Evaluation of Genetic Diversity in 19 *Glycine max* (L.) Merr. Accessions Included in the Czech National Collection of Soybean Genotypes

Miroslav Baránek, Miroslav Kadlec, Jana Raddová, Miroslav Vachůn and Miroslav Pidra

Mendeleum, Horticultural Faculty, Mendel University of Agriculture and Forestry in Brno, Lednice na Moravě, Czech Republic

Abstract: The random amplified polymorphic DNA (RAPD) technique was used to evaluate both genetic diversity among 19 soybean accessions included in the Czech National Collection of Soybean Genotypes and their potential as a new source of genetic variations for soybean breeding programs. Only 22 of all the 40 random primers used in RAPD reactions showed polymorphism acceptable for an effective characterisation of these accessions. Altogether 122 highly reproducible RAPD fragments were generated, 55 of them were polymorphic (46%). However, because of the previously observed low degree of RAPD polymorphism in the case of *Glycine max*, fragments with low level of informativeness were evaluated, too. Presented results enable the selection of genetically distinct individuals. Such information may be useful to breeders willing to use genetically diverse introductions in soybean improvement process.

Keywords: soybean; molecular genetics; genetic diversity; dendogram; RAPD

Identification and utilisation of diverse germplasm is the central issue in plant breeding. Thorough knowledge of genetic diversity of the crop is necessary for parental selection that maximises genetic improvement. More accurate and complete description of genotypes and patterns of genetic diversity could help determinate future breeding strategies and facilitate introgression of diverse germplasm into the current commercial soybean genetic base.

Molecular markers have provided a powerful new tool for breeders to search for new sources of variation and to investigate genetic factors controlling quantitatively inherited traits. Many types of molecular markers have already been used (RAPD, SSR, RFLP, AFLP) and their niche in plant breeding has now been found.

The range of genetic diversity within the *Glycine* species is unknown, but almost every report on soybean genetic variation has concluded that diversity is lower in comparison with other self-pollinated species. This is attributed to the narrow genetic base of gene pool available for breeding (Apuya *et al*. 1988). The first type of molecular marker applied to soybean method was restriction fragment length polymorphism (RFLP). Unfortunately, owing to the lack of polymorphism, this approach was not as successful in soybean as in other crop species (Tanksley *et al*. 1989). The use of polymerase chain reaction (PCR) and PCR-based markers, i.e. random amplified polymorphic DNA (RAPD) and microsatellite sites study (SSR), could be alternative methods in the search for polymorphism in soybean. Microsatellites have been widely applied in mammalian (Weisenbach *et al*. 1992) and plant genomes (Thomas *et al*. 1993; Wang *et al*. 1994). They are reported to show a high level of polymorphism and have been successfully used in soybean, too (Morgante *et al*. 1994; Akkaya *et al*. 1995). Their major drawbacks are high cost of its development and higher technical laboriousness in comparison with RAPD.

RAPD markers are mainly used for diversity analysis because of their low cost, low technical difficulty and high throughput capabilities (Williams *et al*. 1990). A great asset of RAPDs is the large number of available

Supported by the Ministry of Education, Youth and Sports of the Czech Republic (Project No. MSM 435100002).

random primers and the large data sets that can be generated from them. RAPD markers were successfully applied to evaluate genetic diversity within soybean species too but these works were focused on stock of North America soybean cultivars mainly (THOMPSON *et al*. 1998; THOMPSON & NELSON 1988).

The aim of this work was to evaluate the genetic diversity within 19 soybean genotypes included in the Czech National Collection of Soybean Genotypes by RAPD method. The goal of the selection of genotypes was to obtain a mixture of North American cultivars and accessions originated in Europe. Of course, geographic origin is not an adequate indicator of potential similarity between accessions because many North American and European genotypes often originated from the same sources in the Asia (BROWN-GUEDIRA *et al*. 2000). For all that, some interesting facts were registered in this work.

