# INFORMACION MORFOMETRICA EN ESTUDIOS INTERPO-BLACIONALES: UNA COMPARACION DE TECNICAS BASADAS EN PROPORCIONES Y COORDENADAS PROCRUSTES

Lumila P. Menéndez\* y S. Ivan Perez

División Antropología. Facultad de Ciencias Naturales y Museo. Universidad Nacional de La Plata. La Plata. Argentina Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Buenos Aires. Argentina

PALABRAS CLAVE morfometría tradicional; morfometría geométrica; estadística multivariada descriptiva

RESUMEN Tradicionalmente se ha abordado el estudio de la variación morfológica mediante técnicas estadísticas uni y multivariadas, aplicadas principalmente a proporciones calculadas sobre distancias lineales entre puntos anatómicos (*landmarks*). Más recientemente, se ha propuesto el empleo de coordenadas cartesianas superpuestas (coordenadas Procrustes) de landmarks y de puntos sobre contornos y superficies (*semilandmarks*). El objetivo de este trabajo es comparar los resultados obtenidos a partir del empleo de estas aproximaciones a estudios morfométricos de poblaciones humanas. Para esto estudiamos cráneos de individuos adultos de ambos sexos, provenientes de diferentes muestras de Argentina (N=130) empleando coordenadas Procrustes en 3D, proporciones y métodos multivariados. Se compararon los resultados obtenidos a partir del estudio de tres propiedades morfométricas: tamaño, *shape y form.* Los resultados obtenidos sugieren que las técnicas basadas en coordenadas Procrustes son las más eficientes para estudiar variación en *shape* entre poblaciones. Rev Arg Antrop Biol 13(1):105-111, 2011.

KEY WORDS traditional morphometry; geometric morphometrics; multivariate descriptive statistics

ABSTRACT The study of morphological variation has been traditionally approached through univariate and multivariate statistical techniques applied to proportions of linear distances between anatomical points (landmarks). More recently, it has been proposed the use of superimposed Cartesian coordinates (Procrustes coordinates) of landmarks and/or points on contours and surfaces (semilandmarks). The aim of this study is to compare the results obtained from these approaches in morphometric studies of human

La variación morfológica observada a una escala poblacional resulta de factores aleatorios (e.g., deriva génica) y no-aleatorios (e.g., migración y flujo génico, selección y respuestas ecofenotípicas) que actúan dentro y entre generaciones. Por lo tanto, los patrones de variación morfológica resultantes, al igual que otros rasgos del fenotipo, contienen información relevante sobre los factores que les dieron origen. En este sentido, dentro del campo de la antropología biológica se han realizado numerosos esfuerzos tendientes a desarrollar aproximaciones morfométricas que maximicen la información obtenida a partir de tales rasgos. Tradicionalmente estas aproximaciones se han centrado en el empleo de proporciones calculadas a partir de distancias lineapopulations. For this purpose, adult skulls of both sexes coming from different samples of Argentina (N=130) were measured using 3D Procrustes coordinates, proportions and multivariate methods. The results obtained were compared through the study of three morphometric properties: size, *shape* and *form*. These results shown that the techniques based on Procrustes coordinates prove to be more efficient to study the variation in *shape* between populations. Rev Arg Antrop Biol 13(1):105-111, 2011.

les entre puntos anatómicos, en el marco de análisis estadísticos uni o multivariados (Martin y Saller, 1957; Howells, 1973; Jungers et al., 1995). Más recientemente, se ha propuesto el empleo de coordenadas cartesianas de puntos anatómicos (*landmarks*) y/o de puntos sobre contornos y superficies (*semilandmarks*)

E-mail: lumilam@gmail.com

Recibido 18 Febrero 2011; aceptado 05 Julio 2011

Financiamiento: Proyectos Promocionales de Investigación y Desarrollo N002, 2011-2013, UNLP. Proyecto de Investigación Plurianual 112-200901-00132, 2010-2012, CONICET.

