Quantitative Evaluation of Leaflet Shape in *Fragaria vesca* and *F. viridis* Using Image Analysis and Elliptic Fourier Descriptors

JUOZAS LABOKAS
Nature Research Centre, Institute of Botany, Žaliųjų Ežerų str. 49, LT-08406 Vilnius, Lithuania; email juozas.labokas@botanika.lt

Labokas, J. 2013. Quantitative Evaluation of Leaflet Shape in *Fragaria vesca* and *F. viridis* Using Image Analysis and Elliptic Fourier Descriptors. *Baltic Forestry* 19(1): 22–30.

Abstract

The aim of the study was to emphasize the taxonomic utility of the central leaflet shape variation in *Fragaria vesca* and *F. viridis* considering the factors that may influence it. The image analysis and principal component analysis of elliptic Fourier descriptors of leaflet contours as well as leaflet area, perimeter and shape index, all derived from image outline, were employed. Simultaneously, the identification of diversity patterns of the species was pursued. 100 leaflets from each of six populations of each species were sampled across Lithuania. For *F. viridis* a significantly larger leaflet area and more circular leaflet shape, as estimated by shape index, are typical. The analysis of variance of principal component (PC) scores showed that for both species symmetrical shape variation is much more typical than asymmetrical. If estimated by the first five PC, the symmetrical variation accounts for about 60% of total variation in *F. vesca* and 67% in *F. viridis*, while the asymmetrical variation is up to 22% and 18%, respectively. Hierarchical cluster analysis revealed that PC7 and PC8 are good indicators of species-wise grouping of populations. Cluster analysis of shape index provided informative clustering results of populations within and between species.

Key words: *Fragaria vesca*, *Fragaria viridis*, population, shape index, leaflet shape, elliptic Fourier descriptors, image analysis

Introduction

Two strawberry species, the woodland strawberry (*Fragaria vesca* L.) and the green strawberry (*F. viridis* Weston) are the most distributed species of the genus *Fragaria* L. (*Rosaceae* L.) in the spontaneous floras of much of Europe and Northern Asia. In Europe at least, the distributions of *F. vesca* and *F. viridis* appear quite similar to that of musk strawberry, *F. moschata*, although the latter does not range so far to the east (Staudt 1989) and in the Baltic countries occurs as neophyte in specific habitats, like old parks, former estates, etc. Both of the most distributed *Fragaria* species are light-demanding species and much similar to each other by habit. The main distinctive property of *F. vesca* is its bright red fruit (pseudocarp) with a strong distinct aroma and reflexed sepals. The fruit of *F. viridis* is greenish white with pink top and an apple-like aroma, clasped by sepals from which it poorly separates. Flowering periods for both species coincide in May through June as well as fruiting does in June through July. The species are important for wild berry pickers as they produce ripe fruit first in the season as well as attract much by the strong aroma, particularly, *F. vesca*. There are references indicating that both species were cultivated in home gardens several centuries ago (Natkevičaitė-Ivanauskienė 1971) until they were replaced by the large-fruited American strawberries. Nevertheless, *F. vesca* and *F. viridis* remain unique wild relatives of the cultivated strawberries in Europe. Our previous study showed that in the field collection the two species could be distinguished by the length of inflorescence, which in *F. viridis* is significantly shorter than that in *F. vesca* and, in contrast to the latter, often hides under the leaves (Labokas and Bagdonaite 2005). However, in the wild the between-species distinguishing is not always easy, particularly at no flowering and no fruiting time and when habitats of both species overlap, let alone the within-species diversity. Therefore, the main aim of the current study was to test the taxonomic utility of leaflet shape variation considering the factors that may influence it. Simultaneously, the identification of diversity patterns within both species was pursued. Although the quantitative morphological features, such as leaflet area and perimeter were used for the characterisation of the populations, emphasis was placed on the analysis of shape, which is less dependent on environmental effects and therefore could be used as a morphological marker to recognize different genotypes within the same species (Furuta et al. 1995, Hiraoka and Kuramoto 2004). One of the most appropriate methods for the accurate shape evaluation is based on Fourier expansions and called elliptic Fourier method. The elliptic Fourier method is particularly suitable for evaluating closed contour shapes (Kuhl and Giardina 1982, Rohlf and Archie 1984, White et al. 1988). This method has been applied in studies of different organs of various plant species. For example, birch leaf (White et al. 1988), soybean leaflet (Furuta et al. 1995), soybean pod (Truong et al. 2005), buckwheat kernel (Ohsawa et al. 1998), Japanese radish root (Iwata et al., 1998, 2000), citrus leaf (Iwata et al. 2002a, 2002b), corolla and petal of lisianthus (Kawabata et al. 2009) and petal of primula (Yoshioka et al. 2004; 2005). My recent study on *Corylus avellana* (Labokas 2009) also proved good results of the application of the elliptic Fourier method. In addition, as observed by Jensen et al. (2002) the outline analysis provides a tool for taxonomic analysis.

