MOLECULAR CHARACTERIZATION AND GENETIC DIVERSITY STUDY OF SOYBEAN (GLYCINE MAX L.) CULTIVARS USING RAPD MARKERS

Martin Vivodík*, Želmíra Balážová, Milan Chišapek, Zuzana Hromadová, Lucia Nikolášová, Zdenka Gálová

Address(es): Ing. Martin Vivodík, PhD.
Slovak University of Agriculture, Faculty of Biotechnology and Food Sciences, Department of Biochemistry and Biotechnology, Tr. A. Hlinka 2, 949 76 Nitra, Slovakia.

*Corresponding author: martin.vivodik@uniag.sk

ABSTRACT

Soybean (Glycine max L.) is an important crop plant which contains a high amount of oil and protein. The main goal of the work was to use 13 RAPD (Random Amplified Polymorphic DNA) markers to study genetic polymorphism in a set of 28 soybean genotypes and to construct a dendrogram from the obtained results, based on which we will recommend genotypes for further breeding. In total, using 13 RAPD markers, we determined 108 fragments in a set of 28 soybean genotypes with an average number of 8.31 fragments per genotype. The number of fragments varied from 6 (OPB-08, OPE-07) to 12 (SIGMA-D-01). Of the total number of 108 fragments, 70 were polymorphic with an average number of 5.38 polymorphic fragments per genotype. The number of polymorphic fragments varied from 3 (OPE-07, OPF-14) to 9 (SIGMA-D-01). The average percentage of fragment polymorphism was 63.68% and ranged from 42.86% (OPF-14) to 83.33% (OPB-08). Diversity index (DI) values ranged from 0.710 (OPA-03) to 0.846 (OPD-08) with an average value of 0.763. A dendrogram prepared based on hierarchical cluster analysis using UPGMA algorithm separated 28 soybean genotypes into four clusters. The applied RAPD markers proved to be a suitable tool for detection of the genetic diversity of soybean genotypes. Based on the constructed dendrogram, it is possible to select suitable varieties for marker-assisted breeding.

Keywords: soybean, RAPD, dendrogram, polymorphism, UPGMA

INTRODUCTION

Soy (Glycine max L.) is relatively the cheapest source of protein. Soybean seeds contain up to 40-42% protein, the oil content is at the level of 18-20% and soluble carbohydrates are up to 11%. It follows from the above that soybean seeds are nutritionally interesting for the production of various types of food, but also for the production of animal feed (Devi et al., 2012). The wild and cultivated soybeans showed significant phenotypic diversity but the small reproductive difference, and they have very similar genomes in both its size and content (Singh and Hymowitz, 1999). A potential source of protein and oil makes soybeans a large share in the genome of human nutrition, and also improves soil fertility therefore; soybean is also an important crop for research in the field of detecting genetic variability, which is also important for breeding.

Soybean genotypes are used as a model species, such as: phaseolus (Gjorgieva et al., 2012); pea (Choudhury et al., 2017); castor (Vivodík et al., 2014; Vivodík et al., 2015), patchouli (Pandey et al., 2022), carnation (Sharma et al., 2022), cassia (Eldemerdash et al., 2022), maize (Balážová et al., 2017), rye (Petrovícová et al., 2015), wheat (Kut’ka Hložáková et al., 2016), amaranthus (Štefúnová et al., 2015), sorghum (Ruiz-Chután et al., 2019). The main goal of the work was to use 13 RAPD markers to study genetic polymorphism in a set of 28 soybean genotypes and to construct a dendrogram from the obtained results, based on which we will recommend genotypes for further breeding.

MATERIAL AND METHODS

Soybean genotypes (28) (Table 1) were obtained from the Gene Bank in Piešťany, the Slovak Republic. Genomic DNA was isolated from the 14 days leaves with GenEJET Plant Genomic DNA Purification Mini Kit according to the manufacturer’s instructions. Soybean genotypes were grown in a growth chamber from seeds on humus soil. Concentrations of isolated DNA were estimated using UV-VIS spectrophotometer and the final concentration of DNA was adjusted to 50 ng/µl. All the DNA samples were stored at −20 °C.

