REGULAR PAPER

CURVATURE MEASUREMENT AS A TOOL FOR THE ANALYSIS OF MORPHOMETRIC VARIATION USING *Drosophila* WINGS AS A MODEL

Pedro Roberto Rodrigues Prado[1], Luciano da Fontoura Costa[2],
Evandro Marsola Moraes[3], Maura Helena Manfrin[4] and Fábio Melo Sene[1]

[1]Department of Genetics, Faculty of Medicine at Ribeirão Preto and [4]Department of Biology, Faculty of Philosophy, Sciences and Letters at Ribeirão Preto, University of São Paulo (USP), Ribeirão Preto, SP, Brazil.
[2]Cybernetic Vision Research Group, Institute of Physics at São Carlos, University of São Paulo (USP), São Carlos, SP, Brazil.
[3]Federal University of São Carlos (UFSCar), Sorocaba Campus, Sorocaba, SP, Brazil.

ABSTRACT
This article describes the use of geometric measurements of continuous, discrete parameters to study morphometric variation in the wing cells of two sibling species, *Drosophila mercatorum* and *Drosophila paranaensis*. To validate the results, the same wing samples were also analyzed using truss networks and partial warps, in addition to a comparison with the ellipse method. The use of discriminative measurements in conjunction with a Bayesian-based classification method yielded a relatively high number of correct classifications for new individuals. These results compared favorably with those obtained using truss networks, partial warps and the ellipse method. These findings indicate that continuous curvature and arc length measurements may be useful parameters for the morphometric analysis of insect wings and possibly other biological structures and shapes.

Key words: Bayesian classification, curvature, geometric properties, morphometric analysis

INTRODUCTION
Morphological characteristics are an important source of information for many areas of biological investigation, including systematics and taxonomy. Most studies in these areas are done using meristic and morphometric characteristics. Meristic characteristics are generally countable and informative for species, genera and higher taxonomic levels. However, in interspecific and populational studies, these measurements are not informative and it is therefore necessary to obtain information on morphometric characteristics [5]. These characteristics are generally quantitative phenotypic values obtained from continuous measurements and ratios in which classes are often defined based on means and standard deviation.

Descriptive and comparative studies of morphometric characteristics normally use general morphometric analysis based on measurements such as distances or angles [2,5,13,17,18,24,26,27]. These values are obtained from the distance between homologous points or landmarks [19] of a particular structure. The values generated can then be used to quantify morphological variation. The efficiency of this approach depends on the existence and precise definition of landmarks for these structures and on the association of these homologous points with morphological variation. Since landmarks are not always easy to find, alternative descriptive algorithms that allow the analysis of morphometric variation without defining homologous points are being increasingly used. Such algorithms include methods based on outline data [22] in which the form (shape + size) of the structure is obtained from its contour.

Curvature descriptors
Cesar Jr and Costa [7] and Costa and Cesar Jr [8] proposed the use of continuous curvature as an alternative method for analyzing morphometric characteristics, without the need for landmarks. Continuous curvature allows an analysis of the shape of the object and the length of its arc, both of which are related to the size of the object. Curvature

Correspondence to: Dr. Fábio de Melo Sene
Departamento de Genética, Faculdade de Medicina de Ribeirão Preto (FMRP), Universidade de São Paulo (USP), Av. Bandeirantes, 3900, CEP 14049-900, Ribeirão Preto, SP, Brazil. Tel: (55) (16) 3602-3103, Fax: (55) (16) 3633-0069. E-mail: famesene@usp.br

is a particularly important geometric measurement that expresses the rate of change of the angle between the tangent to the curve and the x-axis [8]. This measurement can be used to characterize a curve, with low values of curvature along a portion of the curve indicating a straight region whereas abrupt variation in the curvature corresponds to a vertex. In addition, concavity along the curve can be determined by the sine of the curvature. Two important additional properties of curvature as a geometric measurement are that the original curve can be reconstructed from the curvature values and are invariant for translation and rotation.

Application of curvature analysis

Wing morphology is an important characteristic for identifying insect species in taxonomic and systematic studies [11,15,16]. Indeed, the morphometric analysis of insect wings, mainly in dipterans, has been used to clarify the relationship between closely related taxa [1,9], to discriminate cryptic species [15,17] and to discriminate co-specific populations [12,16].

In this work, we examined the applicability of continuous curvature analysis of wing morphology for discriminating between *Drosophila mercatorum* and *Drosophila paranaensis*. These two sibling or cryptic species of the *Drosophila mercatorum* subgroup and *Drosophila repleta* group, respectively, are indistinguishable by simple qualitative analysis, and only males can be identified morphologically based on small differences in their genital apparatus (aedeagus) [28].

