<AT>Macromorphoscopic Trait Expression in a Cranial Sample from Medellín, Colombia[†]

<AU>Timisay Monsalve Vargas, PhD^{a*} ##Email##hefnerj1@msu.edu##/Email##, Joseph T. Hefner <Deg>PhD</Deg>, D-ABFA^b

<AU>

<AFF>^aDepartamento de Antropología – FCSH, Universidad de Antioquia, Medellín, Colombia

<AFF>^bDepartment of Anthropology, Michigan State University, East Lansing, Michigan

<AFF>[†]This project was supported by Award No. 2015-DN-BX-K012 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice

<ABS-HEAD>Abstract

<ABS-P>Adjusting existing methods of human identification developed by forensic anthropologists in the United States for use with populations not included in the original development of an analytical method requires data collection using contemporary osteological collections from those populations, and an assessment of the within-group variation present. The primary purpose of this research is to document cranial macromorphoscopic trait variation using methods previously developed in the United States in a sample of 244 individuals from Antioquia, Medellín, Colombia. All individuals are of known age, sex, and birth region. The complex population and demographic history of Colombia makes ancestry assessment particularly difficult in that country. To that end, we explore inter-regional variation throughout Antioquia using birthplace to determine whether forensic anthropologists can provide finer levels of detail beyond identifying an unknown set of human remains as 'Colombian' or, more generally, Hispanic. State and local levels of identification resulting from the varied population histories of each state within Antioquia permit finer resolution, but only to a degree of certainty. Artificial neural networks (aNN) correctly classified only 18.6% of a validation sample, following modest classification accuracies of test/tuning (11.6%) and training (82.8%) samples to original birthplace. As with most neural networks, overfitting is an issue with these analyses. To remedy this overfitting and to document the applicability of aNNs to the assessment of ancestry in Colombia, we pooled the sample of Colombian data and compared that to modern American samples. In those analyses, the best aNN model correctly classified 48.4% (validation) of the sample. Given these results, finer levels of analysis in Colombia are not yet possible using only macromorphoscopic trait data.

<KWD>Keywords: Forensic Science; Forensic Anthropology; Macromorphoscopic Traits; Ancestry; Classification Statistics; Medellín, Colombia; Hispanic

Developing population-specific tools for the assessment of ancestry is an important arena in the general development of ancestry assessment methods in forensic anthropology. Variations in the shape of cranial skeleton useful for such assessments are typically examined via metric and nonmetric features [1-2]. Recent methodological advances in nonmetric trait analysis now permit standardized data collection around the world [3-4] and analysis with robust statistical classification methods [5]. The number of skeletal collections available in countries other than the United States permits population-specific data more useful to local forensic anthropologists than the U.S.-specific data used previously. Employing a sample of 244 individuals from the Skeletal Collection of the University of Antioquia, Medellín, Colombia, we document inter- and intra-population variation within Colombia focusing on the macromorphoscopic traits developed and refined by Hefner [3]. The purpose of this research is the documentation of macromorphoscopic variation within the various regions of Antioquia (Figure 1). Using this region-level variation, this research will also present data and analytical methods useful to forensic anthropologists in Colombia for the assessment of ancestry.

Beyond some of the recently discussed issues surrounding Hispanic classifications in the United States [2, 6-13], Colombia faces a seemingly overwhelming task in the identification of mestizo [*sic*], or mixed, populations. The complex population and demographic history of Colombia makes ancestry assessment particularly difficult. The general demographic background derived from census data specifies the following distributions. The Afro-Colombian population accounts for 10.5 percent of the population, the indigenous population is 3.4 percent, and a small but identified Romani population accounts for 0.01 percent. The remaining 86 percent of the population is considered non-ethnic minority, comprising Europeans, or whites, (37%) and mestizos (49%, mixed white European and Amerindian ancestry) [14].

Ethnicity and social race in Colombia, on the other hand, deals almost exclusively with three geographically-situated groups—Indigenous Colombians, Africans, and Europeans. Gene flow between these three groups has persisted throughout the last 500 years, leading some demographers to describe Colombia as one of the most ethnically diverse countries in the Western Hemisphere [15]. Accordingly, most Colombians identify themselves and others according to social race, ancestry, physical appearance, and sociocultural status [16]. Recent genetic research estimates the average Colombian admixture rate is approximately 65% European, 22% Indigenous Colombian, and 13% African ancestry [16-19]. The population in Antioquia is the product of migration and immigration beginning in the 17th century with the arrival of the Spanish bringing together the native AmerIndians of Colombia and European immigrants. In the 18th century, a large contingent of African slaves was brought to the area.

