Journal of Plant Physiology and Breeding

2014, 4(2): 55-65 ISSN: 2008-5168

Evaluation of Genetic Diversity among Soybean (*Glycine max*) Genotypes, Using ISJ and RAPD Molecular Markers

Mozhgan Panjoo, Farhad Nazarian-Firouzabadi, Ahmad Ismaili and Hadi Ahmadi

Received: October 19, 2012 Accepted: December 10, 2014
Agronomy and Plant Breeding Department, Faculty of Agriculture, University of Lorestan, P.O. Box 465, Khorramabad, Iran
*Corresponding author: E-mail: Nazarian.f@lu.ac.ir

Abstract
Soybean (*Glycine max*) is an important crop plant which contains a high amount of oil and protein. To evaluate the genetic diversity of 45 soybean genotypes growing in Iran, two types of molecular markers (RAPD and ISJ) were used. Two sets of different primers were used to compare potential differentiation power of IT (Intron targeting) and ET (Exon targeting) primers. Data obtained from each molecular marker was analyzed separately and in combination. Ten RAPD primers amplified 103 scoreable bands from which 60% were polymorphic, whereas 15 ISJ primes resulted to 129 sharp bands from which 87% were polymorphic. Cluster analysis was carried out based on simple matching coefficient of similarity and UPGMA method. The cophenetic coefficient for RAPD and ISJ markers were 0.95 and 0.81, respectively. Cluster analysis revealed that ISJ molecular markers were better in genetic diversity studies than RAPD markers in soybean genotypes. Furthermore, significant variation was not found between two groups of IT and ET primers ($P \leq 0.01$). Results of this study suggested that ISJ molecular markers especially ET primes are reliable tools to investigate genetic diversity among soybean genotypes**.**

Keywords: Genetic Diversity; *Glycine max*; ISJ; RAPD; Soybean

Introduction

Soybean (*Glycine max*) is an important oil seed containing a high amount of oil (18%-22%) and even higher amount of protein (40%-50%) (Smith and Huyser 1987). Soybean oil contains about 85% unsaturated fatty acids and it is cholesterol free. In addition, it contains macro-nutrients required for good nutrition, complete proteins, carbohydrates and fat, as well as vitamins and minerals, including calcium, folic acid and iron (Poehlman and Sleper 2006). Therefore, it is very useful in human diet (Aslam *et al.* 1995). It is estimated that more than 163 important products are extracted from the soybean seeds (Rastgar 2006).

Genetic diversity is defined as the biodiversity within species and it is the foundation of plant breeding. Genetic diversity is an essential requirement of plant breeding and the basis of genetic development (Ramanujam *et al*. 1974; Hallauer and Miranda, 1988; Kearsey and Pooni 1998). Nearly all plant breeding procedures aimed at selecting superior plants within a diverse population. Therefore, genetic diversity plays a pivotal role in this regard (Powell *et al*. 1996). Arrays of different markers such as morphological characters, proteins, isozymes and molecular markers have been used in genetic diversity studies (Pejic *et al*. 1998).

Although it might be possible to distinguish plant cultivars released from conventional plant breeding methods by morphological traits, these characters due to high similarities among cultivars derived from the same elite parental group cannot effectively be used to classify closely related plant species (Rodrigues *et al*. 2008). To this end, molecular markers seem to play a pivotal role to study genetic diversity (Lakshmikumaran 2000).

These markers are highly efficient for genetic diversity studies, identification of cultivars, phylogenic analyses, ecology and identification of the same cultivars with different names (Kumar 1999). To the best knowledge, several molecular markers have so far been used to evaluate soybean genetic diversity and the advantage of using such molecular markers has been proved (Doldi *et al*. 1997). Random amplified polymorphic DNA (RAPD) markers are widely used to assess genetic diversity among different plant populations. For instance, Xu and Gai (2003) showed that wild soybean had a higher genetic diversity in comparison to cultivated soybean. Although RAPD as a random molecular marker possess many advantages, its low repeatability limits its application. To overcome this drawback, Weining and Langridge (1991) introduced Intron-Exon Splice Junction (ISJ) semi random marker in cereals. ISJ primers are designed based on conserved sequences found in Intron-Exon boundaries. Moreover, a number of nucleotides in an ISJ primer are completely designed randomly to increase the quality of amplified bands and to avoid band smearing (Weining and Langridge 1991). In addition to the higher accuracy, ISJ markers show a higher polymorphism in comparison to RAPD markers (Jaroslaw *et al*. 2002). The present research aims to study the genetic relationships among 45 soybean genotypes released by Iranian National Oilseed Center, using both RAPD and ISJ markers.

Materials and Methods

Forty five soybean genotypes with different maturity type (I-VIII types) originated from USA and Canada were used in this study (Table 1).

