Australian Journal of Crop Science AJCS
AJCS 5(3):283-290 (2011) ISSN:1835-2707

SSR analysis of soybean (*Glycine max* (L.) Merr.) genetic relationship and variety identification in Thailand

P. Tantasawat, J. Trongchuen, T. Prajongjai, S. Jenweerawat and W. Chaowiset

School of Crop Production Technology, Institute of Agricultural Technology, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

***Corresponding author: piyada@sut.ac.th**

Abstract

Simple sequence repeat (SSR) analysis was used to determine the genetic diversity and relatedness among 25 soybean (*Glycine max* (L.) Merr.) genotypes (15 certified varieties, 8 breeding lines and 2 plant introductions) in Thailand. Eleven SSR primer pairs could amplify polymorphic SSRs from all of these genotypes. A total of 53 alleles with an average of 4.82 alleles per locus were detected. The polymorphic information content (PIC) among genotypes varied from 0.13 (Soy satt 285) to 0.88 (Soy satt 173) with an average of 0.60. Pairwise coefficients of genetic similarity between all genotypes ranged from 0.73 to 1.00 with an average of 0.88. These eleven SSR markers successfully distinguished 23 of the 25 soybean genotypes, with the exception of a pair of closely related breeding lines from the same cross. Allelic variation was observed at the two SSR loci associated with agronomic traits. In addition, only the five most polymorphic SSR markers were able to clearly identify all 15 certified varieties and would be useful for DNA fingerprinting. Unweighted pair-group method arithmetic average (UPGMA) analysis allocated the genotypes in 4 major clusters containing 19, 2, 2 and 2 genotypes, respectively. The largest cluster (I) was divided into subclusters Ia and Ib consisting of 13 and 6 genotypes, respectively. The genetic relationships among genotypes generally agreed with known pedigrees. Principal coordinate analysis (PCoA) confirmed the separation of soybean genotypes into four groups comparable to those from UPGMA analysis. However, genotypes in subclusters Ia and Ib were more clearly separated. These results suggest that SSR markers are efficient for measuring genetic diversity and relatedness as well as identifying varieties of soybeans. Genetic diversity and relationship assessments among soybean genotypes in Thailand could provide useful information for efficient utilization of these materials, especially for genetic improvement.

Keywords: DNA fingerprinting; genetic diversity; microsatellite; simple sequence repeat
Abbreviations: AFLP, amplified fragment length polymorphism; PCoA, principal coordinate analysis; PIC, polymorphism information content; RAPD, random amplified polymorphic DNA; RFLP, restriction fragment length polymorphism; SCN, soybean cyst nematode; SNP, single nucleotide polymorphism; SSR, simple sequence repeat.

Introduction

Soybean (*Glycine max* (L.) Merr.) is one of the world's most important economic legume crops. In Thailand major production impediments that cause continuous reduction in cultivation areas include low yield, lack of high quality seeds, low market value compared with other crops, and high cost of production compared with that in other countries (Office of Agricultural Economics, 2008). Therefore, breeding of soybeans for high yield, adaptability, disease resistance, nutritive content, and maturity date are particularly crucial in Thailand. Soybean germplasm in Thailand is still limited consisting of approximately 200 indigenous varieties (introduced into Thailand from various unknown sources), 16 certified varieties and various exotic soybean introductions in different breeding programs (Chowdhury et al., 2002; Pookpakdi, 2003; Chotiyarnwong et al., 2007). Among the 149 indigenous and 11 recommended soybean varieties, Chotiyarnwong et al. (2007) reported only a small genetic differentiation. This was not surprising because most southeast Asian soybean accessions were suggested to be derived from the Chinese germplasm pool by repeated and independent introductions (Abe et al., 2003). The low genetic diversity found was common in soybean and emphasized the importance of parental selection to avoid genetic relatedness and maintain genetic diversity in breeding programs (Thompson et al., 1998; Chotiyarnwong et al., 2007). An intensive soybean breeding effort in Thailand began in 1960 by selection of lines introduced from Japan and Taiwan, resulting in the release of SJ 1, SJ 2, and SJ 3 varieties which gave higher yield than two popular local varieties, SB 60 and Pakchong (Pookpakdi, 2003). Since then there has been continuous release of novel varieties, including SJ 4, SJ 5, Nakhon Sawan 1 (NS 1), Chiang Mai 1 (CM 1), CM 2, CM 3, CM 4, CM 60, Sukhothai 1 (ST 1), ST 2, ST 3, Chakkrabhandhu 1 (CB 1) and Khonkaen University 35 (KKU 35). The sixteen varieties mentioned above represent all Thai certified varieties that are grown extensively. At present, these soybean varieties are divided into 3 groups by maturity (Pookpakdi, 2003); (i) early maturing soybean (70-80 d) including NS 1 and CM 2, (ii) medium maturing soybean (90-100 d) including SB 60, SJ 1, SJ 2, SJ 4, SJ 5, ST 1, CM 60, ST 2, ST 3, CM 3 and CM 4 and (iii) late maturing soybean (110-120 d) including KKU 35 and CB 1.