Material and methods

Six populations of each species, *Fragaria vesca* and *F. viridis* were selected for leaf sampling across Lithuania and documented in the Institute's of Botany database of useful plants (Table 1). In order to ensure right classification of the species and avoid inclusion of possible interspecific hybrids of *Fragaria×biferta*, only those populations were selected which had been carefully pre-observed during flowering/fruiting time, when the between-species segregation is the most feasible by inflorescence, flower and fruit properties as well as position of sepals on fruit as described by Staudt et al. (2003). Samples of leaves, 100 per population, were collected in June through September 2009, from fully developed plants. The sampling sites were exposed to the similar light conditions, not shaded or only slightly shaded by trees. The geographical coordinates (WGS-84) were recorded for each population with the Garmin eTrex Vista GPS receiver (Table 1).

From the collected trifoliate leaves, the central leaflets were plucked off and dried in a herbarium press at an ambient room temperature. The air-dried leaflets were scanned with Microtek ScanMaker 3600 scanner on a white background to obtain full colour (24-bit) bitmap images. All leaflets were placed on the scanner glass abaxial (underside) side up.

The images were processed with the SHAPE ver. 1.3, a software package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors (Iwata and Ukai 2002). In brief, closed contours were obtained from the images and were chain-coded (Freeman 1974). The coefficients of the elliptic Fourier de-

Table 1. Sampling data of *Fragaria vesca* and *F. viridis* leaves

Population No.	Population designation	Habitat characteristics	Latitude N Longitude E	Nearest[1] population (km)	Date of sampling
F. vesca					
1	Barčiai	Lakeside meadow, sandy loam	54°16'21" 23°40'41"	Braziūkai 73	20.VI.2009
2	Braziūkai	Pinewood glade, sand	54°54'51" 23°26'07"	Barčiai 73	16.IX.2009
3	Juodkrantė	Seaside pinewood glade, sand	55°32'42" 21°06'40"	Braziūkai 163	14.VII.2009
4	Palūšė	Lakeside meadow, sandy loam	55°19'43" 26°06'22"	Verkiai 82	23.VI.2009
5	Rokantiškės	Birchwood glade, sandy loam	54°41'39" 25°20'39"	Verkiai 7	17.VI.2009
6	Verkiai	Mixed forest glade, sandy loam	54°45'06" 25°18'31"	Rokantiškės 7	15.VI.2009
F. viridis					
1	Aukštadvaris	Hillside meadow, sandy loam	54°35'19" 24°31'32"	Merkinė 53	11.VII.2009
2	Betygala	Riverside slope meadow, sandy loam	55°21'07" 23°22'13"	Veliuona 31	12.IX.2009
3	Merkinė	Riverside floodplain meadow, sand	54°08'42" 24°12'22"	Aukštadvaris 53	03.IX.2009
4	Rambynas	Riverside floodplain meadow, sand	55°03'43" 22°02'41"	Veliuona 80	09.VI.2009
5	Veliuona	Riverside slope meadow, sandy loam	55°04'30" 23°17'56"	Ž. Panemunė 13	16.IX.2009
6	Ž. Panemunė	Riverside slope meadow, sandy loam	55°02'34" 23°29'26"	Veliuona 13	16.IX.2009

[1] Distance calculation based on http://www.csgnetwork.com/longlatdistance.html

scriptors (EFDs) were calculated from the chain-code data and normalized to be invariant with respect to the size, rotation and starting point of tracing leaflet contour with the procedure based on the ellipse of the first harmonic, the method presented by Kuhl and Giardina (1982). The normalized coefficients of the EFDs (20 harmonics were used) were subjected to the principal component analysis in order to obtain summarized information of the variations contained in the coefficients. The 77 coefficients calculated were classified into two groups related to symmetrical (*a* and *d*, total 39 coefficients) and asymmetrical (*b* and *c*, total 38 coefficients) variations (Iwata et al. 1998) and used for the subsequent analyses. The variation of leaflet shape as represented by a particular principal component was illustrated by reconstructing its contour using the coefficients of the EFDs calculated for mean ± 2 S. D. (standard deviation) of a particular principal component scores, letting the rest of principal component scores be equal to zero, a method called inverse Fourier transformation.

Leaflet area in square millimetres, or half-surface area of a leaflet, was estimated by multiplying the total number of pixels per leaflet image to the pixel size in square millimetres provided by the ChainCoder program of the SHAPE package. Leaflet perimeter was calculated by the following procedures. First calculated length of a chain element (CE) by the formula: $CE = 1 + 0.5\ (\sqrt{2} - 1)\ (1 - (-1)^{c})$, where *c* is the code of the chain element (from 0 to 8). Then calculated perimeter in pixels by summing up all chain element values of the contour, and finally calculated perimeter in millimetres as a product of perimeter in pixels multiplied by a square root of a pixel size in square millimetres.

Shape index (SI) of a leaflet was calculated using the formula proposed by Keefe and Draper (1986): $SI = 4\pi\ (\text{area})\ /\ (\text{perimeter})^{2}$. The shape index calculated in this way indicates how close the shape of a given contour is to that of a circle, the shape index of which is equal to 1.

For the calculations of leaflet perimeter the MS Excel 2007, supporting large number of columns (2830 needed in this study), was used. For the statistical procedures, the analysis of variance, correlation and cluster analyses were employed with the statistical package PASW Statistics. Cophenetic correlation coefficients were computed with the MultiDendrograms 2.1.0 software (Fernández and Gómez 2008).