MATERIAL AND METHODS

Plant material

A combination of 19 soybean genotypes bred in North America (NA) and Europe (E) was selected. These genotypes are included in the Czech National Collection of Soybean Genotypes. In the case of some genotypes identification numbers are added in the list of accessions. In such cases these genotypes are included in USDA Soybean Germplasm Collection and some next descriptions of them are available on the web site address: http:/genome.cornell.edu/cgibin/WebAce/weba-ce?db=soybase.

Genomics DNA extraction

DNA was extracted from freshly newly expanded leaves (field material) using protocol adapted from DELLAPORTA (1983). To obtain high-purity DNA, RNase digestion, Proteinase K digestion and phenol extraction steps were performed subsequently.

Quantification of the isolated DNA was achieved by comparing intensity of the DNA bands from the extract with the intensity of the bands from the uncut lambda DNA standards.

RAPD analysis

The PCR reaction mixtures had a total volume of 25 µl. The mixture contained 0.5 U of DyNAzyme II DNA polymerase (recombinant enzyme, Finnzyme com.), 0.4 µM primer (Operon Technologies), 0.2 mM of each dNTP (Promega), 1× appropriate reaction buffer (10 mM Tris-HCl, pH 8.8; 1.5 mM $MgCl_2$; 150 mM KCl and 0.1% Triton X-100) and 20 ng of template DNA. Reactions were performed in UNO II thermal cycler (Biometra) programmed as follows: 3 min at 94°C for initial denaturation, 40 cycles 1 min at 94°C (denaturation), 1 min at 36°C (annealing) and 1.5 min at 72°C (extension). A final extension step at 72°C for 9 min followed.

A total number of forty 10-mer primers with arbitrary sequences (kits E and M from Operon Technologies) were tested for PCR amplification. The amplification products were visualised on 1.5% agarose gels stained with ethidium bromide using standard methods (SAMBROOK *et al*. 1989). The size of specific RAPD fragments was estimated with 100 base pair DNA ladder (New England Biolab) with 15 DNA fragments ranging from 100 to 1500 base pairs. No primer requiring more than two amplifications to confirm repeatable banding pattern was used for next evaluation.

Data analysis

Only the bands that were present in both replications were considered and scored for each genotype and primer. Results were scored manually and transfered onto a present-absent scale (1 or 0 for each allele and genotype). This data were subjected to analysis using POPGENE 32 software (YEH & BOYLE 1997). Amplified

Table 1. List of the studied genotypes

1 = Wilkin (PI 548501) – NA	11 = Eszter – E
2 = Dippes Fruhegelbe – E	12 = Mlochovska (PI 423713) – E
3 = Zefir (PI 556836) – NA	13 = Mandin Kajon – unknown
4 = Jubilejnaja – E	14 = Belosnezka (PI 507674) – E
5 = Semu 8107 – E	15 = Miriam – E
6 = Fiskeby 5 (PI 360955) – E	16 = Acme (PI 548498) – NA
7 = Charkovskaja – E	17 = Clay (PI 548534) – NA
8 = Vilnensis (PI 189867) – E	18 = Maple Amber (PI 548592) – NA
9 = Wanda – E	19 = Mc Call (PI 548582) – NA
10 = Kunitz (PI 542044) – NA	

products were analysed by comparison of genotypes based on the percentage of common fragments and similarity matrix (NEI 1978). Dendograms were constructed by means of unweighed pair group method using arithmetic averages (UPGMA) and drawn using TREEVIEW software (PAGE 1996).

RESULTS AND DISCUSSION

Among 40 initially screened primers, RAPD patterns of 35 primers were considered to be reproducible and scored for analysis (banding patterns for OPE5, OPE10, OPE11, OPE13 and OPM19 primers were not repeatable). In all, 122 amplified fragments were scored (average number of loci per primer is 3.48) and 55 of them (45%) were found to be polymorphic. Other similar publication indicated 34% of polymorphic RAPD products (THOMPSON *et al*. 1998). Higher degree of diversity in our study is probably inflated because genotypes included in the study were selected to represent the great amount of diversity and do not represent typical material found in soybean breeding programs.