Table 1. Soybean genotypes utilized to assess genetic diversity by ISJ and RAPD molecular marker

Genotypes	
Chippewa X	Pickett
A100	Roanoke
Lindarin	Coker
Chippewa	UK150
Faur X	Dortchsoy 2A
Ba 608	Hill CE
Wilson	T3
Kanrich	Crowford
Mandegar	Forrest
Harosoy 2	Introduction(in)
Chippewa 64	UK holandi
Habbit	Cocer2
A3237	Cokerhamton 266
A3935	Ayora
Anava	Wayne
Faur	Roanoke 2
Clark 63	Dorsan
Bethel	SRF
Dair	USA 8/1
Bossier	Lee
Hill	Davis
Columbus	Gorgan 3
	Games

DNA isolation

Total genomic DNA was extracted from soybean genotype leaves, using CTAB method (Williams *et al*. 1990). The quality and quantity of extracted DNA were measured by electrophoresis and spectrophotometry methods.

RAPD analysis

Twenty five micro liter reaction mixture contained 10ng genomic DNA, 1× reaction buffer **(**100mM KCL, 20 mM Tris HCL pH: 8, 0.5% Tween-20, 50% Glycerol, 0.1mM DDT), 2 mM $MgCl_2$, 0.2mM dNTPs mixture, 0.4 ppm random primer (Table 2) and 0.04 U Taq DNA polymerase enzyme (Sinagene, Tehran, Iran). PCR cycling was carried out in a DNA thermal cycler (Master Cycle, gradient 5331) programmed to run the following temperature profile: one time pre-denaturation for 5 minutes at 94°C, 40 cycles of 1 minute denaturation at 94°C, 30s annealing at 37.5°C and 2 minutes extension at 72°C, following 10 minutes final extension at 72°C.

Table 2. RAPD and ISJ fingerprints data for the molecular characterization of soybean genotypes. ET (Exon targeting), IT (Intron targeting)

Primers		Primer sequence (5'-3')	Reference	Polymorphism (%)
RAPD	B06	TGCTCTGCCC	Trindade et al., 2008	75
	B15	GGAGGGTGTT	Trindade et al., 2008	61
	D10	GGTCTACACC	Trindade et al., 2008	50
	D13	GGGGTGACGA	Trindade et al., 2008	67
	D15	CATCCGTGCT	Trindade et al., 2008	69
	E19	ACGGCGTATG	Trindade et al., 2008	61
	D03	GTCGCCGTCA	Trindadeet al., 2008	78
	E20	AACGGTGACC	Trindadeet al., 2008	50
	OPH 16	TCTCAGCTGG	Sunaret al., 2009	36
	OPB5	TGCGCCCTTC	Sunar et al., 2009	50
ET	ET 12-28	AGCAGGTCGAAG	Przetakiewiczet al., 2002	100
	ET 12-29	AGCAGGTCGTGA	Przetakiewiczet al., 2002	100
	ISJ 10	ACTTACCTGCATCCCCCT	Gawel*et al*., 2002	100
	ET 18-1	ACTTACCTGAGGCGCGAC	Sawicki and Szczecinska2007	57
	ET 15-35	ACTTACCTGCCGCAG	Przetakiewiczet al., 2002	89
	ET 15-31	ACTTACCTGGGCCAG	Przetakiewiczet al., 2002	57
	ET 15-32	ACTTACCTGGGCACG	Przetakiewiczet al., 2002	100
	ET 15-36	ACCTACCTGGGGCTC	Gawelet al., 2002	100
	ET 15-33	ACTTACCTGGCCGTG	Przetakiewiczet al., 2002	89
IT	IT 18-2	GCAGAGGGCCAGGTAAGT	Gawelet al., 2002	57
	IT 15-35	CGGCATCAGGTAAG	Gawelet al., 2002	100
	ISJ 4	GTCGGCGGACAGGTAAGT	Sawicki and Szczecinska 2007	100
	ISJ 8	GACCGCTTGCAGGTAAGT	Sawicki and Szczecinska 2007	67
	IT 15-32	GACTCGCCAGGTAAG	Gawelet al., 2002	100
	IT 15-31	GAAGCCGCAGGTAAG	Gawelet al., 2002	57

ISJ analysis

PCR reaction for ISJ markers was performed in two steps according to Williams *et al*. (1990). Low temperature in a few PCR cycle provides an opportunity to reproduce appropriate fragments, then the annealing temperature of the primer increases slightly to increase the production capacity of band avoiding band smearing. Therefore, after denaturation of sample DNA at 95 °C, 7 cycles were performed at 95 °C for 1 min, primer annealing at T_m+2 for 45 seconds and extension for 2 min at 72 °C following by 32 cycles at 95 °C for 1 min, T_m+6 annealing for 45 second and 72 °C for 2 min, and 72 °C final extension for 10 min. PCR master mix was prepared in a 55 micro liters volume including 1× KCL incomplete reaction buffer (500 mM KCL, 100 mM Tris HCL pH 8.8, 0.1% Tween-20, 25 mM $MgCl_2$), 0.2 mM dNTPs mixture, 0.3 pmol random primer (Table 2) and 0.05 U Taq DNA polymerase enzyme (Bioron, Germany). PCR products were electrophoresed on 1.5 % agarose gels in 1× TAE buffer at 100V.

