Jomard
Publishing

DELINEATING GENETIC DIVERSITY IN SOYBEAN (*GLYCINE MAX* L.) GENOTYPES: INSIGHTS FROM A-PAGE ANALYSIS OF GLOBULIN RESERVE PROTEINS

ORCID **Orkhan Bayramli**[*]

Institute of Genetic Resources, Ministry of Science and Education, Baku, Azerbaijan

Abstract. This study explores the genetic diversity of soybean (*Glycine max*) genotypes using polyacrylamide gel electrophoresis (A-PAGE) to analyze globulin reserve proteins. The research utilized 51 soybean genotypes from various origins, including Canada, Turkey, Uzbekistan and others. The A-PAGE method was employed to extract and analyze the globulin storage proteins in the grains of soybean samples. The proteins were categorized into ω, γ, β and α zones, revealing distinct electrophoretic patterns among the genotypes. A total of 22 spectra and 107 patterns were identified, demonstrating polymorphism based on the frequency of occurrence. The genetic diversity index, calculated using Nei's formula, indicated varying levels of genetic diversity across the different zones.

1. Introduction

Soybeans, scientifically known as Glycine max, are versatile legumes with a rich history in human nutrition and agriculture, China is the birthplace of soybean (G. max (L.) Merr.), a valuable source of protein for people as well as premium animal feed (FAO, 2003). Furthermore, increased demand for soybean production is a result of expanding consumption and the significant dietary supplements included in soybeans. With over 23,000 varieties across Asia, soybeans were first domesticated in China before being brought to the USA and Brazil (López-López *et al*., 2010). A concise chronicle of the global soybean distribution may be found by Rodríguez-Navarro et al. (2011). One of the main global suppliers of vegetable oil and animal protein feed is soybean (Sugiyama *et al*., 2015). Furthermore, soy protein has the ability to reduce obesity and muscular tiredness (Agyei *et al*., 2015). One of the main crops in five South American nations, soybeans account for almost 63% of all cropland (Wingeyer *et al*., 2015). Its growth reduced the amount of land used for other crops and natural vegetation while increasing soybean yield at a pace of about 6% each year. A larger production area, coupled with a lesser increase in grain output, was the primary cause of the rise in soybean yield (Wingeyer *et al*., 2015). It's possible that the word "soybean" alludes to the bean used to

