

# Geometric Morphometric Study of the Regional Variation of Modern Human Craniofacial Form

Robin J. Hennessy<sup>1</sup> and Chris B. Stringer<sup>2\*</sup>

<sup>1</sup>*Department of Clinical Pharmacology, Royal College of Surgeons in Ireland, Dublin 2, Ireland*

<sup>2</sup>*Department of Palaeontology, Natural History Museum, London SW7 5BD, UK*

**KEY WORDS** craniometry; principal components; regression

**ABSTRACT** The regional variability of the modern human craniofacial form is of importance to debates about human origins. The study of craniofacial form has generally been carried out either by interlandmark distance measurement and analysis or by observation and character scoring. In this study of four modern human groups (Eskimo/Inuit, African, Australian, and Romano-British), nine craniofacial landmark coordinates were recorded by extraction from laser scans. The coordinates were studied by geometric morphometrics, and a regression analysis

was used to investigate the dominant variability in shape within and between groups. Statistical tests of shape difference between groups were carried out. By these methods, the statistical patterns of shape variability and their geometric interpretations were studied on a common basis. The results were found to be in agreement with the classic studies of Howells ([1989:189] *Pap Peabody Mus 79*), and show the potential of this approach for future research. *Am J Phys Anthropol* 117:37–48, 2002.

© 2002 Wiley-Liss, Inc.

This paper seeks to address an old problem, the characterization of the regional variability of modern human craniofacial shape, with a new praxis, geometric morphometrics. In this paper we both describe and quantify the patterns of facial shape variation, using a methodology in which the numerical and visualization aspects share a common basis. This is possible because geometric morphometrics allows the results of the statistical analysis of shape to be visualized as deformations of shapes (typically the mean shape). Since this procedure provides an unambiguous geometrical interpretation of the statistical analysis, it counters a recent criticism of traditional anthropometric analyses, in that they supposedly do not correspond to what an experienced morphologist sees (e.g., Wolpoff and Caspari, 1997).

Since geometric morphometrics (Bookstein, 1997; Rohlf and Marcus, 1993; Rohlf, 1998; Richtsmeier et al., 1992; O'Higgins, 2000) is relatively new, at least in its applications to anthropological questions, we briefly summarize its features and contrast them with traditional morphometrics. Geometric morphometrics distinguishes the form of an object (shape plus size) from the shape (form with scale removed) by scaling to unit size. In traditional multivariate morphometrics, the form of a biological object is typically recorded as a set of measurements of distance and, possibly, angle. In an ordination study, the investigator might use principal components analysis (PCA) and attempt to relate the first few principal components to the principal form and shape variation across the sample. Geometric mor-

phometrics, by contrast, links the set of measurements with the shape of the object. The form of the object is recorded as the coordinates of defining features, i.e., landmarks. Henceforth, the set of scaled landmark coordinates, though a severe abstraction from the complex shape of the real object, is regarded as approximating its shape. All analyses are performed on the suitably transformed landmark coordinates, with the geometric relations preserved.

The landmark coordinates for a set of objects are typically transformed into points in the shape space of Kendall (1984), via scaling and alignment procedures known as generalized Procrustes analysis (Goodall, 1991; Rohlf, 1999a). For each object, the transformed coordinates represent a single point in shape space. The dimensionality, for three-dimensional (3D) data, is  $3k - 7$ , where  $k$  is the number of landmarks. This shape space has well-understood properties (Dryden and Mardia, 1998; Le and Kendall, 1993; Rohlf 1999a). It is possible to define a mean shape, the Procrustes mean, and the set of points cluster around the mean in shape space. In this study, each point is the set of transformed landmarks of a single cranium. For tight distributions about the Procrustes mean, as are likely in biological studies, the points can be projected onto the linear tangent space at the Procrustes mean, with minimal

\*Correspondence to: Professor Chris Stringer, Department of Palaeontology, the Natural History Museum, Cromwell Road, London SW7 5BD, UK. E-mail: C.Stringer@nhm.ac.uk

Received 15 February 1999; accepted 2 August 2001.

distortion of their configuration. These projected points, the Procrustes tangent coordinates, are adequately approximated by the Procrustes residuals in most biological studies (O'Higgins and Jones, 1998).

Geometric morphometrics also provides an important measurement of shape difference called the Procrustes distance. If two sets of landmarks are scaled and aligned by least-squares optimization, the sum of squared distances between equivalent landmarks is a good approximation to the squared Procrustes distance.

Multivariate analysis, such as PCA, of the Procrustes tangent coordinates, which are described above, can be carried out to investigate the main shape variations. Two groups of shapes can be compared by finding the combination of PCs which best distinguish them. This combination can be computed by multiple regression of PC scores on group membership (Penin and Baylac, 1999). Sufficient PCs can be included in the regression model to account for the majority, say 90%, of the total variance.

The geometric meaning of the shape variation described by principal components can be visualized, since the landmark configuration of a hypothetical specimen can be displayed by computer graphics. The shape of a hypothetical specimen positioned along a PC axis can be calculated by adding the weighted eigenvector to the Procrustes mean coordinates (O'Higgins and Jones, 1998). The weighting factor is determined by the position that the hypothetical specimen occupies along the PC axis. As specimens further along the PC axis are visualized, the coordinates of the mean configuration are progressively displaced from their initial configuration. These displacements can be represented in various ways. In two dimensions (2D) they are conveniently depicted by thin plate splines (O'Higgins and Dryden, 1993), and in 3D by vectors (Slice, 1996), warping a wireframe or rendered representation (Penin and Baylac, 1999) or transformation grids (O'Higgins and Jones, 1998). The set of PCs from a regression model can be similarly visualized by weighting the contribution of each eigenvector to the displacement by the appropriate regression coefficient (Penin and Baylac, 1999).

An alternative approach to the statistical analysis of landmark coordinates, Euclidean distance matrix analysis (EDMA), has been developed. This method avoids registration and analyzes the interlandmark distances derived from the coordinates (Lele and Richtsmeier, 1991). EDMA, however, does not allow the direct visualization of results of the statistical analysis, and recent work has criticized it for its lack of statistical power (Rohlf, 2000) and its potential to produce misleading ordinations (Rohlf, 1999b). Its principal advantage over generalized Procrustes analysis, i.e., avoidance of the arbitrariness of choice of registration method, is expected to be relatively unimportant for the current study, since reasonable registration methods should give similar biological

results (O'Higgins, 2000). For these reasons, we employ a registration-based technique in this work.

In summary, the principal advantages of the geometric morphometric praxis over the traditional approach are that it provides a shape space in which geometry is preserved and which can be interpreted statistically. In applying this approach to primate craniometry, researchers have used a variety of methods to generate shape coordinates, perform statistical analyses, and represent the results graphically. Examples of such studies are, in 2D, sexual dimorphism of African and Romano-British groups (Lynch et al., 1996); comparison of African, Australian, Chinese, and Australian human groups (Lynch et al., 1996); characterization of a Neanderthal cranium (Yaroch, 1996); and sexual dimorphism in *Pan*, *Gorilla*, and *Pongo* (O'Higgins and Dryden, 1993). 3D studies include comparisons of *Pan* and *Pongo* (Penin and Baylac, 1999); shape difference between two American human groups (Ross et al., 1999); growth of *Cercocebus torquatus* (O'Higgins and Jones, 1998); and human craniofacial growth (Hennessy et al., 1997).

In this paper, we seek to demonstrate that the principal features of craniofacial variability among four modern human regional groups can be both identified, and visualized objectively, supported by statistical analysis. We also test for sexual dimorphism within one group. To allow comparison with the results of traditional studies, we located and used standard craniofacial osteometric landmarks in this study. We also note the potential for pitfalls, and document the applicability of the approach to the study of variation in human craniofacial form.

