Euphytica **102:** 143–149, 1998.

Evaluation of variation of root shape of Japanese radish (*Raphanus sativus* L.) based on image analysis using elliptic Fourier descriptors

Hiroyoshi Iwata[1], Satoshi Niikura[2], Seiji Matsuura[2], Yasushi Takano[1] & Yasuo Ukai[1]
[1] *Department of Agricultural - Environmental Biology, Graduate School of Agriculture and Life Sciences, University of Tokyo, Yayoi 1-1-1, Bunkyo, Tokyo 113, Japan;* [2] *Tohoku Seed Company, Nishihara 1625, Himuro, Utsunomiya 321-32, Japan*

Received 27 June 1997; accepted 16 January 1998

Key words: Japanese radish, root shape, elliptic Fourier descriptors, image analysis, principal component analysis

Summary

Variation was of root shape in Japanese radish, due to genotypes, soil types and growth stages, were quantitatively evaluated by principal components scores based on elliptic Fourier descriptors. Photographic images of sampled roots on 35mm color reversal films were converted into digital images. After image processing, the contour of each root was expressed as chain-code and then described by 77 coefficients of elliptic Fourier descriptors. After normalization about size, rotation, and starting point of the contour, two groups of the coefficients, which are related to the symmetrical and asymmetrical variations of shape, were analyzed separately, since artificially determined direction of curvature of the root may influence the results. Principal component analysis of the coefficients showed that the major part of the symmetrical (A) and asymmetrical (B) variations were summarized by at most 5 components. The cumulative contribution was 95.2% and 97.1%, respectively. Analysis of variance of each component indicated that the variety effect was highly significant for the 1st, 2nd and 3rd principal components derived from group A coefficients, which were related to the aspect ratio, bluntness of the distal part of the root, and swelling of the middle part, respectively. This suggests that these traits are heritable and can be effectively selected through quantified measures based on elliptic Fourier descriptors presented in this report. Direction and degree of curvature of root could be analyzed independently of the symmetrical variation.

Introduction

Shape of plant organs such as root, grain, leaf and fruit is an important target for crop plant improvement, since the shape of such organs is directly or indirectly related to the quantity and quality of agricultural products. In Japanese radish (*Raphanus sativus* L.) shape as well as size of root have long been genetically improved to meet different consumer demands (Nishiyama, 1958). Recent diversification of consumers taste to the vegetable and change in the transporting system made the desirable shape much more variable and delicate than before. Clearly, therefore an accurate quantitative description of root shape is required for improvement of this vegetable. The length, maximum thickness and diameter at the proximal part of the root are the main indices used in selecting lines with a desirable root shape. However, sole combination of these indices is insufficient for expressing subtle variations due to genotypes and/or cultural conditions. Moreover, measurement of these indices were taken manually and required much time and labor.

Several methods have been proposed for evaluating the shape of materials quantitatively, some of which have been successfully applied to the analysis of shape of living organisms. The shape description methods, such as moment invariants (Hu, 1962) and Fourier descriptors, possess one feature in common, namely that they employ information about the location of all the points on the contour, thereby making description of the entire shape possible. In particular, Fourier descriptors have been used effectively in the morphological analysis of shape of various organs in plants and animals. Several types of Fourier descrip-

