Evaluation of genetic and morphological distances between soybean (*Glycine max* L.) cultivars

Larysa PRYSIAZNIUK, Yuliia SHYTIKOVA, Iryna DIKHTIAR, Nataliia MIZERNA

Ukrainian Institute for Plant Variety Examination
15 Henerala Rodymtteva, 03041 Kyiv, Ukraine
E-mail: prysiaznhiuk_1@ukr.net

Abstract

The study focuses on the polymorphism in soybean (*Glycine max* (L.) Merr.) cultivars based on molecular and morphological marker traits. The evaluation of soybean cultivars by DNA markers was carried out with the aid of polymerase chain reaction (PCR) analysis using four microsatellite markers (Satt 228, Satt 726, Satt 063 and Satt 114) and 20 morphological traits. The frequency of identified alleles varied from 0.02 to 0.28; polymorphism index by the markers under study averaged 0.89. Genetic distances between the cultivars by simple sequence repeat (SSR) markers and morphological markers were determined using cluster analysis. According to the obtained distribution by microsatellite markers, the largest distance (3.87) was between cultivars ‘Alaska’ and ‘Alinda’. The most related were cultivars forming the same cluster with the value of 2.00, namely DH 530 and ‘Abelina’, ‘OAC Leikviu’ and ‘Monarkh’, ‘SG SR Picor’ and ‘Hieba’. Cluster analysis of soybean cultivars by morphological traits showed that ‘Abelina’ was the most distant from the group of investigated cultivars with a value of genetic distances ranged from 2.8 to 11.0. Cultivars ‘Amadeus’ and DH 530, which formed a cluster and were at a distance of 1.4, appeared the most morphologically similar. Analysis of genetic distances by SSR markers and morphological traits showed a positive correlation by Mantel test. Description of morphological traits and microsatellite markers is useful for the identification of soybean cultivars, building-up collections of well-known cultivars and determination of the differences between them.

Key words: cluster analysis, Mantel test, molecular and genetic polymorphism, simple sequence repeat.

Introduction

Soybean (*Glycine max* (L.) Merr.) is an important leguminous crop and a major source of oil and protein. Recent investigations found that genetic diversity of elite soybean germplasm is limited (Mulato et al., 2010; Zhang et al., 2016; Orazaly et al., 2018; Shi et al., 2018). Along with the increase in the number of cultivars having minimal differences and the possibilities of phenotypic variability, examination of cultivars by DNA markers is timely (Tasma et al., 2011; Volkova et al., 2015). Studies of soybean based on simple sequence repeat (SSR) polymorphism are used to determine the relationship between genotypes, the purity of commercial cultivars and to assess the genetic diversity (Tantasawat et al., 2011; Dong et al., 2014; Chakraborty et al., 2018; Sun et al., 2018). This type of analysis provides sufficient stability to determine the distinctiveness of cultivars and build-up collections of well-known cultivars for distinctness, uniformity and stability (DUS) examination (UPOV, 1998; Volkova, 2014; 2015).

The issue of the relationship between morphological traits and DNA markers has been studied for over 30 years (Burstin, Charcosset, 1997; Geng et al., 2016; Valliyodan et al., 2016). To explore the effect of heterosis, the studies were carried out to determine the correlations between genetic distances of morphological traits and SSR, inter simple sequence repeat (ISSR) and amplified fragment length polymorphism (AFLP) markers in sunflower and corn (Burstin, Charcosset, 1997; Hudcovicova, Kraic, 2003; Tantasawat et al., 2011; Darvishzadeh, 2012; Abugalieva, 2013; Akond et al., 2013; Goncharov et al., 2016; Samanfar et al., 2017). Leading European plant expertise institutes, such as GEVES (France) and Naktuinbouw (The Netherlands), use the correlation between genetic distances based on SSR marker and morphological trait polymorphism to evaluate new cultivars and to compose their collections for DUS-test (Riday et al., 2003; Karuri et al., 2010; Valliyodan et al., 2016; Shen et al., 2017; Sun et al., 2018).