## MATERIALS AND METHODS

Only crania determined as adult from records or dental examination were analyzed, all from the skeletal collections of the Natural History Museum, London (NHM). Four regional groups were studied: 68 European crania (from the NHM Poundbury Romano-British series), 35 Australian crania (from various parts of Australia, and NHM and Oxford collections), 35 African crania (from West Africa, NHM collection) and 29 Inuit or Eskimo crania, here termed Inuit (predominantly from Canada and Greenland, NHM and Oxford collections). Some of the specimens had been sexed by a variety of methods, using cranial or postcranial features. These sex determinations for Poundbury (36 male, 23 female, and 9 unclassified) were utilized to test the possible effect of sexual dimorphism on the analyses.

### Digitization

A variety of 3D digitization technologies are now practiced in physical anthropology. Handheld digitizers are generally used to record landmarks (O'Higgins and Jones, 1998; Ferrario et al., 1997), but photogrammetry (Shaner et al., 1998), computed tomography (Valeri et al., 1998), and optical surface

(laser) scanning (Wood et al., 1998) are also employed. The skull surfaces in this study were digitized using the laser scanner developed by the Medical Graphics and Imaging Group at University College London (UCL) (Moss et al., 1989). This allows the digitization of the surface of each skull into ~40,000 3D coordinates in ~10 sec. The skulls are rotated while illuminated by a vertical stripe of light from a low-power diode laser. The profiles are captured by a CCD camera and converted into surface coordinates by calibration based on triangulation. This type of laser scanner has the advantages of speed and ease of use, and provides adequate accuracy (~1 mm) and resolution (~0.5 mm). It does not, however, capture the upper and lower extremities of objects. Fortunately, these shortcomings do not impact severely on this study. The start and end positions of each scan were ~1 cm lateral to the zygomaticotemporal suture. The angular spacing of the profiles was 0.8°. Two photographs were also taken of each skull, in one of which the landmarks were located by small (~1 mm diameter) spherical markers.

#### Landmark extraction

The 3D coordinates of nine landmarks which characterize facial shape were extracted from the laser scans. These were chosen to correspond to those used by Howells (1973). Three were midsagittal at the nasion, nasospinale, and prosthion. Three pairs were bilateral at the zygomaxillare, orbitale (defined as the most superior point of the zygomaticomaxillary suture), and frontomolare anterior (defined as the most anterior point of the zygomaticofrontal suture).

The laser scans were visualized and the landmarks were located using software written at UCL. The 3D coordinates of the landmarks extracted from the laser scans are illustrated in Figure 1. Three features of the software were used to aid in the location of landmarks. Firstly, the software allows a point of interest to be selected and moved over the surface while horizontal and vertical profiles through the point are drawn. Landmarks can be recorded at the extremes of these profiles. This was helpful in locating all landmarks, but it was particularly helpful in locating accurately midsagittal landmarks. Secondly, the object can be tilted to any viewing direction, and this facility was used to locate the frontomolare anterior landmarks while viewing the skull from the side. Thirdly, the direction of illumination can be varied. Low lighting was used to locate the nasion in cases where it was uncertain, and side lighting was used to help locate the orbitale. The latter was generally the hardest landmark to locate accurately, but careful illumination and close comparison with the photographs enabled it to be located with reasonable certainty. After the landmarks were placed, the skull was viewed from varied viewpoints to doublecheck their positioning.

## ANALYSIS

### Details of subgroups

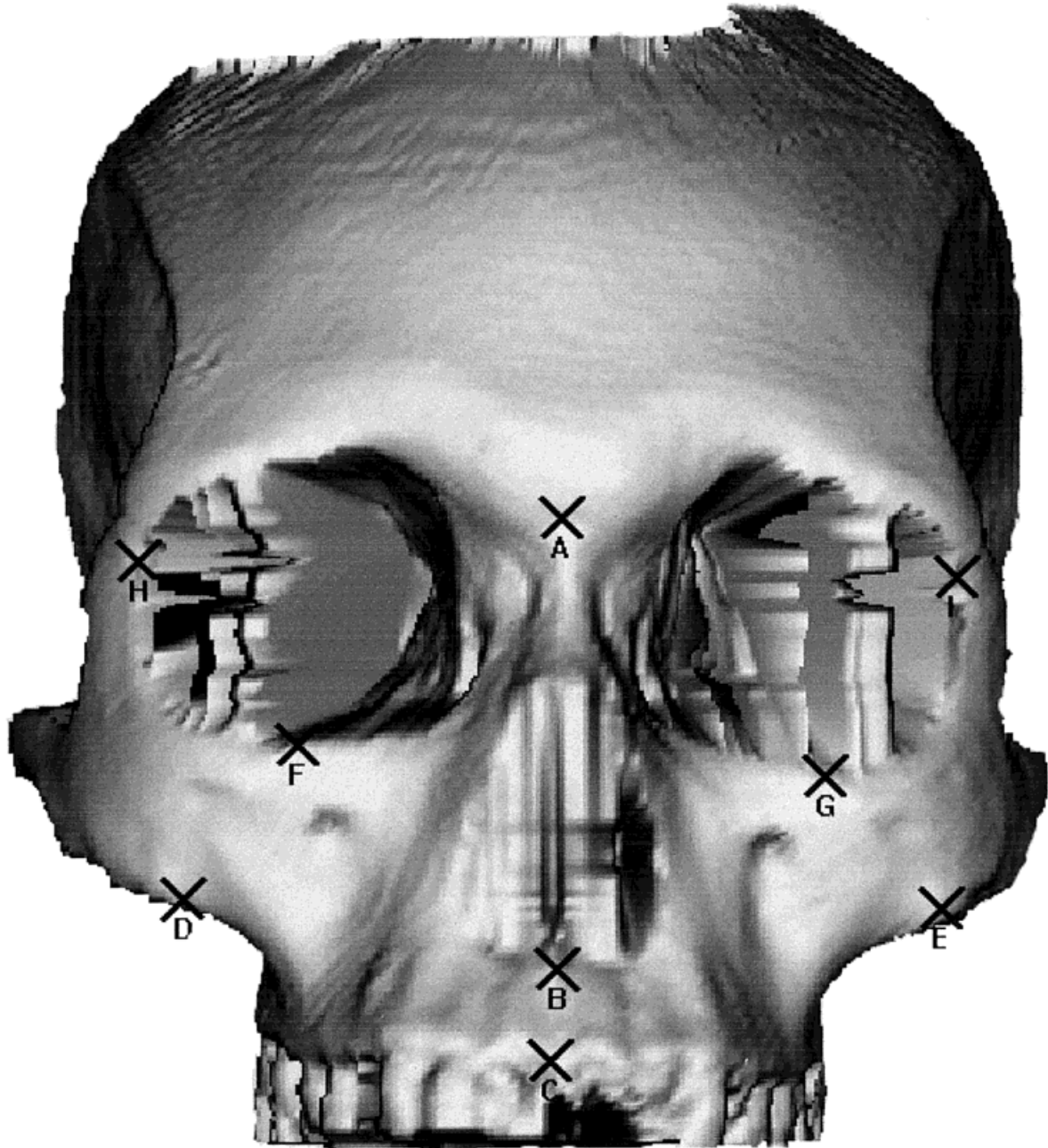
The entire dataset was divided into three subgroups, namely, the sexed crania of the Romano-British group; the six possible combinations of two regional groups; and all the crania together.