Antioquia experienced relative population isolation during the 19th century [20], only to face another large wave of immigration in the 20th century [21]. These migrations were used by Carvajal et al. [22] to posit that approximately 94% of the *Y* chromosomes in Antioquia derive from European, with the remainder from African (5%) and Amerindian (1%). These percentages significantly vary within and between regions, reflecting that diverse population history and leading to difficulties in the identification of ancestry in Colombia. These difficulties are further confounded when Colombian forensic anthropologists are using methods and reference samples specific to the United States.

Currently in Colombia there are no population specific methods for the assessment of ancestry using cranial macromorphoscopic features. Several studies on population affiliation exist, but these focus predominantly on pre-Hispanic populations and pre- or proto-historic indigenous groups [23]. Research into cranial macromorphoscopic trait variation in Colombia

using contemporary and recent skeletal collections permits the adjustment of existing U.S.-based methods of human identification for application to these local populations. The value of such population-specific adjustments is most evident when the humanitarian issues in Colombia are understood.

In Colombia there are approximately 78,200 missing persons [24], although underreporting and issues with data on forced disappearances suggest this number could be much higher [24]. Nearly 46,000 of the missing in Colombia are considered "forced disappearance," the result of internal conflict, population control, political opposition, and suppression. In Antioquia forced disappearances account for 10,880 individuals registered as missing [24].

There are also a large number of unidentified human remains not associated with political or state violence, but rather medico-legal cases lacking identification. The Interior of Ministry in Colombia reported 28,195 unidentified remains recovered from 295 public and private cemeteries [25]. Antioquia leads that list with 5,528 currently unidentified remains. The National Prosecution Specialized in Transitional Justice in Colombia documented 5,140 mass graves around Colombia containing a total of 6,553 unidentified remains. Of those, 1,138 were recovered in Antioquia [24, 25]. These individuals remain unidentified in many instances because of sub-standard identification tools and the implementation of U.S.-based analytical methods for assessing biological distance and ancestry.

To address this gap in method development for Colombian populations, we explore the range in variation within a sample of skeletal material from Medellín, Antioquia, Colombia. Working under the premise that biologically related groups will share morphological similarities reflected in the skeleton, we develop population-specific frequency distributions using previously defined macromorphoscopic traits conventionally used as osteological indicators of

ancestry. Next, we utilize social classifications based on Colombian-defined geographic regions as *a priori* group labels to determine the finest level of ancestry analysis possible in Colombia. Finally, we explore a pooled sample of Colombians and compare that data to various other populations.

Materials and Methods

Data Collection

Macromorphoscopic data (Table 1) were collected using the computer program Macromorphoscopic [v. 1.6], which is designed following *Osteoware* (Smithsonian Institution, 2010). Nine macromorphoscopic traits previously defined and illustrated in Hefner [3] were selected for analysis. Possible outcome values for each trait include: 1) 0 to 1 (a binary variable; the presence or absence of a post-bregmatic depression); 2) 1, 2, 3 (a nominal variable; the shape of zygomaticomaxillary suture); and, 3) 1, 2, 3 (an ordinal variable; the breadth of the interorbital space). Following data collection, each individual represents a 9-dimensional vector suitable for analysis.

Sample selection

Colombian Data

Table 2 presents the Colombian sample composition, by sex and regional composition. Of the ca. 400 individuals currently housed in the osteological reference collection at the University of Antioquia, Medellín, Colombia, macromorphoscopic data were available for 244 adult individuals. Data for the other individuals were not possible to collect due to several factors, including, but not limited to taphonomic damage to the cranium, the non-adult ages of some individuals, and/or trauma & pathology potentially affecting macromorphoscopic trait expression. The collection at the University represents individuals with birth years ranging from the mid-twentieth century into the current century. These skeletal remains were recovered from the San Pedro Cemetery Museum and the Universal Cemetery of the city of Medellín. The skeletal collection currently houses over 400 remains, ranging in age from fetal to 99 years, and representing individuals from throughout Antioquia. The individuals comprising this reference collection were previously interred for four years in above-ground vaults or buried in individual graves in one of the two aforementioned cemeteries. At the end of the four years, family members may either claim the remains or donate the decedent to the University of Antioquia.

The forensic anthropology laboratory at the university processes the remains (maceration, cleaning, etc.). Following standards outlined in the Isaza and Monsalve [26], the remains are prepared for curation and research (e.g., inventory, photography, labeling, and storage).