However, genetic relationships have been evaluated only among eleven of them (Chotiyarnwong et al., 2007). Genetic relationships among accessions are helpful for designing future breeding efforts for yield, quality and pest resistance improvement (Wang et al., 2006a). Complete description of existing certified soybean varieties and patterns of genetic diversity could facilitate introgression of diverse germplasm into the current commercial soybean genetic base (Tara Satyavathi et al., 2006). Soybean genetic diversity and relationships can be assessed by the differences in morphological and agronomic traits, pedigree information, geographic origins, isozymes, and DNA markers (Nelson et al., 1987, 1988; Juvik et al., 1989; Perry and McIntosh, 1991; Gizlice et al., 1994, 1996; Sneller, 1994; Griffin and Palmer, 1995; Bernard et al., 1998; Dong et al., 2004; Guan et al., 2010; Wang et al., 2010). However, some factors affect these methods, e.g. the influence of environmental factors on morphological and agronomic traits, uncertain or incomplete data and possible errors in the pedigree information and origins of accessions, and the limitation of data provided by isozymes (Li and Nelson, 2001; Wang et al., 2006a). Among different DNA markers, restriction fragment length polymorphisms (RFLPs), random amplified polymorphic DNAs (RAPDs), amplified fragment length polymorphisms (AFLPs), single nucleotide polymorphisms (SNPs) and microsatellites or simple sequence repeats (SSRs) have been used extensively in soybeans, each with its own advantages and limitations (Keim et al., 1992; Maughan et al., 1996; Powell et al., 1996; Thompson et al., 1998; Narvel et al., 2000a; Ude et al., 2003; Li et al., 2010). SSRs have been shown to produce the highest polymorphisms compared to RFLPs, AFLPs and RAPDs, and have much greater ability to identify unique alleles in elite and PI soybean germplasm than other marker systems (Maughan et al., 1996; Narvel et al., 2000a; Morgante et al., 2002; Seitova et al., 2004; Wang et al., 2006a). Using SSR and SNP to measure genetic diversity among 5 Korean, 8 Thai and 3 wild soybeans, Tanya et al. (2001) found that SSR was more informative than SNP. Similarly, Li et al. (2010) found SSR to have higher resolving power for detecting population structure and estimating genetic diversity than SNP in a sample of 303 accessions of domesticated soybean and its wild progenitor *G. soja*. In particular, SSRs have been used successfully in estimation of genetic diversity and relationships among soybean genotypes in different populations (Narvel et al., 2000a; Abe et al., 2003; Sutakom, 2004; Wang et al., 2006a,b; Guan et al., 2010; Wang et al., 2010). In Thailand, RAPD and morphological markers were used to assess genetic relationships among 48 soybean introductions for selection of potential parents (Chowdhury et al., 2002). Recently, SSR has also been used to estimate genetic diversity among 149 Thai indigenous and 11 certified soybean varieties (Chotiyarnwong et al., 2007). However, five certified varieties including CM 1, CM 4, KKU 35, ST 2 and ST 3, some of which are more recent varieties with high yield, protein and oil content, or resistance to several diseases, have not been evaluated. These five certified varieties were analyzed in this study along with ten other popular certified varieties, eight elite breeding lines and two soybean introductions useful for breeding purposes using SSR markers distributed across different linkage groups. We also explored allelic variation at two agronomic trait-related SSR loci. The objectives of this research were (i) to measure the genetic diversity of 15 Thai certified soybean varieties and 8 breeding lines, and 2 soybean accessions from plant introduction (PI) by SSR analysis, (ii) to study the genetic relationships among these genotypes, (iii) to generate molecular fingerprints of certified varieties currently used commercially in Thailand for variety identification and (iv) to evaluate the allelic variation at 2 agronomic trait-related SSR loci. This information should be useful for soybean breeding programs and genetic studies.

Materials and methods

Plant materials

Twenty five soybean genotypes consisting of 15 certified varieties (CB 1, CM 1, CM 2, CM 3, CM 4, CM 60, KKU 35, NS 1, SJ 1, SJ 2, SJ 4, SJ 5, ST 1, ST 2, ST 3), 8 breeding lines (KKU 65, KKU 69, KKU 74, KKU 120, KKU 137, KKU 215, KKU 486, KKU 863) and 2 plant introductions (PIs; Long juvenile 2 [LJ 2], Prolina) were used in this analysis. These genotypes covered fifteen of sixteen certified soybean varieties used commercially in Thailand, elite breeding lines with maturity less than 85 days, and plant introductions with potential use for future breeding programs. LJ 2 is a PI with long juvenile (vegetative) growth, giving potentially high yield from high pod numbers and seed numbers per pod (Thitiporn Machikowa, personal communication). Recently, Machikowa and Laosuwan (2009) showed that breeding of soybean for yield improvement in Thailand could be accomplished by extension of days to flowering. Therefore this PI will be very useful for increasing yield of early and medium maturing soybean varieties. Prolina is a PI developed by the USDA-ARS, in cooperation with the North Carolina Agricultural Research Service, and is notable for high protein contents (approximately 46.1%; Burton et al., 1999).

DNA isolation

Genomic DNA was isolated from bulk young leaves of 10 plants from each genotype following the cetyl trimethyl ammonium bromide (CTAB) method of Saghai Maroof et al. (1984). DNA was quantified by comparing the intensity of ethidium bromide-stained DNA bands on 0.8 % agarose gels with those of known concentrations of uncut λ DNA to adjust final concentration to 50 ng $\mu L^{-1}$for use in PCR analysis.

Simple sequence repeat (SSR) analysis

A total of 12 SSR markers that were previously mapped on 10 linkage groups (LGs) of soybean (Cregan et al., 1999a) and had been shown to be highly polymorphic among soybean genotypes (Narvel et al., 2000a,b; Wang et al., 2006a) were chosen for the analysis (Table 1). Two of these SSR loci are associated with agronomic traits, soybean cyst nematode (SCN) resistance and fasciation mutation (Cregan et al., 1999b; Karakaya et al., 2002). The distribution of these markers has been described by Narvel et al. (2000b). Most of the SSR markers that are included had an (ATT)n motif due to their abundance and polymorphic nature in soybeans and their easily interpretable allele patterns (Narvel et al., 2000b). These markers have been shown to be effective for the assessment of soybean diversity (Narvel et al., 2000a,b). DNA was amplified by PCR in a total volume of 10 µl containing 50 ng template DNA, 1X buffer (75 mM Tris.HCl (pH 9.0), 50 mM KCl, 20 mM $(NH_4)_2SO_4$), 2 mM $MgCl_2$, 200 µM of each dNTP, 5 pmol SSR primers, and 1 unit BIOTOOLS DNA polymerase (BIOTOOLS B&M Labs, SA). PCR reactions were carried out in a Thermo Hybaid

Table 1. Primer sequence, linkage group, number of alleles and polymorphism information content (PIC) of SSR primers