#### Specific methods for comparing two groups

Four methods, all based on geometric morphometrics, were employed to analyze the six regional group pairings and the male/female Romano-British pairing. These methods provide three tests of statistical significance of intergroup shape difference, two classifications of group membership, and a visualization of the shape difference between group pairs. The robustness of results to the choice of method can be assessed by comparing their various results. The methods are Goodall's  $F$  test, Hotelling's  $T^2$  test, linear discrimination function analysis, and multiple regression.

Goodall's  $F$  test (Goodall, 1991; Bookstein, 1997) was used to test for intergroup shape difference. This tests for overall shape difference between groups, and takes into account all the sample variance. The statistic is based on the ratio of the squared Procrustes distance between the means of each group, to the sum of the squared Procrustes distance of each specimen to its group mean. The test assumes that the landmarks are symmetrically distributed around the Procrustes mean. Observations of 3D scatterplots of these datasets indicate that this is a reasonable assumption (Bookstein, 1997). The test generates  $F$  and  $P(F)$  values. The calculations were carried out using SPlus macros written by Prof. Ian Dryden (School of Mathematical Sciences, University of Nottingham, UK).

Hotelling's  $T^2$  two-sample test of the projected Procrustes residuals was carried out as an alternative test of group difference. The raw landmark coordinates were converted to coordinates in Kendall's shape space using the GRF-ND (generalized rotational fitting of  $n$ -dimensional data) program (Slice, 1996). The generalized least-squares option was used to align the forms. The Procrustes residuals were projected onto the tangent plane at the Procrustes mean, using the orthogonal projection operation of the Numerical Taxonomy and Multivariate Analysis Program (NTSYS) (Rohlf, 1993; Rohlf and Bookstein, 1987). NTSYS was used to calculate the covariance matrix and eigenvectors of the projected points and to project the shape coordinates onto the principal components axes. Hotelling's  $T^2$  two-sample test of the projected Procrustes residuals was carried out using the LINDA (Linear Discrimination Analysis) program (Cavalcanti, 1999), which calculates the Mahalanobis distance, Hotelling's  $T^2$ , the  $F$  ratio, and  $P(F)$ . The procedure involves inverting the sample covariance matrix, and this requires the rank of the covariance matrix to be no greater than the dimensionality of the tangent



**Fig. 1.** Australian specimen with landmarks (A, nasion; B, nasospinale; C, prosthion; D, E, zygomaxillare; F, G, orbitale; H, I, frontomalare anterior).

space. Since the dimension of the tangent space is 20, only the first 20 elements of the Procrustes residuals were included in the analysis (Dryden and Mardia, 1998). The procedure was also carried out on the Procrustes residuals to investigate whether the orthogonal projection affected the results significantly.

Linear discrimination function analysis was carried out on the projected Procrustes residual coordinates to assess the utility of the method in classifying unknown specimens between each pair of groups.

Multiple regression analysis of PC scores of the Procrustes residuals was carried out. This provides

a geometric representation of the shape difference between groups, as well as an F statistic and classification statistics. The regression model is represented by the equation

$$Y = \alpha + \beta_1 PC_1 + \dots + \beta_n PC_n \quad (1)$$

where Y, the dependent variable, codes for group membership as 0 or 1,  $PC_{1..n}$  are the n PC scores of the Procrustes residuals, and  $\beta_{1..n}$  are the n regression coefficients. Sufficient PCs were included in the model to account for 90% of the total variance. The numerical analysis was carried out by the APS (Applied Procrustes Software) program (Penin and

TABLE 1. Details of  $F$  tests for paired group difference: Goodall's  $F$  test, Hotelling's  $T^2$ , and multiple regression

Paired groups	F value, degrees of freedom, $P$ value <sup>1</sup>		
	Goodall's $F$	Hotelling's $T^2$	Multiple regression
Males $\geq$ females	1.2, [20, 1,140], 0.25	1.7, [20, 38], 0.08	1.6, [12, 46], 0.12
African $\geq$ Inuit	37.6, [20, 1,240], <0.001	15.2, [20, 43], <0.001	28.6, [8, 55], <0.001
Inuit $\geq$ Australian	30.5, [20, 1,240], <0.001	13.4, [20, 43], <0.001	25.4, [9, 54], <0.001
Inuit $\geq$ Romano-British	29.0, [20, 1,900], <0.001	15.3, [20, 76], <0.001	24.7, [10, 86], <0.001
Australian $\geq$ African	6.7, [20, 1,360], <0.001	2.9, [20, 49], <0.001	5.8, [9, 60], <0.001
Australian $\geq$ Romano-British	12.4, [20, 2,020], <0.001	7.7, [20, 82], <0.001	9.2, [11, 91], <0.001
Romano-British $\geq$ African	19.2, [20, 2,020], <0.001	10.9, [20, 82], <0.001	15.2, [10, 92], <0.001

<sup>1</sup> Degrees of freedom are in brackets.

Baylac, 1999; Penin, personal communication, 2000). The numerical output from the regression model was recorded, i.e.,  $R^2$ ,  $F$ , and  $P(F)$  values and classification statistics.

The visual display of the regression model, i.e., the regression coefficients, was also provided by APS. If the mean coordinates are represented as a 27 length vector  $\mathbf{x}_{\text{mean}}$ , the coordinates of a hypothetical specimen,  $\mathbf{x}_h$ , lying along the discrimination vector are computed as

$$\mathbf{x}_h = \mathbf{x}_{\text{mean}} + c\beta_1\gamma_1 + \dots + c\beta_n\gamma_n \quad (2)$$

where  $\beta_{1..n}$  are the regression coefficients,  $\gamma_{1..n}$  are the eigenvectors, and  $c$  is a weighting factor which determines how far along the discrimination vector the hypothetical specimen lies.

APS visualizes Equation (2) by displaying the mean shape of the combined groups,  $\mathbf{x}_{\text{mean}}$ , as a 3D graphic. The hypothetical specimen,  $\mathbf{x}_h$ , is superimposed. As the weighting factor is adjusted, the landmarks of the hypothetical configuration are progressively displaced from the mean. This displacement represents the shape difference that distinguishes the two groups. These displacements were carefully observed using the 3D graphics facility, and views of the graphic display were recorded.

### Method for comparing all four groups

The four regional datasets were pooled, and a discriminant analysis was carried out to find the shape variability, within the pooled dataset, that differentiated the four groups. The Procrustes residuals and Procrustes mean were calculated, using APS, and stored. The Procrustes residuals were then subjected to PCA, using NTSYS. The eigenvectors and percentages explained by the PCs were stored to disc. The Procrustes residuals were then projected onto the eigenvectors using NTSYS, and the first 10 PC scores, which explained 91.4% of the total variance, were analyzed via discriminant analysis. This procedure generates three canonical discriminant functions. These are linear combinations of the PC scores, as in Equation (1), that best discriminate between the groups. All 10 PC scores were entered into the model. The discriminant function coefficients, eigenvalues and scores were recorded. Classification statistics were also recorded.

TABLE 2. Classification statistics for discriminant and regression analyses of paired groups

Paired groups	Percentage misclassified <sup>1</sup>			
	Discriminant analysis		Regression analysis	
	Group 1	Group 2	Group 1	Group 2
Males $\geq$ females	11.1	8.7	22.2	8.7
African $\geq$ Inuit	0.0	0.0	0.0	0.0
Inuit $\geq$ Australian	3.4	2.9	3.4	2.9
Inuit $\geq$ Romano-British	3.4	1.5	3.4	1.5
Australian $\geq$ African	14.3	14.3	8.6	17.1
Australian $\geq$ Romano-British	8.6	5.9	5.7	19.1
Romano-British $\geq$ African	0.0	5.7	13.2	2.9

<sup>1</sup> Percentage of samples misclassified by discriminant analysis and the regression analysis. Groups 1 and 2 refer to first and second groups in pairings.