make soy sauce. East Asian in origin, soybeans (Glycine max (L.) Merr.) are a legume crop that are currently produced all over the world due to their high oil and protein contents. In 2010/2011, its global output amounted to 263.7 million metric tons, more than twice as much as in 1992/1993 (Figure 1). In order to meet the expanding global population's requirements for food and fuel, there has been an increase in agricultural output, which has led to an increase in soybean production (Orf *et al*., 2010). The United States leads the world in soybean output (34%), followed by Brazil (29%), Argentina (19%), China (6%), India (4%), Paraguay (3%), Canada (2%) and other countries (4%) (United States Department of Agriculture FAS, 2011). Agriculture is a part of the economy of Azerbaijan and accounts for approximately 6% of the country's GDP (Alizade *et al*., 2023a; Alizade & Mammadova, 2023b). The participation of protein-rich feeds such as rapeseed and soybeans in the daily feed ration of livestock-raising farms, especially in the direction of fattening and dairy farming, plays an important role in increasing meat and milk production. Soybeans are used by Asian cultures to make a variety of traditional meals, including tempeh, miso, tofu, natto, soymilk and soy sauce. Soybeans are mostly processed into seed oil and soybean meal in Western cultures. 54% of US-produced soybeans in 2011–2012 were crushed for the country's oil industry, 4% were used for seed, feed and other uses and 42% of the soybeans were exported (Lu *et al*., 2014). Soybeans are crushed by processors into full-fat flakes, which are then extracted using organic solvents to yield oil and defatted flakes. Additionally, full-fat flakes can be processed into full-fat flour and used as culinary ingredients or they can be utilized as parts of animal feed. Defatted flakes are crushed into soybean meal, which is then used to produce texturized vegetable protein, soy concentrate and soy isolates. Soy meal is a premium source of protein for animal feed. As value-added food components, soy protein, isolates and concentrates are utilized in baby formulae, meat and meat-like products, baked goods, whipped toppings, frozen desserts, protein drinks, soup bases and other recipes. The majority of soybean oil produced—roughly 55%—is used for cooking and salad dressings, 24% for baking and frying fats and oils, 4% as a component of margarines, 7% for various culinary and industrial applications and 1% as a feedstock for the synthesis of biodiesel.
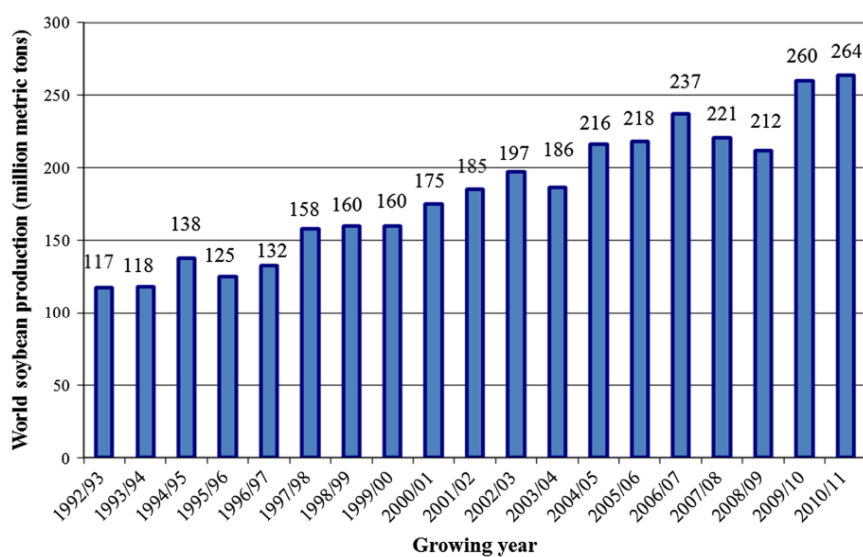

Figure 1. Historical soybean production

Numerous physiologically active substances found in soybeans have been linked to a number of important health advantages. Food producers are being encouraged to find new applications for soybeans in food systems and increase the amount of soy that is consumed by consumers as a result of an increasing number of publications verifying these advantageous benefits. As a result, it is predicted that soybeans will become less of a staple in cattle feed and more of a component in Western societies' normal diets. From a processing and nutritional perspective, it is essential to comprehend the elements that affect the chemical makeup of soybeans and the structure of individual soybean components. Continuous alterations in the characteristics of soybeans necessitate modifying the parameters of processing and product formulation, potentially resulting in modified product quality. This article provides a detailed explanation of the components of soybean seeds and highlights the important influences of environment and genotype on the structural, compositional, functional, and nutritional qualities of soybeans (Sugano *et al*., 2006). Among agricultural crops, soybean (Glycine max (L.) Merrill.) holds a prominent place as the primary source of both vegetable oil and high-quality concentrated proteins. For many generations, people in Asia and other areas of the world have utilized soybean seeds to produce a wide range of fresh, fermented and dried dishes (Probst *et al*., 1973). Nutritious food items made from soy, such miso, tofu, milk and sauce, have been produced for human consumption, while oil-extracted soy meal is utilized as a wholesome animal feed. In addition to being used domestically, soy oil has a wide range of applications in the manufacturing of chemicals, plastics, papers, inks, paints, varnishes, insecticides and cosmetics.