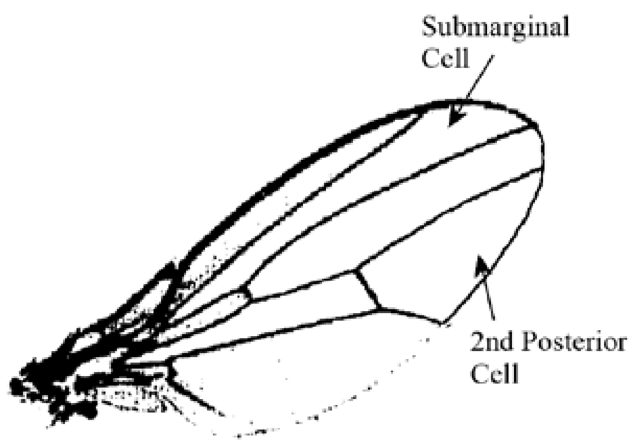

Figure 1. Digital photograph and description of the cells and veins of a *Drosophila* wing (**a**). The wing cell outlines were enhanced using an image editor.

MATERIAL AND METHODS

Samples

Seventy six female (38 of each species) and 80 male (40 of each species) flies from a natural population of each of the two species were used. The wings were mounted on slides as described by Moraes *et al.* [15] and digital images (300 dpi) were obtained. Wings from male and female flies of each species were analyzed separately. The submarginal cells of the wings were examined in detail because of their elongated shape and because they covered most of the wing; the second posterior wing cells were also analyzed to allow comparison with results from a previous study with this same species [15] (Fig. 1).

To obtain the desired measurements, the cells were grouped in pairs (x,y) corresponding to the pixels of the digitized images. Four landmarks, which were generally used to delimit regions of interest in the shape being studied, were identified manually for each pair, and the distance between pairs of landmarks, i.e., the arc, was determined (Fig. 2).

Mathematical and computational aspects

From the curve deconstructed on the x and y axes (Fig. 3), it was possible to calculate the curvature (*k*) from the first and second order derivatives using the equation

$$k = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{\sqrt{(\dot{x}^2 - \dot{y}^2)^3}}$$

where $\dot{x}$ and $\dot{y}$ are first order derivatives and $\ddot{x}$ and $\ddot{y}$ are second order derivatives.

Figure 4 shows the curvature before (*k*) (4a) and after (*Keq*) (4c) standardization with a sigmoid. (4b), which was used to amplify the points of lower curvature [8]. This standardization was required because the curvature tended to be too sensitive at more abrupt variations of the original contour. Without this correction, divergent curvature values may be obtained for two slightly different versions of the same contour point. Note that the peaks of the equalized curve in Figure 4 have a more uniform height.

Using the approach described above, we obtained measurements (Table 1) that were considered jointly with other measurements (described below). The abbreviations used in this work are shown in Table 1. The measurements obtained were:

1) Continuous curvature. The mean value and standard deviation were calculated from the continuous curvature and the standardized continuous curvature (*Keq*). For each fly species and group (females and males), the mean value and the standard deviation were calculated and stored for each wing. The same procedure was used to calculate the curvature of the arcs between landmarks.

2) Discrete curvature. This measurement was related to the angle between vectors determined by successive

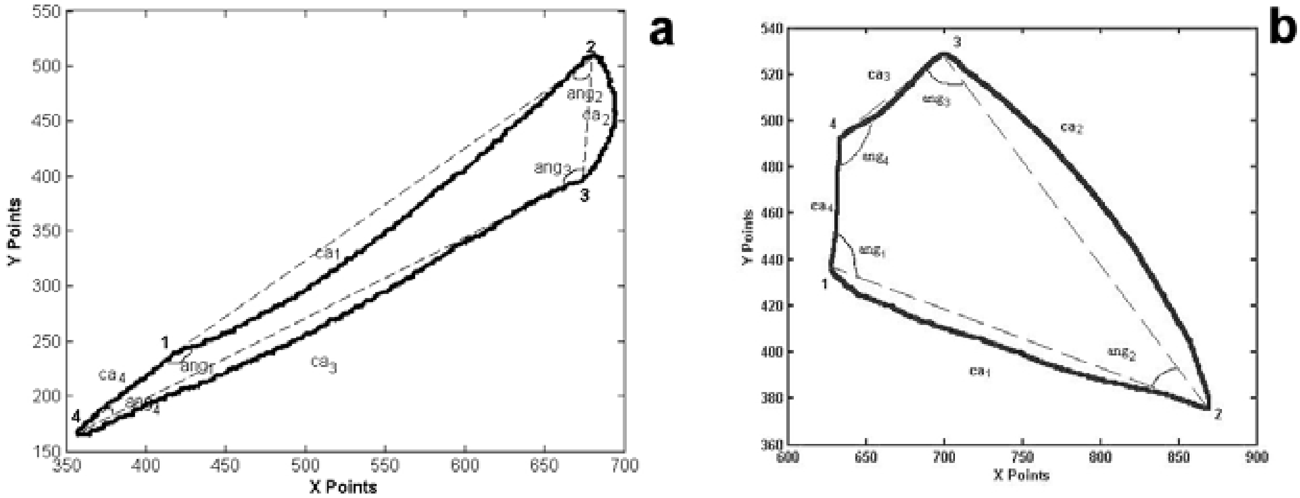

Figure 2. Standardization of the submarginal (**a**) and second posterior (**b**) wing cell, showing the location of four landmarks, the position of the arc length (**ca**), and the position of the discrete curvature (**ang**).

