Revista Española de Paleontología, ISSN 0213-6937
N.º HOMENAJE AL DR. GUILLERMO COLOM, 26-30. 1995

ANALYSIS OF GEOGRAPHIC SHAPE-VARIATION IN *Vargula hilgendorfii* (OSTR., CRUST.)

Richard A. REYMENT

Department of Historical Geology and Palaeontology, Institute of Earth Sciences, University of Uppsala, Sweden.

ABSTRACT

A means of determining the geographical validity of shape patterns based on landmark observations, using geometric morphometry and multivariate statistical analysis, is illustrated by data on the bioluminescent ostracod species *Vargula hilgendorfii* (Müller) from two marine localities (Karatsu and Tateyama) in Japan. Standard multivariate statistical analysis of the shape-data shows that there is a significant difference in means between sites. The latent vectors of the covariances of the data-matrices for shape-change were tested for collinearity under the hypothesis that external causes might influence covariances and hence the statistical distance between samples. There is a significant difference in the orientation of the ellipsoids for total shape-scatter. The reason for the differentiation in the shape-phenotype (for example, expressed as polymorphism) might have a microclimatological background. In addition to the difference in average shape, there is also a divergence in size between the samples from the two sites.

Keywords: Morphometrics, ostracods, multivariate statistical analysis, Japan, shape-analysis.

RESUMEN

Se propone un método que permite contrastar la existencia de fenómenos de variación geográfica de carácter morfológico, considerando la posición relativa de puntos equivalentes, mediante el uso de la morfometría geométrica y la estadística multivariante. El método se ilustra con datos sobre el ostracodo bioluminiscente *Vargula hilgendorfii* (Müller), provenientes de dos localidades marinas (Karatsu y Tateyama) en Japón. Los análisis convencionales, usando métodos multivariantes sobre los datos morfológicos, muestran la existencia de una diferencia significativa entre las medidas de ambas poblaciones. La colinearidad de los vectores propios de las matrices de covarianzas para el cambio morfológico se contrasta bajo la hipótesis de que pudieran influir causas externas sobre las covarianzas y motivar con ello la distancia estadística entre las muestras. Se ha apreciado una diferencia significativa en la orientación de los elipsoides para el morfoespacio global. La razón de las diferencias fenotípicas (p. ej., expresadas por un polimorfismo) podría tener una explicación microclimática. Además de la diferencia en formas medias, se constata una divergencia en el tamaño de ambas poblaciones.

Palabras clave: Morfometría, ostracodos, análisis multivariante estadístico, Japón, análisis de formas.

INTRODUCTION

The recent development of the theory and practice of the analysis of variation in shape by geometric methods (Bookstein, 1989, 1991) has opened up new avenues of research along the lines envisaged by Thompson (1917), the possibilities of which were almost unsuspected until just a few years ago. One of these prospects lies with utilizing such multivariate adaptations for studying evolution in the phenotype as are appearing in evolutionary biology as developed by Falconer (1960) and Lande (1976, 1979); others are concerned with the integration of geometric methods into biometry. The topic considered in the present note concerns the assessment of stability in the shape phenotype observed at two or more localities. One reason for wishing to do this is that regional validity of the phenotype in multivariate quantitative genetic analyses (Lande, 1979) is a mandatory requirement of the methods.

There are two factors to be accounted for. Firstly, morphological differences may be expressed in terms of a difference in means and, secondly, and subordinately, as a test for differences in covariances. The rationale for the first component is reasonably obvious in that it is of interest to establish differences in average shape between sites. The second component is of concern if it is suspected that factors such as persistent microclimatic differences, salinity, redox potential, dissolved gases, temperature, pollution etc. could be affecting the shape of the carapace. The influence of salinity on the shape of the ostracod carapace is well known. Moreover, polymorphism in shape may be induced by environmental cues (Clark, 1976).

