**Evaluation of soybean seed shape by elliptic Fourier descriptors**

*Maki Eguchi¹, Seishi Ninomiya¹,²*

¹ University of Tsukuba, Japan
² National Agriculture and Food Research Organization, Japan

**Abstract**

Round seeds are desirable for food-type soybean [Glycine max (L.) Merr.], and the genetic control of seed geometry is of scientific interest. The aim of this research was to: (1) compare the seed shape variation of 165 soybean cultivars and to quantify the characteristics of the cultivars, and, (2) to evaluate the seed shape of the parental F1 and F2 populations derived from Tamahomare × Peking and Peking × Tamahomare, and to estimate the heritability of the shape factors. For experiment 1, 10 seeds of each cultivar were photographed from the hilum (H) and lateral (L) directions, and the shape from each direction was independently analyzed. For experiment 2, 30 seeds from each of P1 and P2, 100 seeds from F1, and 10 seeds from each of 239 F2 lines were analyzed in the same manner as for experiment 1. The images of the seeds captured were then analyzed based on EF-PCA (Elliptic Fourier-Principal Component Analysis) by using the shape analysis software package, SHAPE (Iwata et al., 2003). The EF-PCA analysis describes seed shape mathematically by transforming seed contour coordinates into elliptic Fourier descriptors, and subsequently summarizes these descriptors by principal component analysis to evaluate shape with comparatively lower number of principal components. We could separate the shape variation into symmetric variation and asymmetric variation against the center axis of the soybean seed. PC1 (1st principal component) and PC2 of the symmetric variation of the seed shape observed from the H direction indicated the aspect ratio and roundness, respectively, while PC1 of the symmetric variation observed from the L direction indicated the aspect ratio. For the symmetric variation, the contribution ratio of PC1 was over 90%, showing that most of the seed shape variation could be explained by the aspect ratio. The broad sense heritabilities were 0.894, 0.836, and 0.405, for PC1-H, PC2-H, and PC3-H, respectively, and 0.892, 0.839, and 0.877, for PC1-L, PC2-L, and PC3-L, respectively.

**Keywords**: Soybean, Seed shape, Genetics, Image analysis, elliptic Fourier descriptors

**Introduction**

Soybean is one of the most important crops in the world. In specialty soy food products, including tofu, natto, miso, and edamame, seed size and shape are important traits (Salas 2006). Round seed is often desirable for food-type soybean, while desirable seed size ranges from large, for tofu and miso, to small, for natto production (Wilson, 1995). Nelson and Wang (1989) determined that seed size and shape in soybean were variable among a collection of varieties and that these traits were stable in time. Sakai and Yonekawa (1991) used digital image analysis for three-dimensional analysis of soybean seeds. The variation of seed shape in the small size ranges was larger than that in the large size ranges. In other words, the large seed was more uniform in shape than the small seed (Sakai and Yonekawa, 1991); although, this
three-dimensional analysis was limited to one cultivar. Correlations between seed shape and seed size were generally not significantly different from zero (Cober et al, 1997).

Cober et al.(1997) estimated the heritabilities for seed size and shape, obtaining moderate to high values, 59-79% for seed shape and 19-56% for seed size. They say that since a soybean seed has a uniform elliptical shape, three-dimensional seed shape can be quantified with two aspect ratios (ratios of minimum to maximum seed diameter) which include the three measurements of height, length and thickness used by Nelson and Wang (1989) to describe seed shape. But the seed shape cannot be quantified with only aspect ratio.

Digital image analysis is one way to compensate for the weakness of the current qualitative evaluation of continuous variation in various characteristics. With recent improvements in computer performance, digital image analysis has been applied to quantitative evaluation in various agricultural and biological research, such as disease assessment (Martin and Rybicki, 1998; Niemira et al.,1999; Olmstead and Lang, 2001), measurement of various plant canopies (Adamsen et al., 1999;Olmstead et al., 2004; Yoshioka et al., 2004a, 2004b, 2005). In the evaluation of petal shape, a combination of elliptic Fourier descriptors and principal-components analysis (EF-PCA) seemed to be most effective among several methods suggested. EF-PCA describes overall petal shape mathematically by transforming coordinate information concerning its contour into elliptic Fourier coefficients (Kuhl and Giardina, 1982), and summarizes these coefficients by means of principal-components analysis (Rohlf and Archie, 1984).