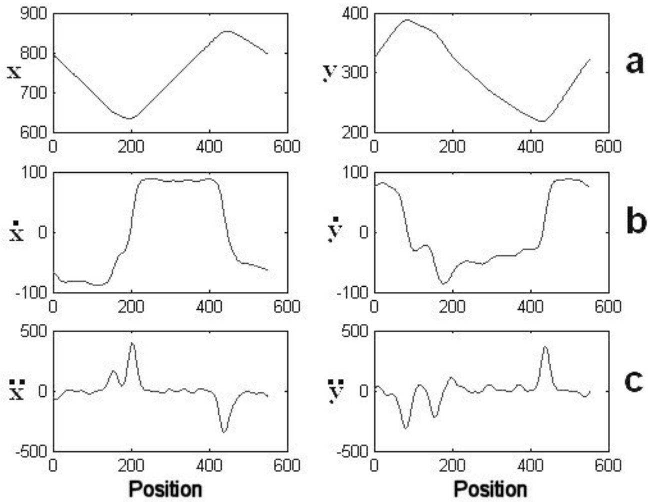

Figure 3. X and Y coordinates (**a**) and the respective first (**b**) and second (**c**) order derivatives.

pairs of points or landmarks along the curve, and was closely related to the continuous curvature, hence the name discrete curvature. The difference was that, whereas the continuous curvature expressed the variation in an angle within a small (actually infinitesimal) neighborhood of a given point, the discrete curvature indicated the rate of angle variation between two consecutive landmarks, as illustrated in Figure 2. Each vector between two landmarks was therefore determined by the coordinates of the points (X_a, Y_a), (X_b, Y_b). The angle was calculated based on the vectorial product of the vectors, as follows:

$$ang = \arccos\left(\frac{(X_a X_b - Y_a Y_b)}{\sqrt{X_a^2 + Y_a^2}\,\sqrt{X_b^2 + Y_b^2}}\right)180/\pi$$

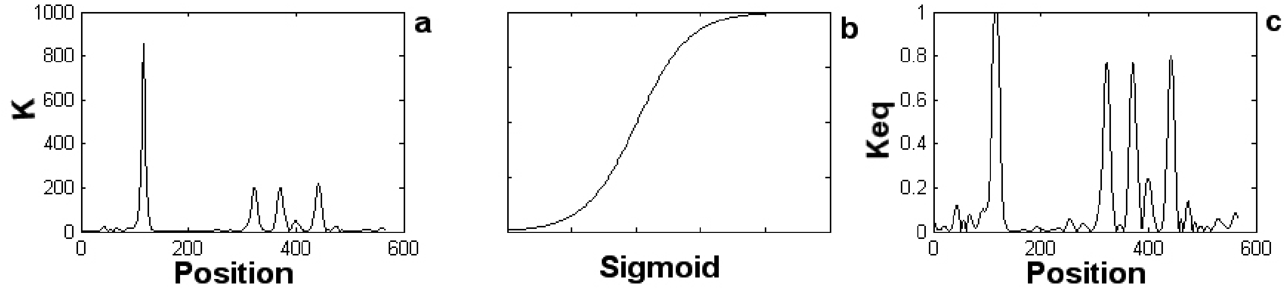

Figure 4. Continuous curvature (**a**), sigmoid transformation to amplify the vertex (**b**) and standardized/equalized continuous version (**c**).

Table 1. Measurements made using the technique described in this work and their respective abbreviations.

Measurement	**Abbreviation**
Continuous curvature	k or curvature
Equalized continuous curvature	Keq or curveq
Equalized continuous curvature of arcs	karc (karc1, karc2, karc3, karc4)
Area	area or A
Normalized area	Na or normarea
Arc length	Ca (ca1, ca2, ca3,ca4)
Relative arc length	relarc (relarc1, relarc2, relarc3)
Discrete curvature	angle or ang. (angle1, angle2, angle3, angle4)

3) Arc length. This measure, which expresses the accumulated distance between consecutive points (Figure 2), was obtained by adding the distances between each pair of points along the curve between two given landmarks, i.e.,

$$ca = \sum_{i=l1}^{l2-1} \sqrt{(p_{x(i+1)} - p_{x(i)})^2 + (p_{y(i+1)} - p_{y(i)})^2}$$

where l_1 is the position of landmark 1, l_2 is the position of landmark 2, $p_{x(i)}$ is the i-th position along the x-axis of a point between l_1 and l_2, and $p_{y(i)}$ is the i-th position along the y-axis of a point between l_1 and l_2.