SSR primer	Linkage group	Primer sequence 5'→ 3'	Number of alleles	PIC
Soy satt 001	K	TGTGCAATGATAGTACATAGATAT GTGCTGATTGAACTATTTGTAGT	4	0.578
Soy satt 005	D1b	TATATCCTAGAGAAGAACTAAAAAA GTCGATTAGGCTTGAAATAATAC	5	0.635
Soy satt 148	I	TTAAGGATTAATTGAGACAAAATCA CTAAAGCATCACAAAACAGAGC	Multiple/ complex	Not available
Soy satt 160	F	ACATCAAAAGTTTATAACGTGTAG CTCCCACACAGTTTTCATATAAT	5	0.649
Soy satt 171	F	TTGAGGGCTCCCACACAGTT CAAAAGTTTATAACGTGTAGATTAA	4	0.616
Soy satt 173	O	CCGGTCCAATCTTTATTCAAAC CCAAGCGAAATCACCTCCTCT	6	0.878
Soy satt 183	J	CACCCTAGGATCTAGAACACC CTCATAAAACTACACACTTTCAG	5	0.616
Soy satt 185	E	CATATGAATAGGTAAGTTGCACT TGTCACTATAAATGGTACCTATTA	7	0.737
Soy satt 285	J	GCGACATATTGCATTAAAAACATACTT GCGGACTAATTCTATTTTACACCAACAAC	3	0.131
Soy satt 307	C2	GCGCTGGCCTTTAGAAC GCGTTGTAGGAAATTTGAGTAGTAAG	4	0.559
Soy satt 309	G	GCGCCTTCAAATTGGCGTCTT GCGCCTTAAATAAAACCCGAAACT	3	0.347
Soy satt 409	A2	CCTTAGACCATGAATGTCTCGAAGATA CTTAAGGACACGTGGAAGATGACTAC	7	0.859
Total			53	6.604
Average			4.82	0.600

Px2 thermocycler (Thermo Fisher Scientific, Inc., MA). Cycling parameters were initial denaturation step at 95°C for 10 min, followed by 95°C, 25 sec, 58°C, 25 sec and 72°C, 25 sec. This cycle was repeated 35 times, followed by 60 min extension at 72°C (Narvel et al., 2000a). The final extension was used to correct for nontemplate addition by Taq polymerase of a nucleotide, primarily adenosine, to the 3' end of amplification products (Smith et al., 1995). The amplified products were separated on 6% denaturing polyacrylamide gels and detected by silver nitrate according to Sambrook and Russell (2001). Allele sizes were estimated in comparison with 25 bp DNA ladder (Invitrogen Corporation, CA).

Data scoring, cluster and principal coordinate analysis

The bands appearing without ambiguity were scored as 1 (present) and 0 (absent) for each primer. Similarity coefficients between various varieties, in a pairwise comparison, were computed using Jaccard's coefficient and the resulting similarity matrix was further analyzed using the unweighted pair-group method arithmetic average (UPGMA) clustering algorithm; the computations were carried out using NTSYSpc version 2.2 (Rohlf, 2000). The goodness of fit of the genotypes to a specific cluster in the UPGMA cluster analysis was determined by Mantel's cophenetic correlation test (Mantel, 1967). The polymorphism information content (PIC), a measure of the allelic diversity at a locus, was determined as PIC = 1-ΣPi^2 where Pi is the frequency of the i^{th} allele in the examined test lines. NTSYSpc version 2.2 (Rohlf, 2000) was also used to perform principal coordinate analysis (PCoA) to show multiple dimensions of the distribution of the genotypes in a scatter-plot (Keim et al., 1992). This multivariate approach was used to complement the information obtained from cluster analysis because it is more informative regarding distances among major groups (Tar'an et al., 2005).

Results and discussion

Rate of polymorphism

A total of 12 SSR primer pairs, distributed on 10 of 20 linkage groups of soybean (Cregan et al., 1999a; Table 1), were used to amplify specific SSR loci from bulked DNA of each soybean genotype. Among these SSR primers, Soy satt 148 amplified multiple complex fragments and was withdrawn from further analysis, leaving eleven primer pairs which produced clear single-locus polymorphic bands for the analysis (Table 1). The high percentage of polymorphic SSR loci (91.7%) detected in this study was consistent with previous studies (Maughan et al., 1995; Rongwen et al., 1995

