Universidad de Guantánamo

Ciencia y Progreso
Publicada en línea: enero, 2022, RNPS: 2477, ISSN: 2707-7098
http://cienciayprogreso.cug.co.cu

Morfometría geométrica aplicada al estudio de la hoja de Theobroma cacao

Geometric morphometry applied to the study of Theobroma cacao leaves

Frank Ernesto Soto- Gómez; Igor Bidot- Martínez
Universidad de Guantánamo
Correo(s) electrónico(s)
frankernesto@cug.co.cu
ibidot@cug.co.cu
ORCID: https://orcid.org/ **0000-0002-4379-4395**
https://orcid.org/**0000-0002-1262-7674**

Recibido: 19 de abril de 2021
Aceptado: 6 de junio de 2021

Resumen

El propósito consiste en analizar los patrones de variación de la forma de la hoja de cacao mediante la morfometría geométrica para identificar los genotipos tradicionales, dada la necesidad de identificar los genotipos tradicionales de cacao basado en la forma de la hoja. Se fotografiaron 490 hojas de siete genotipos y se colocaron 14 puntos clave, se les realizó análisis procrustes, un ajuste con el análisis de componentes principales y un análisis canónico. Se observó variabilidad morfológica distintiva en cuanto a la forma del cacao tradicional y los genotipos comerciales analizados, lo cual permite una diferenciación preliminar entre ellas.

Palabras clave: Componentes principales; Análisis canónico; Morfometría geométrica; Puntos clave.

Abstract

The purpose is to analyze patterns of variation in cocoa leaf shape using geometric morphometrics to identify traditional genotypes, given the need to identify traditional cocoa genotypes based on leaf shape. Four hundred and ninety leaves of seven genotypes were photographed and 14 key points were positioned, procrustes analysis, a fit with principal component analysis and a canonical analysis were performed. Distinct morphological variability was observed in the shape of traditional cocoa and the commercial genotypes analyzed, allowing a preliminary differentiation between them.

Keywords: Principal components; Canonical analysis; Geometric morphometry; Key points.

Introduction

With the use of geometric morphometry, the most variable characters in the systematic of plants are evaluated from a new perspective, according to Siegel & Benson (1982). The development of morphometrics in recent decades has reached areas of biology traditionally devoted to descriptive studies, such as the morphological sciences. Morphology, with the new geometric morphometric tools, allowed the objective quantitative evaluation of morphological changes, as postulated by Zelditch, Swiderski, Sheets &Fink (2004). This was made possible with biometric techniques, instruments and computer programs that capture and analyze data from morphocoordinate matrices representing the geometry of a specimen. These techniques are not limited to obtaining linear measurements, which lack the precision and level of information of geometric data, according to Adams, Rohlf & Slice (2004).

Geometric morphometrics is the set of geometrically based methods and procedures for the analysis of shape patterns, where measurements are usually derived from contours and key morphological points. Geometric morphometric analysis consists of three fundamental stages: (1) obtaining morphometric data, (2) obtaining pure shape information, and (3) exploratory and confirmatory analyses of shape covariation and causal factors. Knowledge of geometric morphometry allows its application to the solution of problems in which morphology plays an essential role. The ability to extract information about the spatial location of morphological variation, its magnitude, position and spatial extent in the organism, is one of the most important benefits of geometric morphometry, as described by Bookstein (1991).

However, despite its broad potential, the application of geometric morphometrics in economically important crop plants is still not widespread. In systematic plant, one of the main points of location and analysis taken into account is leaf morphology, which is studied mainly by means of traditional morphometry. In the case of leaf shape, this offers a number of variants to consider in order to make a biometric analysis of the parts with greater variability. Examples of these are the apex, base, margin or linear dimensions (width and length), some of the qualitative and quantitative characters most frequently used in the classification of species, subspecies and varieties of plants, according to Henderson (2006).

One of the crops that has a wide morphological variability, so that geometric morphometry can potentially be applied is cocoa, Theobroma cacao L. This is a crop tree that is primarily used to produce chocolate from fermented seeds.