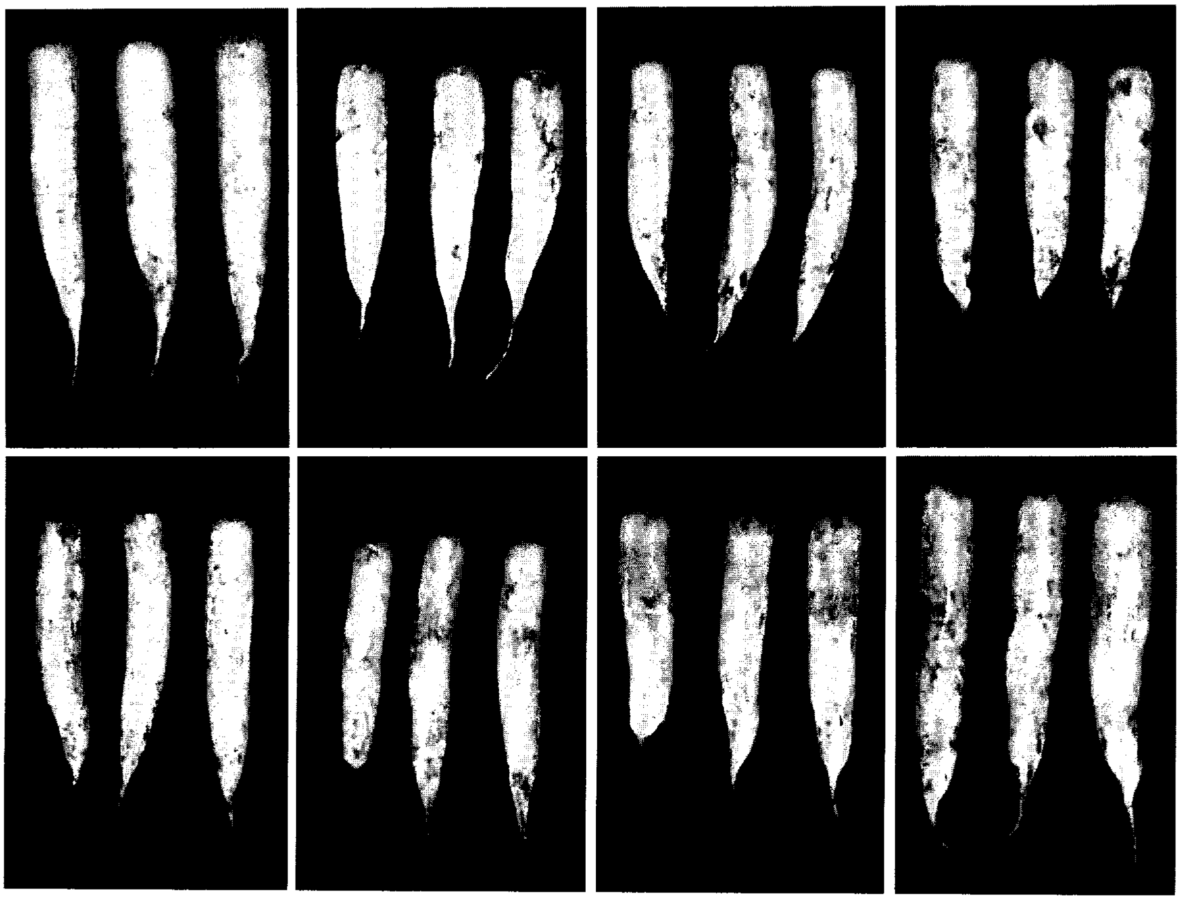

Figure 1. Root images of eight F_1 varieties of Japanese radish (*Raphanus sativus* L.) evaluated in this study.

tors have been proposed up to the present. Among these Fourier descriptors, elliptic Fourier descriptors described by Kuhl and Giardina (1982) have suitable properties for the morphological analysis of a shape contour. Rohlf and Archie (1984) made a comparative study of several Fourier descriptors, and concluded that elliptic Fourier descriptors were the best for quantitative expression of the mosquito wing shape. They also introduced the principal component analysis for summarizing the information contained in these descriptors. Following this study, the shape of various animal organs has been effectively evaluated by elliptic Fourier descriptors (Ferson et al., 1985; Bierbaum & Ferson, 1986; Diaz et al., 1989; Gubanyi, 1996). Application of elliptic Fourier descriptors to plant organ shape was first attempted by White et al. (1988), who analyzed *Betula* leaf shape using several shape description methods. They concluded that elliptic Fourier descriptors had performed best. Furuta et al. (1995) applied elliptic Fourier descriptors to leaf shape of soybean varieties and suggested that the principal component scores of the descriptors can be used as indices of genetic variation in shape. The same method has been employed for the analysis of leaf shape in a perennial tree species (Yamamoto et al., 1994; Hadipoentyanti et al., 1996).

The purpose of the present study is to determine whether elliptic Fourier descriptors are effective for the description of root shape in Japanese radish and to assess the contribution of genotypes to the variation of root shape as compared with other sources of variation, such as soil conditions and growth stage.