From the point of view of primers evaluation, 10 of 35 primers giving reproducible results gave monomorphic spectra. Finally, among 40 primers initially used in this study 25 were polymorphic (62%). In other similar work (DOLDI *et al*. 1997) only 12 from the initial 33 ran-

Table 2. Coefficients of diversity for every evaluated allele

Allele	h*	Allele	h*	Allele	h*	Allele	h*
OPE1-1	0.0000	OPE14-5	0.0000	OPM2-4	0.0000	OPM11-1	0.4321
OPE1-2	0.0997	OPE15-1	0.0000	OPM2-5	0.0000	OPM11-2	0.0000
OPE1-3	0.0000	OPE15-2	0.0000	OPM3-1	0.4321	OPM11-3	0.0000
OPE1-4	0.0000	OPE15-3	0.0000	OPM3-2	0.4986	OPM11-4	0.2659
OPE1-5	0.0000	OPE16-1	0.2659	OPM3-3	0.4321	OPM12-1	0.0000
OPE2-1	0.0997	OPE16-2	0.0997	OPM4-1	0.4321	OPM12-2	0.0000
OPE2-2	0.0000	OPE16-3	0.0000	OPM4-2	0.2659	OPM12-3	0.3324
OPE3-1	0.0000	OPE17-1	0.0000	OPM4-3	0.4321	OPM12-4	0.0000
OPE4-1	0.0000	OPE18-1	0.0000	OPM4-4	0.4321	OPM13-1	0.0000
OPE4-2	0.0000	OPE18-2	0.1884	OPM4-5	0.0000	OPM13-2	0.0000
OPE4-3	0.0000	OPE18-3	0.0000	OPM5-1	0.0000	OPM14-1	0.0997
OPE4-4	0.0000	OPE18-4	0.0000	OPM5-2	0.0000	OPM15-1	0.0000
OPE6-1	0.0000	OPE19-1	0.3878	OPM5-3	0.0000	OPM15-2	0.0000
OPE7-1	0.0997	OPE19-2	0.0000	OPM6-1	0.0000	OPM16-1	0.0000
OPE7-2	0.0997	OPE19-3	0.0000	OPM6-2	0.3878	OPM16-2	0.1884
OPE7-3	0.0997	OPE19-4	0.0000	OPM6-3	0.0000	OPM16-3	0.0997
OPE7-4	0.0997	OPE19-5	0.3878	OPM6-4	0.4654	OPM16-4	0.0997
OPE7-5	0.0000	OPE19-6	0.0000	OPM6-5	0.4986	OPM16-5	0.0000
OPE7-6	0.0997	OPE19-7	0.0000	OPM7-1	0.0997	OPM16-6	0.0000
OPE8-1	0.0997	OPE19-8	0.0000	OPM7-2	0.2659	OPM16-7	0.0997
OPE8-2	0.0997	OPE20-1	0.1884	OPM7-3	0.1884	OPM17-1	0.4321
OPE8-3	0.1884	OPE20-2	0.0997	OPM8-1	0.0000	OPM17-2	0.3324
OPE9-1	0.0000	OPE20-3	0.1884	OPM8-2	0.4321	OPM17-3	0.0000
OPE9-2	0.0000	OPE20-4	0.4654	OPM8-3	0.0000	OPM18-1	0.1884
OPE12-1	0.0000	OPE20-5	0.0000	OPM9-1	0.0997	OPM18-2	0.0000
OPE12-2	0.0000	OPM1-1	0.1884	OPM9-2	0.4875	OPM20-1	0.0000
OPE12-3	0.1884	OPM1-2	0.0000	OPM9-3	0.4986	OPM20-2	0.0000
OPE14-1	0.0997	OPM1-3	0.0000	OPM9-4	0.2659	OPM20-3	0.0000
OPE14-2	0.0000	OPM2-1	0.0000	OPM9-5	0.0000	OPM20-4	0.0000
OPE14-3	0.4654	OPM2-2	0.4875	OPM9-6	0.0997		
OPE14-4	0.0000	OPM2-3	0.4875	OPM10-1	0.3324		