Data analysis

Amplified fragments were scored as 10 for presence and absent of bands, respectively. NTSYS-pc 2.1 software was used for the cluster analysis (Rohlf 1997). Three similarity coefficient matrix including Dice (Dice 1945), Jakard (Gower 1971) and simple matching (Sneath and Sokal 1973) were prepared. Simple matching method measures the amount of similarity between different pairs (Dij= 1-Sij= 1-(a+b/a+b+c+d)). In which "a" is the number of 1-1 modes, "b" and "c" are the number of 0-1 and 1-0 modes and "d" is the number of 0-0 modes, respectively. Similarity matrix was analyzed as a cluster diagram, using NTSYS-pc computer software. Algorithm and the method of integration of pair groups were calculated, using UPGMA method in which the integration was made through connection of group arithmetic mean.

Results and Discussion

Ten RAPD primers (Table 2) produced a total number of 103 scoreable bands including 62 polymorphic and 41 monomorphic bands. The percentage of polymorphism bands among primers ranged from 36% to 78% with average of 60%. The total average figures of polymorphic and monomorphic bands for individual RAPD primer were 10.3 and 6.2, respectively. B15, E19 and D15 primers each with 13 amplified bands and D10 and OPY10 primers each with 6 bands produced the highest and the lowest polymorphic bands, respectively (Table 2). As it can be seen from Figure 1A, there is no clear difference among genotypes with respect to number of polymorphic bands.

We also used 15 semi-random primers in the present study which were chosen based on previous studies (Table 2). Bands appeared between 150 bp to 2 kb were considered and coded. A total number of 129 scoreable and reproducible bands were scored, of which 113 (87%) and 16 (12%) bands were polymorphic and monomorphic, respectively. The average polymorphic bands for ISJ primer was 8.6. IT15-35 and ISJ4 primers each with 13 amplified bands produced the highest and few ISJ primers each with 6 bands produced the lowest polymorphic

bands, respectively (Table 2). Results of this study showed that ISJ primers produced 27% more polymorphic bands among genotypes than RAPD primers. As it can be seen from Figure1B, there is a high genetic diversity among genotypes considering the pattern of reproduced bands.

Difference between ET and IT primers

ISJ primers consisted of 6 IT(Intron Targeting) and 9 ET (Exon Targeting) primers. IT group primers produced a total number of 56 bands, including 47 polymorph (86%) and 8 monomorph (14%) bands. Nine ET primers produced a total number of 73 bands, including 65 polymorph (89%) and 8 monomorph (11%) bands. IT and ET primers produced 8 and 7.22 polymorph bands. No significant difference (P>0.01) between two groups of primers was observed by t-test. However, IT primers produced more clear bands than ET primers. This may be due to the fact that the 3' parts of IT primers is designed based on conserve sequence boundaries between Intron and Exon regions, whereas in ET primers the 5' part of the primers is conserve which may well influence the quality of amplified bands (Weining and Langridge 1991).