This study aims to: (1) clarify if soybean seed shape is genetically controlled by comparing the seed shape variation of 165 soybean cultivars, and, (2) to estimate heritability of the shape factors, using the material of parents, their F1 and F2 populations, by utilizing such a powerful tool to analyze plant shape, elliptic Fourier-Principal Component Analysis.

**Materials and methods**

**Plant materials**
To find if there is significant varietal difference among soybean seed shape, and to evaluate characteristics of each soybean cultivars’ seed shape, we used 10 grains of each of 165 cultivars. For the second heritability analysis, we used cultivar Tamahomare (6 individuals, 5 seeds each) and Pekin (6 individuals, 5 seeds each) and their F1 (5 seeds each for right and back cross progeny, total 10 seeds each for 10 individuals) and F2 population (239 strains, 10 seeds each). All the materials were grown and photographed by slide film from both hilum side and lateral side in 1994 at the Chushin Agricultural Experimental Station of Nagano prefecture, Japan.

**Image processing and contour recording**
We converted analog slide film images to digital images by using Canoscan 9950F for scanning and Canon Remote Capture version software ver.2.7.5.27. and finally produced binary black silhouettes images in BMP format using Paint Shop Pro software. From these images, we extracted the closed contours of the seeds and chain coded them (Freeman,1974) for the further analysis.

**Elliptic Fourier descriptors and Principal Component Analysis**
The coefficients of elliptic Fourier descriptors were calculated by the discrete Fourier transformation of the chain-coded contour through the procedure proposed by Kuhl and Giardina (1982). In this study, we approximated the shape by the first twenty harmonics. Thus, 77 (4×20-3) coefficients of normalized elliptic Fourier descriptors were derived from the
contour of seed shape, and were used for the analysis. Then, we conducted principal component analysis to summarize the variation of seventy seven coefficients into lower number of the variables and to determine the effect of each principal component (PC) on seed shape by recalculating the Elliptic Fourier coefficients, assuming that the score for a particular PC equal the mean ± 4SD (standard deviations) while keeping the scores of the remaining components equal to the means. The analysis was conducted for symmetric variation and asymmetric variation separately (Iwata and Ukai, 2002). We performed a series of these analyses (EF-PCA) using the SHAPE software package developed by Iwata and Ukai (2002).

**ANOVA and Heritability Estimation**

In this study, principal components scores are used as shape factors. Varietal difference in the shape factors was tested using ANOVA and broad sense heritability were estimated based on the data from F1 and F2.

**Results and Discussion**

We conducted EF-PCA including all the materials. Cumulative contribution ratios of the first three principal components of symmetric variation were 95.9% for lateral and 98.3% for hilum, respectively. This result indicates that the shape variation of the soybean seed is mostly explained by the first three principal components of symmetric variation. Fig.1 shows the seed shape reconstructed assuming ± 4SD variation of the first principal component score (PC1), the second principal component (PC2) and the third principal component (PC3) of both lateral view and hillum view. The results indicate that PC1 of L (lateral view) and H (hilum view) represent aspect ratio and PC2 of H represents roundness.

![Fig.1. Emphasized effect (-4s.d. - +4s.d.) of each principal component on soybean seed shape. The numbers 1,2, and 3 indicate PC1, PC2, and PC3 respectively. L: lateral view. H: hilum view.](image)

The results of ANOVA for 165 soybean cultivars (Tab.1) showed that PC1, PC2, and PC3 of symmetrical and asymmetrical components in lateral view were significantly different, while those for only symmetrical components in hilum view are significantly different. The results strongly indicate genetic control on soybean seed shape. Fig.2 shows four examples of mean seed shape of lateral view.