Results and discussion

First, an attempt was made to analyse the area and perimeter of the central leaflet of *Fragaria vesca* and *F. viridis*. The obtained results show that these quantitative features vary within species significantly as indicated by high *F* ratios at 1% significance level (Table 2). All the studied populations of *F. vesca* fall into four groups (homogeneous subsets) (a, b, c and d) and *F. viridis* – into three groups (a, b and c) by the leaflet area at a 5 % significance level. The analysis of leaflet variation in perimeter in *F. vesca* also produced four groups, while in *F. viridis* – five groups of populations. The intervals between maximum and minimum values indicate that *F. vesca* exhibits a wider morphological variation than *F. viridis* in both leaflet area (830.85 - 443.10 = 387.75 mm^2 versus 960.07 - 635.59 = 324.48 mm^2) and leaflet perimeter (172.33 -114.06 = 58.27 mm versus 168.75 - 127.53 = 41.22 mm). The coefficients of

Table 2. Variations of leaflet area and perimeter and correlation between them in *Fragaria vesca* and *F. viridis*

Population No.	*F. vesca* leaflet Area (mm^2)	leaflet perimeter (mm)	correlation coefficient *r*	*F. viridis* leaflet area (mm^2)	leaflet perimeter (mm)	correlation coefficient *r*
1	673.67 bc[1]	150.94 b	0.921	844.13 b	163.84 de	0.944
2	443.10 a	114.06 a	0.911	960.07 c	168.75 e	0.967
3	611.80 b	157.42 bc	0.952	659.53 a	140.54 b	0.914
4	752.62 cd	166.38 cd	0.946	833.31 b	157.79 cd	0.962
5	722.87 c	164.38 cd	0.930	635.59 a	127.53 a	0.948
6	830.85 d	172.33 d	0.920	792.90 b	149.21 bc	0.957
Pooled	672.49	154.25	0.926	787.59	151.28	0.944
SD	248.65	33.96		257.03	27.96	
Min-Max	443.10-830.85	114.06-172.33	0.911-0.952	635.59-960.07	127.53-168.75	0.914-0.967
CV (%)	37.0	22.0		32.6	18.5	
F ratio	38.33**	56.02**	3591.93**	27.58**	40.43**	4882.95**
DF	599	599	599	599	599	599

[1] Values within columns followed by different letters differ significantly (Tukey HSD test, <0.05); ** – significant at the 1 % level; *SD* – standard deviation; *CV* – coefficient of variation; *DF* – degrees of freedom

variation also confirm this (37.0 versus 32.6% and 22.0 versus 18.5%, correspondingly). This pattern of variation is caused mainly by the population No. 2 of *F. vesca*, which produced the absolutely smallest leaves, probably due to the environmental effects – poorest sand soil and late season collecting time. The coefficients of correlation (r) between leaflet area and perimeter indicate strong linear relationships in all populations. In *F. vesca* r varies from 0.911 to 0.952 (pooled r = 0.926, F = 3591.9) and in *F. viridis* – from 0.914 to 0.967 (pooled r = 0.944, F = 4882.9).

However, as the area and perimeter are quantitative morphological features, they are subject to change under the influence of environmental conditions. This is noted by many researchers with different species and was indirectly proved by our previous study on *Fragaria* species (Labokas and Bagdonaite 2005). Therefore, relative values were calculated to obtain a qualitative feature, shape index of leaflet (Table 3). Although there are suggestions reported (Iwata et al. 2002) that in case of highly correlated values of area and perimeter, shape index becomes synonymous with the reciprocal of the perimeter or area and cannot give additional information about shape characteristics, an attempt was made to check if this is true on *Fragaria* species. Thus, analysis of variance indicates that shape index varies in both species significantly with the F ratios of 78.67 in *F. vesca* and 63.46 in *F. viridis*, significant at 1% level, creating four homogeneous subsets in each species. These subsets do not coincide in most cases with those of the area and perimeter. However, the variation of shape index if judged by its Max/Min ratio (1.374 in *F. vesca* and 1.251 in *F. viridis*) (Table 3) is lower than that of area (1.875 in *F. vesca* and 1.511 in *F. viridis*) and perimeter (1.511 in *F. vesca* and 1.323 in *F. viridis*) (Table 2). Also it is interesting to note, that the smallest leaflets (both by area and perimeter) are relatively most circular, shape indices of which are 0.426 in *F. vesca* population No. 2 (Braziūkai) and 0.488 in *F. viridis* population No. 5 (Veliuona). In general, *F. viridis* produces more circular shape of leaflets than *F. vesca* as indicated by pooled shape indices 0.428 and 0.354, respectively. The only exception is *F. vesca* population No. 2, which by shape index is much closer to *F. viridis*.

Correlation analyses of leaflet area, perimeter and shape index with latitude and longitude of sampling sites of leaves were performed in order to test the influence of a geographic factor. As the climate in Lithuania gets more continental going eastwards from the Baltic Sea, i.e. deeper into the continent, it may influence the size and shape of leaflets. However, the correlations computed revealed weak or no relationships between the morphological properties of strawberries and sampling site coordinates. The only medium correlation (r = 0.302) was observed between the area of leaflet and the longitude of sampling site in *F. vesca* (Table 4) which could be attributed to the continentality of the climate. However, more observations are needed to prove this.