In Ukraine, the evaluation of soybean polymorphism is used to arrange breeding process and mark economic and valuable features; therefore, the investigation of the relationships between genotypes with the involvement of DNA polymorphism evaluation for the examination of soybean cultivars is relevant.
The purpose of this study was to evaluate the correlation between genetic distances based on SSR marker and morphological trait polymorphism of soybean cultivars.

**Materials and methods**


In order to evaluate the polymorphism by simple sequence repeat (SSR), 30 genotype samples of each cultivar were taken. DNA extraction was carried out using cetyltrimethylammonium bromide (CTAB). Chloroform-isoamyl alcohol was used for the first purification and ethanol solution for the second one (Akkaya et al., 1992; Song et al., 2004; Poik et al., 2010; Jun et al., 2011; Prysiazhniuk et al., 2017).

The molecular and genetic polymorphism of soybean cultivars was studied by four microsatellite loci with specific primers: Satt 114, Satt 228, Satt 726 and Satt 063 (Song et al., 1999; Shi et al., 2009) (Table).

**Table.** Characteristics of simple sequence repeat (SSR) loci primers

<table>
<thead>
<tr>
<th>Microsatellite locus</th>
<th>Forward primer sequence (5´3´)</th>
<th>GC (%) reverse primer sequence (5´3´)</th>
<th>Motif</th>
<th>Hybridization temperature °C</th>
<th>Expected amplicon size, bp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satt 726</td>
<td>gcttttttagttaggataaatttt</td>
<td>gegaagggacgaaggtgt</td>
<td>(ATT)$_2$</td>
<td>55</td>
<td>170–280</td>
</tr>
<tr>
<td>Satt 063</td>
<td>aatgattacatggtttatat</td>
<td>actgtccatgtaaatcc</td>
<td>(ATT)$_2$</td>
<td>50</td>
<td>95–210</td>
</tr>
<tr>
<td>Satt 114</td>
<td>gggttatcctcccaata</td>
<td>ataggggatgataagtgaa</td>
<td>(ATT)$_1$</td>
<td>55</td>
<td>75–130</td>
</tr>
<tr>
<td>Satt 228</td>
<td>tctaaagcaagtatgtaaaa</td>
<td>cattataagaagactgtaaa</td>
<td>(ATT)$_1$</td>
<td>60</td>
<td>200–270</td>
</tr>
</tbody>
</table>

Reaction mixture was supplemented with 100 ng of DNA. The final concentration of the components was as follows: 1X buffer (10 mM Tris-HCl, pH 9.0, 50 mM KCl, 0.01% Triton X-100), 1.5–2.5 mM of MgCl$_2$, 200 μM of deoxynucleotide triphosphate (dNTP), 0.2–0.5 μM of each primer and 1 unit of Taq polymerase (Thermo Fisher Scientific, USA). The total volume of the mixture was 20 μL. Polymerase chain reaction (PCR) was carried out with the aid of amplifier TC- Y (Crea Con Technologies, The Netherlands). For each pair of primers, the following amplification parameters were used: initial denaturation at 94°C (2–3 min), denaturation 93°C (30 s), hybridization of primers 50–60°C (60 s), elongation at 72°C (60s), the number of cycles (35) and final elongation at 72°C (3 min). Visualization of PCR products was carried out by electrophoresis in a 2% agarose gel using 0.5× TBE (triborate buffer solution), according to Poik et al. (2010), Ramazanova (2016), Li et al. (2017) and Pagar et al. (2017) at the electric field intensity of 5 V cm$^{-1}$. The size of amplicons was determined using software STATISTICA, version 12 (trial version) (StatSoft Inc., USA). The grouping of cultivars by SSR markers was carried out by the method of unweighted average bonds and the grouping by morphological traits by unit bonds with the calculation of Euclidian distances (Fortin et al., 2002; Drozdov, 2010; Everitt et al., 2011; Dong et al., 2014).

The determination of the correlation between SSR markers and the morphological traits was carried out based on genetic distances using the Mantel test and software XLSTAT, version 2018 (Legendre, Fortin, 2010; Diniz-Filho et al., 2013).