* h = Nei's (1973) gene diversity

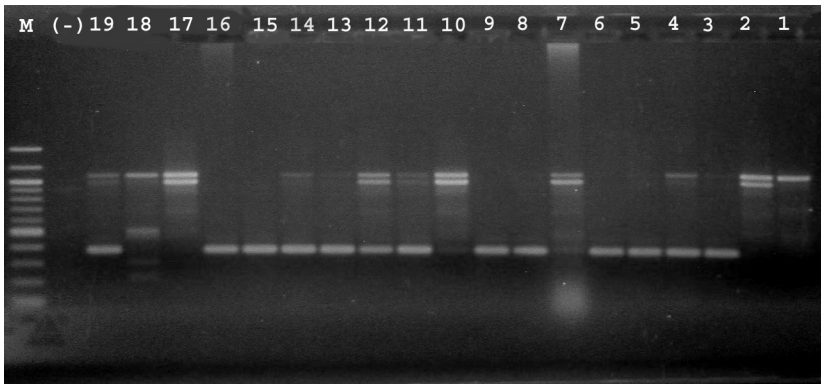

Fig. 1. RAPD patterns of 19 Glycine max genotypes obtained with primer OPM 3. Genotype order is the same as in Table 1 (M – DNA size marker, 100 bp ladder)

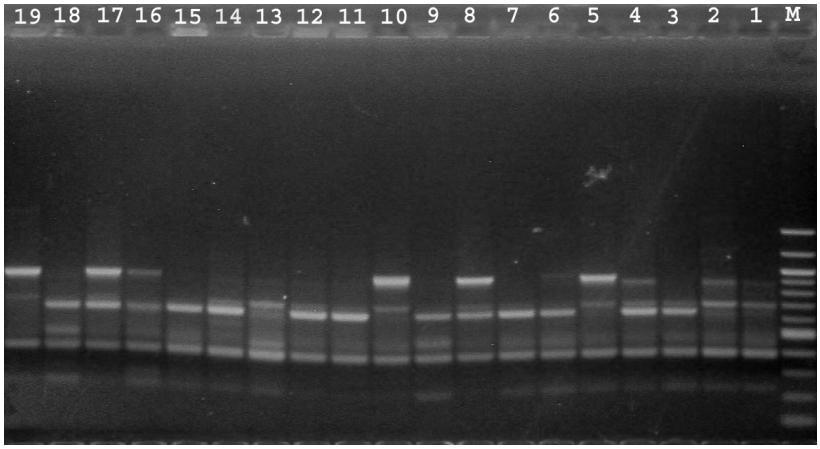

Fig. 2. RAPD patterns of 19 Glycine max genotypes obtained with primer OPM 4. Genotype order is the same as in Table 1 (M – DNA size marker, 100 bp ladder)