Table 3. Variation of leaflet shape index in *Fragaria vesca* and *F. viridis*

Population	*F. vesca*		*F. viridis*	
No.	leaflet shape index[2]	interpretation of shape index	leaflet shape index	interpretation of shape index
1	0.368 c[1]		0.390 a	least circular leaflets
2	0.426 d	most circular leaflets	0.421 bc	
3	0.310 a	least circular leaflets	0.420 b	
4	0.336 b		0.416 b	
5	0.336 b		0.488 d	most circular leaflets
6	0.349 b		0.436 c	
Pooled	0.354		0.428	
Min-Max	0.310-0.426		0.390-0.488	
Max/Min	1.374		1.251	
F ratio	78.67**		63.46**	
DF	599		599	

[1] Values within columns followed by different letters differ significantly (Tukey HSD test, <0.05);
** – significant at the 1 % level; *DF* – degrees of freedom.
[2] shape index = $4\pi(\text{area})/(\text{perimeter})^2$

The shape index is a kind of one-dimensional descriptor which does not allow to judge about a particular shape pattern of the object. In other words, objects with the same shape index values could be of different shape. On the other hand, the elliptic Fourier method allows a much deeper insight in variation of this qualitative morphological feature. In order to better elucidate shape variation, Fourier coefficients were analysed separately for symmetrical and asymmetrical variations, as recommended primarily for the objects, having no clear top and bottom sides (Iwata et al. 1998). After the principal component analysis of the coefficients, for the initial evaluation, ten principal components (PC1 through PC10) were taken. They represent 93.61 % of symmetrical and

Table 4. Correlations between leaflet area, perimeter and shape index and geographical coordinates of sampling sites of *Fragaria vesca* and *F. viridis* leaves

Leaflet	*F. vesca*		*F. viridis*	
descriptor	Latitude N	Longitude E	Latitude N	Longitude E
Area	-0.068	0.302**	0.227**	-0.075
Perimeter	0.047	0.229**	0.131**	-0.012
Shape index	-0.272**	0.025	0.204**	-0.162**

** – significant at the 1 % level

88.90 % of asymmetrical leaflet shape variation in *F. vesca* as well as 95.18 % and 88.38 % in *F. viridis*, respectively (Table 5). In *F. vesca,* the symmetrical variation revealed by the first ten principal components makes up 66.30 % of total variation, while that in *F. viridis* amounts to 72.34 %. The variation percentage revealed by the asymmetrical dataset is several times lower, 25.94 and 21.20 %, respectively. In general, this shows that the symmetrical variation of leaflet shape is much more typical of the studied *Fragaria* species, particularly, for *F. viridis*. In symmetrical dataset of each species, more than ¾ of variation is covered by the first three PCs, while in asymmetrical dataset five PCs are employed in covering of approximately the same degree of variation. For this reason, and in order to simplify the data processing, it was concentrated with further analyses into the highest level PCs, taking into account the effective PCs (those which contribute >1/39 and >1/38 in group A and group B, respectively) (Table 6).

It must be noted that no PC should be considered an important marker of morphological variation based on its contribution percentage only. The analysis of variance shows that some PCs with lower contribution percentage differ significantly between populations and are more reliable morphological markers because of their high *F* ratio (Table 6). Therefore, the PC6 with the contribution of 2.79 % (Table 5) is more important than PC5 with 4.36 % or even PC4 with 4.51 % contribution in symmetrical variation and PC3 with 11.74 % is more important than PC2 with 12.67 % contribution in asymmetrical variation of *F. vesca* leaflet shape. Similarly, in *F. viridis* the PC7 takes over the PC6 and the PC4 over the PC3 in symmetrical and asymmetrical variations, respectively.

The reconstructed contours of leaflet shape variation based on symmetrical and asymmetrical datasets can be easier explained visually, particularly those approximated by a higher level PCs (Figure 1 and Figure 2).

Table 5. Eigenvalues and contributions of the first ten principal components of elliptic Fourier descriptors estimated for symmetrical (Group A) and asymmetrical (Group B) variation of leaflet shape in *F. vesca* and *F. viridis*

Principal component	*F. vesca*			*F. viridis*		
	Eigenvalue ($\times 10^{-6}$)	Proportion (%) of group	Proportion (%) of total	Eigenvalue ($\times 10^{-6}$)	Proportion (%) of group	Proportion (%) of total
			Group A			
PC1	3810.4	49.46	35.03	2871.3	46.44	35.30
PC2	1426.9	18.52	13.12	1418.8	22.95	17.44
PC3	565.3	7.34	5.20	805.2	13.02	9.90
PC4	347.7	4.51	3.20	230.2	3.72	2.83
PC5	336.2	4.36	3.09	130.0	2.10	1.60
PC6	214.7	2.79	1.97	122.8	1.99	1.51
PC7	158.8	2.06	1.46	99.5	1.61	1.22
PC8	144.9	1.88	1.33	82.0	1.33	1.01
PC9	116.5	1.51	1.07	74.0	1.20	0.91
PC10	90.7	1.18	0.83	50.8	0.82	0.62
Total		93.61	66.30		95.18	72.34
Total variance[1]	7704.6			6182.9		
			Group B			
PC1	1058.4	33.36	9.73	762.2	39.06	9.37
PC2	402.0	12.67	3.70	336.4	17.24	4.14
PC3	372.4	11.74	3.42	144.2	7.39	1.77
PC4	346.7	10.93	3.19	136.5	7.00	1.68
PC5	178.2	5.62	1.64	113.0	5.79	1.39
PC6	145.3	4.58	1.34	73.1	3.75	0.90
PC7	131.0	4.13	1.20	55.5	2.84	0.68
PC8	76.6	2.41	0.70	38.6	1.98	0.47
PC9	61.7	1.95	0.58	34.2	1.76	0.42
PC10	48.0	1.51	0.44	30.6	1.57	0.38
Total		88.90	25.94		88.38	21.20
Total variance[1]	3172.2			1951.1		