The following criteria were considered during collection sampling for this research. First, confirmation of an individual's sex, age-at-death, and geographic origin (certificate of citizenship found on the death certificate) through sound documentation were required. Only adult individuals were selected for analysis. Although recent research has confirmed macromorphoscopic trait analysis using subadult remains [27], we elected to focus exclusively on adults for this study. Finally, given the nature of macromorphoscopic data analysis for a reference sample, crania in poor condition or with pathological conditions/trauma affecting the midface and/or vault were excluded from the analysis.

To explore how phylogeographic populations and social classifications are used in Colombia, the entire sample was further sub-divided into birthplace cohorts. These were defined following the region/municipality in Antioquia in which the individual was born. Ten regions/municipalities within Antioquia were identified (see Table 2). These states were then clustered according to phylogeographic social classifications identified within Antioquia to assess how these identifiers correlate to macromorphoscopic trait variation (see Table 2). Subsequent analyses identify these sub-groupings as birthplace cohort (region/municipality) and socially defined units, respectively. Because no macromorphoscopic data have been previously collected for populations from Colombia, there are no expectations concerning the frequency and distribution of these traits within this sample.

Non-Colombian Data

A randomly generated sample of American Black, American White, Asian, Mexican, Guatemalan, and Pacific Islander crania were obtained from the Macromorphoscopic Databank

(MaMD) for comparison with the Colombian crania. The MaMD is a databank of cranial macromorphoscopic trait scores from around the world, obtained under funding from the National Institute of Justice (NIJ 2015-DN-BX-K012). Hefner [28] provides a thorough description of the dataset and outlines the various populations contained in the MaMD and the institutions or individuals responsible for data collection. For the purpose of this research, data from the United States was predominately used for comparisons. A brief description of those samples is provided below.

The American Black and White sample comprises modern individuals with 20th century birth years obtained from the William M. Bass Donated Skeletal Collection housed in the Anthropology Department at the University of Tennessee, Knoxville, Tennessee. These individuals are of known age, sex, and ancestry. The Asian sample comprises individuals from Japan and China. Descriptions of these data are provided in Hefner [3]. The Mexican sample comprises modern individuals recovered and identified by the Pima County Office of the Medical Examiner, Tucson, Arizona. These data were previously described in several publications [4, 7]. The Guatemalan data are from the MaMD and comprise identified individuals from *La Verbena* Cemetery in Guatemala City, Guatemala. These data are described fully in Hefner et al. [7]. Finally, the Pacific Island data from the MaMD comprise multiple populations from the Pacific Island region fully described in Ratliff [29]. These data were collected from the University of Pennsylvania, Museum of Archaeology and Anthropology and the National Museum of Natural History, Smithsonian Institution. This island sample includes data from the Philippine Islands, Indonesia, Malaysia, Borneo, Sumatra, Fiji, and the Solomon Islands.

Data Analysis

Hefner and Ousley [5] provide multiple classification statistics appropriate for macromorphoscopic trait analysis. For our purposes, we elected to use an artificial neural network (aNN) to explore classification accuracies within the Colombian crania. Prior to that, however, we use several exploratory statistics to document trait variability expression and to determine how the distributions within Colombia compare to Hispanic populations documented in the United States.

Frequency distributions, inter-trait correlations, and descriptive statistics were calculated to explore trait variability within the Colombia dataset. Statistical significance was assessed using Fisher's exact test (with Yate's continuity correction) for independence of all table rows and columns or nonparametric tests for statistical significance in the distribution of macromorphoscopic trait expression between birth cohorts using Kruskal-Wallis ANOVA. A singular value decomposition performed on the standardized residuals matrix was used to plot two dimensions from each table and display these graphically in a bi-plot. This method is multiple correspondence analysis (mCA) and can be used to demonstrate clustering/ordination in table data. Although the distances between points of different clouds in these plots have little meaning statistically, the relative closeness (correspondence) of points in one cloud to points in the other cloud(s) is important.

Hefner and Ousley [5] outline the applicability of artificial neural networks (aNN), describing these machine learning search algorithms and their implementation in forensic anthropological research, particularly as it relates to the classification of an unknown cranium into one of several reference groups using macromorphoscopic trait scores. To examine how well these models work with crania from Colombia, we implemented several different aNN models. As Hefner and Ousley indicate, "machine learning methods must employ more rigorous cross-

validation methods because they will otherwise report optimistically-biased results" [5: 886]. Therefore, all models were initially trained using a random sample (60%) of the total data. After constructing each model, 20% of the remaining sample is used to test (tune) the newly constructed model and, if necessary, adjust the parameters to maximize classification accuracies while avoiding overfitting and data specific models. Finally, the last 20% of the original dataset is used as an independent, hold-out sample to validate the final aNN models.