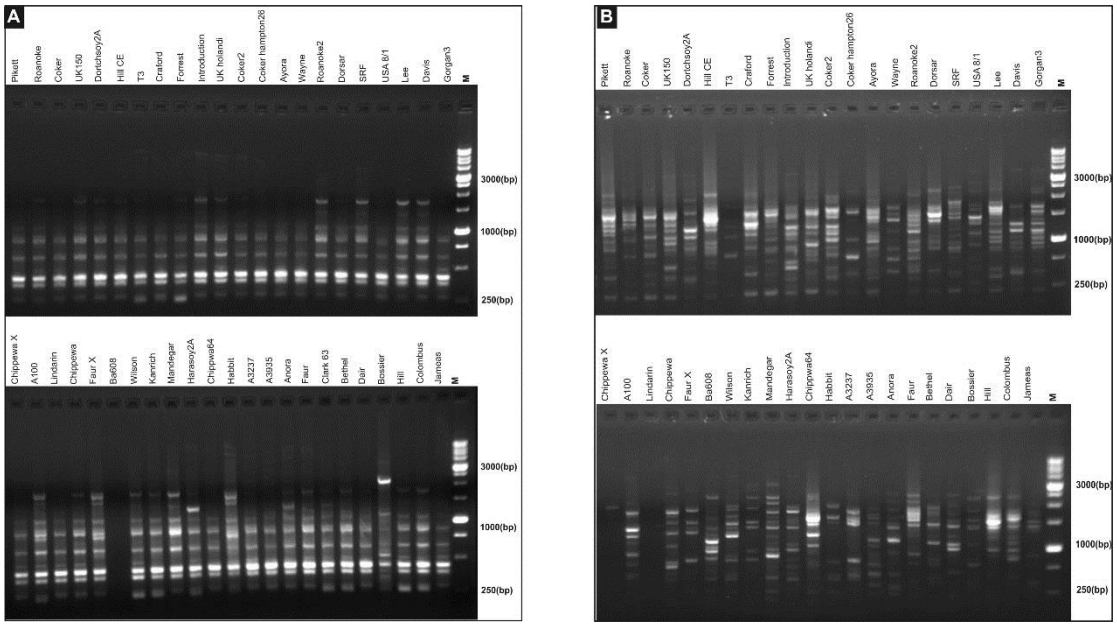

Figure 1. OPH16 (A) and IT15-35 (B) primer products on a 1.5% agarose gel. M: 1Kb leader marker; C-: negative control

Marker analysis

In order to evaluate the genetic distance in autogamous plants such as soybean, it is often suggested to use simple matching method (Correa *et al*. 1999). In this study, cluster analysis was performed based on both simple matching and UPGMA algorithm methods. The cophenetic coefficient values for RAPD and ISJ markers were 0.95 and 0.81, respectively. The similarity coefficient of genotypes for RAPD markers ranged from 0.7 to 0.99 (Figure 2A). In line with few other studies, dendrogram revealed no clear differentiation of soybean genotypes due to high similarity among them (Devos and Gale 1992;

Powell *et al*. 1996). In spite of established efficiency of RAPD markers in some crop plants, RAPD primers were not able to effectively differentiate soybean genotypes in this study. High homozygosity and weak gene pool (Doldi *et al.* 1997) may explain these results. Powell *et al* (1996) showed that RAPD markers in comparison to other DNA markers, underestimates genetic distances among species (Devos and Gale 1992; Powell *et al*. 1996). Accordingly, the researchers do not recommend using this marker for taxonomical studies.

The similarity coefficient of genotypes for ISJ markers ranged from 0.59 to 0.83 (Figure 2B). The highest similarity coefficient (84%) was found between SRF and Ayora as well as Lee and Davis genotypes. Our results showed that Cocer2 and A3237 with 47% similarity coefficient had the lowest similarity among all 45 genotypes. As it can be seen from Figure 2B, ISJ markers classified the soybean genotypes in seven groups. Interestingly, comparison between RAPD and ISJ markers showed that ISJ markers had a relatively higher differentiation power than RAPD markers, indicating that ISJ markers to a great extent, could reveal higher level of polymorphism among genotypes.

Although according to various reports, RAPD markers are not reliable to identify polymorphs, results of this study were consistent with soybean maturity stage grouping. Doldi *et al*. (1997) integrated RAPD and SSR markers data in soybean and found that integrated information was much better consistent with genealogy information. In a study involving 21 Iranian and foreign soybean cultivars, researchers came to the conclusion that DAF markers identified higher relationships among cultivars than RAPD markers. Furthermore, integration of DAF and RAPD data showed that Iranian cultivars had the highest similarity with each other than foreign counterparts (ShahNejat Boushehri 2002).

To compare the efficiency of two types of ISJ primers, IT and ET dendrograms were separately generated and compared (Figure 3). Based on this comparison, ET dendrogram showed better similarity to the integrated ISJ markers than IT dendrogram (data not shown).When four dendrograms were verified together, it was evident that introduction genotype (In) with unknown name may have been obtained from a cross between Dortchsoy 2A and Croeford genotypes. Furthermore, Habbit, Faur and Hill genotypes tended to group together as a single group in all four dendrograms. Interestingly, our conclusion further strengthened the reports of other researchers (ShahNejat Boushehri *et al.* 2000; ShahNejat Boushehri 2002). Each of the Boissier and A100 genotypes formed a separate group in most dendrograms. From morphology