THE MATRIX OF SHAPE-CHANGE

Bookstein's (1991) geometric morphometry analyzes within-populational and between-populational shape-change, based on landmark data. Landmarks are observations made at diagnostically important sites on an organism at corresponding points identified by their coordinate-pairs in a Cartesian system. Such geometric variables contain much essential information on size and shape of the organism, but, of course, not all.

The sequence of steps employed here for the determination of regional validity of the shape-phenotype is.

1. Extract a suitable multivariate geometric expression of shape from a data matrix of landmark observations. This can be done conveniently on the pooled sample made up of the observations available from all sites, in the present example, two.
2. Compute the principal components for the samples from each of the sites, using a matrix of weights, as defined in Rohlf (1993), and the generalized statistical distance between the samples, using the weight matrix formed from the pooled observations.
3. Test the principal axes so obtained for collinearity by the method of Anderson (1963) and Reyment (1969).

Shape variation encompasses two components, an affine part and a non-affine part (Bookstein, 1991). The mathematical term *affine* means uniform change such as obtains when a square is deformed into a parallelogram or a circle into an ellipse. In affine change, orthogonality of principal axes is preserved and parallel lines remain parallel. Mathematically, these are maps, i.e. the projection of a point from one surface to a second surface. The term *non-affine* refers to the non-

https://doi.org/10.7203/sjp.24214

uniform pattern of deformation produced, for example, when an initially flat object is twisted or warped. It is the residual of size-free change that remains after the difference due to any affine change has been subtracted from the total change in shape.

The analysis of within-populational morphometric variation based on landmark data proceeds by a method which consists of fitting an interpolating function to the set of *x-y*-coordinates of the landmarks located on each of *N* specimens in a sample. The variation among the specimens is described in terms of variance in the parameters of the fitted functions in relation to a *reference configuration*, which is often taken to be the mean configuration of landmarks after some suitable alignment of specimens. The relative warps are the principal components in a multivariate space in which, according to Bookstein's (1989) original presentation, each point corresponds to a specimen and the axes are the inversely weighted principal warps of the bending energy matrix which is defined by the reference configuration of landmarks. The relative warps depict the main features of variation in shape among specimens as deformations (i.e. warps) in shape. The relative warps can be given an appealing graphical presentation by the thin-plate splining technique (Rohlf, 1993). For details of the algorithms for computing relative warps, I refer to Bookstein (1991; section 7.6.2) and Rohlf (1993). The essential output from these computations for present purposes is a new data matrix, the weight-matrix. The calculations for finding relative warps are rather like canonical variate analysis with the role of the pooled within-groups covariance matrix being assumed by the bending energy matrix.

The elements of the *weight matrix* represent each specimen as a linear combination of principal warps, the latent vectors computed from the bending energy matrix (Bookstein, 1989). This weight-matrix, suitably partitioned, can be used in the same manner as any data matrix for many methods of multivariate statistical analysis (Reyment and Jöreskog, 1993). For many purposes, two weight-matrices are of interest. One of these is obtained for the non-uniform case, there being 3 principal warps less than the number of landmarks. The second weight-matrix, expanded so as to encompass uniform variation, is augmented by two columns (Rohlf, 1993). This latter array is used in the present analysis.

TESTING FOR REGIONAL STABILITY IN COVARIANCES

Blackith and Reyment (1971) demonstrated the empirical value of the angles between latent vectors in morphometric work. This procedure is, however, useful only in an advisory capacity, since it cannot be connected to a statistical test. Interest in collinearity of the principal axes of multivariate dispersions lies with the possibility offered for probing agreement in patterns of variation at geographical sites. The latent roots and vectors computed from the covariance matrix formed from the weight matrix can be used for assessing regional stability of the shape phenotype (cf. Lande, 1979) by the following procedure.