The diversity of leaves and the complex distribution in cocoa cultivars constitute the fundamental evidence of leaf variation, a hypothesis to test with such techniques and multivariate analysis. Despite their probable phytogeographic result, studied in other species by geometric morphometry, it is noteworthy that no published studies were identified in cocoa where the extent of differences among cocoa genotypes or in regions with variable edaphoclimatic characteristics are analyzed, although similar studies exist in other plant species.

Based on this background, the present research focuses on how to identify traditional genotypes of Theobroma cacao based on leaf shape. Based on this problem, the following hypothesis is proposed: If the patterns of variation of Theobroma cacao leaf shape are analyzed using geometric morphometry, it would be possible to identify traditional genotypes.

To evaluate this hypothesis, the general objective was: To analyze the patterns of variation in the shape of the cocoa leaf using geometric morphometry to identify traditional genotypes.

To achieve the objective, the following specific objectives were met:

1. to define the key points to be located on the Theobroma cacao leaf for shape analysis.
2. To obtain the consensus shape of Theobroma cacao leaves of the evaluated genotypes from the germplasm bank.
3. Confirm the differences between Theobroma cacao genotypes evaluated from the germplasm bank.

The scientific novelty of this research is that, for the first time in Cuba, the extent of the variation in the shape of Theobroma cacao leaves is evaluated and geometric morphometry is used to buy its morphological variability.

Development

Morphometric data acquisition

The samples analyzed were taken at the cocoa germplasm bank located at the Agroforestry Research Institute, Science and Technology Unit of Baracoa. The area has high accumulated annual rainfall, with an average of more than 2200 mm, the highest in the country. The topographic characteristics have a 95% of mountainous relief.

Two groups of plants were taken: two traditional genotypes (K 250 and K 251) and five commercial genotypes (UF 12, UF 29, UF 221, UF 613 and UF 296). At random, 75 leaves were selected by individuals from mature leaves without margin damage, which were digitized with a camera with 18

Mpx resolutions, on a background with graph paper and a ruler as a reference for size. The leaves were oriented with the axial part towards the observer and the apex towards the bottom of the image, according to Hearn (2009).

The images were processed with the TPS series programs, according to Rohlf & Marcus (1993). With the TpsUtils program the project was created and with the TpsDigs program the key points were located in the image and their Cartesian coordinates were determined. In each image of the leaves 14 key points were placed, two of anatomical type corresponding to the apex and the base of the leaf and 12 of type 2, according to Bookstein (1991), distributed uniformly on the edge of the leaves (Figure 1).

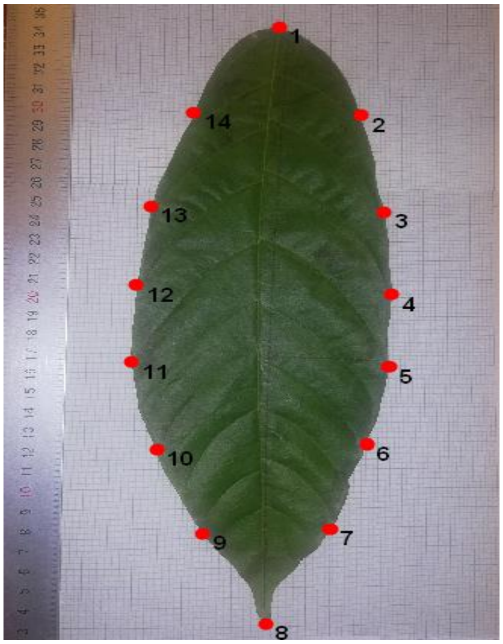


Figure 1: Location of key points on cocoa leaves.

Source: Own elaboration

Obtaining information describing pure form

Procrustes analysis was used to obtain the consensus shape of the cocoa leaves and the set of points representing the central tendency of the key points of the cocoa leaves. In this analysis, the key points arranged at the edge, base and apex of the cocoa leaves were adjusted by the least squares method and their respective matrix coordinates.