<sup>\*</sup>Correspondencia a: Lumila P. Menéndez. División Antropología. Museo de La Plata. Paseo del Bosque S/N. La Plata 1900. Buenos Aires. Argentina.

después de una Superposición Procrustes Generalizada (*i.e.*, coordenadas Procrustes) debido a que permitirían capturar mayor cantidad de información al preservar la localización espacial de los puntos anatómicos y describir formas curvas (Bookstein, 1991, 1996a, b, c, 1997; Adams et al., 2004; Mitteroecker y Gunz, 2009).

Es esperable que la información obtenida a partir del modelado de los rasgos morfológicos varíe en función de la aproximación morfométrica empleada y por lo tanto los patrones generados a partir de los mismos podrían diferir significativamente. Esto es relevante en la medida en que los patrones generados podrían llevar a distintas explicaciones acerca de los factores involucrados en la diferenciación morfológica poblacional. En este sentido, el objetivo de este trabajo es comparar los resultados obtenidos a partir del empleo de proporciones y coordenadas Procrustes de landmarks y semilandmarks en 3D como variables en estudios morfométricos multivariados. Particularmente, se evalúa si estos dos tipos de variables generan diferentes patrones de variación entre las muestras a partir de la comparación de tres propiedades morfométricas: tamaño, shape y form.

### MATERIAL Y METODOS

Estudiamos un total de 130 individuos adultos de ambos sexos, provenientes de cinco muestras depositadas en el Museo de La Plata (Lehmann-Nitsche, 1910). Las muestras fueron asignadas al Holoceno tardío (4000 años AP hasta momentos históricos) en función del contexto arqueológico y fechados radiocarbónicos (Bernal et al., 2008) y proceden de Buenos Aires (BA) (N=19), San Juan (SJ) (N=17), Río Negro (RN) (N=16), Chubut (CH) (N=40) y Pampa Grande (PG) (N=38). Se seleccionaron cráneos completos y sin deformación craneana. El sexo y la edad de los individuos estudiados fueron determinados a partir de caracteres morfológicos del cráneo (Buikstra y Ubelaker, 1994).

Se registraron 84 coordenadas cartesianas

en 3D describiendo diferentes rasgos anatómicos del cráneo (Fig. 1), empleando un brazo mecánico MicroScribe G2X. Las mismas corresponden a 42 landmarks y 42 semilandmarks que describen la totalidad de la superficie del cráneo. A partir de las coordenadas se generaron tres conjuntos de datos: 1) landmarks (L): definidos como coordenadas cartesianas de puntos anatómicos homólogos entre individuos (Bookstein, 1991). Estos puntos fueron seleccionados siguiendo a Howells (1973) y Buikstra y Ubelaker (1994) (Fig. 1a y Apéndice online 1); 2) landmarks y semilandmarks (LySL): este conjunto de datos contiene las coordenadas cartesianas de los landmarks y coordenadas de semilandmarks o puntos definidos para describir superficies y contornos (Bookstein, 1997; Adams et al., 2004) (Fig. 1a y Apéndice online 1) y 3) Proporciones (P) de las distancias lineales entre landmarks (DL). Se calcularon 24 distancias lineales, las cuales fueron estimadas como las distancias Euclideanas entre las coordenadas de landmarks siguiendo parcialmente a Howells (1973), Buikstra y Ubelaker (1994) y Pucciarelli (2008). Las P se calcularon como el cociente entre las DL y la media aritmética de la totalidad de las distancias por individuo (Darroch y Mosimann, 1985; Jungers et al., 1995) (Fig. 1b y Apéndice online 2).

Los semilandmarks fueron registrados como puntos en contornos y superficies de los diferentes huesos delimitados entre landmarks. Este procedimiento para ubicar semilandmarks parte del supuesto de correspondencia biológica u homología entre huesos y suturas (Gunz et al., 2005). Los semilandmarks fueron tratados como puntos igualmente espaciados entre landmarks siguiendo a Williams y Slice (2010). Además de la traslación, escala y rotación que se les efectúa a los landmarks, los puntos correspondientes a los semilandmarks son deslizados a lo largo de una curva hasta que coinciden lo mejor posible con las posiciones de los puntos correspondientes a lo largo de un contorno en una configuración de referencia (Adams et al., 2004). Esto se realiza porque las curvas o con-