Following the above result, we estimated broad sense heritability of PC1s for both lateral and hilum views, using Tamahomare, Peking, and those of F1 and F2 derived from the parents (Fig.3). Contribution of the first few principal components calculated only with this particular data set (Table 2) was again as high as that for all the data sets. That is; the contribution ratios of PC1 for lateral and hilum views were 91% and 98%, respectively, and the cumulative total
contributions of the first three principal components for lateral and hilum views were 98% and 99% respectively (Tab.1). Then we compared PC1 of Tamahomare, Peking and those of F1, F2 (Fig.3) and calculated broad sense heritabilities (Tab.1).

<table>
<thead>
<tr>
<th>Lateral View</th>
<th>Hilum View</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetric</td>
<td>Asymmetric</td>
</tr>
<tr>
<td>PC1</td>
<td>0.0001&gt;</td>
</tr>
<tr>
<td>PC2</td>
<td>0.0001&gt;</td>
</tr>
<tr>
<td>PC3</td>
<td>0.0001&gt;</td>
</tr>
</tbody>
</table>

Pekin-L   Harosoy-L  Enrei-L  Tamahomare-L

Fig.2. Typical shape of lateral view of representative 4 cultivars of soybean.

As for Fig.3., PC1 distributions of F1 of L and H are located between those of parents, and those of F2 showed unimodal distributions. Fig.3. also shows strong absence of transgressive
distortion. Broad sense heritability values were both 0.89 in lateral and hilum views (Tab.2). As well, in both L and H, there was not a significant difference between reciprocal matings in F1 and F2.

Table 2 Heritability and contribution ratio of PC1, PC2 and PC3 for the analysis based on Peking, Tamahomare, F1 and F2

<table>
<thead>
<tr>
<th></th>
<th>Lateral View</th>
<th></th>
<th></th>
<th>Hilum View</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Symmetric</td>
<td>Asymmetric</td>
<td></td>
<td>Symmetric</td>
<td>Asymmetric</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heritability</td>
<td>Contribution</td>
<td>Heritability</td>
<td>Contribution</td>
<td>Heritability</td>
<td>Contribution</td>
</tr>
<tr>
<td>PC1</td>
<td>0.89</td>
<td>0.91</td>
<td>0.58</td>
<td>0.73</td>
<td>0.98</td>
<td>0.90</td>
</tr>
<tr>
<td>PC2</td>
<td>0.84</td>
<td>0.05</td>
<td>0.59</td>
<td>0.11</td>
<td>0.98</td>
<td>0.04</td>
</tr>
<tr>
<td>PC3</td>
<td>0.88</td>
<td>0.1</td>
<td>0.49</td>
<td>0.06</td>
<td>0.43</td>
<td>0.13</td>
</tr>
<tr>
<td>Total</td>
<td>0.98</td>
<td>0.90</td>
<td>0.99</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As we mentioned above, most of the variance of soybean seed shape is explained by the first principal components of Elliptic Fourie coefficients both L direction and H direction (Tab.2), and it was revealed that the trait was controlled genetically through the ANOVA conducted for 165 soybean cultivars. And the heritability derived clearly proved this. For example, because the contribution of PC1 in symmetrical variation for hilum view and its heritability are 0.98 and 0.89 respectively, we can understand that 87% (0.98 x 0.89) of the total seed shape variation in hilum view can be explained genetically by PC1. In the same manner, 81% of the total shape variation in lateral view is explained genetically by PC1. This indicates that aspect ratios of soybean seed shape in both directions are genetically controlled extensively, considering that the effects of the PC1s on shape for both directions correspond to aspect ration. It was also indicated that such as small variations in PC2 and PC3 that have not been visually identifiable before, are also genetically controlled.

As a conclusion, we revealed that there was diverse variation among soybean seed shape and it is strongly controlled genetically. Although there were some former studies, we first succeeded in quantitatively evaluating soybean seed shape and quantify genetic control on the shape. This study has opened a way to breed soybean seed shape effectively not only about the very clear variation in aspect ratio represented by PC1 but also such as roundness that was not detected before. For a further study, a combination of the result with molecular markers and QTL analysis is expected to identify genes to control the shape.
References


Cober ER, HD. Voldeng and J A. Fregeau-Reid (1997) Heritability of seed shape and seed size in soybean. Crop Sci 37:1767-1769


