Characterization of soybean varieties based on protein and oil content

Vesna Perić, Ana Nikolić, Snežana Mladenović Drinić, Natalija Kravić, Dragan Kovačević
Maize Research Institute „Zemun Polje”, Belgrade, Serbia, Slobodana Bajica 1, vperic@mrizp.rs

Abstract

The aim of this study was to evaluate 135 soybean accessions (cultivars and experimental lines from different geographical areas) obtained from the Gene Bank of the Maize Research Institute Zemun Polje, on the basis of their protein and oil concentration, and to identify the genotypes with the best protein: oil ratio as a potentially new sources of genetic variation for the soybean breeding program in the future. Seed samples of 135 entries were analyzed for protein and oil content by a near infra-red reflectance spectroscopy (NIRS) using Infratec 1241 Grain analyzer. Oil content ranged from 12.28 to 22.19%, while the protein content ranged from 32.18 to 48.20 %. Cluster analysis has been performed among 135 genotypes based on the data combining two traits. The dendrogram based on combined traits shows that all genotypes are grouped into two main clusters, first consisted of 5 and second consisted of 2 subclusters. Within subclusters, different categories of genotypes are identified according to their protein and oil content (low protein-medium oil, low protein-low oil, low protein-high oil, medium protein-high oil, medium protein-low oil, high protein-lower oil and high protein-medium oil). The majority of the varieties were classified as intermediate category (medium protein-medium oil), while only few genotypes were identified as high protein-medium oil and high oil-medium protein type, providing a valuable source of germplasm for the improvement of the new cultivar development.

Key words: cluster, oil, protein, soybean

Introduction

Seed protein and oil content are the two main seed quality traits in soybean. Soybean contains 38 to 40% good quality protein and approximately 19% oil, expressed on dry matter basis (Wilson, 2004), which make this crop major protein and oilseed source in many regions worldwide. The differences in seed composition are inherent among soybean cultivars, but it has also been observed significant environmental impact on protein and oil concentration (Miladinović et al., 1996, Yin and Vyn, 2005, Balešević-Tubić et al., 2009). Soybean breeding programs, commonly, do not look for the improvement of isolated traits, but for the genetic enhancement in set of the correlated traits. Numerous studies have shown that seed protein and oil concentration is generally negatively correlated (Voldeng et al., 1997, Morrison et al., 2000). According to results obtained by Balesevic-Tubic et al., (2011), this correlation is influenced by environmental factors and negative relationship was observed only in high yielding environments. The highly negative correlation between oil and protein content (Wilcox and Cavins, 1995, Cober and Voldeng, 2000) as well as the negative correlation between protein and yield is a problem in obtaining high yielding and high protein cultivars. Genetic success at increasing both yield and oil concentration has maintained protein concentration, but high protein cultivars tend to have low oil concentration (Haq and Mallarino, 2005). Greater opportunities for improvement of these traits come from crosses between genetically distant parents. Agronomic and morphological characters submitted to cluster analysis allow unifying multiple information from a set of characters.
that result in information about the diversity of breeding material. The objective of this study was to quantify the oil and protein content in 135 soybean accessions (cultivars and experimental lines from different geographical areas) from the Gene Bank of the Maize Research Institute Zemun Polje, and to identify the genotypes with the best protein: oil ratio as a potentially new sources of genetic variation for the soybean breeding program in the future. The NIR spectroscopy is quick and reliable method for determination of protein and oil concentration in intact seed (Balešević-Tubić et al., 2008).

**Materials and methods**

A total of 135 soybean accessions from maturity groups 000-IV (commercial cultivars and experimental lines from different geographical areas) were selected to represent Maize Research Institute soybean collection. Genotypes were grown in year 2010 at the experimental fields in Zemun Polje, in a randomised complete block experiment with two replicates and two rows per genotype, including one row of isolation between experimental plots to eliminate the effects of varietal competition. After harvesting, seed samples of 135 entries (two samples of each) were analyzed for protein and oil content with a near infra-red transmittance spectroscopy using instrument FOSS Infratec 1241 Grain analyzer. Seed protein and oil content are expressed in percentage (%) on a dry matter basis. The nearest neighbour method for clustering was applied, using Squared Euclidean Distance as similarity measure among soybean cultivars. A relationship among 135 genotypes based on combined traits is illustrated by dendrogram. Statistical program SPSS 15.0 for Windows Evaluation Version was used to cluster the data.

**Results and discussion**

Protein content varied from 32.18 to 48.20 %, while oil content ranged from 12.28 to 22.19%, which is in agreement with previous studies on soybean oil and protein concentration (Miladinović et al., 2008, Taški-Ajduković et al., 2010). Dendrogram was constructed on the basis of the Squared Euclidean Distance similarity matrices, estimated from two combined quantitative traits analyzed in 135 soybean accessions. Cluster analysis revealed two main groups, one formed by 117 (cluster A) and the other by remaining 14 genotypes (cluster B). Both groups were further classified into a several sub-clusters at different similarity coefficient levels (Figure 1).

**Table 1.** Classification of soybean accessions in sub-clusters based on protein and oil content

<table>
<thead>
<tr>
<th>Sub-cluster</th>
<th>Phenotypic categories</th>
<th>Average protein (%)</th>
<th>Average oil (%)</th>
<th>Number of genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>low protein-high oil</td>
<td>35.78</td>
<td>19.66</td>
<td>53</td>
</tr>
<tr>
<td>b</td>
<td>low protein-low oil</td>
<td>35.08</td>
<td>17.36</td>
<td>3</td>
</tr>
<tr>
<td>c</td>
<td>low protein-high oil</td>
<td>34.45</td>
<td>21.32</td>
<td>6</td>
</tr>
<tr>
<td>d</td>
<td>medium protein-high oil</td>
<td>37.19</td>
<td>21.30</td>
<td>3</td>
</tr>
<tr>
<td>e</td>
<td>medium protein-medium oil</td>
<td>38.07</td>
<td>19.16</td>
<td>58</td>
</tr>
<tr>
<td>f</td>
<td>medium protein-low oil</td>
<td>38.65</td>
<td>16.66</td>
<td>4</td>
</tr>
<tr>
<td>g</td>
<td>high protein-medium oil</td>
<td>40.58</td>
<td>17.91</td>
<td>10</td>
</tr>
</tbody>
</table>

Cluster A comprised 5 sub-clusters (Table 1), including sub-cluster “e”, consisted of the majority of the varieties (42.96 %). These varieties are classified as intermediate category (medium protein-medium oil). The low protein-high oil category contains 53 lines (39.26% of the genotypes observed). Genotypes Mini Soja and Nato are classified to high protein-low oil group, with average protein content 46.38%. These genotypes, along with genotypes PI 416 and Pod Cluster Cuuse (low protein-medium oil category), are independently associated with the main cluster.
Characterization of soybean varieties based on protein and oil content

Figure 1. Dendrogram of similarity patterns based on Squared Euclidean Distance estimated from combined traits (protein and oil concentration) analyzed in 135 soybean genotypes
The most important for the breeding program are the varieties from the groups "d" (medium protein-high oil) and "g" (high protein-medium oil). Those varieties are identified as a valuable source of breeding material for improvement of seed quality and will be used as parents for the hybrid combinations.

In conclusion, our study has shown that soybean genetic resources from the Maize Research Institute Gene Bank clearly showed a wide range of variation in protein and oil content. It is considered that many varieties hold great potential and would be useful as materials in breeding cultivars for industrial processing.

References