4. Area. Since the surface area increased proportionally with the size of the wing cells, this parameter was normalized by dividing the area by the square of the diameter, i.e., the largest distance between any pair of points, of the wing cell being studied. The area was calculated as the double sum for the specified region:

$$area = \sum \sum x(t)\dot{y}(t)$$

Bayesian classification

This method is based on probability density functions (related to relative frequency histograms), normalized by mass probability, for measurements estimated for each species [12,27]. Provided that these functions can be properly estimated, the Bayesian approach to classification is theoretically optimal because it minimizes the number of misclassifications. In the general case involving *k* classes, the classification of an object characterized by a specific set of measurements is defined by the density function with the highest value. The following approaches for the estimation of density functions are normally applied:

1) Non-parametric: the density function is obtained through procedures used to estimate densities, such as relative frequency histograms or convolutions with area-preserving kernels, such as a normalized Gaussian distribution (Parzen windows approach) [8,10].

2) Parametric: in this case, a putative analytical density is assumed to apply and involves parameters that are estimated by traditional approaches (maximum likelihood) [10]. In this report, we use one-dimensional Gaussian distributions.

Analysis criteria

The geometric measurements obtained in this work were divided into shape (continuous curvature and standardized continuous curvature) and size (arc length) measurements. The three criteria used to compare the methods described here were measures of shape, size and a combination of shape and size. This approach provided information on the best group of measurements for differentiating individuals. These results were then divided into two groups of measurements, one for submarginal cells and the other for second posterior cells. Each cell and criterion was then analyzed by (a) canonical variate analysis (for measurements defined by each criterion) and (b) probability density functions (based on the scores for these variables).

Truss network and "partial warps" methodologies

To allow comparison with other methods reported in the literature, the data obtained from the curvature measurements were compared with those obtained from the same set of female *Drosophila* wings using truss network methods [25] and "partial warps" [4].

Truss network analysis

The truss network method, described by Strauss and Bookstein [25], uses the distances between specific landmarks. For the wings used here, the landmarks were defined in a similar fashion to those used by Klaczko and Bitner-Mathé [13] (Fig. 5). The wings were digitized and the position of the landmarks were parameterized. The distance between the landmarks was calculated using the Euclidean distance between the two points.

Partial warps analysis

The partial warps method proposed by Bookstein [4] yields information on the shape of the structure from the superposition of landmarks. The landmarks used here were the same as those defined by the truss network, and the scores were obtained by energy matrix shape deformation. Landmark superposition is done after the effects of size, orientation and position have been eliminated. The shape variables were studied using the invariance principle proposed by Lele and Richtsmeier [14], which deals with size, orientation and position. The configuration of anatomical landmarks does not vary when its components are multiplied by a scale [4] or when rotated, translated or scaled [3].

The shapes in question were analyzed after application of the generalized orthogonal least-squares method described by Rohlf and Slice [23]. In all, nine landmarks were defined on female wings (Fig. 5).

Programs used to obtain the measurements.

The measurements and analyses described here, including the truss network, were done using the Matlab 6.5 – Release 13 software. For partial warps, the programs TPSDIG [20] and TPSRELW [21] were used to generate the files with the position of the landmarks and to obtain the scores by the partial warps method. The statistical methods used to analyses the partial warps and for to measure the wings cells were similar.

RESULTS

Continuous curvature

Of the various measurements obtained from wing cells (Table 1), the best differentiation was obtained for continuous curvature and arc length. Hence, subsequent analyses concentrated on this group of measurements, which were obtained from the two cells of the wings. Table 2 summarizes the results obtained for the classification of individuals based on measurements of continuous curvature and arc length. Figure 6 shows the results of the analysis of submarginal cells in females (Fig. 6, lines 1 and 2) and males (Fig. 6, lines 3 and 4). Lines 1 and 3 of Figure 6 shows the distribution of the measurements whereas lines 2 and 4 show the probability density functions. Column A shows the measurements for shape and size, column B shows the measurements for shape and column C shows the measurements for size. Figure 7 shows the results for the second posterior cells (presented using the same arrangement as in Figure 6).

Comparative analyses

Truss network

The landmark pairs (d1 to d11), which define the straight lines used in the truss network analysis, are shown in Figure 5. An analysis of canonical variables applied to the measurements obtained yielded a density probability function for the first canonical variable, and it was possible to discriminate among individuals by using the Bayesian classifier. The percentage of reclassified individuals is shown in Table 3. The values for the first canonical variable and its respective density probability function are shown in Figure 8.

Partial warps

The results obtained by partial warps analysis [4] of the landmarks of female *Drosophila* wings are shown in Table 4, along with the respective percentages of reclassified individuals. Figure 9 shows a statistical test of canonical variable analysis used to determine the variability between species based on the first canonical variable and the respective density probability function. Table 5 summarizes the percentage of correct classifications obtained from the analysis of *D. mercatorum* and *D. paranaensis* wings by the curvature technique used here compared to other methods reported in the literature.

Table 2. The classification of flies by the curvature technique described in this work.