Subsequently, the dimensionality of the data was reduced by principal component analysis, which allowed obtaining the consensus shape of the cocoa leaf. The results were represented by a figure with the vectors of movement between the consensus shape and the mean of the change in leaf shape, a figure with the mean of the principal component analysis of the key points of the cocoa leaves, and a two-dimensional scatter plot with the principal components that explained the greatest percentage of the variance.

From the results of the principal component analysis, the separation between genotypes was maximized by canonical analysis. The results obtained were represented by two scatter plots of the genetic groups differentiated according to leaf shape and a figure to compare the consensus shape of the cocoa leaf with the canonical analysis adjustment. All analyses were performed with MorphoJ software, as reported by Zelditch et al. (2004).

Exploratory and confirmatory analyses of shape covariation and causal factors

To maximize the differences, the mean of the procrustes, principal component and canonical analyses obtained in the previous section was calculated using 10000 permutations of Markov chains and the Mahalanobis distance, according to Jolliffe (2002). The mean of these analyses was compared with the results of the same analyses and the results were plotted as explained under the heading obtaining pure form. All analyses were performed with the MorphoJ program, according to Zelditch et al. (2004).

Discussion of morphometric data acquisition

The cocoa leaf consensus shape of the genotypes analyzed and the set of points representing the key points used were obtained, as expressed by Rohlf & Slice (1990) and Robinson, Blackwell, Stillman & Brook (2001).

In the principal component analysis, the first components showed a high percentage of correlation as shown by the Pearson correlation results. In the principal components from 1 to 10, the degree of correlation was higher than 0.7; a result above which they are considered highly correlated, according to Zelditch et al. (2004). With these first ten principal components, the vectors of movement between the leaf consensus shape and formation of the vectors of movement obtained in the result of the principal component analysis and the displacement of the points of type 1 and 2 of the sample of the analyzed leaves were obtained, as argued by Lattin, Carroll, & Green (2003). This result allows estimating the horizontal and vertical displacement of each region in the image.

The principal component analysis allows comparing the obtained consensus shape (blue line) with the mean of the leaf shape change (red line) (Figure 2).

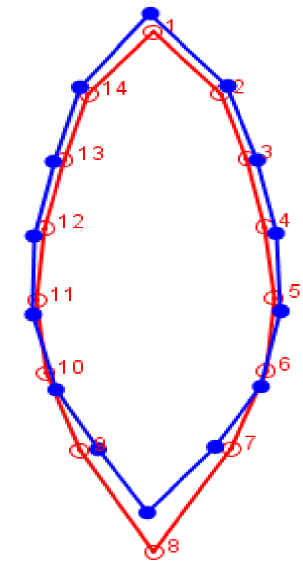


Figure 2: Mean of principal component analysis of key points of cocoa leaves.

Source: Own elaboration

The results obtained in the principal component analysis were represented by a scatter plot of the different genotypes evaluated with respect to leaf shape (Figure 3).

The results obtained show a high degree of similarity between groups, with some differences between groups, for example, group K 251 with respect to the others. Overlap between cocoa leaf shapes was observed, although with a low level of variability between genotypes.

Due to the small differences observed, a canonical analysis of the variables was performed as analyzed by Lele & Richtsmeier (2001). The six principal components that contributed most to the variation had a significance greater than 0.05. The results of the canonical analysis with these six components (blue line) were compared with the consensus form (red line) (Figure 4).