Several aNN models were generated using the following algorithms: linear, radial basis function, and MultiLayer Perceptron, or MLP. These are all feedforward, back propagation aNNs that map input data (in this case the character state scores of each individual) onto a set of appropriate outputs (in this instance population, ethnicity, or ancestry) using a number of hidden layers to properly weigh the input variables. In total, we tested 1000 networks, retaining (and herein reporting) only the best five, a distinction given to those models with the lowest classification error rates generated using the validation (hold-out) sample.

All statistical analyses were computed using R [30] or Statistica [31].

Results

Joint frequency distribution data for the Colombian sample are provided in Table 3, by location of birth and for the pooled Colombian sample. Corresponding data for the other populations were previously reported [2-4, 7, 29] and will not be repeated. No statistically significant differences were observed in the expression of macromorphoscopic traits when the Colombian sample was separated by birth cohort (Table 4), so subsequent analyses comparing this data to other world populations will be pooled. Intra-Colombian analyses will still be conditioned on birth place cohort. A polychoric correlation matrix for the Colombian (pooled) sample is presented in Table

5. Significant, but weak, correlations were noted. This is particularly true for nasal aperture width (NAW), which was correlated to six of the nine traits.

Figure 2 provides a bi-plot derived from a multiple correspondence analysis for the character state manifestations and birthplace cohort within the Colombian sample. The first two principal axes explain 40% of the variation in the sample and 26% of the inertia for the overall mCA model ($\chi^2 = 491.94$). The relative dispersion of the birth place cohorts suggests three main ordination clusters. The first axis delineates Occidente, Otre, Surdeste, and Uraba; the second axis separates Valle and Nordeste, along with Oriente and Norte positioned in Quadrant III (-,-). The dispersion of character states around the birthplace cohorts suggests wider nasal apertures, smaller anterior nasal spines, more rounded nasal bone contour, and a postbregmatic depression among individuals from Occidente, Otre, Surdeste, and Uraba, while more narrow nasal apertures and interorbital breadth, more triangular nasal bones, no postbregmatic depression and larger anterior nasal spines correspond to Norte, Oriente, Valle, and to a lesser degree Nordeste.

Although no significant differences were detected in the frequency distribution of macromorphoscopic trait expression among the birthplace cohorts, we chose to assess classification accuracies for these groups using artificial neural networks. One thousand models were built using nine macromorphoscopic traits. The five models with the lowest classification error were retained and presented in Table 6. The results of this analysis are disappointing, but are described below for clarity. Figure 3 visualizes the best multilayer perceptron model, which contained 10 hidden layers and correctly classified 18.6% of the validation sample, following modest classification accuracies of the test/tuning (11.6%) and training (82.8%) samples. Clearly the training sample is overfitting. Given the relatively low classification accuracies of the

validation and testing samples, the aNN classifying birth cohorts is not performing adequately enough for forensic anthropological purposes, despite accuracies above random allocation.

When birthplace cohorts were pooled into the socially defined units observed by Colombians—which interestingly parallel the results from the mCA analysis—the results do not significantly improve. The best aNN model (again an MLP) using the groupings derived from the mCA had the following accuracies: training = 88.2%; testing = 25.6%; and, validation = 18.1%. The most parsimonious explanation for this high error rate is gene flow, shared environmental selective pressures, and the small sample sizes for some of the samples.

The poor performance of the classification models using birthplace cohorts and socially defined units led to additional analyses. The final phase of this analysis explores inter-population classification accuracies using a pooled Colombian sample and US reference samples from the MaMD in multiple aNNs. Table 7 presents model summaries for the five retained models. The best model, a multilayer perceptron with 11 hidden layers, correctly classified 48.4% (validation) of the holdout sample. Table 8 provides the classification matrix from the aNN analysis to illustrate the distribution of incorrect classifications. Many of the incorrect classifications within the Colombian sample were American Black (10%), American White (16.67%), or Hispanic (10%). The relative position of these populations in multivariate space may explain these incorrect allocations. Figure 4 clarifies the multivariate distribution of group centroids derived from a canonical analysis of the principal coordinates. Note the position of the Colombian sample relative to the American Black, American White, and Hispanic samples.

Discussion

The various expressions of macromorphoscopic traits in a sample of identified (known age, sex, and ancestry) modern Colombians were documented to ascertain how well ancestry assessment methods, developed in the United States, classify unknown individuals from Colombia. Ancestry assessment accuracies were documented on three levels: within birth place cohorts, among socially defined units, and using a pooled Colombian sample for comparison to modern American samples.