Amplification of RAPD fragments was performed according to Gajeraa et al., (2010) using decamer arbitrary primers (Table 2). A total volume of 25 μl of the reaction mixture contained 100 ng of DNA, 12.5 μl of Master Mix (Genei, Bangalore, India) and 10 pmol of primer. DNA amplification was performed in a thermocycler (Biometra, Germany) programmed as follows: initial DNA denaturation at 94°C for 5 min, followed by 42 cycles of denaturation at 94°C for 1 min, primer annealing at 38°C for 1 min, synthesis of the new DNA strand at 72°C for 1 min and a final step at 72°C for 5 min. Amplified DNA products were separated by horizontal gel electrophoresis in 1.5% agarose gel in 1 x TBE buffer at a constant voltage of 100 V for approximately 1 hour. Ethidium bromide was used as an intercalating agent in the gel. Evaluation of the gels was performed using a UV lamp using the UVP DigiDoc-It system and evaluated by the program GelAnalyzer.
### RESULTS AND DISCUSSION

In total, using 13 RAPD markers, we determined 108 fragments in a set of 28 soybean genotypes with an average number of 8.31 fragments per genotype (Figure 1), (Table 3). The number of fragments varied from 6 (OPB-08, OPF-07) to 12 (SIGMA-D-01). Of the total number of 108 fragments, 70 were polymorphic with an average number of 5.38 polymorphic fragments per genotype. The number of polymorphic fragments varied from 3 (OPF-07, OPF-14) to 9 (SIGMA-D-01). The average percentage of fragment polymorphism was 63.85% and ranged from 42.86% (OPF-14) to 83.33% (OPB-08). Diversity index (DI) values ranged from 0.710 (OPA-03) to 0.846 (OPF-08) with an average value of 0.763. For all 13 RAPD markers, the DI values were higher than 0.700, which indicates a high degree of polymorphism of the analyzed soybean genotypes. Polymorphic Information Content (PIC) values ranged from 0.639 (SIGMA-D-14) to 0.754 (OPA-07) with an average value of 0.698. In this case, 8 RAPD markers had a calculated PIC value higher than 0.700 and 5 RAPD markers had a PIC value lower than 0.700. The PIC values also speak of the good usability of RAPD markers for soybean identification and differentiation. Probability of identity (PI) values ranged from 0.004 (OPA-03) to 0.070 (OPF-08) with an average value of 0.028 (Table 3).

Table 2. List of RAPD primers (Gajeraa et al., 2010)  
| Primers | Primer sequence (5'-3') | Molecular weight range (bp) |
|---------|-------------------------|-----------------------------|
| OPA-02  | TGCCGAGCTGT             | 400-2800                    |
| OPA-03  | AGTCCAGCAC              | 330-870                     |
| OPA-13  | CAGCACCCAC              | 370-1800                    |
| OPB-08  | GTCCACACGGG             | 530-1550                    |
| OPD-02  | GAGCACAACC              | 280-1850                    |
| OPD-07  | TGGCCACGGG              | 360-1440                    |
| OPD-08  | GTGCCACCCA              | 260-1700                    |
| OPD-13  | GGCTTGACGA              | 160-1800                    |
| OPE-07  | AGATGACGCC               | 300-1940                    |
| OPF-14  | TGCTCGAGGT              | 190-1850                    |
| SIGMA-D-01 | AAACCGCCGC              | 280-1350                    |
| SIGMA-D-14 | TCTCCTCCTCA             | 350-900                     |
| SIGMA-D-P | TGGACCGGTTG             | 300-3000                    |

### Data analysis

The presence (1) or absence (0) of polymorphic reproducible markers was assessed in the obtained DNA profiles. The resulting binary matrix was used to construct a dendrogram. The representation of the genetic relatedness of the analyzed varieties in the form of a dendrogram was carried out based on a hierarchical cluster analysis using the UPGMA algorithm (Unweighted Pair Group Method using arithmetic averages) according to the Jaccard coefficient of genetic similarity using a specific SPSS Professional Statistics module version 17, a statistical package of SPSS programs for Windows. Genetic diversity was expressed in relative values in the range 0-25. For the assessment of the polymorphism between genotypes ricin and usability RAPD markers in their differentiation we used diversity index (DI) (Weir, 1990), the probability of identity (PI) (Paetkau et al., 1995) and polymorphic information content (PIC) (Weber, 1990).
85.80%. OPA-19 contains the lowest percentage of polymorphism, generating only 1 polymorphic band out of 3 of all DNA fragments. On the other hand, 11 primers (OPA-2, OPA-07, OPA-10, OPA-11, OPA-12, OPA-13, OPA-14, OPA-15, OPA-16, OPA-18, OPA-20) showed 100% polymorphism. The authors recommend calculating PIC, EMR and MI values for DNA polymorphism detection and to identify the primer that is most informative.