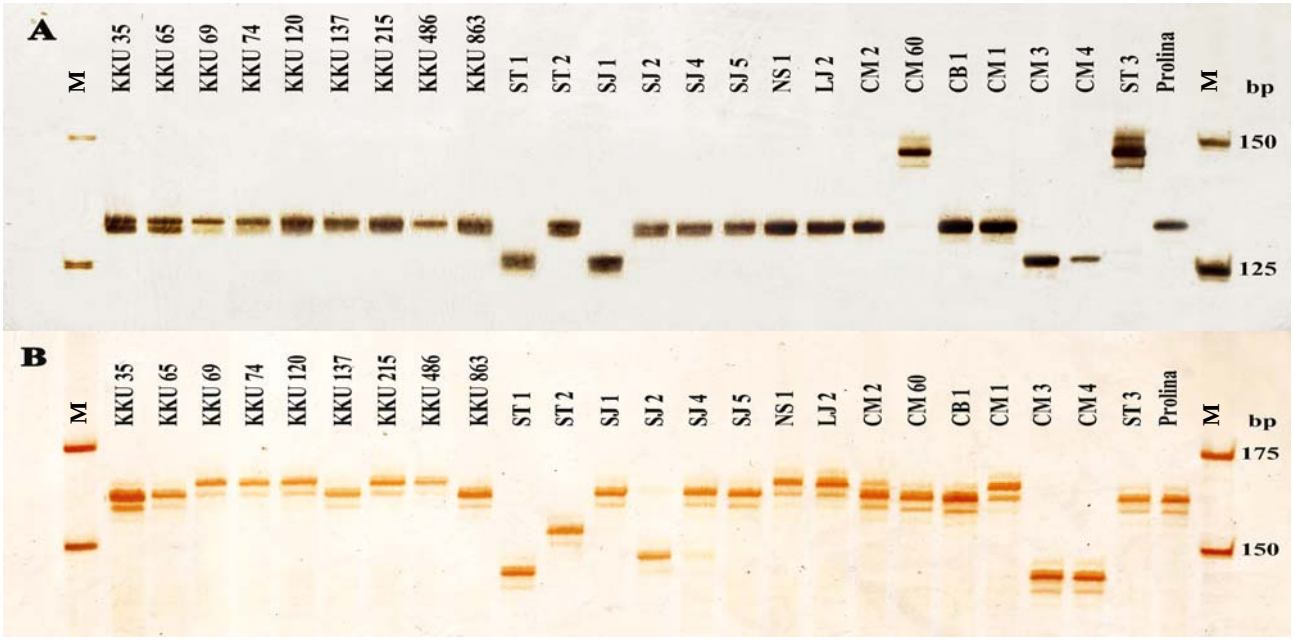

Fig 1. The SSR profiles of 25 soybean genotypes showing allelic variation at the Soy satt 309 (A) and Soy satt 005 (B) loci. Lane M, molecular mass marker (25 bp DNA ladder).

; Diwan and Cregan, 1997; Narvel et al., 2000a; Kumar et al., 2009; Khan et al., 2010; Singh et al., 2010). A total of 53 alleles from 11 SSR markers were detected across all 25 genotypes. The number of alleles per primer pair (locus) ranged from 3 for Soy satt 285 and Soy satt 309 to 7 for Soy satt 185 and Soy satt 409 with an average of 4.82 (Table 1). Figure 1 showed an example of DNA profiles at the Soy satt 309 and Soy satt 005 loci with 3 and 5 distinct alleles among different soybean genotypes, respectively. There were 69.8% of the alleles at a frequency of 0.25 or less. Only 3.8% of the alleles occurred at a frequency of 0.75 or higher. Unique alleles specific for CM 1, CM 4, LJ 2, Prolina, SJ 1, SJ 2, ST 1, ST 2 and ST 3 were observed in this study and may be useful for DNA fingerprinting. PIC values, a measure of the allelic diversity, ranged from 0.13 in Soy satt 285 to 0.88 in Soy satt 173 with an average of 0.60. Three of the eleven SSR markers (Soy satt 173, 185 and 409) had PIC values greater than 0.7 and high allele numbers (6-7), and hence were the most informative for distinguishing among the soybean genotypes. These markers occurred on 3 separate LGs, indicating that molecular polymorphism was spread across different regions of the genome. Significant correlation between the number of alleles and PIC values was observed (r = 0.98; p < 0.01). Most of the SSR markers (10/11) used in this study had PIC values ≥ 0.3, the value that has been used to determine usefulness of RFLP, RAPD and AFLP markers for genetic discrimination in previous soybean germplasm diversity studies (Keim et al., 1992; Lorenzen et al., 1995; Thompson and Nelson, 1998; Ude et al., 2003). The SSR diversity detected among soybean genotypes in this study was moderate compared to that from most previous reports. Diwan and Cregan (1997) detected an average of 10.10 alleles per locus and an average marker diversity of 0.80 when 20 SSR markers were used to distinguish the 35 North American soybean genotypes. Narvel et al. (2000a) calculated an average of 10.20 alleles per locus among 39 soybean elite genotypes and 40 PIs from seven different countries using 74 SSR markers. Similarly, analysis of 60 SSR markers on 129 Chinese soybean accessions had an average of 12.20 alleles per locus and an average PIC value of 0.78 (Wang et al., 2006a). Among 149 Thai indigenous and 11 recommended soybean varieties, the 18 SSR markers used had an average of 11.83 alleles per locus with an average genetic diversity index of 0.83 (Chotiyarnwong et al., 2007). The analysis of allelic profiles at 20 SSR loci produced an average of 11.9 alleles and a mean genetic diversity of 0.782 in 131 soybean accessions introduced from 14 Asian countries (Abe et al., 2003). Furthermore, SSR analysis of 244 Chinese and Japanese soybeans using 46 SSR markers yielded an average allele number of 16.2 per locus (Guan et al., 2010). Allelic variation at 2 SSR loci, Soy satt 005 and Soy satt 309, reported to be associated with agronomic traits was evaluated. The locus Soy satt 309 is on linkage group G less than 0.4 cM from SCN resistance gene *rhg 1*. Four alleles (125, 131, 134 and 149 bp) were detected at this locus in American soybean varieties and some PIs while as many as eight alleles were detected in the Chinese and Japanese soybean germplasm (Cregan et al., 1999b; Karakaya et al., 2002). Among the 56 American soybean varieties and PIs with known SCN resistance, two alleles (131 and 149 bp) were found specifically in susceptible genotypes, the 134 bp allele was confined only to resistant genotypes, while the 125 bp allele was found in both resistant and susceptible genotypes (Cregan et al., 1999b). In our study three alleles (128, 134 and 149 bp) were found among various genotypes with 134 bp as the most frequent one (Fig. 1A). The 128 and 134 bp alleles were also found to be the preponderant types in Chinese and Japanese soybeans, respectively (Guan et al., 2010). Although we do not know the levels of SCN resistance in all genotypes, the polymorphism found among resistant (ST 1 [128 bp]) and susceptible (SJ 4 [134 bp] and CM 60 [149 bp]) varieties suggests that this SSR marker might be useful for marker-assisted selection in SCN resistance breeding programs. However, its association with the SCN resistance gene needs to be verified in specific segregating populations. Soy satt 005 on linkage group D1b is related to the *f* locus conditioning fasciation mutation which has pleiotropic effects on plant development and pattern formation. The phyllotaxy and plastochron of soybeans were changed in fasciation