Fig. 1. Puntos y distancias registrados en el presente trabajo. A) *Landmarks* (cuadrados negros) y *Semilandmarks* (círculos grises), B) Distancias lineales: generales (líneas continuas) y locales (líneas discontinuas). Se registraron medidas generales tradicionales que abarcan varias estructuras y medidas locales de determinados huesos [similar el enfoque craneofuncional de Pucciarelli (2008)] con el propósito de abarcar el cráneo en su completitud. El listado completo de puntos y medidas se encuentran en los Apéndices Online 1 y 2 respectivamente.

tornos deben ser homólogos entre los distintos individuos, aunque sus puntos individuales no necesariamente lo sean (Bookstein et al., 2002). Dado que no existe un acuerdo acerca de cual es el mejor criterio para usar cuando se estudian *semilandmarks*, reanalizamos la base de datos utilizando el criterio mínimo de Procrustes (Bookstein, 1997; Gunz et al., 2005). Este criterio para deslizar los *semilandmarks* es implementado mediante una serie de rutinas en R, las cuales fueron escritas y compartidas por Dean Adams y Erik Otarola-Castillo.

A partir de los tres conjuntos de datos, estudiamos y comparamos los resultados obtenidos de tres propiedades morfométricas: shape, tamaño y form (en este trabajo emplearemos los términos en inglés). Las variables de shape fueron obtenidas mediante las P (Corruccini, 1987) y las coordenadas que fueron ajustadas mediante el método de Superposición Procrustes Generalizado para los conjuntos L y LySL (Bookstein, 1991; Zelditch et al., 2004). El método de Superposición Procrustes Generalizado alinea las configuraciones de puntos eliminando la información sobre el tamaño, posición y orientación original de las coordenadas (Rohlf, 1990; Rohlf y Slice, 1990; Bookstein, 1991, 1997). Las P fueron calculadas como el cociente entre cada una de las distancias y la media aritmética de la totalidad de las distancias por cada individuo (Darroch y Mosimann, 1985; Jungers et al., 1995). El tamaño fue definido como el logaritmo natural del tamaño centroide (Log TC; la raíz cuadrada de la suma de las distancias cuadradas desde cada coordenada al centroide de la configuración; Bookstein, 1991; Mitteroecker y Gunz, 2009) para los conjuntos de L y LySL y como la media geométrica (MG; la enésima raíz del producto de las n distancias; Jungers et al., 1995) para el conjunto P. El tamaño fue calculado a partir de las coordenadas superpuestas y de las distancias lineales de todos los individuos de la muestra. Finalmente, form fue definida como el resultado de la sumatoria del tamaño y shape (Zelditch et al., 2004). Por tanto, las propiedades de form de las estructuras analizadas fueron descritas mediante la sumatoria de las coordenadas alineadas de landmarks y semilandmarks y el logaritmo del Tamaño Centroide para los conjuntos de L y LySL y mediante la sumatoria de las P y la media geométrica para el conjunto de P. Los análisis morfométricos fueron realizados mediante el empleo de los programas MorphoJ (Klingenberg, 2011) y R (R Developmental Core Team, 2011).

Las variables de *form* y *shape* obtenidas a partir de los tres conjuntos de datos fueron resumidas mediante análisis de componentes principales tradicional y componentes principales entre grupos (Zelditch et al., 2004; Mitteroecker y Bookstein, 2011). Los componentes principales tradicionales (CP) son proyecciones ortogonales de los datos de todos los individuos (proporciones o coordenadas Procrustes) sobre los autovectores de la matriz de covarianza total. Los CP son vectores que describen el cambio global a partir del cambio local entre *landmarks* y *semilandmarks* (Bookstein, 1991; Rohlf, 1993, 1996).

Por otro lado, el análisis de componentes principales de la matriz entre grupos (Mitteroecker y Bookstein, 2011) consiste en la proyección de los datos de todos los individuos sobre los autovectores de la matriz de covarianza de las medias poblacionales. De esta manera, se obtienen ejes ortogonales (CPeg) que describen mejor las diferencias en forma entre grupos que los CP y que pueden ser calculados incluso cuando los datos no son de rango completo o cuando las matrices de covarianza de los grupos difieren (Mitteroecker y Bookstein, 2011).