Nkongolo et al. (2020) studied the level of genetic variation among soybean (G. max) accessions from different countries using RAPD markers. Genomic DNAs from 108 soybean accessions from 11 different gene pools were analyzed using five RAPD primers (OPA 11, Pinus 23, UBC 377, UBC186, Grasse 8). The average level of polymorphic loci detected with the RAPD primers was 35%. The highest polymorphic index among accessions was 48.00% in France, followed by accessions from Hungary with 43.30%. Over all, the lowest polymorphic index was 29.90% (Canada), 26.29% (Netherlands) and 24.74% (China). At the primer level, primers UBC 186 and Grasse 8 generated the most number of bands (41) followed by UBC 377 with 40 bands. The lowest number of amplified product was observed with primer Pinus 23. The highest level of polymorphic loci (46.30%) was detected with primer UBC 186 and the lowest with UBC 377.

A dendrogram prepared based on hierarchical cluster analysis using UPGMA algorithm separated 28 soybean genotypes into four clusters (Figure 2). Cluster 1 contains 6 soybean genotypes, 2 of which come from the USA (Anoka, Holt) and one each from Canada (Cesar), Slovakia (Zora), France (Kador) and the Czech Republic (Dacota). Cluster 2 consists of 3 genotypes, of which 2 come from North America, one from the USA (Maverick), one from Canada (Bristol) and the Comet genotype is of unknown origin. Cluster 3 consists only of the Ugo genotype, which originates from Canada. Cluster 4 is the largest of all 4 clusters and contains a total of 18 soybean genotypes, which are divided into 2 subclusters (4A and 4B).

Subcluster 4A contains 2 soybean genotypes from Canada (Baron and Gaillard), one genotype from the Czech Republic (Chmelarova Brenska) and one genotype from Japan (Ishigo Wase). Subcluster 4B is the most diverse and contains 4 genotypes from Canada (Primus, Mario, Korada and Cardiff), 3 genotypes from the Czech Republic (Sluna, Polanka and Jihomoravská Zluta), 3 genotypes from France (Fred, Armor and Recor), one genotype from the USA (Canton) and the ZSSR (Khabarovskaja) and 2 genotypes of unknown origin (Anko and Lokus). In subcluster 4 B, we could not distinguish between 2 Chinese genotypes (Polanka and Jihomoravská Zluta), which probably have the same genetic origin, using 13 RAPD markers. To distinguish them, it would be necessary to use more RAPD markers or other DNA markers. On the basis of the obtained results, we can conclude that with the help of 13 RAPD markers, we failed to distinguish the genotypes based on their genetic origin. To better distinguish soybean genotypes, it would be advisable to use other DNA techniques, such as SSR, SCoT, TRAP and EBAP technique.

Similarly, many authors were able to differentiate genotypes of soybean using RAPD markers. Jain et al. (2017) constructed a dendrogram in which the 24 soybean genotypes analyzed divided into 3 major clusters with a similarity coefficient of 0.29. The RAPD methods displayed genetic variation among 24 soybean genotypes and phylogenetic tree was showing a relationship among them. Wahyudi et al. (2020) constructed a dendrogram in which the analyzed soybean genotypes were divided into 2 main clusters. The wild-type soybean became the first group, which act like an outgroup. The second group was consist of all soybean mutants, which can be separated into 3 subgroups. The similarity index and clustering method were performed to ascertain the degree of genetic relationship among soybean mutants and wild-type.