Let $\mathbf{S}_1$ denote the covariance matrix of the weight matrix of one of two samples (usually taken as that with the larger sample size) and $\mathbf{S}_2$ that of the other. The first of these matrices is designated as being the *reference matrix*. Obtain the latent roots ψ_i of matrix $\mathbf{S}_1$ and latent vectors $\mathbf{B}$ of the second matrix, $\mathbf{S}_2$. Anderson's (1963) large-sample test for collinearity of latent vectors adapted by Reyment (1969) for application to biometrical problems is then applied to the data. The property of collinearity of latent vectors implies that the covariance matrices are equal with respect to the orientation of dispersions. The test is carried out by computing the approximate chi- square relationship

$$(N_1-1)[\psi_i \mathbf{b}^T_i \mathbf{S}_1^{-1} \mathbf{b}_i + \psi_i^{-1} \mathbf{b}^T_i \mathbf{S}_1 \mathbf{b}_i - 2] \qquad (1)$$

where N_1 is the size of the first matrix and i can be given any value from 1 to k, where k denotes the number of warps represented in the weight matrix. The relationship (1) is approximately distributed as χ^2 with k-1 degrees of freedom. The test is performed separately for as many of the latent vectors as are of interest.

The procedure relies rather strongly on the input matrices being "well behaved", by which is meant that the datamatrices should not deviate too markedly from multivariate Gaussian. It is therefore advisable to scan the data for atypical and influential observations by some such procedure as cross-validational principal component analysis (Krzanowski, 1987; Reyment and Jöreskog, 1993), in addition to a preliminary screening by univariate statistical methods. This enables the analyst to isolate specimens that diverge significantly from their fellows. Such specimens become then candidates for deletion from the analysis, but should, of course, be subjected to subsequent biological scrutiny.

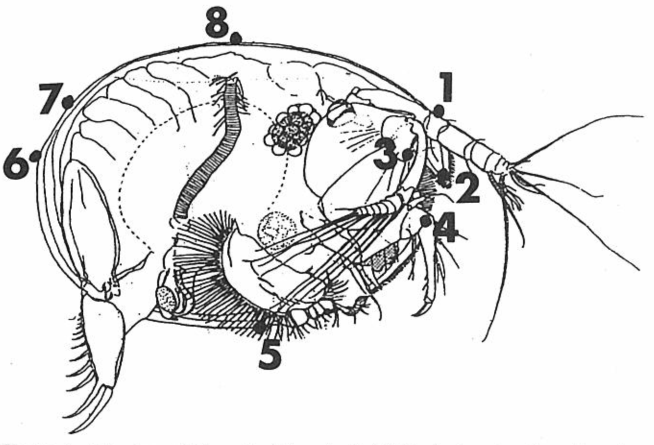

Figure 1. Drawing of *Vargula hilgendorfii* (Müller) showing the sites chosen as landmarks for analysis. The rostrum is defined as the area encompassed by landmarks 1 through 4.

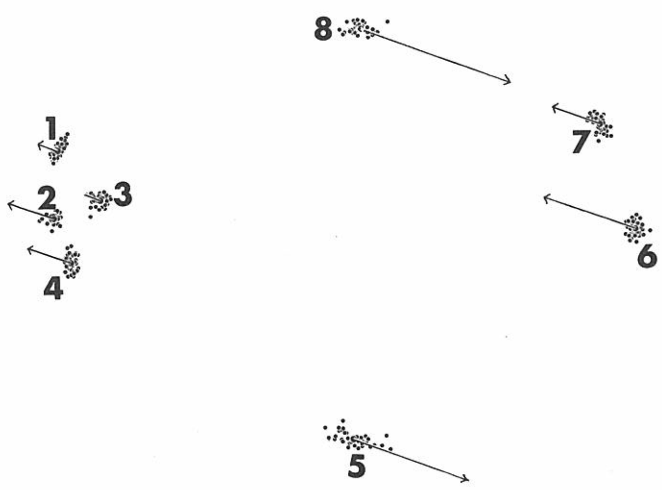

Figure 2. The graphical representation for the first relative warp loadings using $\alpha = 1$. The magnitudes of the arrows denote the relative importance of the loading apportioned to each landmark in this projection. The scatter of the 44 valves superimposed on the reference configuration (based on shape-coordinates) is shown.