A new avenue for renewable energy sources for industrial applications has recently been made possible by the use of soy oil to make biodiesel. Because soybeans are legumes, they can use atmospheric nitrogen through a process called biological nitrogen fixation, which reduces their reliance on artificial nitrogen fertilizers. Given its many uses, there is strong evidence to support its significant participation in significant crop development initiatives throughout the globe. Soybeans are members of the Papilionaceae subfamily and the Leguminosae family. In 1948, Ricker and Morse advocated that the right name for the cultivated soybean be G. max (L.) Merrill. (Ferraz de Toledo *et al*., 1994). There are two subgenera within the genus Glycine: Soja (annuals) and Glycine (perennials). According to Hymowitz et al. (2004), there are 22 recognized species of perennials and two species of annuals: G. max L. Merrill. (Cultigen) and G. soja Sieb. & Zucc. (Wild relatives and progenitor of G. max). In the highly self-pollinated annual G. max, natural cross-pollination often occurs at a rate of less than 1%, however it can occasionally reach 2%–3%. According to reports, Glycine argyrea and Glycine clandestina perennial species can exhibit up to 60% out-crossing (Brown *et al*., 1986). With 2 n = 2 x = 40, both cultivated and wild soybeans are paleopolyploids and are fully cross-compatible (Hymowitz *et al*., 2004). According to Arumuganathan and Earle (1991). The genome of soybean is relatively big (1.12×10^9 bp), with highly repetitive sequences making up around 55% of the genome (Danesh *et al*., 1998). Soybean is a crop that is grown all over the world and many scholars have written extensively on its history. (Hymowitz, 1970; Guo, 1993; Singh, 1999; Guo *et al*., 2010).

2.   Material and Methods

The research was conducted at the Genetic Resources division of the Ministry of Science and Education of Azerbaijan spanning the years 2022 to 2023. A total of 51

varieties, namely Kanata, CHU-11, Ген-8, БК-98, Kyota, Л-1, Л-12, Krasnodar-68, Л-3, Alexa, CHU-10, Л-4, Л-7, Л-9, Opus, CHU-14, Л-8, Anjelica, Л-2, Кол-6, Л-11, БК-82, Л-10, Кол27, Sinara, Asuka, Antonia, REGALE, Biyson, Umanskaya, Bravo, Kanada 4, Kanada 5, Kanada-6, Kanada 7, Л-5, БК-88, Л-6, БК-83, БК-104, Kanada 1, CHU-1, Arisa, Renta, Aqroyol, Kofu, Slava, Fateh, Merjan, Vilana and Lan, were utilized as the subjects of the study. The experimental material comprised 51 distinct genotypes.

Electrophoretic analysis of globulin reserve proteins was carried out at the "Biochemical Genetics and Technology" department of the Genetic Resources division of the Ministry of Science and Education of Azerbaijan. The extraction and electrophoretic analysis of globulin storage proteins in soy grain samples were conducted using a novel method, refined from the modification of Poperelya's technique, employing polyacrylamide gel electrophoresis (A-PAGE).

Table 1. Comprehensive Catalogue of 51 Soybean (*Glycine max* L.) Genotypes Utilized in AZGR Experimentation: Nomenclature and Provenance

	Genotypes	AZGR name	Origin		Genotypes	AZGR name	Origin
1	Kanata	13566	Canada	27	Л-7		Uzbekistan
2	CHU-1	13888	Turkiye	28	Л-5		Uzbekistan
3	CHU-11	13898	Turkiye	29	CHU-10	13897	Turkiye
4	Л-1		Uzbekistan	30	Kanada 1	13880	Canada
5	Л-4		Uzbekistan	31	Kanada 4	13883	Canada
6	Bravo	13879		32	Umanskaya	13906	
7	Kanada 5	13884	Canada	33	Opus	13568	
8	Arisa	13567		34	REGALE	13216	Italia
9	Renta	7172		35	Kyota	13565	
10	Л-11		Uzbekistan	36	Kanada 7	13886	Canada
11	Aqroyol	13887		37	БК-104		Uzbekistan
12	Л-9		Uzbekistan	38	Merjan	14091	Azerbaijan
13	Л-6		Uzbekistan	39	Л-1		Uzbekistan
14	БК-88		Uzbekistan	40	Vilana	4563	Russia
15	БК-98		Uzbekistan	41	Кол-6		Uzbekistan
16	БК-83		Uzbekistan	42	Anjelica	13214	Australia
17	Sinara	13904	Austria	43	Л-3		Uzbekistan
18	Л-12		Uzbekistan	44	Krasnodar-68	13902	
19	Kofu	13569		45	Кол-27		Uzbekistan
20	Л-2		Uzbekistan	46	Biyson	13878	
21	БК-82		Uzbekistan	47	Asuka	13570	Japonia
22	Slava	4564	Ukraina	48	Antonia	13215	Australia
23	Fateh	13862	Azerbaijan	49	Kanada-6	13885	Canada
24	Alexa	13241	Australia	50	Lan	7171	
25	Л-8		Uzbekistan	51	Ген-8		Uzbekistan
26	CHU-14	13901	Turkiye				