![Figure 1 RAPD profiles of bands in soybean genotypes (OPA-02). M- Quick-Load® 2-Log DNA ladder and 1-19 are analyzed genotypes of soybean (Table 1)](image)

**Table 3 The statistical characteristics of the 13 RAPD markers used in 28 soybean genotypes**

| Primer | Total number of fragments | Number of polymorphic fragments | Percentage of polymorphism | DI | PIC | PI |
|--------|--------------------------|---------------------------------|-----------------------------|----|-----|----|
| OPA-02 | 8                        | 5                               | 62.50                       | 0.729 | 0.659 | 0.015 |
| OPA-03 | 7                        | 4                               | 57.14                       | 0.710 | 0.700 | 0.004 |
| OPA-13 | 7                        | 4                               | 57.14                       | 0.780 | 0.721 | 0.010 |
| OPB-08 | 6                        | 5                               | 83.33                       | 0.715 | 0.701 | 0.008 |
| OPD-02 | 9                        | 6                               | 66.66                       | 0.816 | 0.730 | 0.038 |
| OPD-07 | 7                        | 4                               | 57.14                       | 0.714 | 0.659 | 0.015 |
| OPD-08 | 9                        | 6                               | 66.66                       | 0.846 | 0.711 | 0.070 |
| OPD-13 | 11                       | 8                               | 72.73                       | 0.810 | 0.709 | 0.010 |
| OPE-07 | 6                        | 3                               | 50.00                       | 0.825 | 0.754 | 0.062 |
| OPE-14 | 7                        | 3                               | 42.86                       | 0.812 | 0.745 | 0.059 |
| SIGMA-D-01 | 12 | 9                               | 75.00                       | 0.731 | 0.693 | 0.039 |
| SIGMA-D-14 | 10 | 7                               | 70.00                       | 0.711 | 0.639 | 0.019 |
| SIGMA-D-P | 9 | 6                               | 66.66                       | 0.717 | 0.649 | 0.021 |
| Total   | 108                      | 70                              | 63.68                       | 0.763 | 0.698 | 0.028 |
| Average | 5.31                     | 5.38                            | 63.68                       | 0.763 | 0.698 | 0.028 |
The main goal of the work was to use 13 RAPD markers to study genetic polymorphism in a set of 28 soybean genotypes and to construct a dendrogram. In total, using 13 RAPD markers, we determined 108 fragments in a set of 28 soybean genotypes, with an average number of 5.38 polymorphic fragments per genotype. The number of polymorphic fragments varied from 6 (OPB) to 354 (Z09). Based on the obtained results, we conclude that with the help of 13 RAPD markers, we failed to distinguish soybean genotypes based on their genetic origin. To better distinguish soybean genotypes, we would suggest using a combination of several DNA techniques, such as SSR, TRAP, SCOT or EBAP technique. Knowledge regarding the genetic variability of studied soybean genotypes may provide information important for the improvement of existing soybean cultivars in the breeding process, as well as the preservation and maintenance of soybean germplasm resources.

Acknowledgments: This paper was supported by VEGA project No. 1/0291/21 (50%), KEGA project No. 026SPU-4/2021 (15%), KEGA project No. 027SPU-4/2021 (15%) and the Operational Program Integrated Infrastructure within the project: Demand-driven research for the sustainable and innovative food. Drive4SIFood 313011V336, cofinanced by the European Regional Development Fund (20%).

REFERENCES

Balážová, Z., Vivodík, M. and Gálová, Z. (2017). Evaluation of molecular diversity of central European maize cultivars. *Emirates Journal of Food and Agriculture*, 28(2), 93-98. https://doi.org/10.9755/ejfa.2015-05-204.

Deví, K. N., Singh, L. K., Singh, M. S., Singh, S. B., Singh, K. K. (2012). Influence of sulphur and boron fertilization on oil content, yield and stover yield in soybean (*Glycine max* (L.) Merrill). *Acta Botanica Hungarica*, 54(3-4), 251-266.

Eldemerdash, M. M., El-Sayed, A. A., Hussein, H. A., Teleb, S. S. and Shehata, R. S. (2022). Molecular and metabolic traits of some Egyptian species of Cassia L. and Senna Mill (Fabaceae-Caesalpinioideae). *BMC Plant Biology*, 22, 205-220. https://doi.org/10.1186/s12870-022-05343-7.