Los autovalores o los correspondientes porcentajes de variación explicada por los CP, de la matriz de covarianza total de las variables de shape fueron empleados para comparar los conjuntos de datos (Wagner, 1984). Esperamos que en aquellos conjuntos de datos que presentan una estructura de covarianza simple, los primeros CPs resuman gran porcentaje de variación morfométrica (Manly, 1994). Finalmente, los ordenamientos generados por las variables de tamaño (Log TC y MG), form y shape (CP y CPeg) para los tres conjuntos de datos fueron comparados empleando la técnica de PROTEST (Peres-Neto y Jackson, 2001). Este análisis superpone los ordenamientos mediante cuadrados mínimos y calcula una métrica de seudo-correlación  $[m_{12}=(\Sigma 1-ss)^2]$ que fue permutada 10.000 veces para conocer su significación. En este trabajo se utilizó la técnica de PROTEST en lugar del análisis más tradicional de Mantel debido a que el primero considera la dependencia entre los conjuntos de datos (i.e., estudiamos la misma muestra con técnicas diferentes) mientras que el segundo no lo hace (Peres-Neto y Jackson, 2001).

Los análisis estadísticos fueron efectuados usando el programa R 2.13.1 (R Developmental Team, 2011).

#### **RESULTADOS Y DISCUSION**

Los tamaños obtenidos a partir de las coordenadas cartesianas y de las medidas lineales fueron significativamente similares. La correlación entre los mismos para los 3 conjuntos de datos mostró valores elevados (≥0,96; Tabla 1). Con respecto al análisis de *form*, el primer CP calculado sobre las coordenadas superpuestas (más el logaritmo del TC), explica valores más bajos de variación (ca. 35%) en relación al análisis del primer CP sobre las distancias lineales (más la MG) (ca. 75%). A pesar de las diferencias en los valores de variación explicada, las correlaciones (Tabla 1) entre los CP1 para los diferentes datos de form también fueron elevadas ( $\geq 0.90$ ). Esto se debe a que cuando analizamos variación en form la mayor variación está dada por diferencias en tamaño entre los individuos. Estos resultados concuerdan con estudios previos que mostraron que el cálculo de variables de tamaño que difieren en su definición y son estimadas a partir de distintos conjuntos de datos, están altamente correlacionadas (Corruccini, 1987).

TABLA 1.Valores de correlación para las variables de tamaño estudiadas (Media Geométrica y Logaritmo del Tamaño Centroide; triángulo inferior) y el Componente Principal 1 de los tres conjuntos de datos de form (triángulo superior)

	Р	L	LySL
Р	1	0,942	0,902
L	0,97	1	0,973
LySL	0,96	0,98	1

Por el contrario, el análisis de *shape* mostró diferencias entre los conjuntos de datos. Los primeros cinco CPs para cada uno de los tres conjuntos de variables de *shape* explicaron diferentes porcentajes de variación (L: 43,2%; LySL: 41,8%; P: 88,8%), sugiriendo que los conjuntos L y LySL presentan gran cantidad de variación en *shape* entre los individuos que no puede ser resumida en unos pocos ejes de máxima variación (Apéndice online 3). En la Figura 2 se grafican los dos primeros CPs de los tres conjuntos de datos. Los dos conjuntos que presentaron las mayores diferencias son los de P y LySL. Mientras que en el ordenamiento generado por P (Fig. 2a) las muestras se encuentran superpuestas, en el ordenamiento basado en LySL (Fig. 2b) puede visualizarse una mejor separación en-



**Fig. 2.** Distribución de las muestras en los dos primeros CP. A) P, B) LySL, C) L.

tre las mismas. Una situación intermedia se observa en el ordenamiento basado sólo en L (Fig. 2c). El análisis de correlación PROTEST entre los tres primeros componentes principales de los tres conjuntos de datos muestra correlaciones moderadas a bajas (Tabla 2).