Cells	Group	Model	% Match
Submarginal cells	Females	Shape and size	92.1
		Shape	89.5
		Size	61.8
	Males	Shape and size	80.0
		Shape	77.5
		Size	57.5
Second posterior cells	Females	Shape and size	98.7
		Shape	92.1
		Size	64.5
	Males	Shape and size	77.3
		Shape	73.9
		Size	64.8

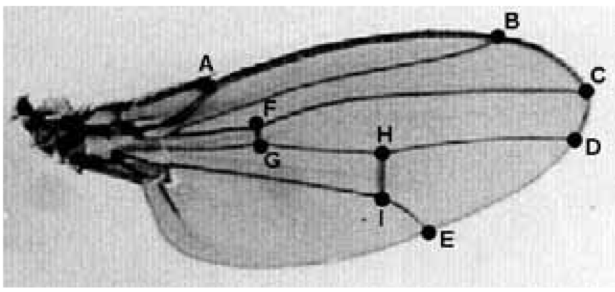

Figure 5. Landmarks (**A** - **H**). Truss network: d1(**A**-**D**), d2(**A**-**B**), d3(**B**-**C**), d4(**C**-**D**), d5(**D**-**E**), d6(**E**-**H**), d7(**A**-**H**), d8(**B**-**E**), d9(**A**-**E**), d10(**C**-**E**) and d11(**B**-**D**).

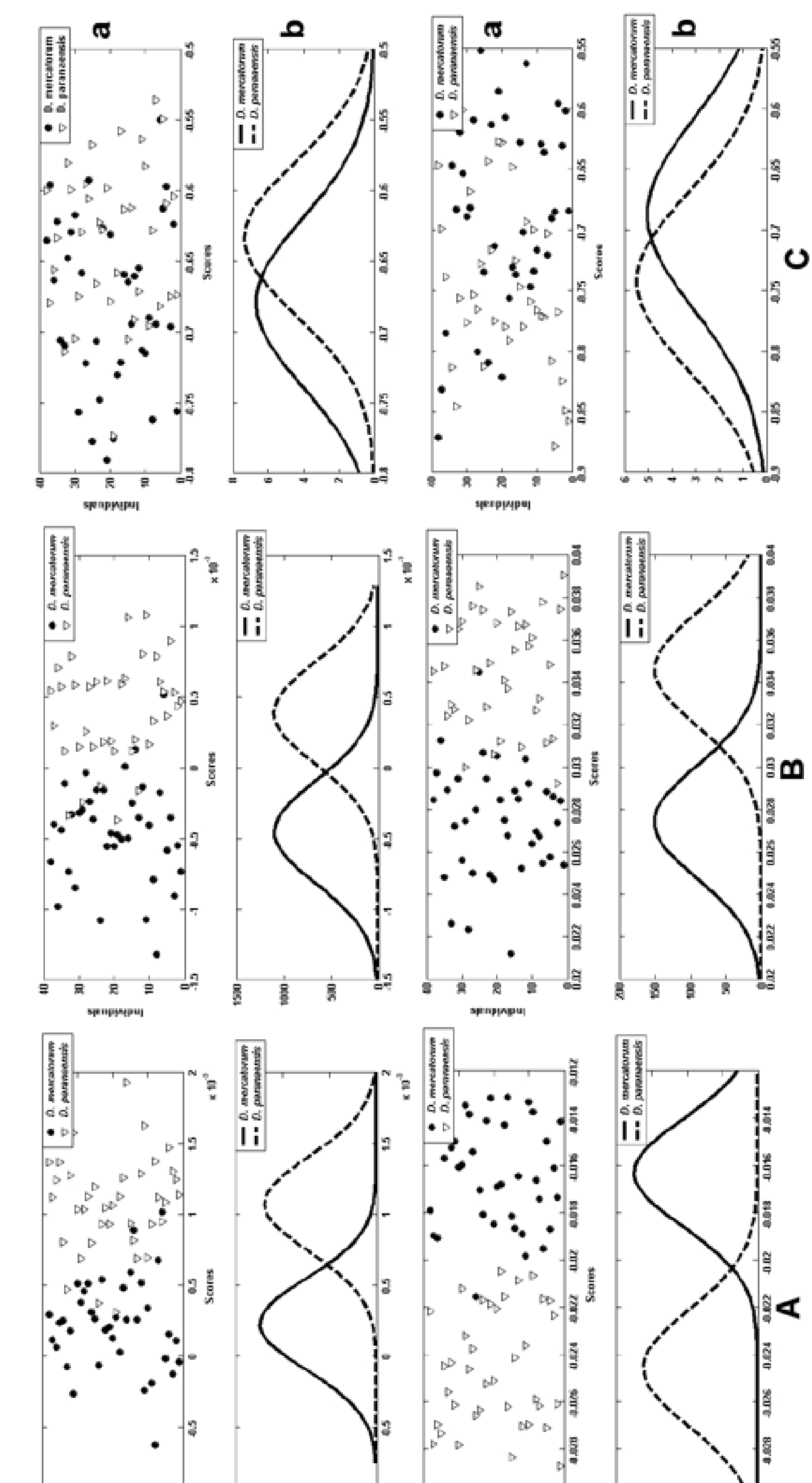

Figure 6. Scores for the first canonical variable (**a**) and the density probability function (**b**) of female wing cells for each group of individuals. Results for the submarginal cell (**1**) and for the second posterior cell (**2**). (**A**) Shape and size measurements grouped together, (**B**) shape measurements and (**C**) size measurements.