Férfi, T., Toth, Z., Csaba, T. (2013). ISJ and RAPD Molecular Markers. *Evaluation of Genetic Diversity among Soybean (Glycine max (L.) Merr.) genotypes using RAPD markers*. Molecular Markers and Crop Evaluation, 67, 1587-1600. https://doi.org/10.1007/s10722-012-00933-4.

Jain, R. K., Joshi, A. and Jain, D. (2017). Molecular Marker Based Genetic Diversity Analysis in Soybean (*Glycine max* (L.) Merril) Genotypes. *International Journal of Current Microbiology and Applied Sciences*, 6(6), 1034-1044. https://doi.org/10.21270/ijcmas.2017.6.6.119.

Kim, D. G., Luu J. I., Lee, M. K., Kim, J. M., Hung, N. N., Hong, M. J., Kim, J. B., Chang-Hyu Bae, Ch-H. and Kwon, S. J. (2020). Construction of Soybean Mutant Diversity Pool (MDP) Lines and an Analysis of Their Genetic Relationships and Associations Using TRAP Markers. *Agronomy*, 10(253), 1-16. https://doi.org/10.3390/agronomy10020253.

Kumar, S. P. J., Susmita, C., Srirajith, K. V., Agarwal, D. K., Pal, G., Singh, A. N., Kumar, S., Rai, A. K., Simal-Gandara, J. (2022). Molecular characterization and genetic diversity studies of Indian soybean (*Glycine max* (L.) Merr.) cultivars using SSR markers. *Molecular Biology Reports*, 49, 2129-2140. https://doi.org/10.1007/s11033-021-00730-4.

Krčka Hložáková, T., Gálová, Z., Gregová, E., Vivodík, M., Balážová, Z., Miháliková, D. (2016). RAPD analysis of the genetic polymorphism in european wheat genotypes. *Potravinářství*, 10(1), 1-6. https://doi.org/10.52195/pot.

Lee, M. E., Grau, C. R., Lukaso, L. A. & Lee, J. M. (2002). Identification of aster yellow phytoplasmas in soybean in Wisconsin based on rflfp analysis of 16s rDNA (ribosomal DNA). *Canadian Journal of Plant Pathology*, 24(2), 125-130. https://doi.org/10.1080/07060606309506986.

Mihajlević, M. Ž., Šarčević, H., Lovric, A., Andrijanić, Z., Sudaric, A., Jukić, G., Pejc, I. (2020). Genetic diversity of European commercial soybean (*Glycine max* (L.) Merr.) germplasm revealed by SSR markers. Molecular Breeding Research, 28, 1-13. https://doi.org/10.1007/s11032-020-00968-x.

Mongara, J., Chudasama, K. and Thaker, V. (2017). ISSR studies on small and large seed varieties of *Glycine max*. *Journal of Pharmacognosy and Phytochemistry*, 6(6), 1652-1656.

Nkongolo, K., Alami, S., Michael, P. (2020). Assessment of Genetic Variation in Soybean (*Glycine max*) Accessions from International Gene Pools Using RAPD Markers: Comparison with the ISSR System. *American Journal of Plant Sciences*, 11, 1414-1428. https://www.scipub.org/journal/ajps.

Parau, D., Calvert, W., Burling, I., Strobel, C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, 4, 347-354.

Pandey, S. K., Baruah, J., Paw M., Sarma, N., Begum, T., Saikia, S. & Lal, M. (2022). Molecular Diversity and Chemical Composition Among Induced Mutation Lines of *Pogostemon cablin* (Blanco) Bent. Using RAPD and ISSR Marker. *Journal of Essential Oil Bearing Plants*, 25(2), 456-466.

Panjoo, M., Nazarian-Frouzabadi, F., Ismaili, A. and Ashmadi, H. (2014). Evaluation of Genetic Diversity among Soybean (*Glycine max*) Genotypes, Using ISL and RAPD Molecular Markers. *Journal of Plant Physiology and Breeding*, 4(2), 55-65.

Perić, V., Nikolić, A., Babić, V., Sudarić, A., Srebrić, M., Đordević, V. and Mladenović Drmić, S. (2014). Genetic relatedness of soybean genotypes based on agromorphological traits and RAPD markers. *Genetika*, 46, 839- 854. https://doi.org/10.2298/genet140839np.