After grinding, soybean samples underwent extraction twice with 500 μl of 70% alcohol, followed by centrifugation at 3500 rpm each time. Subsequently, the samples were washed twice with a 500 μl solution of 0.03% vinegar and acetone and after each

wash, they were rapidly stirred mechanically in a centrifuge rotating at 3500 cycles/min. Following the fourth wash, 500 μl of a 9 molar vinegar-urea solution was added to the extract, which was then analyzed using vertical electrophoresis in a glycine-acetate buffer (pH-3.5).

Based on the Jacquard genetic distance index, a dendrogram illustrating genetic kinship was constructed utilizing the DARwin software package employing the UNJ (unweighted pair group method with arithmetic mean) method.

3. Result and discussion

Extraction and electrophoretic analysis of globulin storage proteins in soybean grains were conducted using polyacrylamide gels (A-PAGE) following the method outlined by Zilman and Bushuk (1979), with modifications introduced by Poperelya (1989).

In the electrophoretic analysis conducted via the A-PAGE method, reserve proteins of globulin are conventionally categorized into ω, γ, β- and α-zones. As a result, the electrophoretic profiles of globulin among the genotypes exhibited marked differences. Numerical designation of patterns was achieved through comparative analysis within each zone, with each pattern receiving a unique number irrespective of repetitions. Consequently, if a particular pattern recurred within the samples, it was not assigned a new number and all patterns were identified according to this criterion. The frequency of occurrence for each pattern within the soft wheat samples was calculated utilizing the Nei (1979) genetic diversity index for all zones.

$$H = 1- \Sigma p_i^2,$$

where H - genetic diversity index; p_i is the frequency of each pattern in the zones.