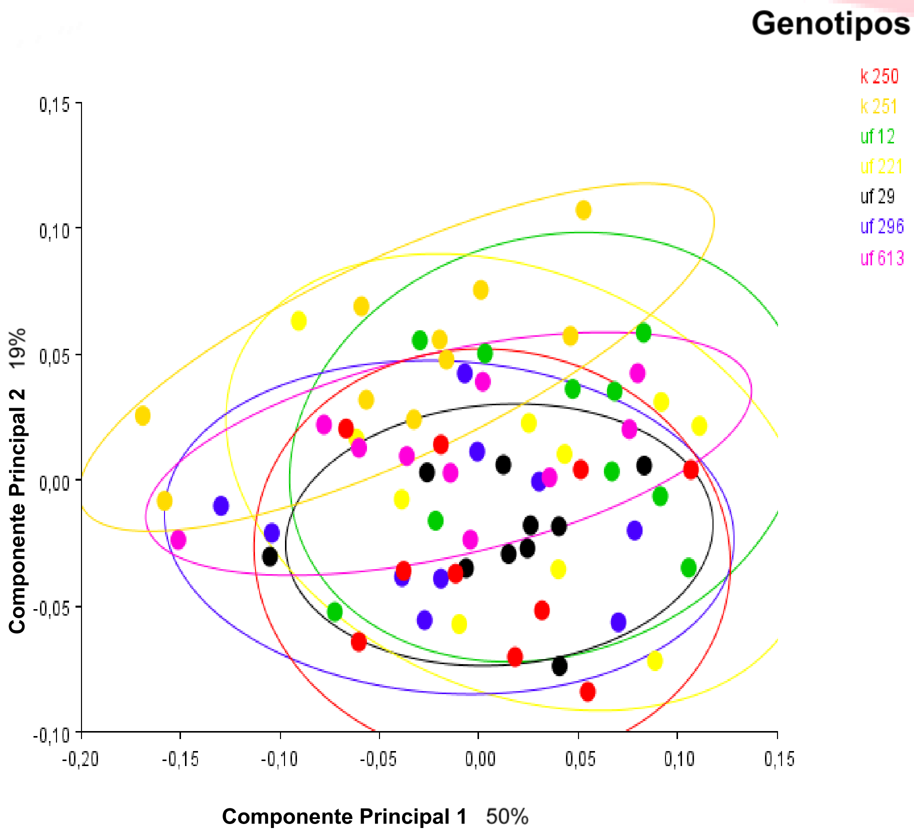


Figure 3: Scatter plot obtained by principal component analysis.
Source: Own elaboration

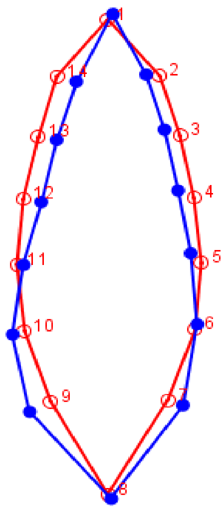
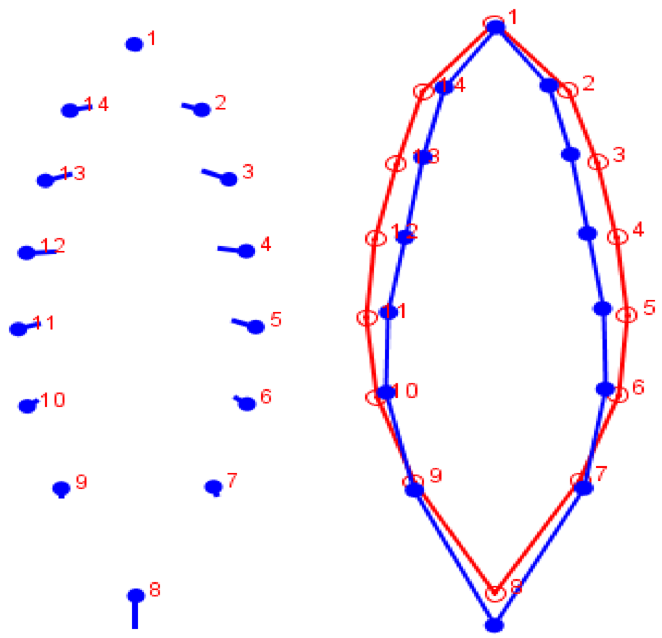


Figure 4: Consensus cocoa leaf shape obtained with canonical analysis.
Source: Own elaboration

Discussion of exploratory and confirmatory analyses of shape covariation and causal factors.