Petrovićová, L., Gálová, Z., Balážová, Z., Vivodík, M., Wójcik-Jagla, M., Rapacz, M. (2015). Assessment of RAPD polymorphism in rye (*Secale cereale* L.) genotypes. *Journal of Microbiology, Biotechnology and Food Sciences*, 4 (special issue 2), 94-97. https://doi.org/10.15414/jmbfs.2015.4.special94-97.

Priestera, J. H., Gea, Y., Micke, R. E., Horsta, A. M., Moritz, S. C., Espinosa, K., Gelbl, J., Walker, S. L., Nisbett, R. M., Any, Y., Schimel, J. P., Palmere, R. G., Hernandez-Mejia, X. and Shehata, S. (2012). Soybean susceptibility to manufactured nanomaterials with evidence for food quality and soil fertility interruption. *Proceedings of the National Academy of Sciences of the USA*, 109, 2451-1456. https://doi.org/10.1073/pnas.1205431109.

Ruíz-Chatún J. A., Salava, J., Janovská, M., Žiarovská, J., Kalousova, M., Fernández, E. (2019). Assessment of genetic diversity in Sorghum bicolor using RAPD markers. *Genetika*, 51(3), 789-803. https://doi.org/10.22989/generika140383rp.

Sharma, P., Nath, A. K., Dhiman, S. R., Dogra, S., Sharma, V. (2022). Characterization of caronanemia (*Dianthus carpophyllus* L.) genotypes and gamma irradiated mutants using RAPD, ISSR and SSR markers. *South African Journal of Botany*, 148, 67-77. https://doi.org/10.1016/j.sajb.2022.04.012.

Sharma, R., Sharma, S., Kumar, S. (2018). Pair-wise combinations of RAPD markers for genetic analysis of reference pea population and single primer RAPD in soybean. *Annals of Agrarian Science*, 16, 243-249. https://doi.org/10.1007/s10709-017-0583-p.

Shuangxia, D., Ramesh, M., Huaguo, Z., Lili, T., Zhongxu, L., Yanxin, Z., Xianlong, Z. (2008). Detection of somaclonal variation of cotton (*Gossypium hirsutum* L.) using cytochrome c oxidase subunit I and nitrate reductase molecular markers. *Plant Cell Reports*, 28, 1303-1316.
Singh, R. J. and Hymowitz, T. (1999). Soybean genetic resources and crop improvement. *Genome*, 42, 605-616.

Štefúnová, V., Bezo, M., Žiarovská, J. and Ražná, K. (2015). Detection of the genetic variability of Amaranthus by RAPD and ISSR markers. *Pakistan Journal of Botany*, 47(4), 1293-1301.

Vivodík, M., Balážová, Ž., Gálová, Z. (2014). DNA analysis of ricin using RAPD technique. *Journal of Microbiology, Biotechnology and Food Sciences*, 3 (special issue), 181-183.

Vivodík, M., Balážová, Ž., Gálová, Z., Chňapek, M., Petrovičová, L. (2015). Study of DNA polymorphism of the castor new lines based on RAPD markers. *Journal of Microbiology, Biotechnology and Food Sciences*, 4 (special issue 2), 125-127. https://doi.org/10.15414/jmbfs.2015.4.special2.125-127

Wahyudi, D., Hapsari, L., Sundari. (2020). RAPD Analysis for Genetic Variability Detection of Mutant Soybean (*Glycine max* (L.) Merr). *Journal of Tropical Biodiversity and Biotechnology*, 5(1), 68-77. https://doi.org/10.22146/jtbb.53653

Weber, J. L. (1990). Informativeness of human (dC-dA)n x (dG-dT)n polymorphism. *Genomics*, 7, 524-530.

Weir, B. S. (1990). Genetic data analysis. Methods for discrete population genetic data. pp. 377. ISBN 0-87893-872-9

Zargar, M., Romanova, E., Trifonova, A., Shmelkova E. and Kezimana, P. (2017). AFLP-analysis of genetic diversity in soybean (*Glycine max* (L.) Merr.) cultivars Russian and foreign selection. *Agronomy Research*, 15(5), 2217-2225. https://doi.org/10.15159/AR.17.066