The three discriminant functions, i.e., the three sets of  $\beta$  coefficients, were then visualized using a program written by the authors. This program takes as input the Procrustes mean coordinates, the eigenvectors, and the discriminant function  $\beta$  coefficients, and visualizes the discriminant coefficients using Equation (2). The procedure for visualizing the discriminant functions was the same as for visualization of the discrimination analysis of paired groups.

## RESULTS

### Comparisons of two groups

Table 1 records the  $F$ ,  $P(F)$ , and degrees of freedom for Goodall's test, Hotelling's  $T^2$  test, and the regression analysis. Table 2 records the classification statistics from the discriminant function analysis and the regression analysis. Table 3 records the  $D^2$  and  $T^2$  values. Table 4 records the number of PCs included in the regression model, the proportion of total variance input to the model, and the proportion of input variance explained by the model.

We visualize the intergroup shape difference by displacing the Procrustes mean coordinates along the discriminating vector in the direction of the second group of the pair. The displacements of landmarks depict how the second group differs from the

TABLE 3. Values of Mahalanobis distance ( $D^2$ ) and Hotelling's  $T^2$  for paired groups

Paired groups	Mahalanobis distance ( $D^2$ )	Hotelling's $T^2$
Males $\geq$ females	3.6	50.1
African $\geq$ Inuit	27.6	437.4
Inuit $\geq$ Australian	24.4	18.8
Inuit $\geq$ Romano-British	18.8	382.6
Australian $\geq$ African	4.5	79.9
Australian $\geq$ Romano-British	8.2	189.7
Romano-British $\geq$ African	11.6	268.6

TABLE 4. Details of regression models for paired groups

Paired groups	PCs <sup>1</sup>	Variance fraction <sup>2</sup>	R squared <sup>3</sup>
Males $\geq$ females	12	0.92	0.35
African $\geq$ Inuit	8	0.92	0.81
Inuit $\geq$ Australian	9	0.92	0.81
Inuit $\geq$ Romano-British	10	0.91	0.75
Australian $\geq$ African	9	0.90	0.49
Australian $\geq$ Romano-British	11	0.92	0.54
Romano-British $\geq$ African	10	0.91	0.63

<sup>1</sup> Number of PCs included in model.

<sup>2</sup> Fraction of total variance described by PCs.

<sup>3</sup>  $R^2$ , fraction of variance explained by the model.

first. These displacements are best viewed with 3D graphics, but we provide an anterior view of a wire-frame diagram of the mean and displaced landmarks, and we also describe the most significant landmark displacements.

#### **Male $\geq$ female in the Romano-British group.**

The F tests all indicate that there is no sexual dimorphism in the Romano-British group which is significant at the 0.05 level. The  $D^2$  value also indicates that the male and female groups are more similar in shape than any of the regional group pairings. Since the regression analysis did not find significant difference, its visualization is not presented.

**African  $\geq$  Inuit.** This paired group has the largest F value for each method as well as the largest  $D^2$  value, and the two classifications are exact. These observations indicate that this is the paired group which differs most in shape. The regression model, which explains 81% of the variance, is given in Figure 2a. The Inuit group differs from the African group in the following respects: the nasion is superior; the posterior, prosthion, and nasospinale are inferior and posterior; the zygomaxillare are anterior and lateral; the frontomolare anterior are medial; and the orbitale are medial and superior. The overall effect is that the Inuit face is flatter and longer, while the zygomaxillare is more laterally located.

**Inuit  $\geq$  Australian.** The F values,  $D^2$  value, and the classifications all indicate that these groups are only slightly less different in shape than the African/Inuit pair. The regression model, which explains 81% of the variance, is given in Figure 2b. The

Australian group differs from the Inuit group in the following respects: the midsagittal landmarks are anterior and the nasospinale is also superior, the zygomaxillare are posterior and medial, the frontomolare anterior are lateral, and the orbitale are lateral. The overall effect is that the Australian face is more prognathic.

**Inuit  $\geq$  Romano-British.** The F values,  $D^2$  value, and the classifications indicate that these groups are nearly as different in shape as the African/Inuit and Inuit/Australian pairings. The regression model, which explains 75% of the variance, is visualised in Figure 2c. The Romano-British group differs from the Inuit group in the following respects: the nasion and nasospinale are anterior and superior, the zygomaxillare are posterior, and the frontomolare anterior are lateral and inferior.

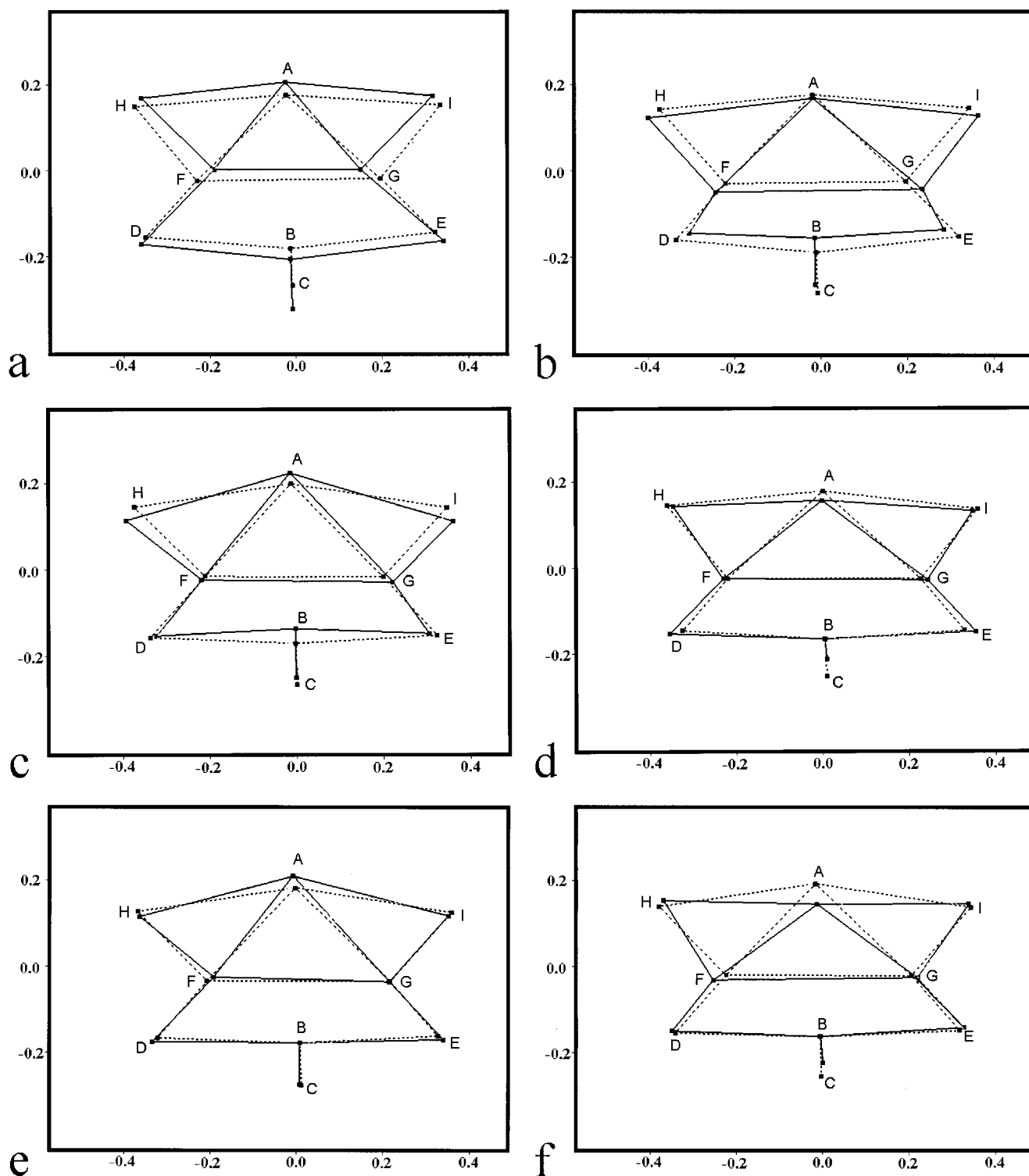
**Australian  $\geq$  African.** The F values,  $D^2$  value, and the classifications all indicate that these regional groups are most similar in shape. The regression model, which explains 49% of the variance, is given in Figure 2d. The African group differs from the Australian group in the following respects: the prosthion is markedly superior, the nasion is inferior, and the zygomaxillare are lateral. The overall difference is mostly in the lower face, particularly in the position of the prosthion.