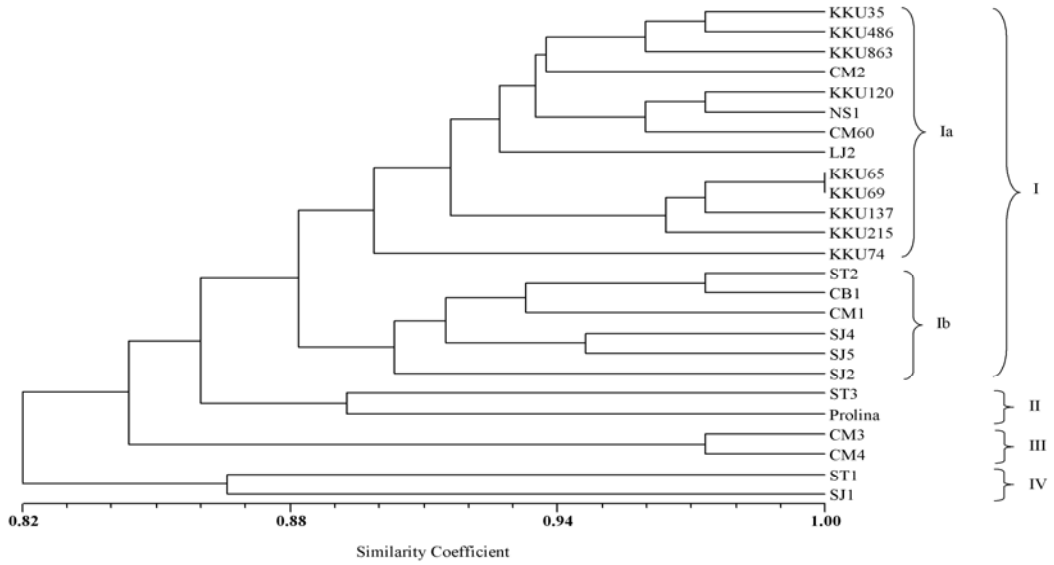

Fig 2. Dendrogram showing similarity coefficients and genetic relationships among 25 genotypes of soybean analyzed by SSR.

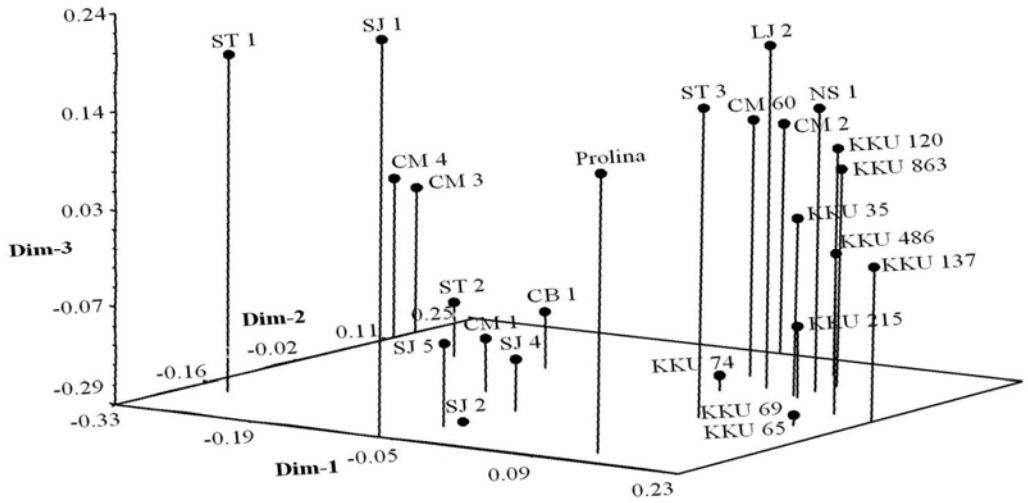

Fig 3. Three-dimensional plot based on the first three principal coordinates from a principal coordinate analysis of 11 SSR markers demonstrating the genetic relationships among 25 soybean genotypes.

mutation due to strong dominance (Karakaya et al., 2002). At this locus we found 5 different alleles (144, 147, 153, 165 and 168 bp) among various genotypes with 165 bp as the most frequent one (Fig. 1B). However, the contribution of each allele to the agronomic traits remains to be determined as has been stated by Guan et al. (2010).

Genetic diversity and relationships among soybean genotypes

All 53 SSR alleles scored were used for the genetic diversity analysis. Jaccard's similarity coefficients were calculated to assess the genetic resemblances among the genotypes and the similarity coefficients matrix was used for UPGMA cluster analysis. The pairwise genetic similarity between soybean genotypes varied from 0.73 (CM 3 vs Prolina) to 1.0 (KKU 65 vs KKU 69) with an average of 0.88. In fact, between the 15 certified soybean varieties under cultivation in Thailand the pairwise genetic similarity ranged from 0.79 to 0.97, 64% of which were ≥ 0.87. This low level of genetic diversity may be ascribed to the emphasis on direct introductions, selection from introduced germplasm and single cross hybrids (some of which shared common parents) in the soybean breeding programs. Therefore, inclusion of more diverse germplasm in the soybean breeding programs may provide the genetic variability necessary to permit continued progress and broad adaptation. A previous report has shown that higher genetic diversity could be found among exotic soybean introductions from different countries (Chowdhury et al., 2002). The dendrogram based on genetic similarities between genotypes showed that the 25 genotypes formed four major clusters (Fig. 2). The Mantel's test with cophenetic correlation coefficient value of 0.81 ($p < 0.01$) indicated good fit for the genotypes to a specific cluster in the dendrogram. Cluster I