En la Tabla 2 se presentan los resultados de las correlaciones entre los tres primeros componentes principales de la matriz entre grupos de los tres conjuntos de datos. El análisis de PROTEST mostró valores elevados de correlación solamente entre los conjuntos L v LySL. En la Figura 3 se grafican los CPeg 1 y 2 de las P, L, LySL. Al igual que para el análisis de CP, en la Figura 3 se observa que en el ordenamiento de P las muestras se encuentran superpuestas, mientras que en el ordenamiento de LySL las muestras se encuentran claramente separadas. También se observa un ordenamiento intermedio para el caso de los L. En el ordenamiento de los LySL, el CPeg 1 separa las muestras de San Juan y Pampa Grande de las muestras de Río Negro, Chubut y Buenos Aires. El CPeg 2 separa la muestra de Buenos Aires de Río Negro y Chubut, mientras las de San Juan y Pampa Grande quedan en un lugar intermedio.

El análisis de coordenadas de puntos orientado a evaluar la variación en *shape* entre las muestras, permitió diferenciar los grupos con mayor eficiencia que las distancias entre puntos empleadas tradicionalmente. Estos resultados son coincidentes con observaciones previas que señalan que el análisis de *shape* basado en coordenadas cartesianas no sólo incluye diferencias proporcionales entre estructuras (*i.e.* tamaño relativo) sino información acerca del desplazamiento espacial de los puntos analizados (Zelditch et al., 2004; Mitteroecker y Gunz, 2009).

Sin embargo, en este trabajo la mayor diferenciación entre las muestras, fue obtenida a partir del empleo conjunto de *landmarks* y *semilandmarks*, que presenta la estructura de covariación más compleja como lo sugiere la comparación de los porcentajes de variación explicada por los CPs. TABLA 2. Valores de correlación PROTEST entre los Componentes Principales calculados sobre la matriz de covarianza total de las variables deforma (triángulo inferior) y los Componentes Principales entre grupos, calculados sobre la matriz de covarianza total entre muestras(triángulo superior)

	Р	L	LySL
Р	1	0,5682	0,4812
L	0,5829	1	0,8923
LySL	0,41	0,4066	1



**Fig. 3.** Distribución de las muestras en los primeros CPeg. A) P, B) LySL, C) L.

Esto indicaría que la mayor información incorporada por los semilandmarks, e.g. aquella referida a la variación en la curvatura de las estructuras craneofaciales (Gunz et al., 2005), resulta relevante en la descripción de los patrones de diferenciación poblacional. En síntesis, los resultados obtenidos sugieren que las técnicas tradicionales, basadas en proporciones y geométricas, basadas en coordenadas Procrustes, difieren significativamente en la cantidad y calidad de información que preservan acerca de las propiedades de las estructuras estudiadas. Por lo tanto, la elección de una aproximación morfométrica que permita modelar adecuadamente la información de rasgos relevantes, dependerá del problema a abordar así como del grado de variación y complejidad de la forma analizada. Particularmente en el estudio de estructuras altamente complejas como el cráneo y que presentan niveles bajos de variación característicos de las poblaciones humanas (Relethford, 1994), las técnicas basadas en coordenadas Procrustes resultarían ser las más eficientes para estudiar variación en shape. La adecuación entre la técnica empleada y el problema a abordar contribuirá a una mejor descripción de los patrones de variación y en consecuencia a mejorar nuestra comprensión de los procesos biológicos subyacentes.

## AGRADECIMIENTOS

Queremos agradecer a los dos revisores anónimos quienes hicieron valiosos comentarios que permitieron mejorar notablemente el presente trabajo.

#### LITERATURA CITADA

- Adams D, Rohlf FJ, Slice D. 2004. Geometric morphometrics: Ten years of progress following the "revolution". Ital J Zool 71:5-16.
- Bernal V, González PN, Perez SI, Pucciarelli HM. 2008. Entierros humanos del noreste de Patagonia: nuevos fechados radiocarbónicos. Magallania 36(2):175-183.
- Bookstein FL. 1991. Morphometric tools for landmark data. Geometry and biology. Cambridge: Cambridge University Press.
- Bookstein FL. 1996a. Biometrics, biomathematics and the

## COMPARACION DE TECNICAS MORFOMETRICAS

morphometric synthesis. Bull Math Biol 58:313-365.