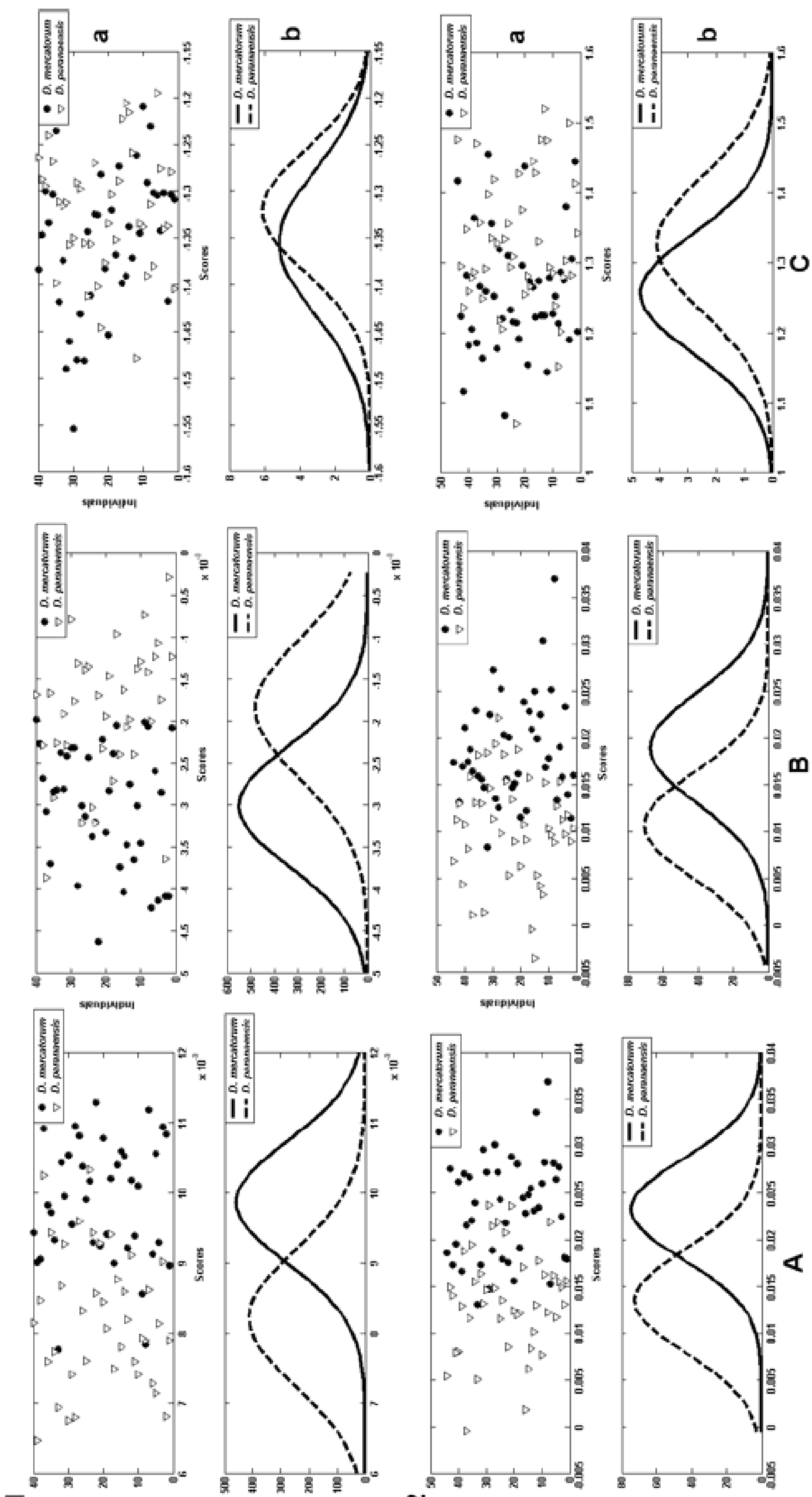

Figure 7. Scores for the first canonical variable (**a**) and the density probability function (**b**) of male wing cells for each group of individuals. Results for the submarginal cell (**1**) and for the second posterior cell (**2**). (**A**) Shape and size measurements grouped together, (**B**) shape measurements and (**C**) size measurements.

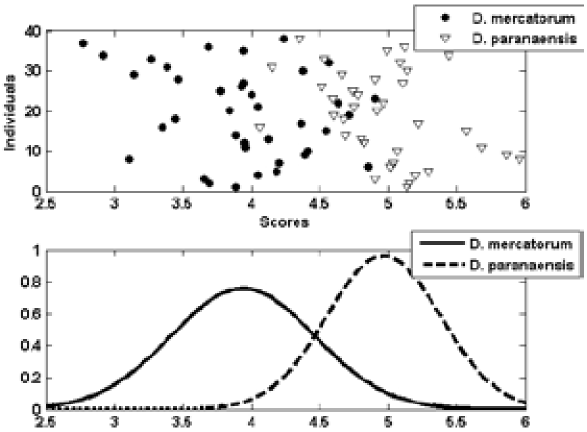

Figure 8. The first canonical variable and the adjusted density probability function for each group of individuals studied.