The comparison of the mean of the procrustes analysis (red line) with the mean of the procrustes analysis (blue line) is shown in Figure 5.



a **b**

Figure 5: Analysis of the motion vectors of the Procrustes mean (a) compared to the Procrustes mean (b).

Source: Own elaboration

The average covariance matrix of the six principal components that contribute most to the variance is shown in Table 1.

Main Components	**Own Values**	**Variance Percentage**
1	0,00105492	57,253
2	0,00042018	22,804
3	0,00024123	13,092
4	0,00007866	4,269
5	0,00002830	1,536
6	0,00001927	1,046

Table 1: Averaged results of the covariance matrix.

Source: Own elaboration

A scatter plot of the analyzed cocoa genotypes was made with the two principal components that contributed most to the variance (Figure 6).

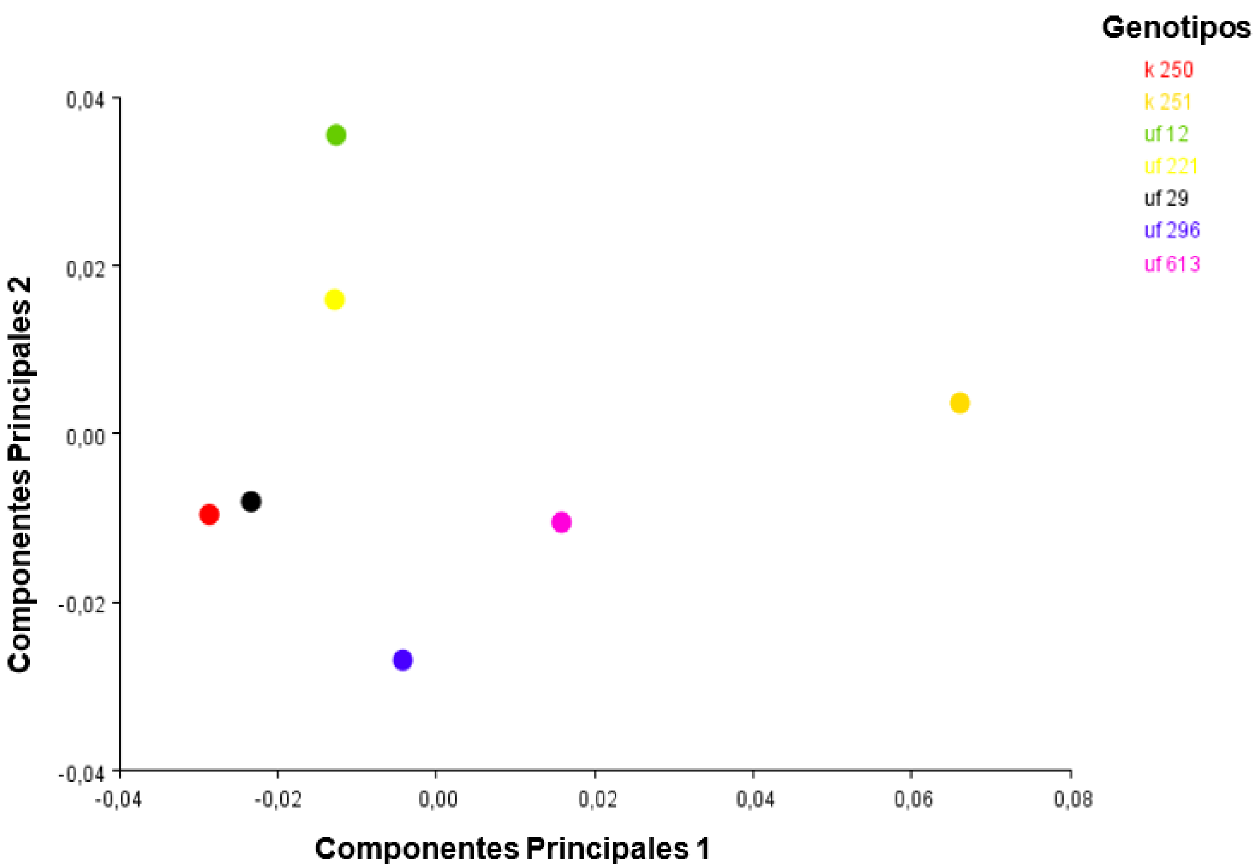
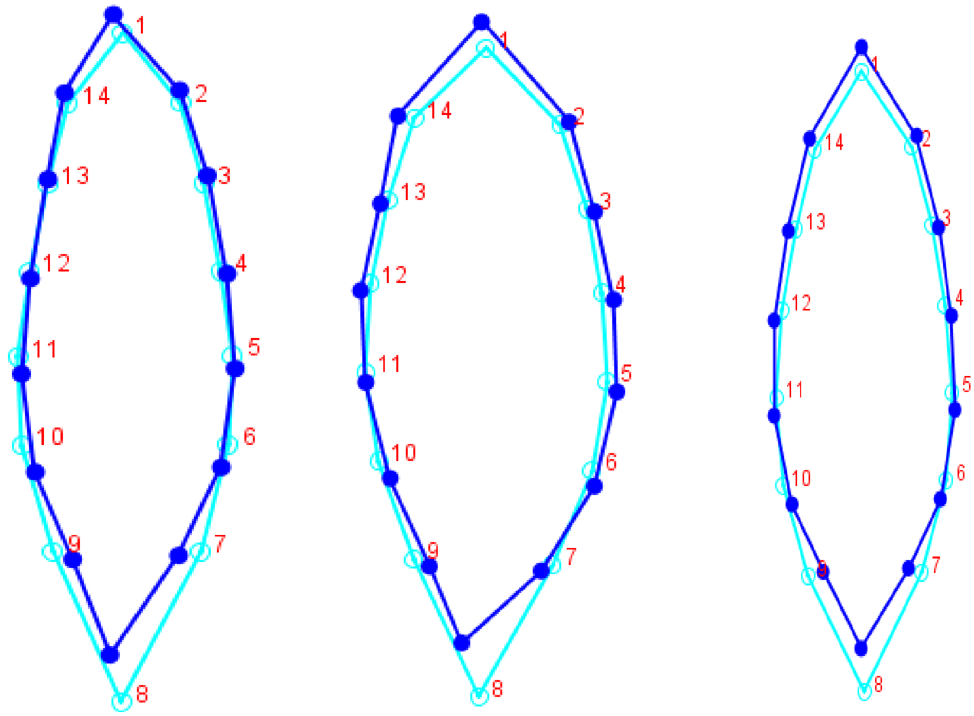