**Australian  $\geq$  Romano-British.** The F values,  $D^2$  value, and classifications all indicate that this group pair is intermediate in shape difference between the Australian/African pair and all Inuit pairings. The regression model, which explains 54% of the variance, is given in Figure 2e. The Romano-British group differs from the Australian group in the following respects: the major differences are that the nasion is superior and the prosthion is posterior and, to a lesser extent, the zygomaxillare are lateral, the orbitale are anterior, and the frontomolare anterior are inferior and medial. This grouping is unusual in that the main differences are located at the extreme superior and inferior positions.

**Romano-British  $\geq$  African.** The F values,  $D^2$  value, and the classifications all indicate that the distinctions between this pairing are similar to those in the Australian/Romano-British pairing. The regression model, which explains 63% of the variance, is given in Figure 2f. The African group differs from the Romano-British group in the following respects: the nasion is markedly inferior, the prosthion is superior and anterior, and, to a lesser extent, the zygomaxillare are lateral. Overall, the shape difference is mostly located at the nasion and prosthion.

#### **Comparison of all samples**

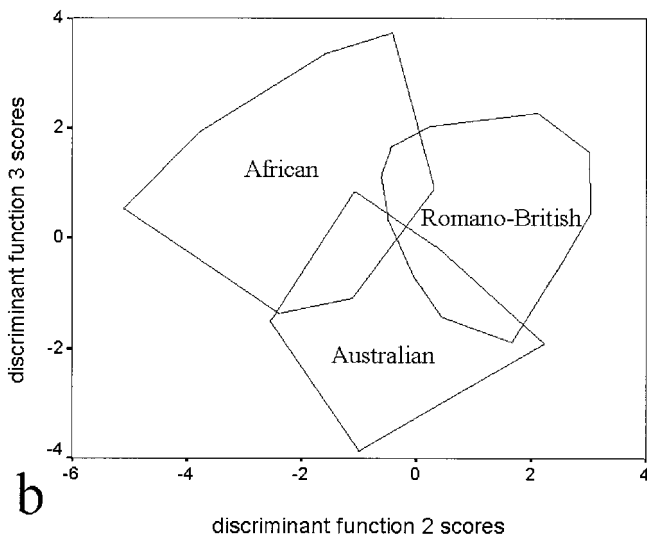
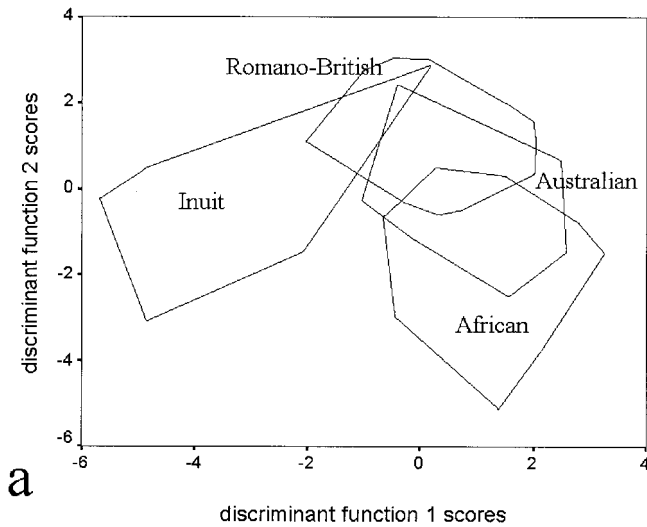
The percentages of the variance explained by the three discriminant functions are 66.9%, 24.4%, and 8.7%.



**Fig. 2.** Visual display of paired group differences. Dots represent landmarks. They are labelled (A–I) according to the scheme of Figure 1. Dotted lines connect landmarks of the mean of the group pair. Solid lines connect landmarks of the hypothetical specimen lying along the discrimination vector (see text for details). **a:** African  $\geq$  Inuit. **b:** Inuit  $\geq$  Australian. **c:** Inuit  $\geq$  Romano-British. **d:** Australian  $\geq$  African. **e:** Australian  $\geq$  Romano-British. **f:** Romano-British  $\geq$  African.

Figure 3a plots discriminant function 1 scores against discriminant function 2 scores. The plot of discriminant function 2 scores against discriminant function 3 scores for all samples revealed that the

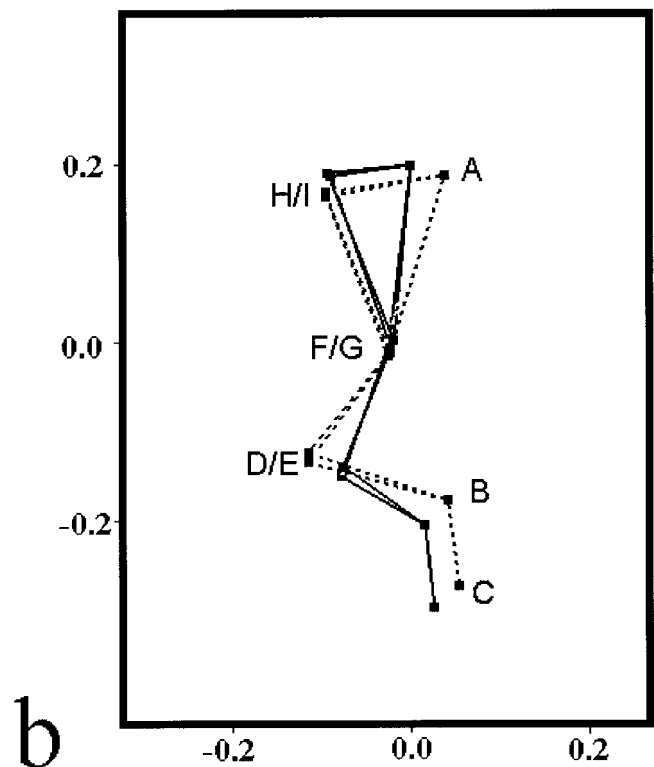
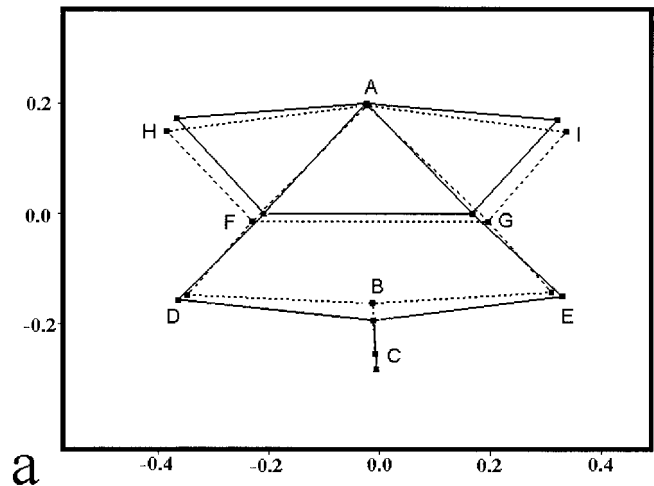
Inuit crania are evenly dispersed, i.e., these discriminant functions do not describe the differences of the Inuit from the rest. For this reason, the plot of discriminant function 2 scores against discriminant