EXEMPLIFICATION OF THE METHOD

THE DATA

In an ongoing study of the morphometry of the bioluminescent marine myodocopid ostracod species *Vargula hilgendorfii* (Müller), being conducted with Katsumi Abe, Department of Geosciences, University of Shizuoka, Japan, a very complex regional morphological situation has been uncovered. Data obtained from two of the sites under study are used here to exemplify the method proposed in this note, namely, a sample taken at Karatsu ($N = 24$), western Honshu (130°,00’: 33°,26’) and Tateyama ($N = 20$), eastern Tokyo Bay (139°,5’: 34°,59’).

The landmarks digitized on the specimens are illustrated in Fig. 1. There are four sites (1-4) located around the rostrum at sharp points of inflection, two at the sites of maximum dorsal and ventral curvature (5 and 8), one at the posterior extremity (7) and one at the posteroventral change of inflexion in curvature (8). The required baseline for the computations was taken between landmarks 2 and 6. Bookstein's (1991) method of shape coordinates which is a procedure for aligning the specimens (Bookstein, 1986) was then used.

FINDINGS

Analysis of shape variation

The relative warps for the pooled set of observations ($N=44$) on females were computed for the exponent $\alpha=1$ (Rohlf, 1993). There are $8 - 3 = 5$ principal warps with latent roots > 0. The first relative warp (Fig. 2) indicates compression between sites 6, 7 and 8 and an anterior displacement of the four rostral landmarks (1-4). The close clustering of the sample points at each landmark is worth noting.

In Fig. 3, the first relative warp is illustrated as a thin-plate spline for a positive displacement. This mode of representation portrays what a specimen would look like if its relative warp score were at an extreme position along the first axis and nought on all other axes. This deformation echoes the information supplied by Fig. 2 and touches mainly on landmarks 1-4 and landmarks 6 and 7. This warp is mainly a function of a single principal warp, number five, with coefficients -0.918 on the x- and 0.380 on the y- coordinates. This is the largest scale principal warp.

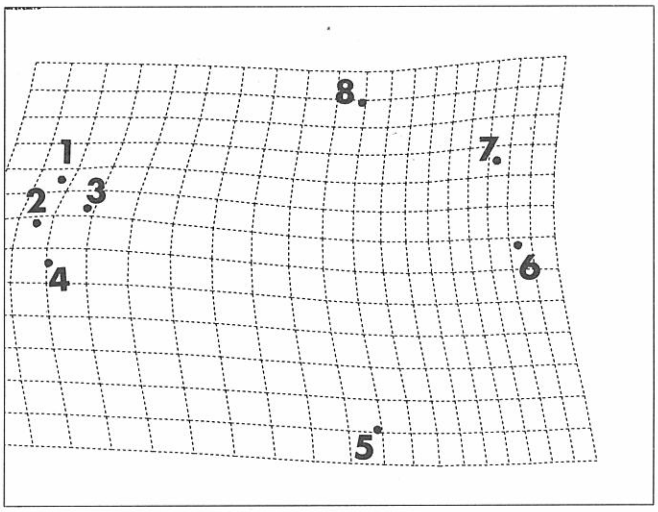

Figure 3. The first relative warp depicted as a thin-plate spline for a positive displacement. This, and the other thin-plate spline portrayals of deformation, are to be viewed as though they were in three dimensions.

The second relative warp depicted as a thin-plate spline for a positive displacement (Fig. 4) indicates compression in the rostral and posterior zones and dilation over the region lying between the dorsal and ventral margins (valve-height). The fifth principal warp is, again, the major function of this relative warp, with coefficients (-0.358, -0.834).