- Bookstein FL. 1996b. Combining the tools of geometric morphometrics. En: Marcus LF, Corti M, Loy A, Naylor GJP, Slice DE, editores. Advances in morphometrics. Nato ASI Series, Series A: Life Science, Vol. 284. New York: Plenum. p 131-152.
- Bookstein FL. 1996c. Applying landmark methods to biological outline data. En: Mardia KV, Gill CA, Dryden IL, editores. Image fusion and *shape* variability. Leeds: University of Leeds Press. p 79-87.
- Bookstein FL. 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline *shape*. Med Im Anal 1:225-243.
- Bookstein FL, Streissguth AP, Sampson PD, Connor PD, Barr HM. 2002. Corpus callosum *shape* and neuropsychological deficits in adult males with heavy fetal alcohol exposure. Neuroimage 15:233-251.
- Buikstra JE, Ubelaker DH. 1994. Standards for data collection from human skeletal remains. Arkansas: Arkansas Archaeological Survey 44.
- Corruccini RS. 1987. *Shape* in morphometrics: Comparative analyses. Am J Phys Anthropol 73:289-303.
- Darroch JN, Mosimann JE. 1985. Canonical and principal component of *shape*. Biométrika 72:241-252.
- Gunz P, Mitteroecker P, Bookstein FL. 2005. Semilandmarks in three dimensions. En: Slice DE, editor. Modern morhometrics in Physical Anthropology. Developments in primatology: progress and prospects. Chicago: University of Chicago. p 73-98.
- Howells WW. 1973. Cranial variation in man: a study by multivariate analysis of patterns of difference among recent human populations. Papers of the Peabody Museum of Archaeology and Ethnology. Cambridge: Cambridge Peabody Museum.
- Jungers WL, Falsetti A, Wall CE. 1995. Shape, relative size and size-adjustments in morphometrics. Yearb Phys Anthropol 38:137-161.
- Klingenberg CP. 2011. MorphoJ: an integrated software package for geometric morphometrics. Mol Ecol Resour 11:353-357.
- Lehmann-Nitsche R. 1910. Catálogo de la Sección Antropológica del Museo de La Plata. Buenos Aires: Universidad Nacional de La Plata.
- Manly BFJ. 1994. Multivariate statistical methods. Londres: Ed. Chapman & Hall.
- Martin R, Saller K. 1957. Lehrbuch der Anthropologie. Bd 1. Sttugart: Fischer G Verlag.

- Mitteroecker P, Bookstein F. 2011. Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. Evol Biol 38:100-114.
- Mitteroecker P, Gunz P. 2009. Advances in geometric morphometrics. Evol Biol 36:235-247.
- Peres-Neto P, Jackson D. 2001. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. Oecologia 129:169-178.
- Pucciarelli HM. 2008. Evolución y diversificación biológica humana desde la perspectiva craneofuncional. México D.F.: Instituto de Investigaciones Antropológicas, Universidad Nacional Autónoma de México.
- R Developmental Core Team 2011. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Relethford JH. 1994. Craniometric variation among modern human populations. Am J Phys Anthropol 95:53-62.
- Rohlf FJ. 1990. Rotational fit (Procrustes) methods. En: Rohlf FJ, Bookstein FL, editores. Proceedings Michigan morphometrics workshop. Special publication n8 2. Museum of Zoology. Michigan: University of Michigan. p 227-236.
- Rohlf FJ.1993. Relative warp analysis and an example of its application to mosquito wings. En: Marcus LF, Bello Rojo E, Garcia-Valdecasas A, editores. Contributions to morphometrics. Madrid: Museo Nacional de Ciencias Naturales. p 131-159.
- Rohlf FJ. 1996. Mophometric spaces, *shape* components and the effects of linear transformations. En: Marcus LF, Corti M, Loy A, Slice D, Naylor G, editores. Advances in morphometrics. Nueva York: Plenum. p 117-129.
- Rohlf FJ, Slice DE. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst Zool 39:40-59.
- Wagner GP. 1984. On the eigenvalue distribution of genetic and phenotypic dispersion matrices: Evidence for a nonrandom organization of quantitative character variation. J Math Biol 21:77-95.
- Williams SE, Slice DE. 2010. Regional *shape* change in adult facial bone curvature with age. Am J Phys Anthropol 143:437-447.
- Zelditch ML, Swiderski DL, Sheets HD, Fink WL. 2004. Geometric morphometrics for biologists: A premier. Nueva York: Elsevier Academic Press.