Figure 6: Comparison between the two principal components and group separations.

Source: Own elaboration

These results highlight significant differences between genotypes. These differences, although existing, are not fully evident until comprehensive analyses are performed. The results indicate that, although low, there is variability in leaf shape, which can potentially be used for genotype identification. However, larger samples need to be taken and combined with analyses of other plant organs, such as the fruit, because of their high variability and importance in the classification of the species.

The overlapping behavior of the leaf shape of the genotypes was again evaluated to demonstrate non-obvious differences between groups. The results indicate that there is variability in leaf shape, more evident with a new analysis of the Procrustes mean, variance adjustment and principal component mean analysis (Figure 7).



a **b** **c**

Figure 7: Leaf size variation based on the mean of the Procrustes adjustment.

Source: Own elaboration

The results indicate that the shape shows high variability. This could be due to the fact that the genotypes studied are differentiated not only by their genetic characteristics, but also by morphological characters associated with adaptive factors, according to Krzanowski (1988). Regarding leaf size, the genotypes analyzed presented marked differences observed in the figure by the consensus shape of K 250, K 251, presented larger leaves, while the other commercial genotypes were grouped with a consensus shape of smaller leaves.

The fact that leaf size was positively associated with leaf length of their genotypes supports the idea that size is variable by each genetic group.

The results indicate the importance of morphometric traits to support interspecific studies of cocoa. The research combines for the first time the analysis of anatomical points by geometric morphometry using morphometric characters. This study allowed a comparison between shape change established by ordination of the canonical analysis of the variable to infer shape change, as analyzed by Kaliontzopoulou, Carretero & Llorente (2007).