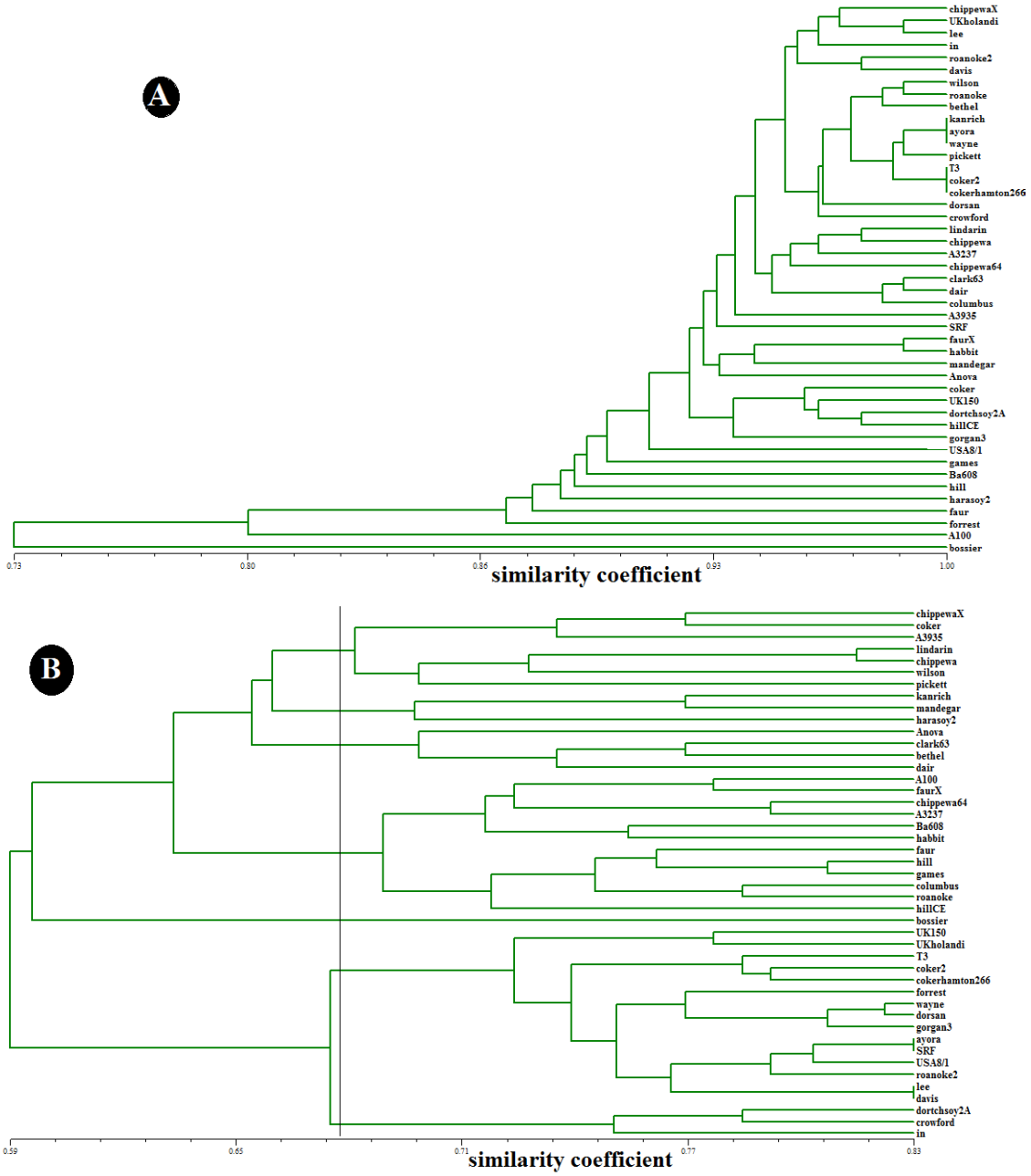

Figure 2. Dendrogram constructed for 45 soybean genotypes, using UPGMA method based on simple matching coefficient of similarity for RAPD (A) and ISJ (B) primers

point of view, the former genotype has white flower and a high bush and it is a late-maturing cultivar (maturity 7), whereas the other genotype is from early maturity group (maturity 1). We think that the results of this study can be used in selection and hybridization programs. To this end