Materials and methods

Plant materials

Seed of eight F_1 varieties of Japanese radish (Figure 1) were sown on the 21th March 1996, and cultivated in three fields with different soil types i.e. andosoil,

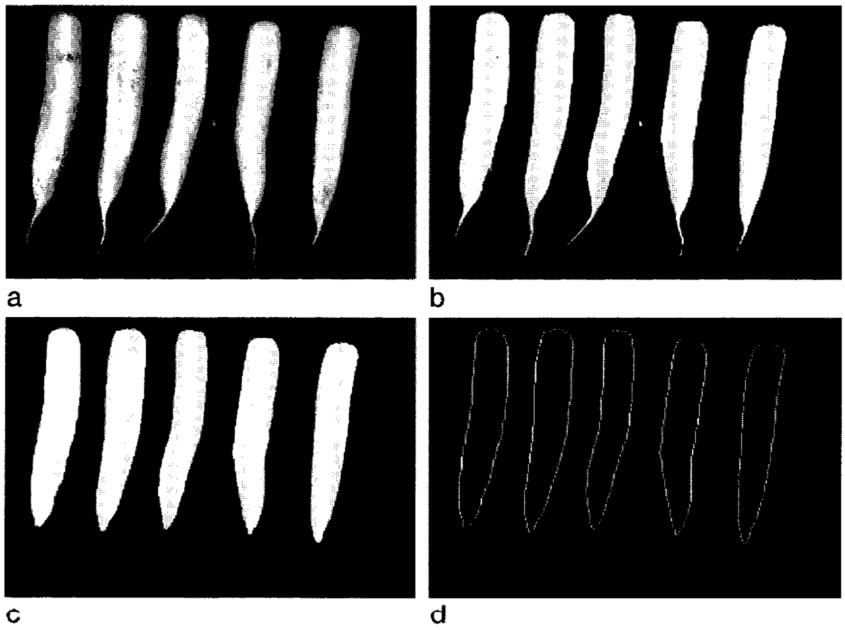

Figure 2. Series of image processing which extracts chain-coded contour from red digital image; a: red digital image of 256 gray levels; b: binary image; c: image after noise reduction, erosion, dilation, and filling holes; d: extracted contours.

clayey soil, and sandy soil. The varieties were arranged in a randomized complete block design with two replications. For each variety the roots were sampled at five different times (18th, 21th, 24th, 28th and 31th May). The fourth date was the usual harvest time of most varieties employed. On each occasion five roots were sampled from each plot.

Image processing and contour recording

After removing all leaves, we took photographs of roots on 35mm color reversal film. Images on the film were converted into analogue video images of NTSC standard composite video with a slide-video converter (PHV-AV, SONY, Tokyo), and then further converted into R (red), G (green) and B (blue) digital images with a real time A/D converter (FA-310, For A Corp, Tokyo). Each digital image has 512 × 512 spatial resolution and 8-bit quantization (256 gray levels). Since red images gave the clearest contrast of roots against the background, it was adopted in the following processing. A color image processor (HRU-TAICHI IV80, Ezel Sharp, Tokyo) controlled by a personal computer (PC9801VX, NEC, Tokyo) was used for processing. Red digital images were converted into binary images with one-bit quantization (Figures 2a, b). After noise reduction, erosion, dilation, and filling holes on the binary images, the closed contours of roots were obtained by edge detection (Figures 2c, d). The contours were described by a chain-code (Freeman, 1975), and then chain-code data were stored in a computer memory unit as the contour information of the root shape.

Elliptic Fourier descriptors

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). The outline of the procedure is as follows:

A contour of the digitized shapes can be represented as a sequence of the x- and y-coordinates of ordered points which are measured counter-clockwise from an arbitrary starting point. Assuming that the contour between two adjacent points is linearly interpolated and the length of the linear segment between $(i-1)$th and the ith points to be Δt_i, then the length of the contour from the starting point to the pth point is $t_p = \sum_{i=1}^{p} \Delta t_i$, and the perimeter of the contour is $T = t_K$, where K is the total number of the points on the contour. Notice that Kth point is equivalent to the starting point. The x-coordinate of the pth point is $x_p = \sum_{i=1}^{p} \Delta x_i$, where Δx_i is the displacement