No significant graphical differences were obtained when the computations were repeated for $\alpha = 0$, which is not unexpected considering the nature and locations of the landmarks (see Rohlf, 1993 for a discussion concerning the effects of the exponent α).

In summary, the indications provided by the relative warps seem to be that there is a large-scale differentiation in shape between the specimens from the two localities. That is, the changes of morphometric importance encompass the entire valve. This aspect was further studied by means of the average configurations for each sample (Bookstein, 1989).

The thin-plate spline between these reference configurations was computed. The graph for uniform change (Fig. 5) shows slight shearing. The total non-uniform change displayed as a splined surface (Fig. 6) indicates slight compression in the rostral area and gentle dilation over the valve-height. There are, therefore, no starkly expressed morphological changes between localities, but differences do occur. It now remains to probe the statistical implications of these.

MULTIVARIATE ANALYSIS OF THE WEIGHT MATRIX

Univariate normality of the warp variables was checked by the usual methods by determining skewness and kurtosis. All of the new variables, for both samples are non-kurtosic. All of the variables of the reference set (Karatsu) are, likewise, conformable with the normal distribution with respect to kurtosis. The second warp-variable of the set for S_2 (Tateyama) shows significant skewness at the 5 % level of significance (*skewness* $= 1.0186$; $t_{\text{skew}} = 1.99$).

The total weight matrix for affine plus non-affine shape changes was analyzed by standard methods of multivariate statistics. A value of $D^2 = 4.43$ was found which for $F_{12,31} = 2.97$ and $P = 0.007$. This result suggests that there is a

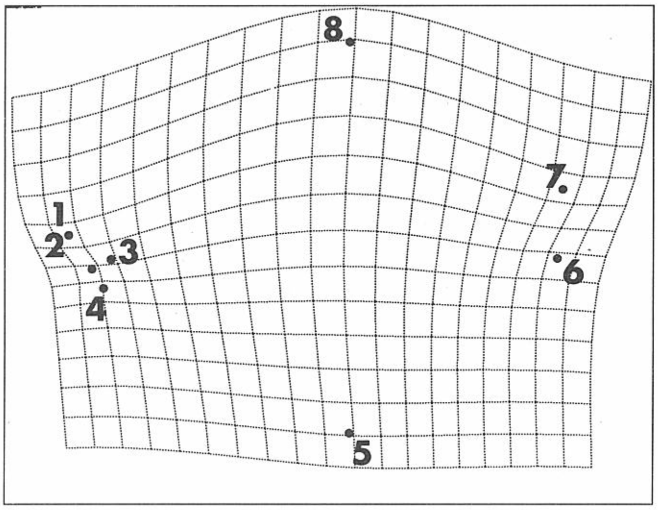

Figure 4. The second relative warp depicted as a thin-plate spline for a positive displacement.

genuine difference in "shape-means" between the sample from Karatsu and that from Tateyama. The discriminant function analysis yielded the result that 12.5 % of the Karatsu individuals are identified as being from Tateyama and 15 % of the Tateyama individuals qualify for admission into the Karatsu sample. There is therefore only slight overlap between the distributions for the shape-variates and hence a remarkable degree of differentiation in shape between the two samples.

The latent roots and vectors of the covariance matrix for $\mathbf{S}_1$ and $\mathbf{S}_2$ were computed. Inspection of the corresponding vector components indicates pronounced differences to exist. Application of formula (1) to these data yields χ^2_{11}-values of 41.38 for the first principal axis and 118.3 for the second principal axis. These are highly significant. It can be interesting to compare the results obtained by equation (1) with the angles computed between corresponding pairs of latent vectors. The angle for the first principal axes is 8.46 radians and that for the second pair, 13.72 radians. These vectors differ in their orientations, but not very strongly. The comparison of the inflations of the scatter ellipsoids gave $\chi^2_{78} = 88.9$, which is not significant.