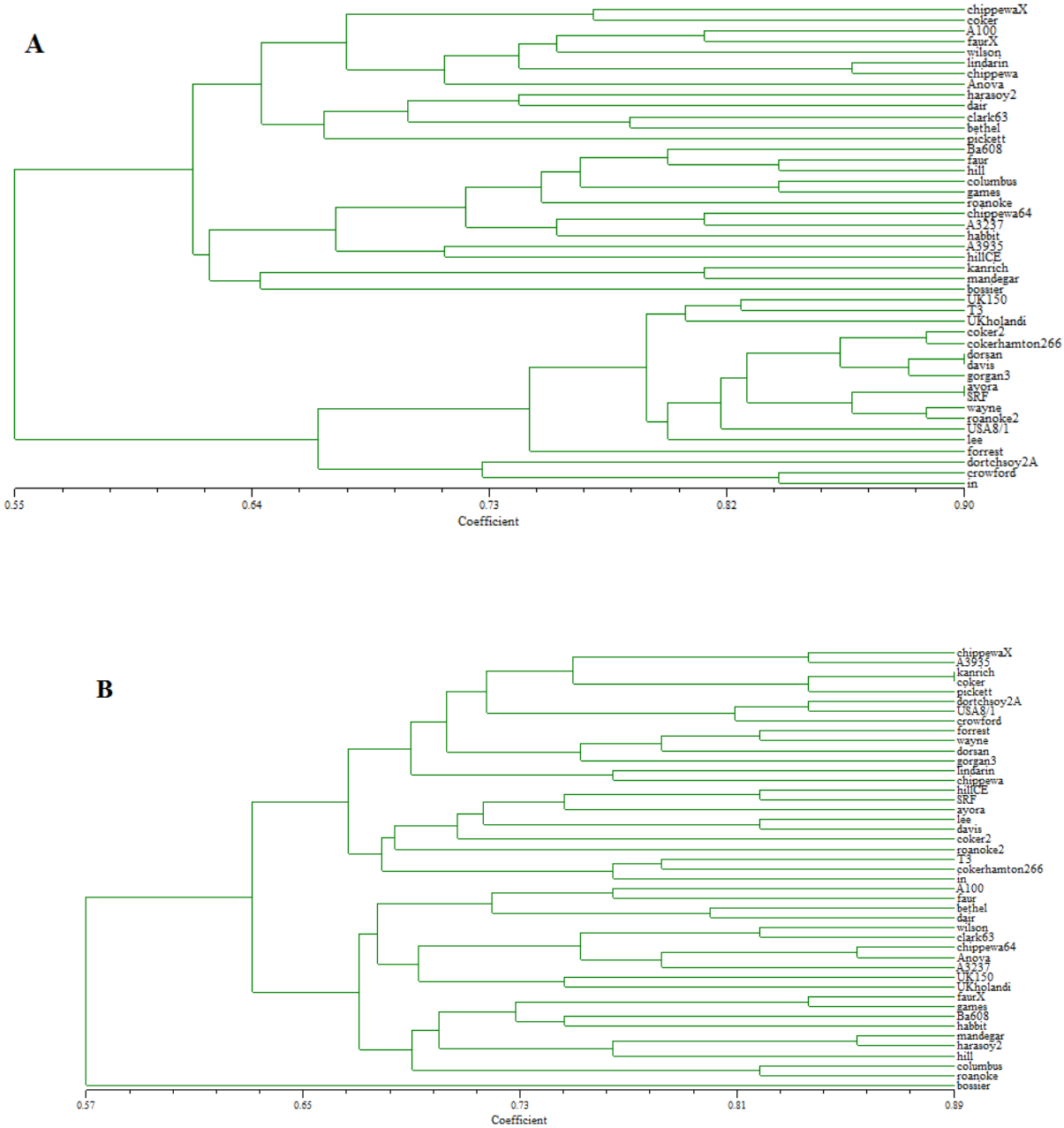

Figure 3. UPGMA tree plot of 45 soybean genotypes constructed based on simple matching coefficient of similarity for ET (A) and IT (B) primers

we recommend Bossier as a parent for future hybridization breeding programs.

Principal component analysis

The main purpose of using principal component method here was actually to determine how the markers are distributed across the whole genome (Vollmann *et al*. 2005). The two-dimensional scatter diagram for all soybean genotypes (RAPD: Figure 4A and ISJ: Figure 4B) was obtained by

the principal component analysis. Accordingly, the contribution of the first three main components was measured in all analyses. The total contribution of the first three components for RAPD, IT and ET markers was 38.72, 28.7 and 40.11, respectively. This figure for the integrated markers was 29.02 (Table 3). ET primers had the largest component contribution, indicating that this marker had lower distribution across the genome. This is consistent with the fact that ET primers specifically amplify gene coding sequences which covers approximately 7% of the whole genome (Udvardi *et al*. 2007). Nevertheless, the lower contribution of IT primers components indicated that this group of markers is widely distributed across the soybean genome.