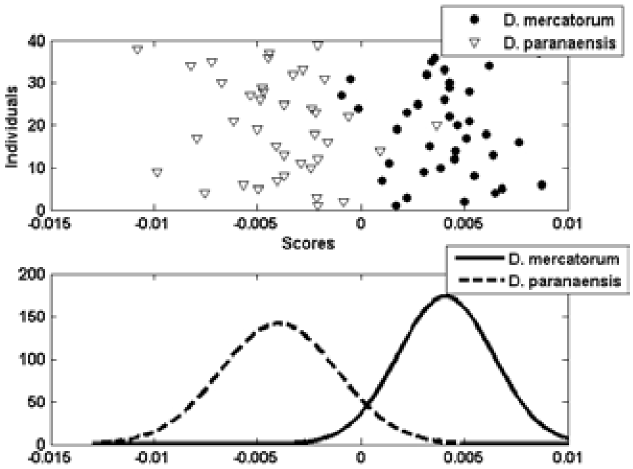

Figure 9. The first canonical variable and the adjusted density probability function for each group of female wings.

Table 3. The number of *D. mercatorum* and *D. paranaensis* females reclassified using the first canonical variable and a Bayesian classifier for the density probability function. Overall, 88.2% of the individuals were classified correctly.

Species	*D. mercatorum*	*D. paranaensis*
D. mercatorum	32	6
D. paranaensis	3	35

Table 4. The number of *D. mercatorum* and *D. paranaensis* females identified using the partial warps scores for wings based on the Bayesian classification of the first canonical variable. Overall, 93.4% of the individuals were classified correctly.

Species	*D. mercatorum*	*D. paranaensis*
D. mercatorum	35	3
D. paranaensis	2	36

DISCUSSION

General morphometric methods [6,17,19] have shown that morphological markers can be used to discriminate between species. In this work, we initially considered the shape of wing cells, particularly the submarginal cell, which has an elongated shape and spans a large part of the wing. Another region, the second posterior region, was also examined because a previous report had also studied this region in the same *Drosophila* species [15]. The results obtained for the second posterior region were better than those for the submarginal cell. Overall, the curvature method described here was easy to use, regardless of the region being analyzed.

Various mathematical tests were used to select the measurements that could differentiate the species. With Bayesian analysis, the percentage of correct classification based on the shape and size of the second posterior cell was 99% for females and 80% for males. Our findings indicate that it is

Table 5. Percentages of correctly classified *D. mercatorum* and *D. paranaensis* based on different methods applied to the same set of wings.

Method		**% match**	
		Males	**Females**
Curvature	Second posterior cell[1]	77.3	98.7
	Submarginal cell[1]	80.0	92.1
Partial warps[2]			93.4
Truss network[2]			88.2
Ellipse method[3]		84.5	98.0

Results obtained by: [1]curvature analysis, [2]other methods, and [3]reported by Moraes *et al.* [15].

possible to discriminate between *D. mercatorum* and *D. paranaensis* by measuring their wing cells, although it remains unclear why male wings were less discriminant than female wings.

To allow a comparison of these results with those obtained by other methods, the same wings from females of the two species were analyzed using truss networks and partial warps. Both of the latter two methods yielded significant reclassification rates that were slightly lower than that obtained by the curvature method described here, especially for the second posterior cell. However, comparison with the ellipses method showed that the results were similar. In this case, the advantage of the curvature method is that the structure being analyzed does not need to be adjusted geometrically to the ellipse. The main advantage of the curvature technique for analyzing shapes is the ease with which measurements can be obtained, without the need for landmarks to delimit specific regions of the organisms or structures being studied.

In conclusion, the continuous curvature approach used here, including the analysis of the arc segments between pairs of specific points, provided valuable information about the shapes being studied and efficiently discriminated between cryptic *Drosophila* species. These findings indicate that continuous curvature and arc length measurements are applicable to the morphometric analysis of insect wings and possibly also other biological structures and shapes.

ACKNOWLEDGMENTS

The authors thank A.C. Laus, R.S. Rosada and S.C. Bonfim for kindly providing the *Drosophila* specimens used in this work. We also thank F.F. Franco and E.C.C. Silva-Bernardi for their critical analysis of the manuscript and P.R. Epifânio for technical assistance. This work was supported by FAPESP (grant nos. 04/08565-8 and 03/05031-0), FINEP (grant nos. 43.860928.00-1987 and 1643/93 (1995-2000)), CNPq (grant nos. 302791/2003-5 and 473296/2003-9) and USP.