Among the soybean samples under study, a total of 22 spectra and 107 patterns were identified, with polymorphism assessed based on the frequency of occurrence of patterns formed by these globin electrophoretic spectra. In the - zone of electropherograms of globin reserve proteins, 9 spectra and 14 different patterns were examined, while the γ-zone revealed 6 spectra and 32 patterns. Within the γ-zone, patterns ω-2, ω-4, ω-8, ω-9, ω-10, ω-12 and ω-13 were found at a frequency of 0.01% and pattern ω-3 occurred at 0.12% in 51 samples. Notably, among the spectra, ω1s exhibited a high occurrence frequency of 68%, ω5S showed a medium frequency of 43.0% and ω2S displayed a low occurrence frequency of 23%. Further analysis of the γ-zone identified 7 spectra and 38 patterns. The γ-1 pattern occurred at a frequency of 0.4%, while the γ-2 pattern was observed at 0.01%. Within the spectra, γ5s showed a high frequency of 66.0%, whereas γ7s had a lower frequency of 50%. In the ß-zone, 5 spectra and 23 patterns were detected. The ß-1 pattern exhibited the highest frequency at 0.81%, while the ß-23 pattern showed the lowest frequency at 0.01%. Notably, the occurrence frequency of ß-2s was 62.0% higher, while that of ß-3s was 49.0% lower.For the α-zone, 4 spectra and 14 patterns were recorded. The α-9 pattern had a frequency of 0.8% and α-12 pattern, α-s2 exhibited a maximum frequency of 50.0%, whereas α-s1 displayed a lower frequency of 43.0%. The genetic diversity index was calculated for all four zones using Nei's formula between genotypes, revealing varying levels of diversity. The ω-zone showed a low genetic diversity index (H=0.731), the γ-zone exhibited a medium level (H=0.856), the β-zone displayed a relatively high index (H=0.947) and the α-zone demonstrated high genetic diversity (H=0.911).
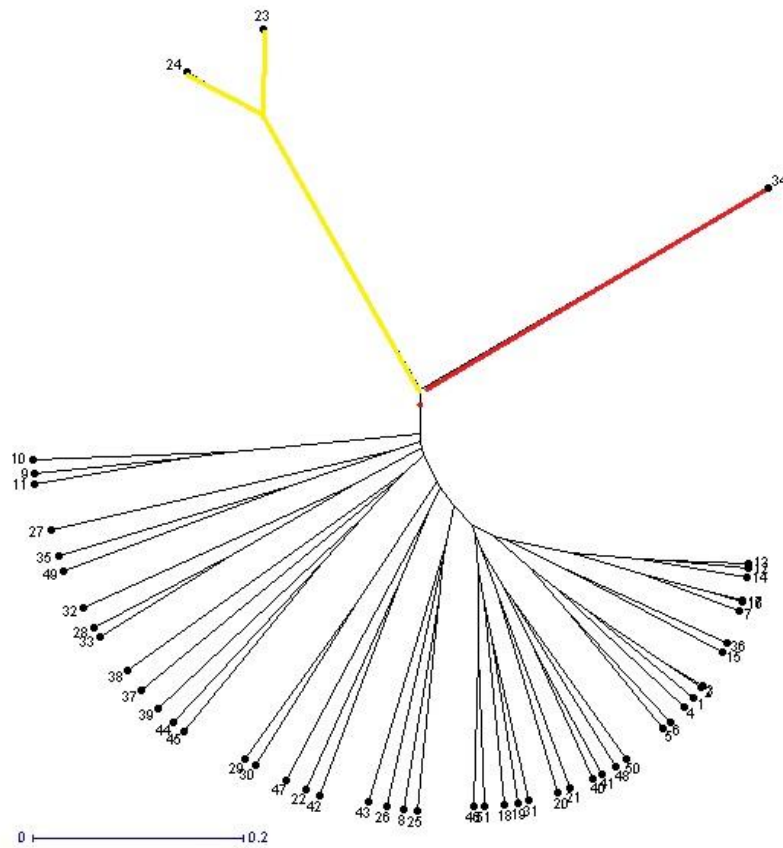
Figure 2. A DARWIN software elaboration performed using the Neighbor Joining method, depicting the estimated relationship among 51 Soybean genotypes

In Figure 2, the dendrogram depicts three distinct color groups representing different genotypes. Genotype REGALE from Italia is highlighted in red, while genotypes Fateh from Azerbaijan and Alexa from Australia are marked in yellow. All remaining genotypes are depicted in black. The numbers at the end of the tree branches indicate the accession numbers (Figure 2).

4. Conclusion

The electrophoretic analysis of soybean globulin storage proteins showed a diverse range of patterns and spectra, with the genetic diversity index reflecting different levels of genetic diversity among the soybean genotypes. The ω-zone exhibited low diversity (H=0.731), the γ-zone medium (H=0.856), the β-zone relatively high (H=0.947) and the α-zone high diversity (H=0.911). This diversity is crucial for understanding the genetic variability and potential for breeding in soybean crops.

References

Agyei, D., Potumarthi, R. & Danquah, M.K. (2015). Food-derived multifunctional bioactive proteins and peptides: Applications and recent advances. In *Biotechnology of Bioactive Compounds: Sources and Applications*, 507–524.

Alizade, S., Mammadova, R. & Sirajli, N. (2023a). Evaluation of morphometric traits of upland cotton genotypes under different concentration of NaCl. *Advances in Biology & Earth Sciences, 8*(3), 301-307.

Alizade, Sh., Mammadova, R. (2023b). Assessment of salt stress resistance of cotton varieties based on different parameters. *Advances in Biology & Earth Sciences, 8*(1), 58-66.

Arumuganathan, K., Earle, E.D. (1991). Nuclear DNA content of some important plant species. *Plant Molecular Biology Reporter*, 9, 208–219.