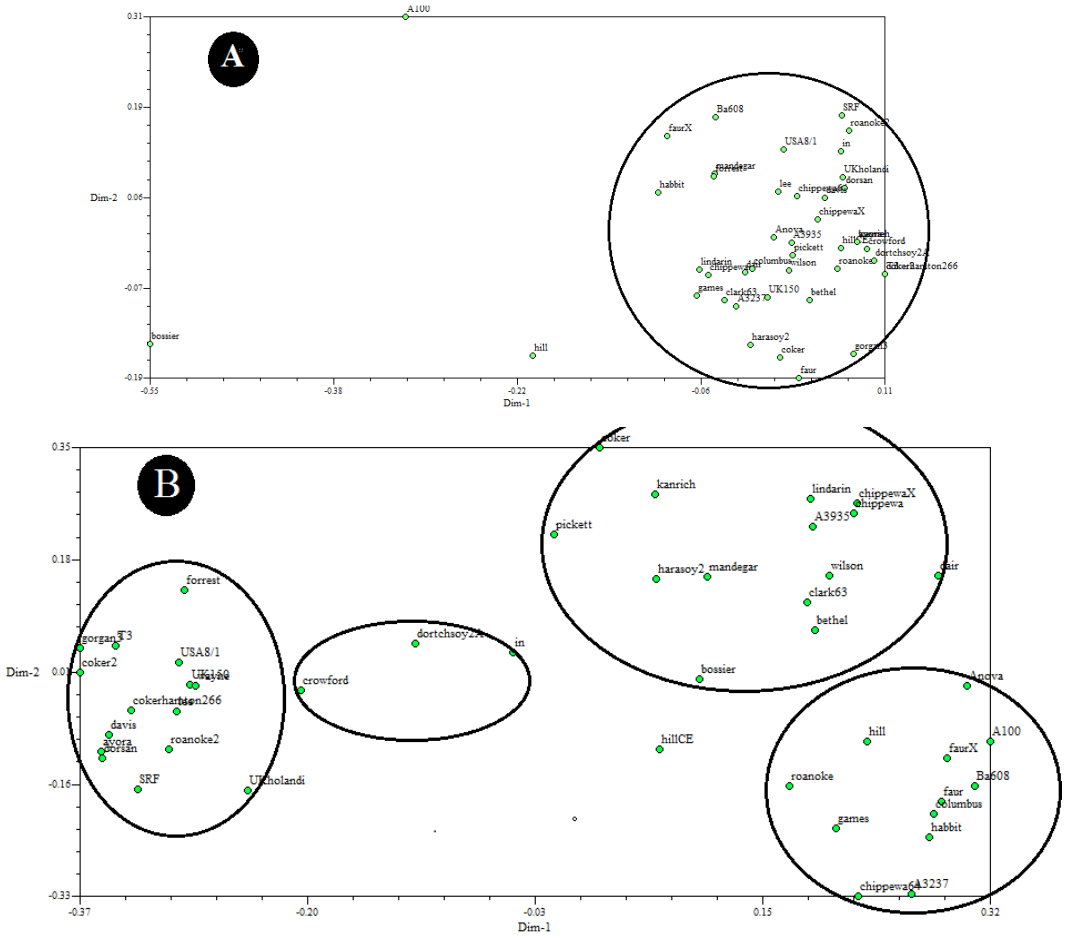

Figure 4. Two-dimensional diagram related to principal component analysis for 45 soybean genotypes based on simple matching coefficient of similarity. A) RAPD and B) ISJ

Table 3. Principal component analysis of 45 soybean genotypes

Components	RAPD	ISJ	IT	ET	RAPD and ISJ
One	**15.73**	**17.75**	**13.4**	**23.74**	**16.3**
Two	**12.9**	**7.75**	**8.8**	**9.32**	**7.01**
Three	**10.09**	**6.28**	**6.5**	**7.06**	**5.89**
Total	**38.72**	**31.78**	**28.7**	**40.11**	**29.02**

Conclusion

The comparison between the number of random polymorphic markers and semi-random markers showed that ET and IT primers identify about a total of 29% and 26% more polymorphism among genotypes than RAPD primers, respectively. This demonstrates that ET and IT primers are more efficient for evaluation of genetic diversity among soybean genotypes. Based on the dendrogram comparison, these primers also classified a wider range of genotypes. However, there were differences between these two groups of primers for the cluster analysis and principal component analysis. Although based on comparison of genetic diversity, no significant difference was found between reproduced segments by ET and IT primers, they showed different similarity coefficients. Furthermore, ET group primers showed higher as compared with total ISJ data. By combining the information of random and semi- random primers, due to weakness of random primers, it can be seen that the range of dendrogram and the amount of similarity between groups were decreased and it seems that this type of classification is more elaborative than information of IT and ET primers, individually.

References

Aslam M, Mirza MS, Shah SM, Shafeeq S and Naeem U, 1995. Crop Production Bulletin No. 6, Pak. Agric. Res. Council, Islamabad, Pakistan.

Correa RX, Abdelnoor RV, Faleiro FG, Cruz CD, Moreira MA and Barros EGD, 1999. Genetic distances in soybean based on RAPD markers. Bragantia 58: 15-22.

Devos KM and Gale MD, 1992. The use of random amplified polymorphic DNA markers in wheat. Theoretical and Applied Genetics 84: 567-572.

Dice LR, 1945. Measures of the amount of ecologic association between species. Ecology 26: 297-302.

Doldi ML, Volmann J and Lelly T, 1997. Genetic diversity in soybean as determined by RAPD and microsatellite analysis. Plant Breeding 116: 331-335.

Gawel M, Wisniewska I and Rafalski A, 2002. Semi-specific PCR for the evaluation of diversity among cultivars of wheat and triticale. Cellular and Molecular Biology Letters 7: 577-582.