Table 1. Eigenvalues and contribution of each principal component to the symmetrical and asymmetrical coefficients (group A and B), respectively

Component	Group A			Group B		
	Eigenvalue	Proportion(%)	Cumulative(%)	Eigenvalue	Proportion(%)	Cumulative(%)
1	67.065	73.92	73.92	108.112	78.93	78.93
2	12.865	14.18	88.10	14.090	10.29	89.22
3	3.579	3.94	92.05	5.621	4.10	93.32
4	1.721	1.90	93.94	3.791	2.77	96.09
5	1.180	1.30	95.24	1.436	1.05	97.14
Total variance		90.725			136.972	

along the x-axis of the contour between $(i-1)$th and the ith points. Then, elliptic Fourier expansion of the sequences of the x-coordinates is

$$x_p = x_{cen} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{2n\pi t_p}{T} + b_n \sin\frac{2n\pi t_p}{T}\right),$$

where

$$a_n = \frac{T}{2n^2\pi^2}\sum_{p=1}^{K}\frac{\Delta x_p}{\Delta t_p}\left(\cos\frac{2n\pi t_p}{T} - \cos\frac{2n\pi t_{p-1}}{T}\right),$$

and

$$b_n = \frac{T}{2n^2\pi^2}\sum_{p=1}^{K}\frac{\Delta x_p}{\Delta t_p}\left(\sin\frac{2n\pi t_p}{T} - \sin\frac{2n\pi t_{p-1}}{T}\right).$$

In the above equations, x_{cen} is the coordinate of the center point, and n is the harmonic order of the coefficients (a_n and b_n). The coefficients for the y-coordinates, c_n and d_n, are found in the same way. The coefficients of elliptic Fourier descriptors can be mathematically normalized to be invariant to size, rotation, and starting point of the contour trace. In the present study, the coefficients were normalized with the procedure based on the ellipse of the first harmonic. After this normalization, three coefficients are constant ($a_1 = 1, b_1 = 0, c_1 = 0$) and can be neglected.

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 root shape, and were used for the analysis.

Results

From a preliminary study of root shape of Japanese radish, it was suspected that the direction of curvature of roots on the image might distort the results of shape analysis. Since this is a mere artifact, it may be better to separate its contribution to the variation from those of other sources. So the following analytical procedures made separately for the a_n and d_n coefficients (which we will call group A hereafter) on one hand and the b_n and c_n coefficients (group B) on the other. It was thought that the former were related to symmetrical, and the latter to asymmetrical variations in shape. To summarize the information contained in the coefficients of each group (39 and 38 coefficients for group A and B, respectively), we performed a principal component analysis based on a variance-covariance matrix.

The first five principal components for groups A and B provide a good summary of the data, accounting for 95.2 and 97.1% of the total variance respectively (Table 1), while all other components contributed less than 1%.

We visualized the shape variation which can be accounted for by each principal component (Figure 3). The coefficients of the elliptic Fourier descriptors were calculated letting the score on a particular principal component be equal to the mean ± 2 (standard deviation), the scores on the remaining components being zero. Then the root shape on each condition can be reconstructed from the calculated coefficients by the inverse Fourier transformation. These reconstructed shapes indicated that the 1st principal component of group A is a good measure of the aspect ratio of the root, which accounts for 73.9% of the symmetrical variation of root (Table 1). The 2nd component was associated with the bluntness of the distal part of the root, and the 3rd component expressed the swelling of the middle part in relation to the distal and proximal part of the root. The fourth and fifth components were related to variation which cannot be ascribed to the first three components, although the meaning of

