polymerase to proceed across regions of heavily damaged DNA, and the accuracy in replication is therefore decreased. In some cases of microorganism such as *Escherichia coli* and *Streptomyces*, EMS may mediate mutagenesis by an error-prone repair pathway (Todd et al. 1981; Stonesifer and Baltz 1985).

EMS affects the encoded proteins in three different ways, missense, truncation, and silencing (Greene et al. 2003). In 192 different *Arabidopsis thaliana* genes, about 50% of EMS-induced mutations were missense, whereas truncation was involved in only 5% (Greene et al. 2003). In the case of the *GmNARK* gene of soybean, four *nts*-allelic mutants and another mutant, En6500, produced the truncated protein (Searle et al. 2003) as well as did the mutant. The only missense mutation was found in the nts1116 mutant (Searle et al. 2003).

SS2-2 and nts246 turned out to be the same mutant, although they definitely have a different genetic background. Their wild types, Sinpaldalkong 2 and Bragg, showed no alteration of amino acid sequences (synonymous change) in *GmNARK*, although six SNPs were detected in the exon region. The high proportion of synonymous SNPs between wild types and the occurrence of the same mutant for the *GmNARK* gene suggests that preferences of selection against mutations lead to changes in specific nucleotide sequences (Zhu et al. 2003). These SNPs may be helpful for investigating the origin of different soybean cultivars in terms of evolution.

More remarkably, the identified mutation at the same position in *GmNARK* was found in both SS2-2 and nts246. This type of mutation rarely occurs, considering the chances of this occurrence. This is in contrast to the random occurrence of mutations in the genome generated by EMS.

In this study, the sequence mutation of A → T in *GmNARK* was associated with a functional difference between wild-type Sinpaldalkong 2 and its mutant, SS2-2, as normal- and supernodulation, respectively. The identified SNP in *GmNARK* could directly affect phenotypic variations for nodulation. For this reason, a SNP directly linked to a specific trait would be very useful as a DNA marker in selection for the trait. Recently, the advent of SNP discovery has led to the introduction of a new concept of DNA markers. Instead of DNA marker classification on their technical properties, as has been done up until now, Andersen and Lübberstedt (2003) divided markers into random DNA markers, gene-targeted markers, and functional markers (FMs), based on the level of functional characterization of the polymorphisms. FMs, in particular, are derived from polymorphic sites within a gene causally involved in phenotypic trait variation (Andersen and Lübberstedt 2003). Therefore, the A/T SNP of the *GmNARK* gene must be considered as a FM. FMs have complete linkage to the corresponding trait. Selection error of FM in natural as well as breeding populations is therefore zero, owing to little or no possibility of recombination at meiosis (Andersen and Lübberstedt 2003).

Nodulation traits in soybean are difficult to assay phenotypically because of being subterranean characters and the requirement of inoculation with *Bradyrhizobium*. We developed a PCR-based SNAP marker on the identified SNP in *GmNARK*. This SNAP marker was able to distinguish between normal nodulation and supernodulation in not only the F_2 population of SS2-2 × Sinpaldalkong 2, but also among several soybean genotypes (Figs. 5, 6). Two other types of widely used PCR markers based on SNPs, cleaved amplified polymorphic sequence [(CAPS) Konieczny and Ausubel 1993] and derived CAPS (Michaels and Amasino 1998; Neff et al. 1998), could only be used on the condition of the existence or introduction of restriction sites at the site of a SNP (Drenkard et al. 2000). Although other types of PCR markers are available, this SNAP marker is the easiest detection system because of using only a simple, agarose gel-based assay and does not require the use of restriction enzymes. In conclusion, our SNAP marker for *GmNARK* represents a functional marker, and should be accessible to any soybean genetic researcher or breeder.

Acknowledgements This research was also supported by a grant (code no. CG3121) from Crop Functional Genomic Center of the 21st Century Frontier Research Program funded by the Ministry of Science and Technology (MOST) and Rural Development Administration (RDA) of the Republic of Korea. We also thank the National Instrumentation Center for Environmental Management at Seoul National University in Korea, and the International Atomic Energy Agency in Austria.

References

Akao S, Kouchi H (1992) A supernodulating mutant isolated from soybean cultivar Enrei. Soil Sci Plant Nutr 38:183–187

Andersen JR, Lubberstedt T (2003) Functional markers in plants. Trends Plant Sci 8:554–560

Ashburner M (1990) *Drosophila*: a laboratory handbook. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Cai H, White PS, Torney D, Deshpande A, Wang Z, Marrone B, Nolan JP (2000) Flow cytometry-based minisequencing: a new platform for high throughput single-nucleotide polymorphism scoring. Genomics 66:135–143

Carroll BJ, McNeil DL, Gresshoff PM (1985a) Isolation and properties of soybean [*Glycine max* (L) Merr.] mutants that nodulate in the presence of high nitrate concentrations. Proc Natl Acad Sci USA 94:773–781

Carroll BJ, McNeil DL, Gresshoff PM (1985b) A supernodulation and nitrate-tolerant symbiotic (*nts*) soybean mutant. Plant Physiol 78:34–40

Chen J, Iannone MA, Li M-S, Taylor JD, Rivers P, Nelson AJ, Slentz-Kesler KA, Roses A, Weiner MP (2000) A microsphere-based assay for multiplexed single nucleotide polymorphism analysis using single base chain extension. Genome Res 10:549–557