[1] Total variance is the sum of eigenvalues of all 39 and all 38 principal components in group A and group B, respectively

Table 6. *F* ratios for the source of variation "Population" in the analysis of variance of each principal component for the symmetrical (group A) and asymmetrical (group B) variation

Principal	*F. vesca*		*F. viridis*	
component	Group A	Group B	Group A	Group B
PC1	27.642**	10.886**	53.784**	12.911**
PC2	38.687**	1.028	148.657**	8.345**
PC3	12.282**	15.257**	116.733**	1.134
PC4	0.501	6.509**	22.502**	7.022**
PC5	1.714	2.516*	15.812**	5.123**
PC6	12.193**	8,897**	0.922	3.642*
PC7	0.944	12,790**	4.501**	4,225**
PC8	3.815*	2.574*	43.419**	8.925**
PC9	1.425	3.037*	28.100**	3.341*
PC10	3.087*	12.582**	11.322**	2.939*
Effective[1]	PC1-PC6	PC1-PC7	PC1-PC4	PC1-PC7
DF	599	599	599	599

[1] Effective PCs are those, which contribute >1/39 in group A and >1/38 in group B.
** – significant at the 1 % level; * – significant at the 5% level; *DF* – degrees of freedom

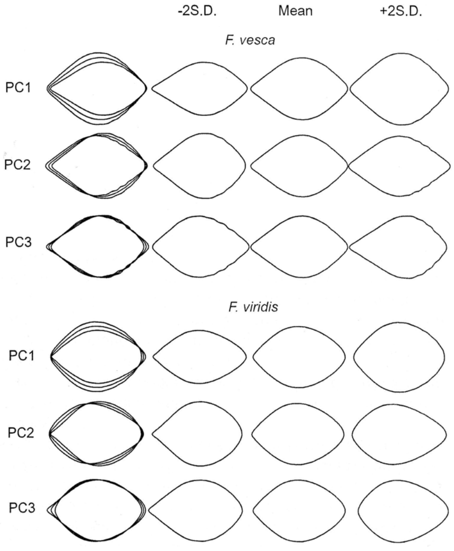

Figure 1. Reconstructed leaflet contours for the visualization of symmetrical shape variation using elliptic Fourier descriptors estimated under three principal component scores (PC1, PC2 and PC3) in *Fragaria vesca* and *F. viridis*. Each column shows that the score takes either mean + 2 S.D., mean or mean – 2 S.D. The leftmost column shows the overlaid drawings of the three cases

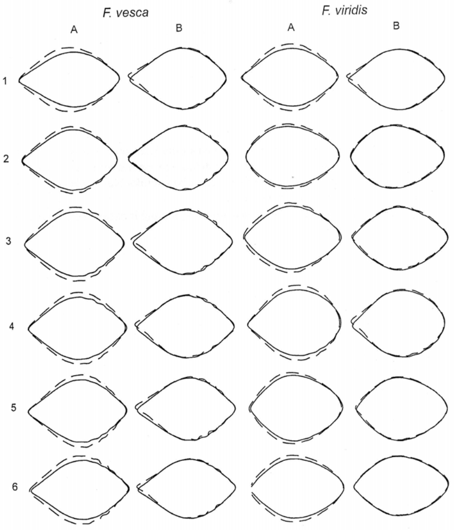

Figure 2. Reconstructed leaflet contours for the visualization of symmetrical (column A) and asymmetrical (column B) shape variation using elliptic Fourier descriptors estimated under the first principal component scores (PC1) in each of the six populations (1-6) of each *Fragaria vesca* and *F. viridis*. Solid, dashed and dotted lines represent mean, mean + 2 S.D. and mean - 2 S.D., respectively. For population numbers see Table 1

The multiple comparisons performed as the post hoc tests of ANOVA indicated that considering the symmetrical variation, the six populations of *F. vesca* could be subdivided into five homogeneous subsets (a = 0.05) by the PC1 including a subset with the only population No. 3 from Juodkrantė, the Curonian Spit. This is the westernmost and remotest population if compared to the rest of the populations studied (see Table 1). The comparisons by the PC2 produced three homogeneous subsets with the population No. 3 distinguishing as a separate subset as well. This means that the population No. 3 singularly takes the marginal positions if analyzed by the PC1 and PC2, which stand for the length to width ratio and bluntness of the distal part of the leaflet, respectively (Figure 1). The PC3, which represents the centroid of the leaflet along its midrib, and PC6 also produced three subsets each. The PC6 is the last effective PC in the symmetrical variation of *F. vesca* contributing to it by 2.79%, i.e., still more than 1/39.

In asymmetrical variation the PC1, PC3 and PC4 are the most important and, particularly, the latter as it generates four homogeneous subsets, although representing less than 11% of the whole asymmetrical variation (Table 5). The rest of the components produce two to three subsets each, except for the PC2, which shows no significant differences between the populations of *F. vesca*, although its contribution is second highest and amounts to 12.67% (Table 5).