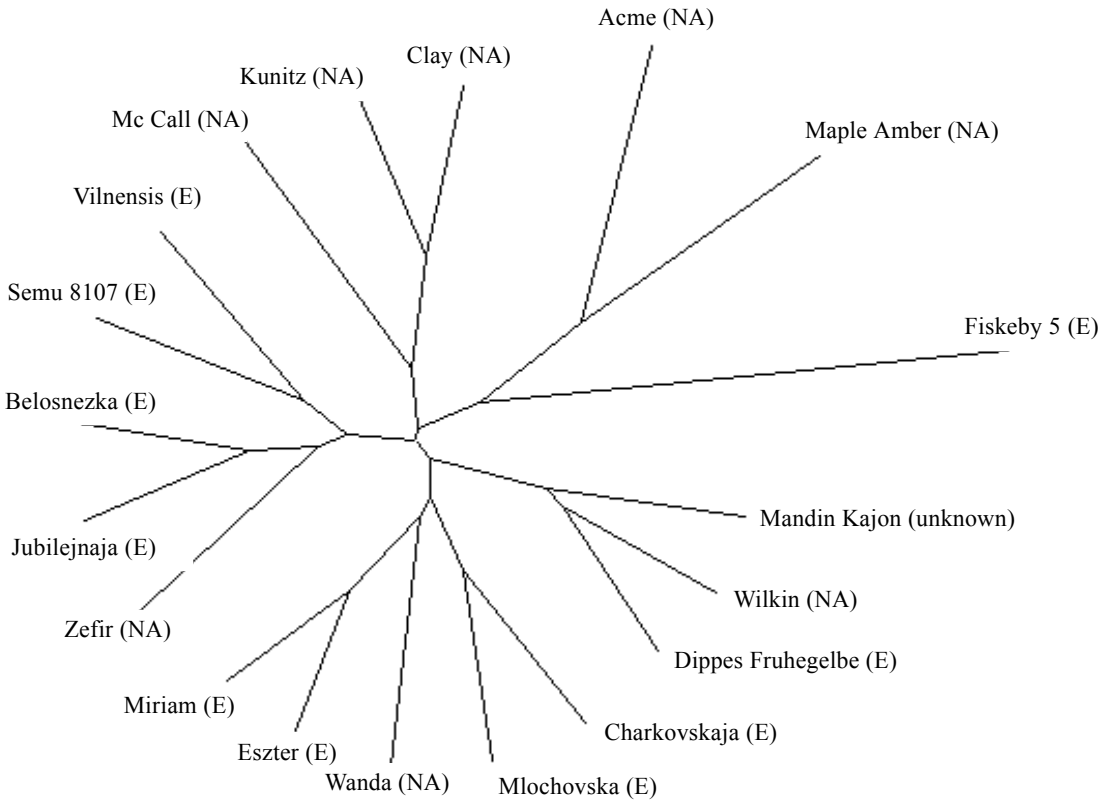

Fig. 3. Dendogram of 19 soybean genotypes based on NEI (1978) similarity coefficients using UPGMA method as the clustering method (unrooted radial form of tree)

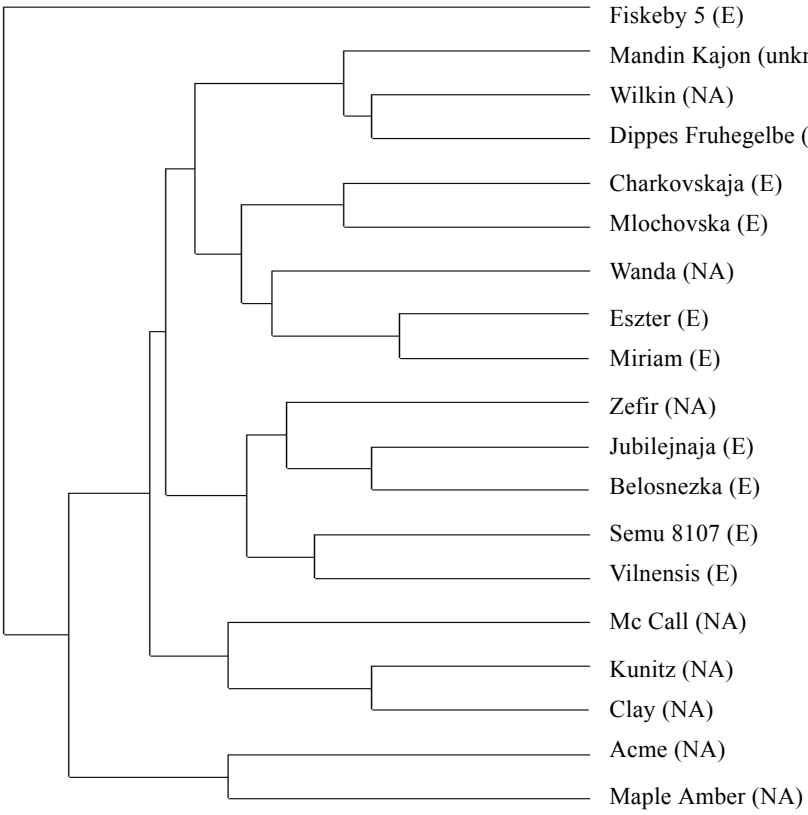

Fig. 4. Dendogram of 19 soybean genotypes based on NEI (1978) similarity coefficients using UPGMA method as the clustering method (phylogram form of tree)