This work documents that the estimation of shape change is more appropriate for phylogenetic approach, for the analysis of coordinate configurations after Procrustes superposition and the calculation of relative deformations, as stated by Viscosi, Lepais, Gerber & Fortini (2009). The results described above indicate that geometric morphometry of cocoa leaves is a suitable tool for genotype discrimination, as well as a method of easy application and statistical validation. With 14 anatomical points on the leaves, commercial and traditional cocoa genotypes are differentiated, although the differences are not large. Due to this result, additional analyses including other genetic groups are still needed to evaluate leaf shape changes for the 10 existing cocoa genotypes.

Although studies using morphometric methods on leaves of several plant species have been published, nothing has been done to date with leaves of cocoa cultivars, the results obtained represent a step towards the identification of genotypes by morphogeometric means. For the first time the appearance of differences in leaf shapes was shown by geometric morphometry.

These results constitute a first approach to the use of geometric morphometry for the identification of cocoa cultivars. This technique has advantages due to obtaining results quickly and at low cost, so it can be a first approach to characterize the biodiversity of a group of species. It can be a useful tool for this purpose, taking into account the difficulties in identifying cultivars, due to the similarities that exist between them. In addition, they can be extended to be used for other species of agricultural or forestry interest.

Conclusions

The 14 key points identified in the cocoa leaves of the genotypes evaluated allowed the analysis of the shape of the leaves. From these key points, the consensus shape of the evaluated cocoa leaves was obtained, where the shape information was summarized. These results showed distinctive morphological variability in the shape of traditional cocoa, distinctive in comparison with the commercial genotypes analyzed, which allowed a preliminary differentiation among

Bibliographic references

Adams, D. C., Rohlf, F. J. & Slice, D. E. (2004). *Geometric morphometrics: ten years of progress following the 'revolution'*. *Ital. J. Zool. 71*: 5-16.

Bookstein, F. L. (1991). *Morphometric tools for landmark data: geometry and biology*. Nueva York: Cambridge Univ. Press.

Hearn, D. J. (2009). *Shape analysis for the automated identification of plants from images of leaves*. *Taxon 58*: 934–954.

Henderson, A. (2006). *Traditional morphometrics in plant systematics and its role in palm systematics. Bot. J. Linn. Soc. 151*: 103-111.

Jolliffe, I. T. (2002). *Principal Component Analysis*. (2daed.). New York: Springer Verlag.

Krzanowski, W. J. (1988). *Principles of Multivariate Analysis: A User's Perspective*. Oxford: Clarendon Press.

Lattin, J., Carroll, J. & Green, P. (2003). *Analyzing multivariate data*. Pacific Grove: Thomson Learning, Inc.

Lele, S. R. &Richtsmeier, J. T. (2001). *An invariant approach to statistical analysis of shapes*. Florida: Chapman & Hall/CRC

Robinson, D. L., Blackwell P. G., Stillman E. C. &Brook A. H. (2001). *Planar Procrustes analysis of tooth shape*. Arch. Oral Biol. *46*: 191–199

Rohlf, F. J. & Marcus, L. F. (1993). A revolution in morphometrics. *Trends Ecol. Evol. 8*: 129– 132.

Rohlf, F. J. & Slice, D. E. (1990). Extensions of the Procrustes method for the optimal superimposition of landmarks. *Syst. Zool. 39*: 40–59.

Siegel, A. F. & Benson, R. H. (1982). A robust comparison of biological shapes. *Biometrics 38*: 341–350.

Viscosi, V., Lepais, O., Gerber, S. & Fortini, P. (2009). Leaf morphological analyses in four European oak species (Quercus) and their hybrids: A comparison of traditional and geometric morphometric methods. *Plant Biosystems 143* (3): 564-574.

Zelditch, M. L., Swiderski, D. L., Sheets, H. D. & Fink, W. L. (2004). *Geometric Morphometrics for Biologists: A primer*. London: Elsevier Academic Press.