The multiple comparisons of the six populations of *F. viridis* show significantly different pattern of shape variations. In symmetrical variation, there are five principal components, the PC1, PC2, PC3, PC8 and PC9, which produce homogeneous subsets with one population each. Just to mention that marginal positions singularly are taken by the following populations: No. 2 by PC2, No. 5 by PC3, No. 1 by PC8 and No. 3 by PC9. The subsets produced also indicate that these four populations could be distinguished by the first three principal components.

The analysis of the asymmetrical variation of *F. viridis* shows that another population, No. 4 distinguishes best of all the rest. Moreover, this can be done by the PC1 alone, representing more than 35% of the total asymmetrical variation of the species.

To better illustrate and summarize the above stated a hierarchical cluster analysis was carried out of all PCs based on squared Euclidean distances between populations, using Between-groups linkage method (Figure 3). This shows clearly that *F. vesca* population No. 3 from Juodkrantė is the most distinct not only within the species but also among the all 12 populations of *Fragaria* studied. The next most distinct is *F. viridis* population No. 4 from Rambynas. The cophenetic correlation coefficient for the dendrogram is 0.866 (with normalized mean squared error 0.0126). If the cluster analysis of all principal component scores does not produce grouping of the populations strictly by species, that of PC7 and PC8 provides a much clearer ordering of the populations (Figure 4). The cophenetic correlation coefficient for this dendrogram is 0.947 (with normalized mean squared error 0.0125). Although contributing to the total variation not significantly, i.e., by 6.33% in *F. vesca* and 3.51% in *F. viridis*, the PC7 and PC8 are one of those few principal components which allow species-wise segregation of *Fragaria* populations. The clustering of all populations by leaflet shape index provides nearly perfect within-species differentiation, except for the population No. 2 of *F. vesca* (Figure 5), which, as mentioned above, is much closer to those of *F. viridis*. The cophenetic correlation coefficient for the dendrogram is 0.633 (with normalized mean squared error 0.1951).

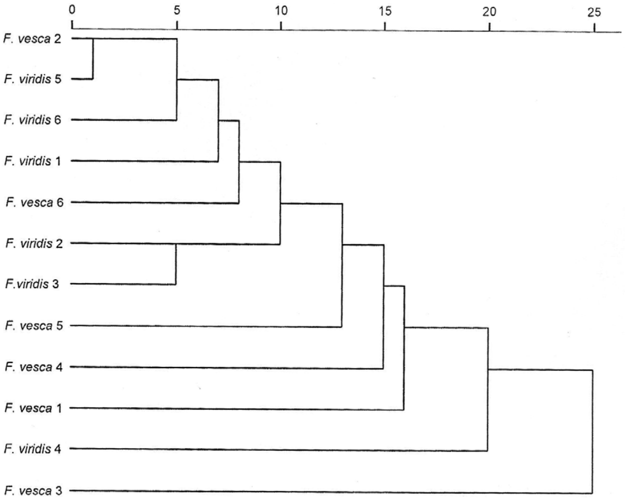

Figure 3. Dendrogram using average linkage between populations of *Fragaria vesca* and *F. viridis* by all principal components (PC1-PC77) of all elliptic Fourier coefficients (a, b, c, d) of leaflet shape. For population numbers, which follow species names, see Table 1

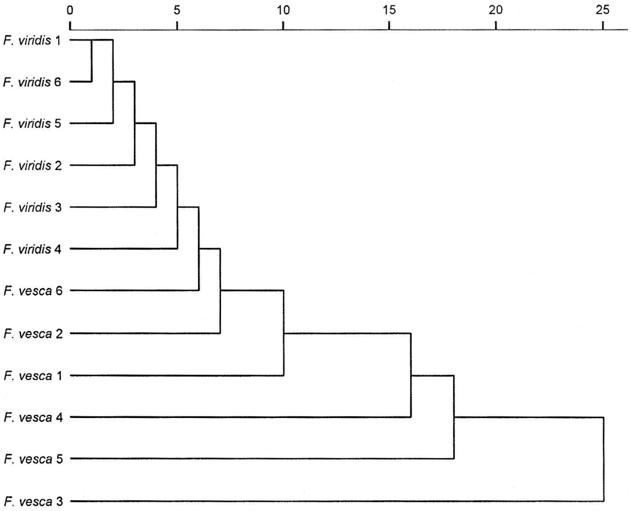

Figure 4. Dendrogram using average linkage between populations of *Fragaria vesca* and *F. viridis* by seventh and eighth principal components (PC7 and PC8) of all elliptic Fourier coefficients (a, b, c, d) of leaflet shape. For population numbers, which follow species names, see Table 1