Brown, A.H.D., Grant, J.E. & Rullen, R. (1986). Outcrossing and paternity in Glycine argyrea by paired fruit analysis. *Biological Journal of the Linnean Society*, 29, 283–294.

Danesh, D., Penuela, S., Mudge, J., Denny, R.L., Nordstrom, H., Martinez, J.P. & Young, N.D. (1998). A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. *Theoretical and Applied Genetics*, 96, 196–202.

Ferraz de Toledo, J.F., Alves de Almeida, L., de Souza Kiihl, R. A., Carrao Panizzi, M.C., Kaster, M., Miranda, L.C. & Menosso, O.G. (1994). Genetics and breeding. In *Tropical Soybean Improvement and Production*, 19–36.

Guo, J., Wang, Y., Song, C., Zhou, J., Qiu, L., Huang, H. & Wang, Y. (2010). A single origin and moderate bottleneck during domestication of soybean (Glycine max): Implications from microsatellites and nucleotide sequences. *Annals of Botany*, *106*(3), 505-514. https://doi.org/10.1093/aob/mcq125

Guo, W.T. (1993). *The History of Soybean Cultivation*. Hehai University Press.

Hymowitz, T. (1970). On the domestication of the soybean. *Economic Botany*, 24, 408–421.

Hymowitz, T. (2004). Speciation and cytogenetics. In *Soybeans: Improvement, Production and Uses*, 16, 97–136.

López-López, A., Rosenblueth, M., Martínez, J. & Martínez-Romero, E. (2010). Rhizobial symbioses in tropical legumes and non-legumes. In *Soil Biology and Agriculture in the Tropics,* 163–184.

Probst, A.H., Judd, R.W. (1973). Origin, US history and development and world distribution. *Agronomy*, 1973.

Poperelya, F.A. (1989). *Gliadin polymorphism and its relationship with grain quality, productivity, and adaptive properties of winter soft wheat varieties*. Moscow, Agropromizdat, 138-149 (in Russian).

Rodríguez-Navarro, D.N., Oliver, I.M., Contreras, M.A. & Ruiz-Sainz, J.E. (2011). Soybean interactions with soil microbes, agronomical and molecular aspects. *Agronomy for Sustainable Development*, 31, 173–190.

Salmanova, M.R., Sadygov, G.B., Kerimov, A.Ya., Sadygova, S.B. & Kalbieva, E.E. (2023). Study of the genetic diversity of globulin proteins in chickpea (*Cicer arietinum* L.) genotypes. *Advances in Current Natural Sciences*, 4, 21–26.

Singh, R.J., Hymowitz, T. (1999). Soybean genetic resources and crop improvement. *Genome*, 42, 605–616.

Sugano, M. (2006). Nutritional implications of soy. In *Soy in Health and Disease Prevention,* 1–16.

Sugiyama, A., Ueda, Y., Takase, H. & Yazaki, K. (2015). Do soybeans select specific species of Bradyrhizobium during growth? *Communications in Integrative Biology*, *8*(4). https://doi.org/10.1080/19420889.2015.992734

United States Department of Agriculture FAS. (2011). *Oilseeds: World market and trade archives*. Full report (10–11). http://www.fas.usda.gov/oilseeds/Current/default.asp

United States Department of Agriculture Foreign Agricultural Service. (2011). *Oilseeds: World market and trade archives*. http://www.fas.usda.gov/oilseeds_arc.asp

United States Department of Agriculture. (2013). *Soybeans: Supply, disappearance and price, US, 1980/81–2012/13*. http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1290

Wingeyer, A.B., Amado, T.J.C., Pérez-Bidegain, M., Studdert, G.A., Varela, C.H.P., Garcia, F.O. & Karlen, D.L. (2015). Soil quality impacts of current South American agricultural practices. *Sustainability*, 7, 2212–2242.

Zilman, R.R., Bushuk, W. (1979). Wheat cultivar identification by gliadin electroforegrams. *Canadian Journal of Plant Science,* 59, 287-298.