**Fig. 3.** Discriminant function scores of all samples. To avoid cluttering, the diagram convex hulls were drawn and the enclosed points are not shown. **a:** Discriminant function 1 scores plotted against discriminant function 2 scores. **b:** Discriminant function 2 scores plotted against discriminant function 3 scores. Inuit specimens were not included (see text for details).

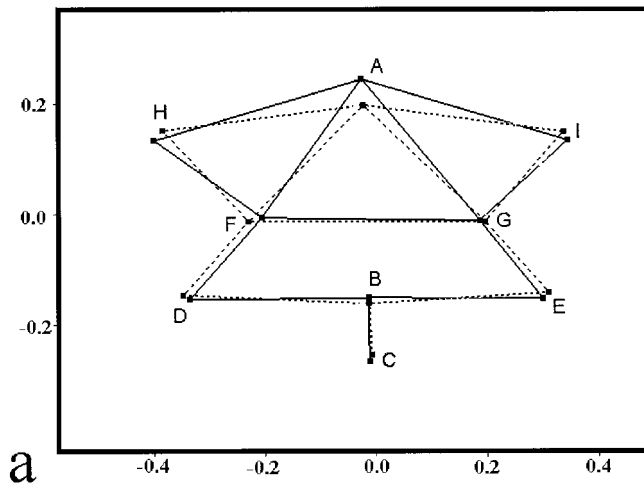
function 3 scores, in Figure 3b, does not include the Inuit crania. To simplify the diagram, convex hulls were drawn around the groups. Discriminant function 1 separates out the Inuit crania from the others. Discriminant function 2 largely separates the Romano-British from the Australians and the Africans. Discriminant function 3 largely separates the Australians from the Africans. Overall, inspection of Figure 3a,b demonstrates that the discriminant function was largely successful in separating the four groups into different regions in the discriminant function space. The percentage misclassifications for the four groups were as follows: Romano-British (23.5%), African (22.8%), Australian (22.9%), and Inuit (6.9%).

The discriminant functions are visualized in Figures 4a,b, 5a,b, and 6a,b. Figure 4a,b shows that the

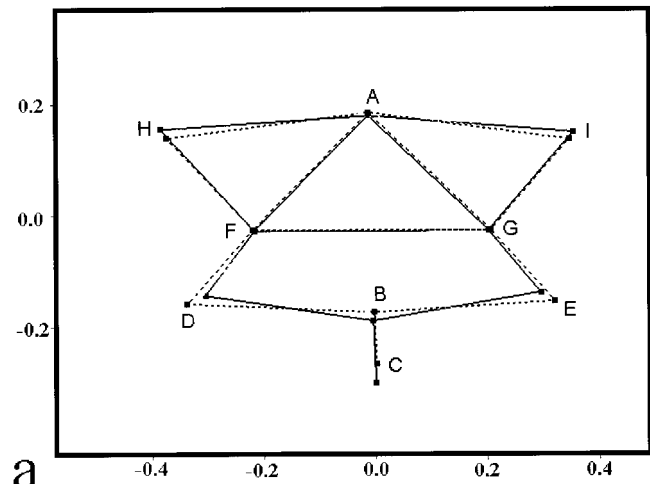


**Fig. 4.** Visual display of discriminant function 1. Dots represent landmarks. They are labelled according to the scheme (A-I) of Figure 1. Dotted lines connect landmarks of the mean of all samples. Solid lines connect landmarks of a hypothetical specimen lying along the discriminant function (see text for details). **a:** Anterior view. **b:** Lateral view.

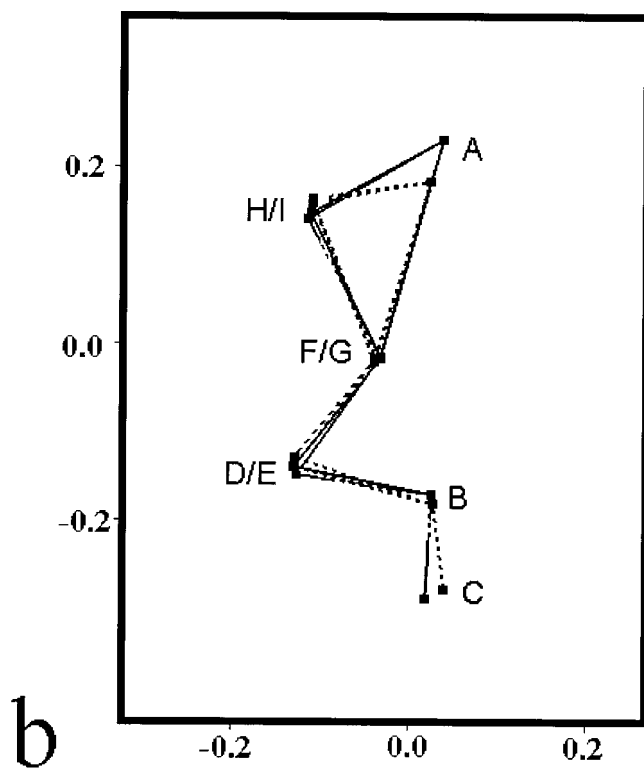
Inuit face is distinguished from the other three groups by its flatness. The flattening described by discriminant function 1 is uniform over the face. Figure 5a,b shows that the Romano-British group is largely distinguished from the Australian and African groups by the following features: the nasion is superior, the prosthion is posterior, the orbitale are medial, and the frontomale anterior are lateral. Figure 6a,b shows that the Australian group largely differs from the African group in the following re-



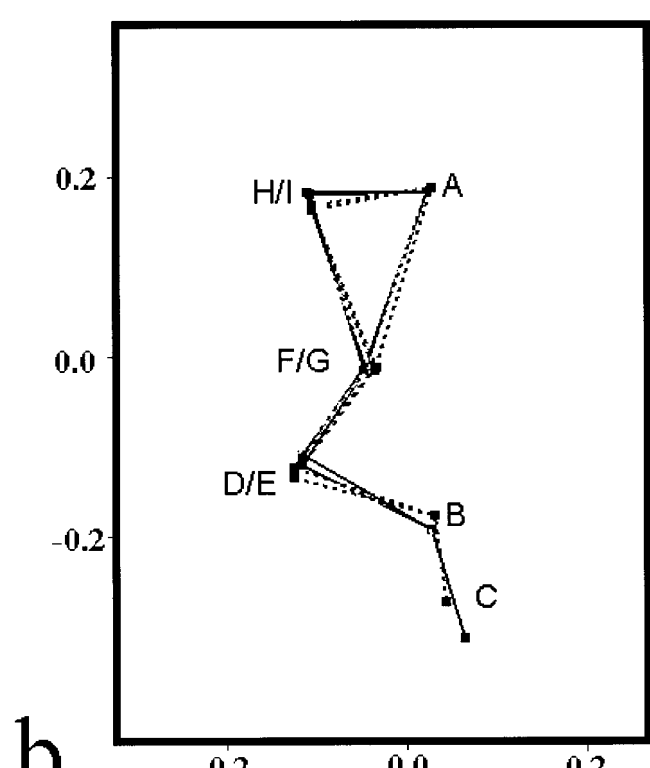
a



a



b



b

**Fig. 5.** Visual display of discriminant function 2 (see legend of Fig. 4 for details). **a:** Anterior view. **b:** Lateral view.