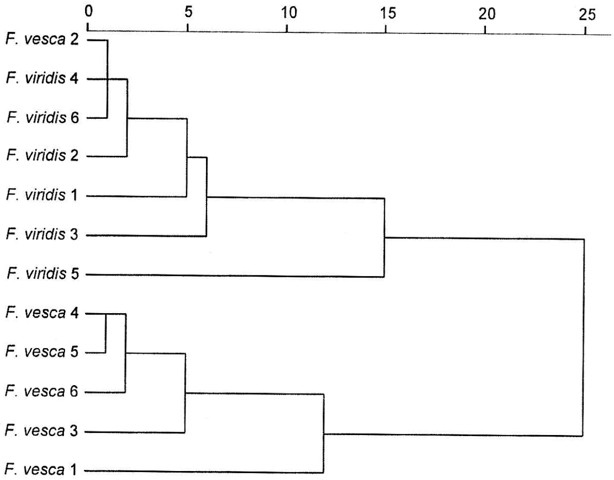

Figure 5. Dendrogram using average linkage between populations of *Fragaria vesca* and *F. viridis* by shape index of central leaflet. For population numbers, which follow species names, see Table 1

Finally, it is worth to observe that shape index values are quite closely related to the percentages of shape variation estimated by the first principal component, particularly, on account of asymmetrical variation. The average numbers respectively are 0.354 (= 35.4%) and 33.79% for *F. vesca* and 0.428 (= 42.8%) and 40.68 for *F. viridis* (Table 7).

Table 7. Comparison of leaflet shape index with shape variation percentages revealed by the first principal component (PC1) on account of symmetrical (A) and asymmetrical (B) variations in *F. vesca* and *F. viridis*

Population No.	*F. vesca*			*F.viridis*		
	leaflet shape index	shape variation (%), A	shape variation (%), B	leaflet shape index	shape variation (%), A	shape variation (%), B
1	0.368	55.58	37.20	0.390	63.34	43.53
2	0.426	44.05	24.45	0.421	39.48	44.18
3	0.310	43.39	35.81	0.420	47.55	39.87
4	0.336	50.53	30.56	0.416	50.31	46.85
5	0.336	49.41	39.53	0.488	49.37	39.52
6	0.349	56.29	35.17	0.436	55.02	30.15
Pooled	0.354	49.88	33.79	0.428	50.85	40.68

Conclusions

The current study revealed that both *Fragaria vesca* and *F. viridis* possess high within species and between species variation as estimated by shape and size characteristics of the central leaflet. However, the between species variation is higher. For the most of *F. viridis* populations the bigger leaflet area and more circular leaflet shape are typical as estimated by shape index. The analysis of variance of principal component scores derived from elliptic Fourier coefficients of leaflet shape is a good means for the detailed shape analysis. It showed that for the both species of *Fragaria* a symmetrical shape variation is much more typical than asymmetrical. If estimated by the first five principal components, the symmetrical variation accounts for about 60% of the total variation in *F. vesca* and 67% in *F. viridis*, while asymmetrical variation 22% and 18%, respectively. Here it is important to observe that the contribution of the principal components to the total variation should be estimated based not only on their percentage input but also with regard to the *F* ratio indicating whether the variation of a particular PC scores is consistent between populations. The hierarchical cluster analysis of PC scores revealed that PC7 and PC8 are good indicators of species-wise grouping of populations if pooled populations of both *F. vesca* and *F. viridis* are analysed. The cluster analysis also indicated that the population No. 3 of *F. vesca* (from Juodkrantė, Curonian Spit, western Lithuania) falls into a separate cluster by most of the principal components analysed. This is the most distinctive population not only within the species but also between them and presumably is an example of geographically isolated genotypic variation of the species. The study also showed that the direct observation characteristics, such as leaflet area and leaflet perimeter, should not be ignored, as the cluster analysis of shape index derived of leaflet area and perimeter provided good clustering results of populations between species.

Acknowledgements

This work was supported, in part, by the Plant Gene Bank of Lithuania. My thanks go to Dr. Edita Bagdonaite for her assistance in sampling plant material of Fragaria species.

References

Fernández, A. and Gómez, S. 2008. Solving non-uniqueness in agglomerative hierarchical clustering using multidendrograms. *Journal of Classification* 25: 43-65.

Freeman, H. 1974. Computer processing of line-drawing images. *Computing Surveys* 6(1): 57-97.

Furuta, N., Ninomiya, S., Takahashi, N., Ohmori, H. and Ukai, Y. 1995. Quantitative evaluation of soybean (*Glycine max* L. Merr.) leaflet shape by principal component scores based on elliptic Fourier descriptor. *Breeding Science* 45: 315-320.

Hiraoka, Y. and Kuramoto, N. 2004. Identification of *Rhus succedanea* L. cultivars using elliptic Fourier descriptors based on fruit shape. *Silvae Genetica* 53(5-6): 221-226.

Iwata, H., Nesumi, H., Ninomiya, S., Takano, Y. and Ukai, Y. 2002a. Diallel analysis of leaf shape variations of citrus varieties based on elliptic fourier descriptors. *Breeding Science* 52: 89-94.

Iwata, H., Nesumi, H., Ninomiya, S., Takano, Y. and Ukai, Y. 2002b. The evaluation of genotype × environment interactions of citrus leaf morphology using image analysis and elliptic Fourier descriptors. *Breeding Science* 52: 243-251.

Iwata, H., Niikura, S., Matsuura, S., Takano, Y. and Ukai, Y. 1998. Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors. *Euphytica* 102: 143-149.

Iwata, H., Niikura, S., Matsuura, S., Takano, Y. and Ukai, Y. 2000. Diallel analysis of root shape of Japanese radish (*Raphanus sativus* L.) based on elliptic Fourier descriptors. *Breeding Science* 50: 73-80.