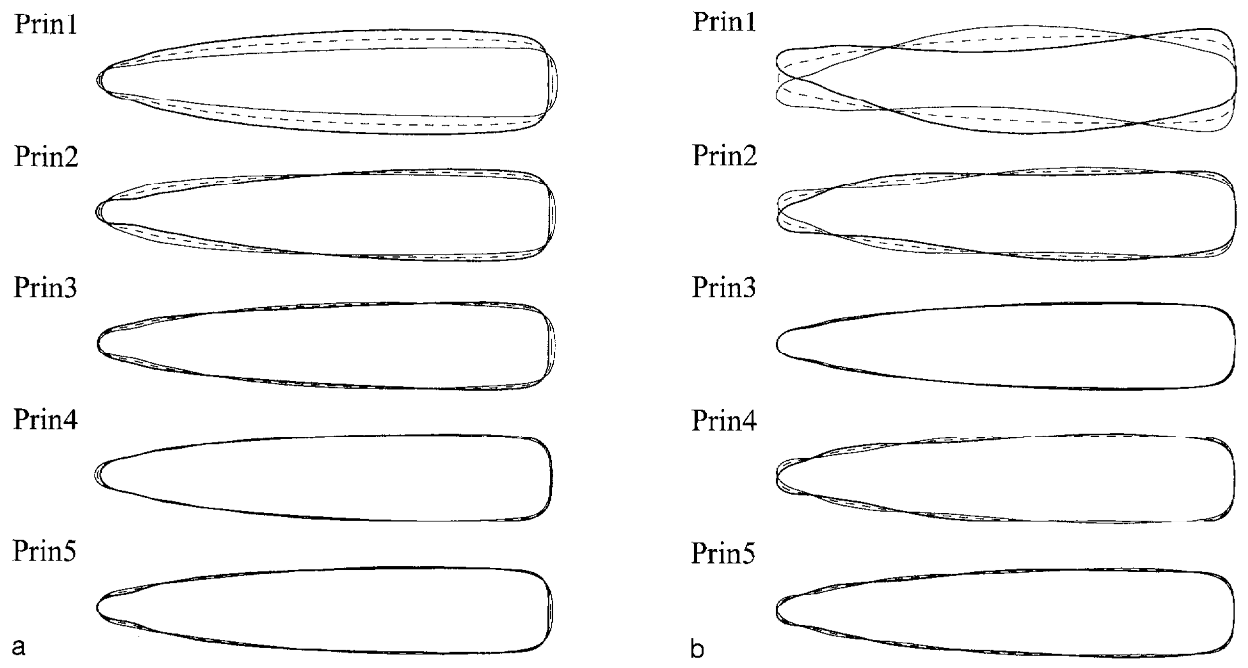

Figure 3. The shape variation which can be accounted for by each principal component. Each shape was reconstructed from the coefficients which were calculated by letting the score on the corresponding principal component be equal to mean with plus or minus two times standard deviation and the scores on the remaining component zero. (a) symmetrical variation from group A coefficients. (b) asymmetrical variation from group B coefficients. Dashed line, thick solid line, and thin solid line stand for mean, mean + 2S.D., mean - 2S.D.

this variation was difficult to explain. The 1st principal component for group B indicated curvature of the middle part and 2nd component curvature of the proximal and distal parts. The 3rd, 4th and 5th components were related to asymmetrical variation which could not be explained by the 1st and 2nd components.

The scores of the first five principal components were averaged over 5 roots within each plot. Error variances were obtained as the pooled interaction variances between variety (V), growth stage (G), and V × G interaction on one hand and blocks on the other. Heterogeneity of error variances between the three soil types was tested with a Bartlett's test. The variances differed significantly at the 5% level for all principal components tested. Clayey soil field showed a much higher variance than the others, which might be ascribed to damage by insect pests which was severe on some plants of this field. For this reason, we excluded the data from the clayey soil field from the following analysis.

To investigate the effects of variety, soil type, and growth stage on shape, the analysis of variance for the scores of each component was performed separately. The interactions between variety, soil type, growth stage on the one hand and block effect on the other were used as estimates of error. Error term was tested for normality with a Shapiro-Wilk W test (Shapiro and Wilk 1965) prior to the analysis of variance, showing no significance at the 5% level except for one principal component in groups A and B. From the analysis it is evident that the variety effect was highly significant for all principal components in group A (Table 2), but particularly for the 1st, 2nd and 3rd principal components. The effect of soil type was significant for the 1st, 2nd and 3rd components and growth stage for the 1st, 2nd, 4th and 5th components. The interactions were non-significant for all principal components except for S × G and V × S × G in the 2nd component. Values for the coefficient of determination (R^2) showed that the statistical model adopted here, which includes main effects of variety, soil type and growth stage and their interactions, fitted well, especially for the 1st and 2nd components.