dom primers showed polymorphism (36%), so our value is significantly higher. On the other hand informativeness of the RAPD fragment spectra produced by individual primers strongly varied (Table 2 – Nei's coefficient of gene diversity; NEI 1973).

For example the spectra of OPE1, OPE2 and OPM14 primers were judged as polymorphic, but only one polymorphic RAPD band was presented by comparison with all genotypes banding pattern. On the other hand, OPM3, OPM4, OPM6 and OPM9 primers were evaluated as the most informative. Examples of RAPD patterns obtained with primers OPM3 and OPM4 are shown in Figs 1 and 2

Resulted banding patterns were scored manually to the present/absent form for all repeatable fragments. The so created matrix was transferred to POPGENE 32 software to generate similarity/dissimilarity matrix among genotypes (NEI 1978), dendograms of similarity among genotypes (UPGMA method) and some other characteristics of loci. Dendograms were drawn by TreeView software in order to ensure higher graphical quality (Figs 3 and 4).

From both dendograms (Figs 3 and 4) it is evident that all 19 accessions are separated into the five distinguished clusters. The fact of difficult classification of soybean genotypes according to their alleged origin has already been mentioned in this paper (BROWN-GUEDIRA *et al.* 2000). From this point of view there is no surprise that some clusters in Figs 3 and 4 contain both European and N. American genotypes. On the other hand, it is clearly visible in Fig. 4 that the usual location of North American originated cultivars is in the lower part of the dendogram, whereas European genotypes are mainly located in the upper part. It suggests enlarged probability of higher genetic distance between these two groups.

From the point of view of new resources of genetic diversity, the lack of distinct association of Fiskeby 5 with any of the other groups may represent its genetic originality. Therefore this genotype should be a useful source of new genetic diversity for soybean breeding programs.

References

AKKAYA M.S., SHOEMAKER R.C., SPECHT J.E., BHAGWAT A.A., CREGAN P.B. (1995): Integration of simple sequnce repeat DNA markers into a soybean linkage map. Crop. Sci., **35**: 1439–1445.

APUYA N.R., FRAZIER B.L., KEIM P., ROTH E.J., LARK K.G. (1988): Restriction fragment length polymorphism as a genetic markers in soybean, *Glycine max* (L.) Merrill. Theor. Appl. Genet., **75**: 889–901.

BROWN-GUEDIRA G.L., THOMPSON J.A., NELSON R.L., WARBURTON M.L. (2000): Evaluation of genetic diversity of soybean introductions and North American ancestor using RAPD and SSR markers. Crop Sci., **40**: 815–823.

DELLAPORTA S.L., WOOD J., HICKS J.B. (1983): A plant DNA minipreparation: version II. Plant Mol. Biol. Rep., **1**: 19–21.

DOLDI M.L., VOLLMANN J., LELLEY T. (1997): Genetic diversity in soybean as determined by RAPD and microsatellite analysis. Plant Breed., **116**: 331–335.

MORGANTE M., RAFALSKI A., BIDDLE P., TINGEY S., OLIVIERI A.M. (1994): Genetic mapping and variability of seven soybean simple sequence repet loci. Genome, **37**: 763–769.

NEI M. (1973): Analysis of genetic diversity in subdivided populations. Proc. Natl. Acad. Sci. USA, **70**: 3321–3323.

NEI M. (1978): Estimation of average heterozigosity and genetic distance from a small number of individuals. Genetics, **89**: 583–590.