Drenkard E, Richter BG, Rozen S, Stutius LM, Angell NA, Mindrinos M, Cho RJ, Oegner PJ, Davis RW, Ausubel FM (2000) A simple procedure for the analysis of single nucleotide polymorphisms facilitates map-based cloning in *Arabidopsis*. Plant Physiol 124:1483–1492

Greene E, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, Enns LC, Burtner C, Johnson JE, Odden AR, Comai L, Henikoff S (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis*. Genetics 164:731–340

Gupta PK, Roy JK, Prasad M (2001) Single nucleotide polymorphisms: a new paradigm for molecular marker technology and DNA polymorphism detection with emphasis on their use in plants. Curr Sci 80:524–535

Gut IG (2001) Automation in genotyping of single nucleotide polymorphisms. Hum Mutat 17:475–492

Ha B-K, Lee S-H (2001) Genetic mapping of hypernodulation in soybean mutant SS2-2. Korean J Crop Sci 46:416–419

Hayashi K, Hashimoto N, Daigen M, Ashikawa I (2004) Development of PCR-based SNP markers for rice blast resistance genes at the *Piz* locus. Theor Appl Genet 108:1212–1220

Horsfall MJ, Gordon AJE, Burns PA, Zielenska M, van der Vliet ME, Glickman BW (1990) Mutational specificity of alkylating agents and the influence of DNA repair. Environ Mol Mutagen 15:107–122

Innis MA, Myambo KB, Gelfand DH, Brow MAD (1988) DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA. Proc Natl Acad Sci USA 85:9436–9440

Kim MY, Ha B-K, Jun T-H, Hwang E-Y, Van K, Kuk YI, Lee S-K (2004) Single nucleotide polymorphism discovery and linkage mapping of lipoxygenase-2 gene (Lx_2) in soybean. Euphytica 135:169–177

Kolchinsky A, Landau-Ellis D, Gresshoff PM (1997) Map order and linkage distances of molecular markers close to the supernodulation (*nts-1*) locus of soybean. Mol Gen Genet 254:29–36

Konieczny A, Ausubel F (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. Plant J 4:403–410

Landau-Ellis D, Angermuller S, Shoemaker R (1991) The genetic locus controlling supernodulation in soybean (*Glycine max* L.) co-segregates tightly with a cloned molecular marker. Mol Gen Genet 228:221–226

Lee HS, Chae YA, Park EH, Kim YW, Yun KI, Lee S-H (1997) Introduction, development, and characterization of supernodulating soybean mutant. I. Mutagenesis of soybean and selection of supernodulating mutant. Korean J Crop Sci 42:247–253

Lee HS, Kim YW, Park EH (1998) Introduction, development, and characterization of supernodulating soybean mutant: Shoot factor regulation of nodule development in supernodulating soybean mutant. Korean J Crop Sci 43:28–31

Livak KJ (1999) Allelic discrimination using fluorogenic probe and the 5′ nuclease assay. Genet Anal 14:143–149

Men AE, Laniya TS, Iturbe-Dramaetxe I, Gresshoff I, Jiang Q, Carroll BJ, Gresshoff PM (2002) Fast neutron mutagenesis of soybean (*Glycine soja* L.) produces a supernodulating mutant containing a large deletion in linkage group H. Genome Lett 3:147–155

Michaels SD, Amasino RM (1998) A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. Plant J 14:381–385

Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. Plant J 14:387–392

Pastink A, Heemskerk E, Nivard MJ, van Vliet CJ, Vogel EW (1991) Mutational specificity of ethyl methanesulfonate in excision–repair-proficient and -deficient strains of *Drosophila melanogaster*. Mol Gen Genet 229:213–218

Pierce M, Bauer WD (1983) A rapid regulatory response governing nodulation in soybean. Plant Physiol 73:286–290

Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning Z, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann N, Zody MC, Linton L, Lander ES, Altshuler D (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933

Searle IR, Men AE, Laniya TS, Buzas DM, Iturbe-Ormaetxe I, Carroll BJ, Gresshoff PE (2003) Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase. Science 299:109–112

Shi MM (2001) Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies. Clin Chem 47:164–172

Shure M, Wessler S, Fedoreff N (1983) Molecular identification and isolation of the *waxy* locus in maize. Cell 35:225–233

Stonesifer J, Baltz RH (1985) Mutagenic DNA repair in *Streptomyces*. Proc Natl Acad Sci USA 82:1180–1183

Syvänen A-C (2001) Accessing genetic variation: genotyping single nucleotide polymorphisms. Nature Rev Genet 2:930–942

Taylor JD, Briley D, Nguyen Q, Long K, Iannone MA, Li M-S, Ye F, Afshari A, Lai E, Wagner M, Chen J, Weiner MP (2001) Flow cytometric platform for high-throughput single nucleotide polymorphism analysis. Biotechniques 30:661–669

Todd PA, Brouwer J, Glickman BW (1981) Influence of DNA repair deficiencies on MMS- and EMS-induced mutagenesis in *Escherichia coli* K12. Mutat Res 82:239–250

Vidal A, Abril N, Pueyo C (1995) DNA repair by Ogt alkyltransferase influences EMS mutational specificity. Carcinogenesis 16:817–821

Ye F, Li M-S, Taylor JD, Nguyen Q, Colton HM, Casey WM, Wagner M, Weiner MP, Chen J (2001) Fluorescent microsphere-based readout technology for multiple human single nucleotide polymorphism analysis and bacterial identification. Hum Mutat 17:305–316

Zhu YL, Song QJ, Hyten SM, Fickus EW, Young ND, Cregan PB (2003) Single-nucleotide polymorphism in soybean. Genetics 163:1123–1134