Gower JC, 1971. A general coefficient of similarity and some of its properties. Biometrics 27: 857-874.

Hallauer AR and Miranda JB, 1988. Quantitative Genetics in Maize Breeding. Iowa State University Press, Ames, Iowa.

Jaroslaw N, Wieslaw P and Dorota N, 2002. Molecular research on the genetic diversity of Polish varieties and landraces of *Phaseolus coccineus* L. and *Phaseolus vulgaris* L. using the RAPD and AFLP methods. Cellular and Molecular Biology Letters 7: 753-762.

Kearsey, MJ and Pooni HS, 1998. The Genetical Analysis of Quantitative Traits, Stanley Thornes (Publishers) Ltd.

Kumar LS, 1999. DNA marker in plant improvement: An overview. Biotechnology Advances 17: 143-182.

Lakshmikumaran M, 2000. Application of molecular markers for assessment of genetic diversity. Proceedings of National Symposium on Prospects and Potentials of Plant Biotechnology in India in the 21st Century and 23rd Annual Meeting of PTCA(I), Oct. 18-21, Jodhpur, pp. 188.

Min W, Run-Zhi L, Wan-Ming Y and Wei-Jun D, 2010. Assessing the genetic diversity of cultivars and wild soybeans using SSR markers. African Journal of Biotechnology 9(31): 4857-4866.

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G and Motto M, 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. Theoretical and Applied Genetics 97: 1248-1255.

Poehlman J and Sleper D, 2006. Breeding Field Crops. 5th edition. Iowa State University Press, Ames, Iowa.

Powell W, Machray GC and Provan J, 1996. Polymorphism revealed by simple sequence repeat. Trends in Plant Science 1:.215-222.

Przetakiewicz J, Nadoolska-Orcezyk A and Orczyk W, 2002. The use of RAPD and semi-random markers to verify somatic hybrids between diploid lines of *Solanum tuberosum* L. Cellular and Molecular Biology Letters 7: 671-676.

Ramanujam S, Tiwari AS and Mehra RB, 1974. Genetic divergence and hybrid performance in mungbean. Theoretical and Applied Genetics 45: 211-214.

Rodrigues DH, Alcântaraneto FD and Schuster I, 2008. Identification of essentially derived soybean cultivars using microsatellite markers. Crop Breeding and Applied Biotechnology 8: 74-78.

Rohlf FJ, 1997. NTSYS.pc numerical taxonomy and multivariate analysis system version 2.1. Setauket, NY.

Sawicki J and Szczecinska M, 2007. Semi-specific intron-exon splice junction markers in *bryophyte* studies. Biodiversity Research and Conservation 58: 25-30.

ShahNejat Boushehri A, 2002. Study of Soybean genetic diversity using DAF and RAPD analysis. Iranian Agriculture Science Journal 34: 625-633 (In Persian with English Abstract).

ShahNejat Boushehri A, Yazdi-samadi B and Abdmishayi S, 2000. Study of soybean genetic diversity with using DAF analysis. Iranian Agriculture Science Journal 31: 535-545 (In Persian with English Abstract).

Smith KJ and Huyser W, 1987. World distribution and significance of Soybean. In: Wilcox JR

Sneath PHA and Sokal RR, 1973. Numerical Taxonomy. Freeman, San Francisco,

Sunar S, Aksakal O, Yildirim N, Agar G, Gulluce M and Sahin F, 2009. Genetic diversity and relationships detected by FAME and RAPD analysis among *Thymus* species growing in eastern Anatolia region of Turkey. Romanian Biotechnology Letters 14(2): 4313-4318.

Trindade H, Costa MM, Pedro LG, Figueiredo AC and Barroso JG, 2008. Genetic diversity and chemical polymorphism of *Thymus caespititius* from Pico, SaoJorage and Terceira islands (Azores). Biochemistry and System Ecology 36: 190-197.

Udvardi MK, Kakar K, Wandrey M, Montanari O, Murray J, Andriankaja A, Zhang J-Y, Benedito V, Hofer JM and Chueng F, 2007. Legume transcription factors: global regulators of plant development and response to the environment. Plant Physiology 144: 538-549.

Vollmann J, Grausgruber H, Stif G, Dryzhyruk V and Lelley T, 2005. Genetic diversity in camelina germplasm as revealed by seed quality characteristics and RAPD polymorphism. Plant Breeding 124: 446-453.

Weining S and Langridge P, 1991. Identification and mapping of polymorphisms in cereals, based on polymerase chain reaction. Theoretical and Applied Genetics 82: 209-216.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA and Tingey SV, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18: 6531-6535.

Xu DH and Gai JY, 2003. Genetic diversity of wild and cultivated soybeans growing in China revealed by RAPD analysis. Plant Breeding 122: 503-506.