PAGE R.D.M. (1996): TreeView: An application to display phylogenetic trees on personal computers. Computer Appl. Biosci., **12**: 357–358.

SAMBROOK J., FRITSCH E.F., MANIATIS T. (1989): Molecular cloning: A laboratory manual. 2nd ed., Cold Spring Harbor Laboratory Press.

TANKSLEY S.D., YOUNG N.D., PATERSON A.H., BONIERBALE M.W. (1989): RFLP mapping in in plant breeding: new tool for an old science. BioTechnol., **7**: 257–264.

THOMAS M.R., MATSUMOTO S., CAIN P., SCOTT N.S. (1993): Repetitive DNA of grapevine: clases present and sequnces suitable for cultivar identification. Theor. Appl. Genet., **86**: 173–180.

THOMPSON J.A., NELSON R.L., VODKIN L.O. (1998): Identification of diverse soybean germplasm using RAPD markers. Crop. Sci., **38**: 1348–1355.

THOMPSON J.A., NELSON R.L. (1998): Core set of primers to evaluate genetic diversity in soybean. Crop Sci., **38**: 1356–1362.

WANG Z., WEBER J.L., ZHONG G., TANKSLEY S.D. (1994): Survey of plant short tandem DNA repeats. Theor. Appl. Genet., **88**: 1–6.

WEISENBACH J., GAYPAY G., DIB C., VIGNAL A., MORRISETTE J., MILASSEAU P., VAYSSEIX G., LATHORP M. (1992): A second-generation linkeage map of the human genome. Nature, **359**: 794–801.

WILLIAMS J.G.K., KUBELIK A.R, KENNETH J., RAFALSKY J.A., TINGEY S.V. (1990): DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., **18**: 6531–6535.

YEH F.C., BOYLE T.J.B. (1997): Population genetic analysis of co-dominant and dominant markers and quatntitative traits. Belgian J. Botany, **129**: 157.

Received for publication September 25, 2001
Accepted after corrections December 14, 2001

Abstrakt

BARÁNEK M., KADLEC M., RADDOVÁ J., VACHŮN M., PIDRA M. (2002): **Zhodnocení genetické diverzity v rámci 19 genotypů zařazených do kolekce pro uchování genetických zdrojů sóji luštinaté v České republice**. Czech J. Genet. Plant Breed., **38**: 69–74.

Pomocí molekulárně-genetické metody RAPD (tzv. metoda náhodně amplifikované polymorfní DNA) byla hodnocena genetická diverzita v rámci 19 genotypů sóji luštinaté, které jsou mimo jiné zařazeny do programu uchování národních genových zdrojů sóji luštinaté pro Českou republiku. Z celkového počtu 40 použitých primerů pouze 22 poskytovalo polymorfní spektra vhodná pro efektivní charakterizaci sledovaných genotypů. Celkem bylo nalezeno 122 dobře reprodukovatelných RAPD fragmentů, přičemž 55 z nich bylo v rámci sledované skupiny genotypů polymorfních (46 %). Vzhledem k tomu, že v případě sóji luštinaté byla v dřívějších publikacích pozorována nízká míra dosahovaného RAPD polymorfismu, byly hodnoceny i fragmenty s nízkým informačním obsahem. Předkládané výsledky umožňují výběr geneticky rozdílných genotypů, což může být přínosem pro šlechtitele, kteří se v rámci procesu zlepšování genetického potenciálu sóji luštinaté snaží do svých šlechtitelských programů zařadit některé geneticky výrazně rozdílné novinky.

Klíčová slova: sója; molekulární genetika; genetická diverzita; dendogram; RAPD

Corresponding author:

MIROSLAV BARÁNEK, MSc., Mendelova zemědělská a lesnická univerzita, Zahradnická fakulta, Mendeleum, Valtická 334, 691 41 Lednice, Česká republika

tel.: + 420 627 34 01 05–7, fax: + 420 627 34 01 59, e-mail: baranek@zf.mendelu.cz