Similar analyses were performed for group B. To remove the influence of the direction of curvature, all the scores of principal component analysis were converted to absolute values prior to analysis. The result

Table 2. F ratios for different sources of variation in the analysis of variance of each principal component for the symmetrical and asymmetrical coefficients (group A and B), respectively. Coefficients of determination (R^2) for the statistical estimator below are also shown.
$Y = \mu + V_i + S_j + G_k + (VS)_{ij} + (VG)_{ik} + (SG)_{jk} + (VSG)_{ijk}$.

Source	DF	Prin1	Prin2	Prin3	Prin4	Prin5
Group A						
Variety (V)	7	56.702 **	54.641 **	20.924 **	4.184 **	9.624 **
Soil type (S)	1	21.793 **	34.630 **	37.497 **	0.796	1.031
Growth stage (G)	4	11.219 **	211.423 **	1.212	3.623 **	10.816 **
V * S	7	1.605	1.797	0.908	1.388	1.026
V * G	28	0.908	1.468	1.267	0.869	1.019
S * G	4	1.193	3.026 *	1.731	0.568	4.623
V * S * G	28	1.130	1.758 *	0.424	0.885	1.402
R^2		0.871	0.945	0.757	0.569	0.719
Group B						
Variety (V)	7	7.928 **	3.500 **	1.569	2.387 *	2.925 **
Soil type (S)	1	4.965 *	4.024 *	0.001	2.552	0.029
Growth stage (G)	4	8.838 **	4.239 **	2.657 *	1.790	0.609
V * S	7	1.448	2.222 *	1.620	3.113 **	1.921
V * G	28	1.882 *	1.188	1.815 *	1.117	0.804
S * G	4	1.188	2.813 *	0.903	2.078	0.198
V * S * G	28	1.398	0.926	1.275	1.300	0.378
R^2		0.717	0.622	0.606	0.608	0.468

** - significant at the 1% level; * - significant at the 5% level.

indicated that the variety effect was significant for the 1st, 2nd, 4th and 5th components. Soil type was significant for the 1st and 2nd components, and growth stage for the 1st, 2nd and 3rd. It is to be noted that the relatively minor component (5th) was highly significant despite the insignificance of the 3rd component. The first order interactions (V × S, V × G, S × G) were significant for 5 out of 15 cases, while the second order interactions were non-significant throughout.

Discussion

Roots in general have few landmarks. Accurate description of root shape based on landmarks has therefore not been previously successful. Previously, in the selection of desirable types of root shape in Japanese radish, the length, thickness and diameter at the proximal part were the only measures utilized. When differences in root shape between varieties or strains used as parents in crossing were large, such a conventional method of measurement proved satisfaction. Nowadays, however, subtle variation in root shape has considerable influence on the market value of the vegetable. Hence a precise description of root shape is required for efficient selection of the desired type. Among several characteristics associated with root shape, straight form at the middle part, bluntness of the distal part and absence of curvature are now the most important targets for selection. The present report indicates that variation in root shape can be accurately described by a principal component analysis of elliptic Fourier descriptors obtained from the contour of the digitized image of the root.

Direction, not degree, of curvature is a mere artifact in that it depends on how the tip of the root is by chance oriented at the time of taking photograph. So it is better to distinguish shape variation related to curvature from other types of variation. It was found possible to distinguish them by dealing the coefficients of elliptic Fourier descriptors which are related to symmetrical and asymmetrical variations, separately. The principal components which were obtained from the latter group of coefficients expressed the direction and degree of curvature. Symmetrical type of variation such as aspect ratio and bluntedness at the distal part and swelling at the middle part could be evaluated without influence of the variation associated with curvature. If all coefficients were dealt in a single principal component analysis, both types of shape variations would be confounded in the results. Unlike root shape, such procedure as to curvature is unnecessary for analysis of shape of leaves for which top and bottom sides are clear.