**Results and discussion**

As a result of PCR by four SSR markers with specific primers, alleles of specific size were obtained. According to the obtained data, some cultivars were marked by intra-cultivar polymorphism, specifically by four SSR markers in ‘Alinda’, by Satt 228 and Satt 726 loci in ‘Arnika’, by Satt 063 and Satt 726 loci in ‘Furio’. A total of 60 alleles were detected with an average of 15 alleles per marker. Using the same markers, Song et al. (1999) and Shi et al. (2009) obtained from 7 to 18 alleles with PIC 0.77–0.82. The frequency of the identified alleles ranged from 0.02 to 0.28, depending on the locus, and the polymorphism index varied from 0.83 to 0.94 (0.89 on average). This indicates that the identified alleles are evenly represented in the soybean cultivars. In order to analyse the polymorphism of 25 soybean cultivars by SSR markers and morphological traits, a cluster analysis
was carried out and genetic distances between the cultivars were calculated (Fig. 1).

![Figure 1. Dendrogram of soybean cultivars based on simple sequence repeat (SSR) markers](image-1.png)

On the basis of the dendrogram we determined nine clusters by SSR markers Satt 063, Satt 114, Satt 228 and Satt 726, which included cultivars ‘Abelina’ and DH 530, ‘Monarkh’ and ‘OAC Lakeview’, ‘Berkana’ and DH 618, ‘Hieba’ and ‘SG SR Picor’, ‘Kano’ and ‘OAC Calypso’, ‘OAC Madok’ and ‘SG Eider’, ‘Alaska’ and ‘Arisa’, ‘Alinda’ and ‘Arnika’. Analysis of genetic distances between the studied soybean cultivars showed that the largest distance (3.87) was between ‘Alaska’ and ‘Arisa’. Along with the increase in the affinity of the cultivars, their genetic distances shorten. In our research, the most related cultivars were those with the value of 2.00: DH 530 and ‘Abelina’, ‘OAC Lakeview’ and ‘Monarkh’, ‘SG SR Picor’ and ‘Hieba’. The distances between the majority of cultivars were 3.61, 3.16 and 2.83 (Prysiazhniuk et al., 2017). According to the distribution data, cultivars ‘Alaska’ and ‘Arisa’ appeared to be Canadian, while ‘Alinda’ and ‘Arnika’ – Ukrainian. However, the data on the distribution of SSR markers between cultivars ‘Abelina’ and DH 530, ‘OAC Madok’ and ‘SG Eider’, ‘Monarkh’ and ‘OAC Lakeview’, ‘Berkana’ and DH 618, ‘Hieba’ and ‘SG SR Picor’, ‘Kano’ and ‘OAC Calypso’, ‘OAC Madok’ and ‘SG Eider’ led us to the conclusion that materials from different countries were used in the breeding process. However, it should be noted that despite the close proximity, the cultivars under study are different.

Shown in Figure 2 is a phylogenetic tree of hierarchical classification of soybean cultivars by morphological traits. The soybean cultivars under study composed nine clusters, which, according to morphological traits, placed in order of decreasing affinity between cultivars in separate clusters. The value of the genetic distance between the ‘Amadeus’ and DH 530 found in one cluster is 1.4. This indicates that they have the most common morphological traits. The next in terms of affinity is the cluster including ‘Alaska’ and ‘Nordika’ with a genetic distance of 1.7; ‘Karra’ is similar to these cultivars since it is adjacent to the specified cluster at a distance of 1.7. The clusters of cultivars DH 618 and ‘OAC Madok’, ‘OAC Calypso’ and ‘SG SR Picor’ are at the same level (distances of 2.2), by the number of identical morphological marker traits. The cultivars ‘Arnika’ and ‘Monarkh’ also share one cluster with a genetic distance of 5.6 and they have the least common traits compared to all previous clusters.