**Fig. 6.** Visual display of discriminant function 3 (see legend of Fig. 4 for details). **a:** Anterior view. **b:** Lateral view.

spects: the prosthion is anterior and inferior and, less significantly, the zygomaxillare are medial and inferior.

## DISCUSSION

### Paired groups

The four methods of analysis are consistent in their assessment of the extent of dimorphism of the six regional group pairings and the male/female pairing. The three sets of  $F$  values, the Mahalanobis distances, and the two sets of classification statistics all indicate that the greatest distinctiveness is shown by the Inuit group paired against any other

group, and the Romano-British group paired against the Australian and African groups. The African/Australian pairing shows markedly less difference, and the sexual dimorphism within the Romano-British group is considerably less again. There is a very clear inverse relationship between the extent of between-group differences, as measured by  $F$  values, and the goodness of fit of the regression model, as measured by  $R^2$  values.

The agreement between the four methods is least good in the case of Romano-British sexual dimorphism. While  $F$  values are all much lower than for all other pairings,  $P(F)$  values differ considerably.

The most reliable value is likely to be Goodall's *F*, because it considers all the variance (unlike the regression modelling) and, unlike the Mahalanobis distance-based statistics, requires no loss of rank. An important conclusion from this finding is that where intergroup shape difference is low, several techniques should be used and compared before conclusions are drawn.

#### **All four groups**

The pattern of shape differences between the paired groups is corroborated by the analysis of the pooled groups. This analysis shows that the major shape variance across all samples is the variation in facial flatness that distinguishes the Inuit from the rest. The distribution of the four groups in discriminant function space corroborates closely with the findings from the paired-group analyses of the extents of dimorphism in the six paired groups. The visualizations of discriminant functions are also consistent with the visualisations of regression models of the paired groups.

In summary, the four analyses of the paired groups and the discriminant function analysis of the pooled crania are consistent in their conclusions. This is true both of the numerical data and the visualizations. The symmetrical movement of the landmarks in the graphic plots indicates that the observed patterns of shape variation are not significantly distorted by the variable precision of landmark location.

#### **Comparison with other geometric morphometric studies**

The most comparable geometric morphometrics studies are of craniofacial sex differences in African and Romano-British crania (Wood and Lynch, 1996) and of differences between modern African, Australian, Romano-British, and Chinese crania (Lynch et al., 1996). This study agrees with the finding of no sexual dimorphism in Romano-British crania. There is no clear confirmation, however, of the finding that African and Australian crania were more prognathic than Romano-British crania. Our study differs in that it is in 3D, uses only facial landmarks, and employs a more sophisticated analysis technique (regression as opposed to two-point registration of mean shapes).

#### **Comparison with Howells (1973, 1989) and others**

Geometric morphometrics provides an accurate characterization of the overall face shape in each sample and of the essential differences between samples. Comparison with more traditional studies shows many similarities in characterization of facial shape for the different samples. For example, in his extensive multivariate studies of cranial variation, the conclusions of Howells (1973, 1989) regarding discrimination between European, African, Eskimo/

Inuit, and Australian crania closely paralleled those found in this study. European faces were characterized as being rather narrow, with retracted cheekbones, long noses, and prominent midfaces. African facial morphology displayed a broad but midsagittally prominent upper segment, with a low nose set in a low face, while Australian crania also had a low, broad face, but one which was midsagittally flatter and inferiorly more prominent. Eskimo/Inuit crania had high faces, noses, and cheeks, with transverse flatness and prominent cheekbones. All of these characteristics are also apparent in the results of the analyses described here, and in particular the similarity between African and Australian facial shape, compared with European and Asian crania (here, Inuit) is paralleled in the results of Howells (1973, 1989), Lahr (1996), and Hanihara (1996, 2000). It is certainly possible that this similarity is a reflection of morphology retained in common from African early modern humans.

Additional comparisons of the relative positions of recording points are also available from geometric morphometrics, although for comparability and validation, we restricted ourselves here to the standard points used by Howells (1973, 1989). For example, his analyses could record the distance between nasion and prosthion because this was directly measured with a sliding caliper, or the projection of the subspinale relative to the two zygomaxillare points, because this was determined with a coordinate caliper, and could be reflected as a subtense value or a computed angle of prominence. However, his measurements did not record the distance between, say, the nasion and the zygomaxillare points, or the frontomales and prosthion. Using geometric morphometrics, all of these points could be simultaneously related to each other in three dimensions, and moreover, the scanned data could be used to generate entirely new comparisons, without returning to the original crania, e.g., in the construction of midsagittal or transverse profiles between the recording points, or even the addition of newly defined landmarks. An additional potential application for laser scanning is in the recording and quantification of cranial asymmetry, where the interplay of genetic and functional influences on morphology is being investigated (see also Winning et al., 1999).

#### **CONCLUSIONS AND FUTURE PROSPECTS**

The detailed observations presented above can now be integrated with other data to provide an overview of this study and a consideration of further applications for laser scanning technology. Laser scanners are now available which are easily portable, and this should encourage their wider application to physical anthropology.

- 1) Geometric morphometrics provides an accurate characterization of the overall face shape in each sample, of the variation within each sample, and of the essential differences between samples.

Comparison with more traditional studies (e.g., Howells, 1973, 1989) shows close similarities in characterization of facial shape for the different samples. However, using geometric morphometrics, additional points can be simultaneously related to each other in three dimensions, and the scanned data used to generate entirely new comparisons, without returning to the original crania. This can include the construction of midsagittal or transverse profiles, the addition of newly defined landmarks, and the quantification of facial asymmetry.

- 2) This study is a useful validation of geometric morphometrics in practice. In order for the advantages of this new method to be established, practical studies are required in which the results of the two approaches can be compared. In the present case, the new method can be seen to bring valuable new insights that extend, rather than conflict with, earlier work. This is demonstrated in relation to the 2D study of Lynch et al. (1996), where the relative importance of prognathism in population comparisons can be further clarified.
- 3) This study confirms the power of laser scanning technology (Wood et al., 1998), especially in the density of laser scan data. We exploited the possibilities of extracting landmark coordinates while varying the scale, viewpoint, and lighting of the graphic image.
- 4) In addition to the data described above, 360° scans of skulls were also taken, and from these further landmarks will be extracted and analyzed. Moreover, it is possible to study other data than landmarks from laser scans, e.g., profiles and surfaces, which might complement traditional methods. For example, Lahr (1996) carried out a series of comparisons between human crania from different populations in order to test the reality of supposed regional characteristics. She used metrics and a series of reference plaques to record and compare the basic data scored. Geometric morphometrics and surface or profile analysis could be used to further this kind of research, as not only metric features characterizing shape, but morphological features such as extent of torus formation, area of zygomatic trigones, and degree of projection of zygomaxillary tuberosities could be scored and compared. These characters could be compared singly within or between samples, or different data could be integrated to provide an overall sample pattern or to observe correlations between the different traits. It is also possible to extract and analyze other shape descriptors such as ridge curves (Dean et al., 1996) and mean and Gaussian curvatures (Besl and Jain, 1986). In this respect, laser scanning compares favorably with direct measurement of landmark coordinates.
- 5) The geographical/ethnic identity of individual cases can be tested. This could be used for fossil

crania, where their possible affiliation to modern populations is being investigated as part of the study of modern human origins (this will be explored in a subsequent paper), or for forensic cases where the ethnic identity of an unknown individual needs to be determined.