REFERENCES

1. Ansorge J (1994) Tanyderidae and Psychodidae (Insecta: Diptera) from the Lower Jurassic of northeastern Germany. *Paläontol. Zeit* **68**, 199-210.
2. Bitner-Mathé BC, Klaczko LB (1999) Size and shape heritability in natural populations of *Drosophila medi-opunctata*: temporal and microgeographical variation. *Genetica* **105,** 35-42.
3. Bookstein FL (1989) “Size and shape”: a comment on semantics. *Syst. Zool.* **38,** 173-180.
4. Bookstein FL (1991) *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press: Cambridge.
5. Bookstein FL, Chernoff B, Elder RL, Humphries JM, Smith GR, Strauss RE (1985) *Morphometrics in Evolutionary Biology*. Academy of Natural Sciences: Philadelphia.
6. Cavalcanti MJ, Monteiro LR, Lopes PRD (1999) Landmark-based morphometric analysis in selected species of serranid fishes (Perciformes: Teleostei). *Zool. Stud.* **38,** 287-294.
7. Cesar Junior RM, Costa LF (1996) Towards effective planar shape representation with multiscale digital curvature analysis based on signal processing techniques. Patt. Recog. 29, 1559-1569.
8. Costa LF, Cesar Junior RM (2000) Shape Analysis and Classification: Theory and Practice. CRC Press Book Series on Image Processing. Cambridge Univerity Press: New York.
9. Diniz-Filho JAF, Malaspina O (1995) Evolution and population structure of Africanized honey bees in Brazil: evidence from spatial analysis of morphometric data. Evolution 49, 1172-1179.
10. Duda RO, Hart PE, Stork DG (2000) Pattern Classification. John Wiley: New York.
11. Francoy TM, Prado PRR, Gonçalves LS, Costa LF, Jong D (2006) Morphometric differences in a single wing cell can discriminate Apis mellifera subspecies. Apidologie 37, 91-97.
12. Haas HL, Tolley KA (1998) Geographic variation of wing morphology in three Eurasian populations of the fruit fly, Drosophila lummei. J. Zool. Lond. 245, 197-203.
13. Klaczko LB, Bitner-Mathé BC (1990). On the edge of a wing. *Nature* **346,** 321.
14. Lele S, Richtsmeier JT (1991) Euclidean distance matrix analysis: a coordinate-free approach for comparing biological shapes using landmark data. *Am. J. Phys. Anthrop.* **86,** 415-427.
15. Moraes EM, Manfrin MH, Laus AC, Rosada RS, Bomfin SC, Sene FM (2004) Wing shape heritability and morphological divergence of the sibling species *Drosophila mercatorum* and *Drosophila paranaensis*. *Heredity* **92,** 466-473.
16. Moraes EM, Sene FM (2004) Heritability of wing morphology in a natural population of *Drosophila gouveai. Genetica* **121,** 119-123.
17. Moraes EM, Spressola VL, Prado PRP, Costa LF, Sene FM (2004) Divergence in wing morphology among sibling species of the *Drosophila buzzatii* cluster. *J. Zool. Syst. Evol. Res.* **42,** 154-158.
18. O’Higgins P (2000) The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. *J. Anat.* **197,** 103-120.
19. Rohlf FJ (1993) Relative warp analysis and an example of its application to mosquito wings. In: *Contributions to Morphometrics* (Marcus LF, Bello E, Garcia

Valdecasas A, eds). pp. 131-159. Museu Nacional de Ciencias Naturales: Madrid.

20. Rohlf FJ (2004) TPSDIG32. Version 1.40. Department of Ecology and Evolution, State University of New York, Stony Brook, NY. Avaliable at: http://life.bio.sunysb.edu/morph/.
21. Rohlf FJ (2004) TPSRELW32. Version 1.32 Department of Ecology and Evolution, State University of New York, Stony Brook, NY. Avaliable at: http://life.bio.sunysb.edu/morph/.
22. Rolf FJ, Marcus LF (1993) A revolution in morphometrics. *Trend. Ecol. Evol*. **8,** 129-132.
23. Rohlf FJ, Slice DE (1990) Extensions of the Procrustes method for the optimal superimposition of landmarks. *Syst. Zool.* **39**, 40-59.
24. Stock SP, Kaya HK (1996) A multivariate analysis of morphometric characters of *Heterorhabditis* species (Nemata: Heterorhabditidae) and the role of morphometrics in the taxonomy of species of the genus. *J. Parasitol.* **82,** 806-813
25. Strauss RE, Bookstein FL (1982) The truss: body form reconstructions in morphometrics. *Syst. Zool*. **31,** 113-135.
26. Timm RM, Genoways HH (2004) The Florida bonneted bat, *Eumops floridanus* (Chiroptera: Molossidae): distribution, morphometrics, systematics and ecology. *J. Mammol.* **85,** 852-865.
27. Ventura J, Lopez-Fuster MJ, Salazar M, Perez-Hernandez R (2000) Morphometric analysis of some Venezuelan akodontine rodents. *Net. J. Zool.* **50,** 487-501.
28. Vilela CR (1983) A revision of the *Drosophila repleta* group (Diptera, Drosophilidae). *Rev. Bras. Entomol.* **27,** 1-114.

Received: April 8, 2006
Accepted: August 9, 2006