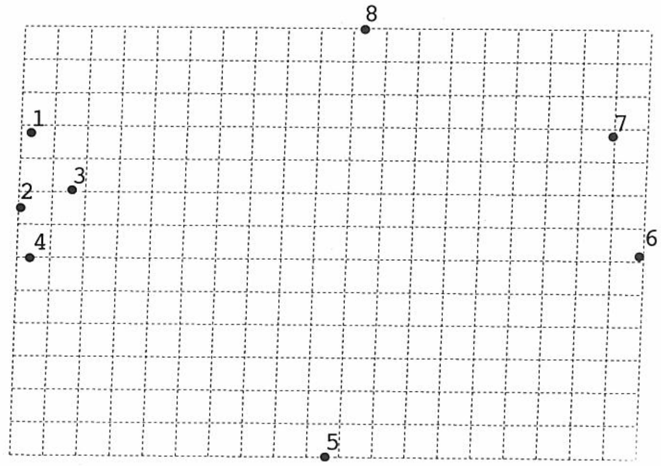

Figure 5. The affine transformation between average configurations for the samples from Karatsu and Tateyama. This is the uniform component of the deformation of the geographically separated *Vargula*.

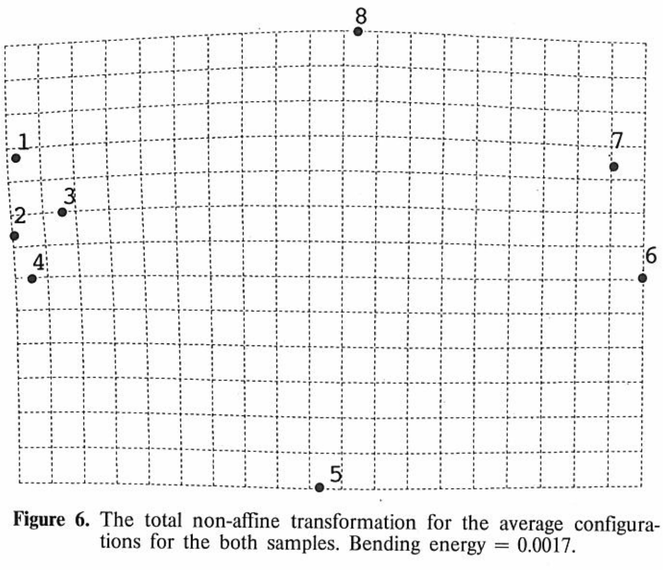

Figure 6. The total non-affine transformation for the average configurations for the both samples. Bending energy = 0.0017.

CONCLUDING REMARKS

The successful application of geometric morphology to the study of biogeographical differentiation requires a useful means of assessing the regional stability of shape variability. The methods employed here are built on multivariate statistical theory that has been used repeatedly in other connexions with a similar underlying model. The example chosen to exemplify the method is drawn from a project currently under way on the regional variability of the ostracod species *Vargula hilgendorfii*. This analysis shows how unsuspected geographic differences in morphology can be successfully brought to light by a combination of geometric morphometric methods and multivariate statistical hypothesis testing. The biological causes underlying the differences in shape could be microclimatological, granted that there is a distance of 900 km between Karatsu and Tateyama. The former site is located on the west coast of Honshu and the latter on its eastern coast, within the confines of Tokyo Bay, both with distinct sedimentological, oceanographical and meteorological regimes with the added possibility of environmentally triggered polymorphism in shape (Clark, 1976). Finally, it should be mentioned that in addition to the shape divergence between samples demonstrated in the foregoing, there is a strongly manifested difference in size. This is illustrated in Fig. 7, which is the plot of centroid size values against the first relative warp. The differentiation between samples occurs along the size-axis.

BIBLIOGRAPHY

Anderson, T. W. 1963. Asymptotic theory for principal components analysis. *Annals of mathematical Statist*, **34**, 122-148.