Analysis of variance of the scores for each principal component indicated that in group A varietal effect was highly significant for the first five components, but particularly for the first three, which corresponded to aspect ratio, bluntness of the distal part and swelling of the middle part respectively. The results suggest that these traits are highly heritable and may be selected for by application of the method reported here. Since the effects of variety, soil type and growth stage showed no significant interaction in most cases, evaluation of root shape can be made simply by adding together the three main effects. In group B varietal effects were significant for 1st, 2nd, 4th and 5th components. F values were much lower than the corresponding components in group A, however it should be noted that varietal effect was non-significant for the 3rd component despite the highly significant 5th component, indicating that components with a larger contribution are not always important sources of genetic variation.

The present method based on elliptic Fourier descriptors has many advantages. Firstly, it requires no landmarks. This feature is advantageous in the analysis of the contour shape of plant organs such as roots, leaves and seeds which are scanty in landmarks. Second, in landmark based methods, an investigator has to make definition of the measures of shape based on landmarks, and sufficient prior knowledge about the shape variation of the object is required. While, in the present method, data are taken automatically as chain-code with no need of prior knowledge. Third, it is possible to visualize the contour information and to reconstruct the original shape. The variation evaluated by each of the principal component can be visualized separately. Forth, the coefficients for the elliptic Fourier descriptors can be mathematically normalized to size, rotation, and starting point of the contour trace, which saves the investigator troublesome manual adjustment of the shape with respect to these factors before analysis. Furthermore it is possible to analyze shape independently of size. Most procedures in the present method can be performed automatically on computer programs which were written on pascal and c language by the first author. This is especially important in the genetic analysis of crop organ shape which generally needs large samples. Diallel crosses of several varieties of Japanese radish has been done using the present method, the result of which will be published elsewhere.

References

Bierbaum, R.M. & S. Ferson, 1986. Do symbiotic pea crabs decrease growth rate in mussels? Biol Bull 170: 51–61.

Diaz, G., A. Zuccarelli, I. Pelligra & A. Ghiani, 1989. Elliptic Fourier analysis of cell and nuclear shapes. Comp Biomed Res 22: 405–414.

Ferson, S., F.J. Rohlf & R.K. Koehn, 1985. Measuring shape variation of two-dimensional outlines. Syst Zool 34: 59–68.

Freeman, H., 1974. Computer processing of line drawing images, Comput Surv 6: 57–97.

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

Gubanyi, A., 1996. Morphometric analysis of microscopic hooks of taeniid tapeworms (Cestoda, *Taeniidae*). In: L.F. Marcus, M. Corti, A. Loy, G.J.P. Naylor & D.E. Slice (Eds). Advances in Morphometrics, pp. 503–510. Plenum Publishing Corporation, New York.

Hadipoentyanti, E., N. Niwa, N. Furuta & S. Ninomiya, 1996. Discrimination between Indonesian clove (*syzygium aromaticum* (L.) Merr. & Perr.) populations by elliptic Fourier descriptors of leaf shape, SABRAO J 28: 25–34.

Hu, M.K., 1962. Visual pattern recognition by moment invariants. IRE Trans Inf Theory IT-8: 179–187.

Kuhl, F.P. & C.R. Giardina, 1982. Elliptic Fourier features of a closed contour. Comp Graph Ima Proc 18: 236–258.

Nishiyama, I., 1958. Nihon no Daikon (Japanese Radish). Nihongakujyutsu Sinkoukai, Tokyo.

Rohlf, F.J. & J.W. Archie, 1984. A comparison of Fourier methods for the description of wing shape in mosquitoes (*Ritera culicidae*). Syst Zool 33: 302–317.

Shapiro, S.S. & M.B. Wilk, 1965. An analysis of variance test for normality (complete samples). Biometrika 52: 591–611.

White, R., H.C. Rentice & T. Verwist, 1988. Automated image acquisition and morphometric description. Can J Bot 66: 450–459.

Yamamoto, C., E. Hadipoentyanti, H. Iwata, M. Niwa, S. Ninomiya & Y. Ukai, 1994. Diversity in Indonesian clove [*Syzygium aromaticum* (L) Merr. & Perr.] populations as surveyed by elliptic Fourier descriptors of leaf shape. Proc 7th Intern Cong SABRAO & WSAA. 53–59.