was the largest and the most diverse cluster consisting of 10 varieties, 8 breeding lines and 1 PI with 90.3% genetic similarity. This cluster was further divided into 2 subclusters; subcluster Ia containing 4 varieties, 8 breeding lines and 1 PI, and subcluster Ib containing 6 varieties. Cluster II included 1 PI and 1 variety with 89.5% genetic similarity. Cluster III and IV each comprised 2 varieties with 97.4 and 86.8% genetic similarity, respectively. Some correspondence between the clustering pattern and the pedigree of soybean genotypes was observed. In subcluster Ia, KKU 35 (Williams X SJ 2) and CM 60 (Williams X SJ 4) had Williams as a common parent. CM 60 was used as a parent in crosses that led to CM 2 (CM 60 X IAC 13) and ST 3 ((Fort Lamy X CM 60) F_3 X CM 60). CM 2 was also grouped in this subcluster, however, ST 3 was grouped separately in cluster II. All 8 early elite breeding lines (KKU 65, KKU 69, KKU 74, KKU 120, KKU 137, KKU 215, KKU 486, KKU 863) derived from a cross between NS 1 (Doteung X Santa Maria) and KKU 35, were also in this subcluster together with their parents. Two of these sister lines (KKU 65 and KKU 69) were very closely related and were unable to be distinguished genetically. KKU 74 was the most divergent among the breeding lines. Both SJ 4 and SJ 5 in subcluster Ib were progenies from crosses with Tainung 4 from Taiwan as one of the parents (Acadian X Tainung 4 and Tainung 4 X SJ 2, respectively). ST 2 (7016 X ST 1) as well as SJ 2, CM 1 and CB 1 (UFV 1 X Santa Rosa) which were introduced from Taiwan, Japan or Brazil were also grouped in subcluster Ib. LJ 2, the only PI in this cluster, which had long juvenile growth and high yield potential, was the most distantly related to CM 4, ST 1 and Prolina (81.58% similarity). Note that most of these varieties are medium maturing varieties the yield potential of which might be increased if the juvenile growth phase was lengthened. Prolina was the only PI in cluster II. This PI was developed by the USDA-ARS, in cooperation with the North Carolina Agricultural Research Service and has especially high protein content (Burton et al., 1999), which is of particular interest in future breeding programs. This PI appeared to be the most distantly related to CM 3, CM 4 and ST 2 (73.7, 76.3 and 79.0% similarity, respectively). ST 3, the other variety in cluster II was a descendent from the (Fort Lamy X CM 60) F_3 X CM 60 cross. Note that both Fort Lamy and Prolina were developed in the United States (Shanmugasundaram, 1988; Burton et al., 1999). It has been shown that the genetics of soybean accessions and ancestral lines from North America are rather distinct from those from China and Japan (Li et al., 2001; Ude et al., 2003). The two varieties (CM 3 and CM 4) in cluster III derived from the same cross (G 9946 X AGS 17) and were very closely related (97.4% similarity). Two varieties, SJ 1 (progeny of a cross from Japan/Taiwan) and ST 1 (Shih Shih X SRF 400 from Taiwan) formed the last separate cluster IV which appeared to be the most distinct from all others. The fact that varieties with common pedigrees such as CM 3 and CM 4 were grouped in the same cluster as well as the grouping of all 8 breeding lines from a single cross in the same subcluster substantiated the conclusion that the constituent genotypes of a cluster share greater genetic homology. Comparison of clustering patterns among the 10 certified varieties (CM 2, CM 3, CM 60, NS 1, CB 1, SJ 1, SJ 2, SJ 4, SJ 5 and ST 1) which were used in both our and Chotiyarnwong et al.'s (2007) studies using different sets of SSR markers showed that most varieties were clustered differently. While we classified these 10 varieties into four clusters (I, II, III, IV) with two subclusters in cluster I (Ia, Ib), Chotiyarnwong et al. (2007) classified them into 2 groups (1, 2) with four subgroups in group 2 (2b, 2c, 2e, 2h). It was found that although SJ 4, SJ 5 and CB 1 were grouped together in both studies (in subcluster Ib and subgroup 2e, respectively), clustering of the remaining seven varieties was different. In their study NSI (subcluster Ia in our study) were grouped with SJ 2 (subcluster Ib), CM 3 (cluster III) and ST 1 (cluster IV) in subgroup 2b. In contrast, CM 60, NS 1 and CM 2, which were clustered together in subcluster Ia in our study, were grouped separately into subgroup 1a, 2b and 2h, respectively in their study. These differences suggested that the same type of marker may provide different genetic relationships among the same subset of plant genotypes, substantiating the importance of other factors for genetic relationship determination such as composition of the entire set of plant materials, the set and number of markers used as well as the distribution of markers in the genome as has been suggested by Souframanien and Gopalakrishna (2004) and Tantasawat et al. (2010). Maximum genetic distance (0.26) was observed between CM 3 (Cluster III) and Prolina (Cluster II). Only a pair of closely genetically related sister breeding lines (KKU 65 and KKU 69) from the same cross had the same genetic profile at these 11 SSR loci. These breeding lines may be identified if additional SSR markers covering other parts of the genome are used. The remaining twenty-three genotypes could be clearly distinguished from one another. These results demonstrated the effectiveness of SSR markers in variety identification in agreement with previous reports (Chotiyarnwong et al., 2007; Singh et al., 2010; Tantasawat et al., 2010). The cost associated with SSR analysis could also be substantially reduced by multiplexing different primers in the same reaction (Narvel et al., 2000a,b). The genetic relationships among soybean genotypes may facilitate the selection of parents in breeding programs with the hypothesis that the more genetically diverse the parents, the more likely they are to possess unique alleles for traits of interest (Narvel et al., 2000a). However, it must be noted that the extent of SSR diversity may not reflect variability of expressed sequences or genomic regions that influence gene expression. Therefore, designing crosses between parents that are diverse based on a random set of SSR markers may not increase genetic diversity for useful phenotypes (Narvel et al., 2000a). PCoA was used to identify multidimensional relationships that describe portions of the genetic variance in a data set. The first three principal coordinates explained 22.6, 17.7 and 14.2% of the total variance, respectively (Fig. 3). Comparable to the UPGMA cluster analysis, PCoA separated the genotypes into four main groups corresponding to clusters I, II, III and IV. However, clearer separation between the subclusters Ia and Ib was observed. The PCoA also generated a good separation of cluster IV from the rest of the clusters, and further showed that most soybean varieties were intermixed into a large group corresponding to cluster I. Our results confirmed the previous reports (Narvel et al., 2000a,b; Wang et al. 2006a,b; Chotiyarnwong et al., 2007) that SSR is particularly useful for conducting diversity analysis, determining pedigree relationships or genotyping for proprietary purposes in soybean. By means of only the five most polymorphic primer pairs (Soy satt 005, 160, 173, 185, 409), all 15 certified soybean varieties currently used commercially in Thailand could be distinguished. These SSR primers, particularly when used together with multiplex technology, could be useful for accurate and cost-effective genotyping for soybean variety identification. The low genetic diversity among Thai certified soybean varieties revealed in the present analysis emphasizes the need to exploit more diverse germplasm in future breeding programs.