Iwata, H. and Ukai, Y. 2002. SHAPE: A computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *Journal of Heredity* 93: 384-385.

Jensen, R. J., Ciofani, K. M. and Miramontes, L. C. 2002. Lines, outlines, and landmarks: morphometric analyses of leaves of *Acer rubrum*, *Acer saccharinum* (*Aceraceae*) and their hybrid. *Taxon* 51: 475-492.

Kawabata, S., Yokoo, M. and Nii, K. 2009. Quantitative analysis of corolla shapes and petal contours in single-flower cultivars of lisianthus. *Scientia Horticulturae* 121: 206-212.

Keefe, P. D. and Draper, S. R. 1986. The measurement of new characters for cultivar identification in wheat using machine vision. *Seed Science and Technology* 14: 715-724.

Kuhl, F. P. and Giardina, C. R. 1982. Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing* 18: 236-258.

Labokas J. 2009. Evaluation of leaf morphology in *Corylus avellana* using image analysis and elliptic Fourier descriptors. *Botanica Lithuanica* 15(4): 227-235.

Labokas, J. and Bagdonaite, E. 2005. Phenotypic diversity of *Fragaria vesca* and *F. viridis* in Lithuania. *Biologija* 3: 19-22.

Natkevičaitė-Ivanauskienė, M. 1971. *Fragaria*. In: M. Natkevičaitė-Ivanauskienė (Editor), Lietuvos TSR flora, vol. 4, Mokslas, Vilnius, p. 187-192 (In Lithuanian).

Ohsawa, R., Tsutsumi, T., Uehara, H., Namai, H. and Ninomiya, S. 1998. Quantitative evaluation of common buckwheat (*Fagopyrum esculentum* Moench) kernel shape by elliptic Fourier descriptor. *Euphytica* 101: 175-183.

Rohlf, F. J. and Archie, J. W. 1984. A comparison of Fourier methods for the description of wing shape in mosquitoes (*Diptera*: *Culicidae*). *Systematic Zoology* 33(3): 302-317.

Staudt, G. 1989. The species of Fragaria, their taxonomy and geographical distribution. *Acta Horticulturae* 265: 23-33.

Staudt, G., DiMeglio, L. M., Davis, T. M. and Gerstberger. P. 2003. Fragaria ×bifera Duch.: Origin and taxonomy. *Botanische Jahrbücher für Systematik, Pflanzengeschichte und Pflanzengeographie. Leipzig* 125: 53–72.

Truong, N. T., Gwag, J. G., Park, Y. J. and Lee, S. H. 2005. Genetic diversity of soybean pod shape based on elliptic Fourier descriptors. *Korean Journal of Crop Science* 50(1): 60-66.

White, R. J., Prentice, H. C. and Verwijst, T. 1988. Automated image acquisition and morphometric description. *Canadian Journal of Botany* 66: 450-459.

Yoshioka, Y., Iwata, H., Ohsawa, R. and Ninomiya, S. 2004. Analysis of petal shape variation of *Primula sieboldii* by elliptic Fourier descriptors and principal component analysis. *Annals of Botany* 94: 657-664.

Yoshioka, Y., Iwata, H., Ohsawa, R. and Ninomiya, S. 2005. Quantitative evaluation of the petal shape variation in *Primula sieboldii* caused by breeding process in the last 300 years. *Heredity* 94: 657-663.

Received 16 July 2012
Accepted 07 June 2013

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ФОРМЫ ЛИСТКА ЗЕМЛЯНИК *FRAGARIA VESCA* И *F. VIRIDIS* С ИСПОЛЬЗОВАНИЕМ АНАЛИЗА КОМПЬЮТЕРНОГО ИЗОБРАЖЕНИЯ И ЭЛЛИПТИЧЕСКИХ ДЕСКРИПТОРОВ ФУРЬЕ

Ю. Лабокас

Резюме

Цель данного исследования – определить таксономическую полезность варьирования формы центрального листа земляники *Fragaria vesca* и *F. viridis*, учитывая влияние разных факторов.

В работе применялись такие методы как анализ основных компонентов эллиптических дескрипторов Фурье контуров листьев, определение площади, периметра и индекса формы листа, полученных путем анализа компьютерных изображениий листьев. Для анализа брали по 100 листьев с каждой из 12 популяций (по 6 в каждом виде), распространенных по всей территории Литвы.

Результаты исследования показали, что для *F. viridis* характерна значительно большая площадь листа и более круглая форма его контура. Дисперсионный анализ чисел основных компонентов показал, что симметрическое варьирование формы листа в обоих видах значительно более характерно, чем асимметрическое. Судя по первым пяти основным компонентам (PC1–PC5), симметрическое варьирование формы в *F. vesca* составляет около 60%, а в *F. viridis* – около 67% от суммарного, тем временем как асимметрическое – соответственно 22 % и 18 %. Иерархический кластерный анализ чисел основных компонентов показал, что PC7 и PC8 являются хорошими идикаторами видоотличающей группировки популяций. Кластерный анализ индексов формы дал наиболее информативный результат как внутривидовой, так и межвидовой группировки популяций.

Ключевые слова: *Fragaria vesca*, *Fragaria viridis*, популяция, форма листка, индекс формы, эллиптические дескрипторы Фурье, анализ изображения.