![Figure 2. Dendrogram of soybean cultivars by a combination of morphological traits](image-2.png)

It is noteworthy that the mentioned cultivars formed a cluster that was separate from other groups of clusters; therefore, it can be concluded that ‘Arnika’ and ‘Monarkh’ are close to each other in terms of morphological marker traits. Cultivar ‘Abelina’ also is placed apart from other cultivars and not included in any cluster. The value of genetic distances between ‘Abelina’ and other cultivars fluctuates between 2.8 and 11.0. The evaluation of soybean cultivars according to their morphological features showed a clear distribution of the cultivars under study according to their countries of origin. Thus, cultivars ‘Amadeus’ and DH 530, ‘Alaska’, ‘Nordika’ and ‘Karra’, DH 618 and ‘OAC Madok’ that formed the same clusters or are adjacent have a common country of origin (Canada). Cultivars ‘Arnika’ and ‘Monarkh’ from a separate cluster are Ukrainian. The ‘Abelina’, which is not included in any cluster, is different from other cultivars, because it is Austrian. Thus, cluster analysis of cultivars based on morphological traits allowed us to clearly track their distribution by the country of origin, while SSR-based analysis made it possible to identify the cultivars for which breeding materials from different countries were used.

To determine the correlation relationships between genetic distances obtained by SSR and morphological markers, Mantel test (linear correlation by Pearson) was used (Figs 3 and 4).

As a result of the analysis, the $p$-value and $r$ (AB) value at the significance level $\alpha = 0.05$ were found, which, according to the interpretation of the test results, allows us to accept the assumption of the presence (Ha) or absence (H0) of correlation. The Mantel test was used by Geng et al. (2016) for Alternanthera philoxeroides. It was found that the molecular marker distance is positively correlated with the dissimilarity of quantitative traits in terrestrial habitat and aquatic habitat. However, they detected no significant correlation between marker distance and dissimilarity of phenotypic plasticity across terrestrial and aquatic habitats.

It is known that the assumption of H0 (absence of correlation) is assumed under the condition of $p > \alpha$. In our study, the calculated value of $p$ (0.009) was lower than the significance level $\alpha = 0.05$; therefore, we should reject the assumption H0 and adopt the alternative assumption Ha regarding the correlation (Burstin,
Correlation coefficient and normality of data distribution by matrixes of genetic distances are illustrated in Figure 4.

Burstin and Charcosset (1997) studied the effect of the polygenic inheritance of the traits used to calculate observed for a protein with quantity controlled by a restricted number of loci. Darvishzadeh (2012) obtained different results in the assessment of the genetic diversity among some sunflower parental lines by agronomic and morphological traits and AFLP markers and then, in the evaluation of the association between parental lines genetic diversity with F1 performance and heterosis under well-watered and water-stressed conditions. There was no strong linear correlation between genetic distance and morphological distance, especially under the conditions of sufficient moisture.

The analysis showed a positive correlation ($r = 0.152$) between the genetic distances between 25 soybean cultivars calculated based on four microsatellite markers: Satt 063, Satt 114, Satt 228 and Satt 726, and their morphological traits. This proves the possibility of using the data obtained from the genetic profiles of cultivars to compose reference collections of well-known soybean cultivars.

Figure 4. Normality of distribution of Mantel test results for soybean cultivars by genetic distances

Conclusions

1. The data of cluster analysis showed that the most affined by Satt 063, Satt 114, Satt 228 and Satt 726 loci were the soybean cultivars with genetic distances of 2.00, namely, DH 530 and ‘Abelina’, ‘OAC Lakeview’ and ‘Monarkh’, ‘SG SR Picor’ and ‘Hieba’. The results of this analysis were used to create a database of molecular and genetic polymorphism of soybean cultivars under study for the purpose of their identification.

2. Using a set of morphological traits, it was found that the least value of genetic distance (1.4) was between cultivars ‘Amadeus’ and DH 530. ‘Abelina’ was the most distant (2.8–11.0) cultivar. Mantel test helped to find out a positive correlation based on genetic distance matrices by simple sequence repeat (SSR) markers and phenotypic distances on the relationship between these distances and heterotic or marker distances. It is consistent with the small correlations observed between marker distance and the distances computed from protein quantities, and the fact that the highest correlation is
3. Consequently, a set of studies on description of morphological traits and microsatellite markers as an additional analysis method is recommended for identifying soybean cultivars, composing collections of well-known cultivars and finding differences. This approach will allow for a more effective examination of distinctness, uniformity and stability (DUS) and will provide additional protection of the breeders' rights.

Received 27 03 2018
Accepted 02 12 2018

References