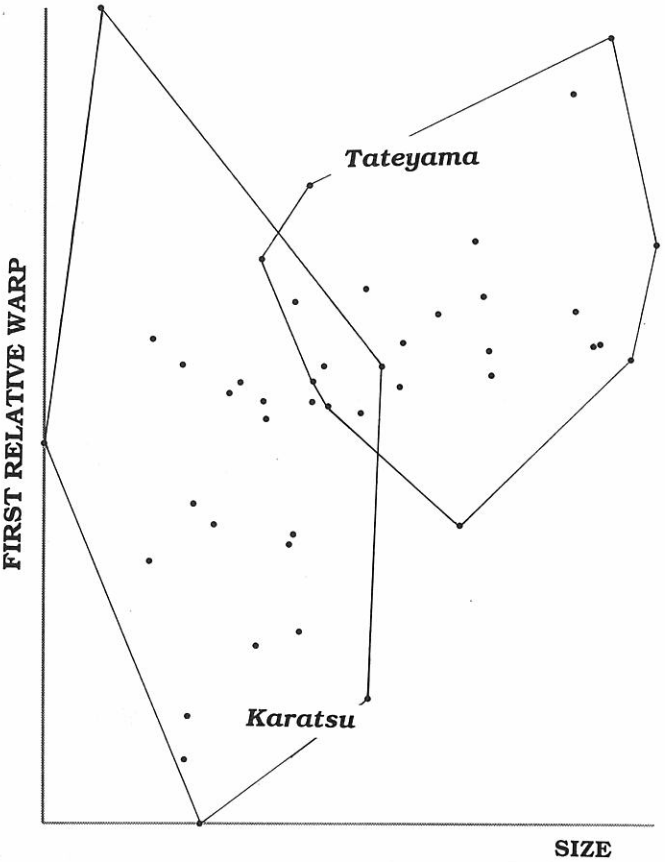

Figure 7. Plot of the scores for the first relative warp against centroid size. The points for each sample are enclosed within their respective convex hulls.

Blackith, R. E. and Reyment, R. A. 1971. *Multivariate Morphometrics*, Academic Press, London.

Bookstein, F. L. 1986. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, **1**, 181-242.

Bookstein, F. L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 567-585.

Bookstein, F. L. 1991. *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press, New York, 435 pp.

Clark, W. C. 1976. The environment and the genotype in polymorphism. *Zoological Journal of the Linnaean Society*, **58**, 255-262.

Falconer, D. S. 1960. *Introduction to Quantitative Genetics*. Oliver and Boyd, 365 pp.

Krzanowski, W. J. 1987. Cross-validation in principal component analysis. *Biometrics*, **43**, 575-584.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution*, **30**, 314-334.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution*, **33**, 402-416.

Reyment, R. A. 1969. A multivariate palaeontological growth problem. *Biometrics*, **22**, 1-8.

Reyment, R. A. and Jöreskog, K. G. 1993. *Applied Factor Analysis in the Natural Sciences*. Cambridge University Press, New York. With an Appendix by Leslie F. Marcus.

Rohlf, F. J. 1993. Relative warp analysis and an example of its application to mosquito wings. *Contributions to Morphometrics. Monografías, Museo Nacional de Ciencias Naturales, Madrid* (Eds. L. F. Marcus, E. Bello and A. García Valdecasas), 256 pp.

Thompson, D. W. 1917. *On Growth and Form*. Cambridge University Press.

Manuscrito recibido: 1 de junio, 1994
Manuscrito aceptado: 9 de noviembre, 1994

* * *

REVISTA ESPAÑOLA DE PALEONTOLOGÍA

VOLUMEN 9
Número 1 **1994**

J. VAN DER MADE. Souidea from the Lower Miocene of Cetina de Aragón (Spain).

H. G. HERBIG. Remark on the late Visean crinoids from the Cantabrian Mountains and Menorca (NW Spain and Balearic Islands).