Acknowledgements

This research was supported by a grant from Suranaree University of Technology, Thailand. The authors express their gratitude to Prof. Dr. Paisan Laosuwan for providing soybean seeds and Asst. Prof. Dr. Thitiporn Machikowa for helpful information. We also thank Dr. Dale E. Taneyhill and Mr. Peter Bint for editing the manuscript.

References

Abe J, Xu DH, Suzuki Y, Kanazawa A, Shimamoto Y (2003) Soybean germplasm pools in Asia revealed by nuclear SSRs. Theor Appl Genet 106: 445-453.

Bernard RL, Cremeens CR, Cooper RL, Collins FL, Krober OA, Athow KL, Laviolette FA, Coble CJ, Nelson RL (1998) Evaluation of the USDA soybean germplasm collection: Maturity groups 000 to IV (FC 01.547-PI 266.807). USDA Tech Bull 1844.

Burton JW, Carter Jr TE, Wilson RF (1999) Registration of 'Prolina' soybean. Crop Sci 39: 294-295.

Chotiyarnwong O, Chatwachirawong P, Chanprame S, Srinives P (2007) Evaluation of genetic diversity in Thai indigenous and recommended soybean varieties by SSR markers. Thai J Agri Sci 40: 119-126.

Chowdhury AK, Srinives P, Tongpamnak P, Saksoong P, Chatwachirawong P (2002) Genetic relationship among exotic soybean introductions in Thailand: Consequence for varietal registration. ScienceAsia 28: 227-239.

Cregan PB, Jarvik T, Bush AL, Shoemaker RC, Lark KG, Kahler AL, Kaya N, vanToai TT, Lohnes DG, Chung J, Specht JE (1999a) An integrated genetic linkage map of the soybean genome. Crop Sci 39: 1464-1490.

Cregan PB, Mudge J, Fickus EW, Danesh D, Denny R, Young ND (1999b) Two simple sequence repeat markers to select for soybean cyst nematode resistance conditioned by the *rhg 1* locus. Theor Appl Genet 99: 811-818.

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

Dong YS, Zhao LM, Liu B, Wang ZW, Jin ZQ, Sun H (2004) The genetic diversity of cultivated soybean grown in China. Theor Appl Genet 108: 931-936.

Gizlice Z, Carter Jr TE, Burton JW (1994) Genetic base for North American public soybean cultivars released between 1947 and 1988. Crop Sci 34: 1143-1147.

Gizlice Z, Carter Jr TE, Burton JW (1996) Genetic diversity patterns in North American public soybean cultivars based on coefficient of parentage. Crop Sci 36: 753-756.

Griffin JD, Palmer RG (1995) Variability of thirteen isozyme loci in the USDA soybean germplasm collections. Crop Sci 35: 897-904.

Guan R, Chang R, Li Y, Wang L, Liu Z, Qiu L (2010) Genetic diversity comparison between Chinese and Japanese soybeans (*Glycine max* (L.) Merr.) revealed by nuclear SSRs. Genet Resour Crop Evol 57: 229-242.

Juvik GA, Bernard RL, Orf JH, Cavins JF, Thomas DI (1989) Evaluation of the USDA soybean germplasm collection: Maturity groups 000 to IV (PI 446.983 to PI 486.355). USDA Tech Bull 1760.

Karakaya HC, Tang Y, Cregan PB, Knap HT (2002) Molecular mapping of the fasciation mutation in soybean, *Glycine max* (Leguminosae). Amer J Bot 89: 559-565.

Keim P, Beavis W, Schupp J, Freestone R (1992) Evaluation of soybean RFLP marker diversity in adapted germplasm. Theor Appl Genet 85: 205-212.

Khan R, Khan H, Farhatullah, Harada K (2010) Evaluation of microsatellite markers to discriminate induced mutation lines, hybrid lines and cultigens in chickpea (*Cicer arietinum* L). Aust J Crop Sci 4: 301-308.

Kumar J, Verma V, Goyal A, Shahi AK, Sparoo R, Sangwan RS, Qazi GN (2009) Genetic diversity analysis in *Cymbopogon* species using DNA markers. Plant Omics J 2: 20-29.

Li Y-H, Li W, Zhang C, Yang L, Chang R-Z, Gaut BS, Qiu L-J (2010) Genetic diversity in domesticated soybean (*Glycine max*) and its wild progenitor (*Glycine soja*) for simple sequence repeat and single-nucleotide polymorphism loci. New Phytol 188: 242-253.

Li Z, Nelson RL (2001) Genetic diversity among soybean accessions from three countries measured by RAPDs. Crop Sci 41: 1337-1347.

Li Z, Qiu L, Thompson JA, Welsh MM, Nelson RL (2001) Molecular genetic analysis of U.S. and Chinese soybean ancestral lines. Crop Sci 41: 1330-1336.

Lorenzen LL, Boutin S, Young N, Specht JE, Shoemaker RC (1995) Soybean pedigree analysis using map-based markers: I. Tracking RFLP markers in cultivars. Crop Sci 35: 1326-1336.

Machikowa T, Laosuwan P (2009) Effects of extension of days to flowering on yield and other characters of early maturing soybean. Suranaree J Sci Technol 16: 169-174.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27: 209-220.

Maughan PJ, Saghai Maroof MA, Buss GR (1995) Microsatellite and amplified sequence length polymorphism in cultivated and wild soybean. Genome 38: 715-723.

Maughan PJ, Saghai Maroof MA, Buss GR, Huestis GM (1996) Amplified fragment length polymorphism (AFLP) in soybean: Species diversity, inheritance, and near-isogenic line analysis. Theor Appl Genet 93: 392-401.