## ACKNOWLEDGMENTS

The authors thank Dr. David Walsh of UCL's Department of Medical Physics for helping with the scanning of skulls, and Robert Kruszynski and Una Vidarsdottir for their general assistance at the Natural History Museum. We also thank four anonymous reviewers whose comments enabled this manuscript to be much improved. The programs APS, GRF-ND, and LINDA can all be accessed at <http://life.biosunysb.edu/morph/>. The SPlus macros, written by Professor Ian Dryden, are also available at <http://www.maths.nott.ac.uk/personal/ild/>.

## LITERATURE CITED

- Besl PJ, Jain A. 1986. Invariant surface characteristics for 3D object recognition in range images. *Comput Vis Image Process* 33:33–80.
- Bookstein FL. 1997. Shape and the information in medical images: a decade of the morphometric synthesis. *Comput Vis Image Understanding* 66:97–118.
- Cavalcanti MJ. 1999. LINDA—linear discriminant analysis and comparison of multivariate samples with randomisation tests, version 5.14. <http://life.biosunysb.edu/morph>
- Dean D, Marcus LF, Bookstein F. 1996. Chi-square test of biological space curve affinities. In: Marcus LF, Corti M, Loy A, Naylor GJP, Slice D, editors. *Advances in morphometrics*. New York: Plenum Press. p 235–251.
- Dryden IL, Mardia KV. 1998. *Statistical shape analysis*. Chichester: John Wiley and Sons.
- Ferrario VF, Sforza C, Poggio CE, Cova M, Tartaglia G. 1997. Preliminary evaluation of an electromagnetic three-dimensional digitiser in facial anthropometry. *Cleft Palate Craniofac J* 35:9–15.
- Goodall CR. 1991. Procrustes methods in the statistical analysis of shape. *J R Stat Soc B* 53:285–339.
- Hanihara T. 1996. Comparison of craniofacial features of major human groups. *Am J Phys Anthropol* 99:389–412.
- Hanihara T. 2000. Frontal and facial flatness of major human populations. *Am J Phys Anthropol* 111:105–134.
- Hennessy RJ, O'Higgins P, Cobb S. 1997. 3D studies of human craniofacial growth. *Genet Res* 70:83.
- Howells WW. 1973. Cranial variation in man. *Pap Peabody Mus* 67: whole volume. Cambridge, MA: Harvard. p 1–259.
- Howells WW. 1989. Skull shapes and the map. *Pap Peabody Mus* 79: whole volume. Cambridge, MA: Harvard. p 1–189.
- Kendall DG. 1984. Shape-manifolds, Procrustean metrics, and complex projective spaces. *Bull Lond Math Soc* 16:81–121.
- Lahr M. 1996. *The evolution of modern human diversity*. Cambridge: Cambridge University Press.
- Le H, Kendall DG. 1993. The Riemannian structure of Euclidean shape spaces. A novel environment for statistics. *Ann Stat* 21:1225–1271.
- Lele S, Richtsmeier JT. 1991. Euclidean distance matrix analysis: a coordinate-free approach for comparing biological shapes using landmark data. *Am J Phys Anthropol* 86:415–427.
- Lynch JM, Wood CG, Luboga S. 1996. Geometric morphometrics in primatology: craniofacial variation in *Homo sapiens* and *Pan troglodytes*. *Folia Primatol (Basel)* 67:15–39.
- Moss JP, Linney AD, Grindrod SR, Mosse CA. 1989. A laser scanning system for the measurement of facial surface morphology. *Optics Lasers Eng* 10:179–190.

- O'Higgins P. 2000. The study of morphological variation in the hominid fossil record : biology, landmarks and geometry. *J Anat* 197:103–120.
- O'Higgins P, Dryden IL. 1993. Sexual dimorphism in hominoids: further studies of craniofacial shape differences in *Pan*, *Gorilla* and *Pongo*. *J Hum Evol* 24:183–205.
- O'Higgins P, Jones N. 1998. Facial growth in *Cercocebus torquatus*: an application of three-dimensional geometric morphometric techniques to the study of morphological variation. *J Anat* 193:251–272.
- Penin X, Baylac M. 1999. Comparaison tridimensionnelle des crânes de *Pan* et *Pongo* par superpositions procrusteennes. *C R Acad Sci [III]* 322:1099–1104.
- Richtsmeier JT, Cheverud JM, Lele S. 1992. Advances in anthropological morphometrics. *Annu Rev Phys Anthropol* 21:283–305.
- Rohlf FJ. 1993. NTSYS-pc—numerical taxonomy and multivariate analysis system, version 1.80. New York: Exeter Software.
- Rohlf FJ. 1998. On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Syst Biol* 47:147–158.
- Rohlf FJ. 1999a. Shape statistics: Procrustes superimpositions and tangent spaces. *J Classif* 16:197–223.
- Rohlf FJ. 1999b. On the use of shape spaces to compare morphometric methods. *Hystrix* 11:8–24.
- Rohlf FJ. 2000. Statistical power comparisons among alternative morphometric methods. *Am J Phys Anthropol* 111:463–478.
- Rohlf FJ, Bookstein FL. 1987. A comment on shearing as a method for “size correction.” *Syst Zool* 36:356–367.
- Rohlf FJ, Marcus LF. 1993. A revolution in morphometrics. *Trends Ecol Evol* 8:129–132.
- Ross AH, McKeown AH, Konigsberg LW. 1999. Allocation of crania to groups via the “new morphometry.” *J Forensic Sci* 44:584–587.
- Shaner DJ, Bamforth JS, Peterson AE, Beattie OB. 1998. Technical note: different techniques, different results—a comparison of photogrammetric and caliper-derived measurements. *Am J Phys Anthropol* 106:547–552.
- Slice DE. 1996. Three-dimensional generalised resistant fitting and the comparison of least-squares and resistant fit residuals. In: Marcus LF, Corti M, Loy A, Naylor GJP, Slice D, editors. *Advances in morphometrics*. New York: Plenum Press. p 179–199.
- SPSS (Statistical Package for Social Sciences). 1997. *Statistical package for social sciences*, version 8.0.0. Chicago: SPSS, Inc.
- Valeri CJ, Cole TM, Lele S, Richtsmeier JT. 1998. Capturing data from three-dimensional surfaces using fuzzy landmarks. *Am J Phys Anthropol* 107:113–124.
- Winning T, Brown T, Townsend G. 1999. Quantifying asymmetry in the human facial skeleton. *Perspect Hum Biol* 4:53–60.
- Wolpoff M, Caspari R. 1997. *Race and human evolution*. New York: Simon and Schuster.
- Wood B, Aiello L, Wood C, Key C. 1998. A technique for establishing the identity of “isolated” fossil hominin limb bones. *J Anat* 193:61–72.
- Wood CG, Lynch JM. 1996. Sexual dimorphism in the craniofacial skeleton of modern humans. In: Marcus LF, Corti M, Loy A, Naylor GJP, Slice D, editors. *Advances in morphometrics*. New York: Plenum Press. p 407–414.
- Yaroch LA. 1996. Shape analysis using the thin-plate spline: Neanderthal cranial shape as an example. *Yrbk Phys Anthropol* 39:43–89.