F. SOTO, I. MÉNDEZ BEDIA y E. FERNÁNDEZ MARTÍNEZ. Construcciones arrecifales del Devónico de la Cordillera Cantábrica (NO de España).

J. GAUDANT, J. LOUISEAU et Ph. OTT d'ESTEVOU. Decouverte d'une frayère fossile de poissons téléostéens dans le Messinien des environs de Campos del Río (province de Murcia, Espagne).

M. T. ANTUNES, B. AZANZA, J. MORALES y D. SORIA. Los rumiantes (Artiodactyla, Mammalia) del Mioceno Inferior del área de Lisboa (Cuenca del Tajo, Portugal).

K. El HAJJAJI, S. POUYET et K. El KADIRI. Biofaciès à Bryozoaires et environnements sédimentaires au Miocène Supérieur du Maroc nord-oriental.

T. M. SÁNCHEZ y C. BABIN. Los géneros *Redonia* y *Catamarcaia* (Mollusca, Bivalvia) de la Formación Suri (Ordovícico temprano, Oeste de Argentina) y su interés paleobiogeográfico.

M. C. MOYA, S. MALANCA, J. A. MONTEROS y A. CUERDA. Bioestratigrafía del Ordovícico Inferior en la Cordillera Oriental Argentina basada en graptolitos.

O. BORDONARO and E. LIÑÁN. Some Middle Cambrian agnostoids from the Precordillera Argentina.

Réplicas de artículos

C. MARTÍN ESCORZA. Comentarios al artículo: *Dimensión Fractal de las suturas del cráneo de Orce*, de Gibert y Palmqvist (1992).

P. PALMQVIST y J. GIBERT. Fractales y suturas craneanas: Nota de Contrarréplica a Martín Escorza.

VOLUMEN 9
Número 2 **1994**

S. H. BASHA and M. S. AQRABAWI, Bajocian Ammonites from Zerqa Wadi, Jordan.

P. PALMQVIST, J. GIBERT y J. A. PÉREZ CLAROS. Morfometría y Paleontología humana: carencias y posibles perspectivas.

M. L. AGUIRRE. Revisión de las especies de *Mactra* Linné (Bivalvia) del Cuaternario bonaerense (Argentina). Aspectos ecológicos y de distribución.

M. R. PALACIOS-FEST, A. S. COHEN y P. ANADÓN. Use of ostracodes as paleoenvironmental tools in the interpretation of ancient lacustrine records.

J. MUÑOZ. Un nuevo Thecidellininae (Brachiopoda) monoseptal del Santoniense de la Serra del Montsec (Surpirineo Central, NE España).

X. MARTÍNEZ-DELCLÒS y A. NEL. Los Gomphidae (Insecta, Odonata, Anisoptera) del Cretácico Inferior de España.

E. MAYORAL, J. C. GUTIÉRREZ MARCO y J. MARTINELL. Primeras evidencias de briozoos perforantes (Ctenostomata) en braquiópodos ordovícicos de los Montes de Toledo (Zona Centro-ibérica Meridional, España).

R. HERRMANN and B. HUBMANN. Devonian udoteacean green algae from the Cantabrian Mountains (Santa Lucía Formation), NW-Spain.

S. GESSA, M. TRUYOLS-MASSONI and M. ROBARDET. Quantitative analysis of *Homoctenowakia bohemica bohemica* (Tentaculitoids) from the Lochkovian of the Valle Syncline, Ossa-Morena Zone (SW Spain).

J. PANTOJA-ALOR, R SCHROEDER, A. CHERCHI, G ALENCASTER and J. M. PONS. Fossil assemblages, mainly foraminifers and rudists, from the Early Aptian of Southwestern México. Palaeobiogeographical consequences for the Caribbean Region.

Seminarios

B. Martínez Navarro, J. Gibert Clols, P. Palmqvist y J. L. Santamaría. Presencia de una asociación de cráneos de grandes mamíferos en Venta Micena.