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30: 194-200.

Narvel JM, Fehr WR, Chu WC, Grant D, Shoemaker RC (2000a) Simple sequence repeat diversity among soybean plant introductions and elite genotypes. Crop Sci 40: 1452-1458.

Narvel JM, Chu WC, Fehr WR, Cregan PB, Shoemaker RC (2000b) Development of multiplex sets of simple sequence repeat DNA markers covering the soybean genome. Mol Breeding 6: 175-183.

Nelson RL, Amdor PJ, Orf JH, Cavins JF (1988) Evaluation of the USDA soybean germplasm collection: Maturity groups 000 to IV (PI 427.136 to PI 445.845). USDA Tech Bull 1726.

Nelson RL, Amdor PJ, Orf JH, Lambert JW, Cavins JF, Kleiman R, Laviolette FA, Athow KA (1987) Evaluation of the USDA soybean germplasm collection: Maturity groups 000 to IV (PI 273.483 to PI 427.107). USDA Tech Bull 1718.

Office of Agricultural Economics (2008) Situation and trend of important agricultural products. Ministry of Agriculture and Cooperatives. http://www.doae.go.th/E-Book/trend2551.pdf.

Perry MC, McIntosh MS (1991) Geographical patterns of variation in USDA soybean germplasm collection: I. Morphological traits. Crop Sci 31: 1350-1355.

Pookpakdi A (2003) Soybean: Golden Plant of Thailand. Kasesart University Press, Bangkok (in Thai).

Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. Trends Plant Sci 1: 215-222.

Rohlf FJ (2000) NTSYS-pc: Numerical taxonomy and multivariate analysis system, version 2.2. Exeter Software, Setauket, New York, USA.

Rongwen J, Akkaya MS, Bhahwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor Appl Genet 90: 43-48.

Saghai Maroof K, Soliman M, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81: 8014-8018.

Sambrook J, Russell DW (2001) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York.

Seitova AM, Ignatov AN, Suprunova TP, Tsvetkov IL, Deineko EV, Dorokhov DB, Shumnyi VK, Skryabin KG (2004) Genetic variation of wild soybean *Glycine soja* Sieb. et Zucc. in the far east region of the Russian Federation. Russ J Genet 40: 165-171.

Shanmugasundaram S (1988) Leguminous vegetable cultivation and seed production. http://203.64.245.61/fulltext_pdf/eam 0149.pdf.

Singh RK, Mishra SK, Singh SP, Mishra N, Sharma ML (2010) Evaluation of microsatellite markers for genetic diversity analysis among sugarcane species and commercial hybrids. Aust J Crop Sci 4: 116-125.

Smith JR, Carpten JD, Brownstein MJ, Ghosh S, Magnuson VL, Gilbert DA, Trent JM, Collins FS (1995) Approach to genotyping errors caused by non-templated nucleotide addition by *Taq* DNA polymerase. Genome Res 5: 312-317.

Sneller CH (1994) Pedigree analysis of elite soybean lines. Crop Sci 34: 1515-1522.

Souframanien J, Gopalakrishna T (2004) A comparative analysis of genetic diversity in blackgram genotypes using RAPD and ISSR markers. Theor Appl Genet 109: 1687-1693.

Sutakom S (2004) Rajamangala 1: The soybean line with high potential for upper Northern region of Thailand. Paper presented at the 9th National Soybean Research Conference, Lampang, Thailand, 3-4 March 2004 (in Thai).

Tantasawat P, Trongchuen J, Prajongjai T, Seehalak W, Jittayasothorn Y (2010) Variety identification and comparative analysis of genetic diversity in yardlong bean (*Vigna unguiculata* spp. *sesquipedalis*) using morphological characters, SSR and ISSR analysis. Sci Hort 124: 204-216.

Tanya P, Srinives P, Toojinda T, Vanavichit A, Ha B-K, Bae J-S, Moon J-K, Lee S-H (2001) Evaluation of genetic diversity among soybean genotypes using SSR and SNP. Korean J Crop Sci 46: 334-340.

Tara Satyavathi C, Bhat KV, Bharadwaj C, Tiwari SP, Chaudhury VK (2006) AFLP analysis of genetic diversity in Indian soybean [*Glycine max* (L.) Merr.] varieties. Genet Res Crop Evol 53: 1069-1079.

Tar'an B, Zhang C, Warkentin T, Tullu A, Vandenberg A (2005) Genetic diversity among varieties and wild species accessions of pea (*Pisum sativum* L.) based on molecular markers, and morphological and physiological characters. Genome 48: 257-272.

Thompson JA, Nelson RL (1998) Utilization of diverse germplasm for soybean yield improvement. Crop Sci 38: 1362-1368.

Thompson JA, Nelson RL, Vodkin LO (1998) Identification of diverse soybean germplasm using RAPD markers. Crop Sci 38: 1348-1355.

Ude GN, Kenworthy WJ, Costa JM, Cregan PB, Alvernaz J (2003) Genetic diversity of soybean cultivars from China, Japan, North America, and North American ancestral lines determined by amplified fragment length polymorphism. Crop Sci 43: 1858-1867.

Wang L, Guan R, Zhangxiong L, Chang R, Qiu L (2006a) Genetic diversity of Chinese cultivated soybean revealed by SSR markers. Crop Sci 46: 1032-1038.

Wang L, Guan Y, Guan R, Li Y, Ma Y, Dong Z, Liu X, Zhang H, Zhang Y, Liu Z, Chang R, Xu H, Li L, Lin F, Luan W, Yan Z, Ning X, Zhu L, Cui Y, Piao R, Liu Y, Chen P, Qiu L (2006b) Establishment of Chinese soybean (*Glycine max*) core collections with agronomic traits and SSR markers. Euphytica 151: 215-223.

Wang M, Li R-Z, Yang W-M, Du W-J (2010) Assessing the genetic diversity of cultivars and wild soybeans using SSR markers. Afr J Biotechnol 9: 4857-4866.