The primary purpose of this research was to determine if Colombian phylogeographic social distinctions are supported by the macromorphoscopic data and if so, can those data be used to classify unidentified individuals in future forensic casework. The various birthplace cohorts and *post hoc* groupings using socially-defined units followed in Colombia serve as proxies for these phylogeographic populations. Unfortunately, we did not find statistically significant differences between these cohorts. This is manifested in the frequency distribution data having similar values with nearly equal distribution among the populations and the poor performance of the artificial neural network classification models. However, the observed distributions are situated between the American Black, Asian, Hispanic, and American White samples from the United States, which could indicate a significant amount of autochthonous variation did survive the colonization influx of the late 1400s. This variation, with a possible geographic component [32], reveals extensive population structure detail and may speak more widely to cultural perceptions of social race and the impact of colonialism and demographic history than forensic applications of ancestry assessment normally consider.

The multiple correspondence analysis (mCA) accounts for a substantial percentage of the variability within the Colombian sample (49.7%) and helps to summarize (and visualize) the group composition of the Colombian sample. Again, the dispersion of character states around the

birthplace cohorts may suggest a higher level of African ancestry among individuals from Occidente, Otre, Surdeste, and Uraba, where wider nasal apertures, smaller anterior nasal spines, more rounded nasal bone contour, and a postbregmatic depression are more frequent. These traits are generally found in higher percentages among individuals of African ancestry. More narrow nasal apertures and interorbital breadths, more triangular nasal bones, no postbregmatic depression, and larger anterior nasal spines also correspond to Norte, Oriente, Valle, and to a lesser degree Nordeste.

The aforementioned census data suggest the latter regions of Colombia are traditionally associated with a larger effective European population; therefore the character state expressions noted among these groups, and identified in the mCA, may reflect greater genetic contributions from European populations. Many of the early European immigrants were Spanish colonizers, who maintained control of the region until Colombian independence in the early 1820s. The history of Colombia is similarly shared by many Central and South American countries, including Mexico and Guatemala, where the majority of the individuals in the Hispanic sample originated. In all likelihood, this historical connection explains the similarities between these two populations and their relative relationship in multivariate space.

Macromorphoscopic traits performed poorly in the discrimination of birthplace cohorts within the Colombian sample. The frequency distributions of each character state suggest various levels of regional homogeneity, a finding supported by mtDNA research on the haplogroup frequencies in Colombia [16-19,23, 33] and more general studies on ethnic variability in the area [34,35]. Although the aNN classification accuracies were better than random allocation, they did not reach the levels of certainty necessary for forensic applications. These results do not support the use of macromorphoscopic trait data for the determination of birthplace. Although

phylogeographic populations and social classifications are used commonly in Colombia, the phenotypic expression used for these social classifications are not as apparent in the bony tissues of the cranium. Our study, and others, clearly illustrates the level of resolution possible when assessing ancestry with cranial macromorphoscopic traits. The homogeneity within the Colombian sample, at least as that variation is expressed in macromorphoscopic character states, is high enough to necessitate a broader level of analysis.

The aNN classification algorithm exploring how well Colombian crania can be distinguished from other world groups is more promising, correctly classifying 60 percent of the Colombian sample. This same algorithm correctly classified nearly 72 percent of the entire sample in a six group analysis that included American Black, American White, Asian, Hispanic, and Pacific Islanders in addition to the pooled Colombian sample. Table 9 provides a distance matrix of these groups and illustrates the close relationship between the Colombian sample and the Hispanic ($D^2 = 26.1$) and the American White ($D^2 = 40.2$) samples. Social and historical aspects of a population's demographic structure must be known prior to undertaking any assessment of ancestry. As is common practice in the United States, when assessing ancestry using craniometric techniques, macromorphometricians should also provide a broad assessment of ancestry (e.g., Hispanic, Colombian, African) first. Then, when possible, a finer level of analysis (e.g., Northern Mexico, Norte, 20th Century American Black) can also be provided, and may prove to be more meaningful to investigators who are searching through missing persons records, family reports of the missing, investigation records, etc.

The results of this research support the current revitalization in multi-group classifications receiving renewed interest in forensic anthropology [36]. Assumptions about the old three-group (White, Black, Other) model's utility in forensic anthropology can have

consequences for forensic anthropology and ancestry assessment studies. Undetected and/or undocumented differences within and between groups beyond the three-group model have an important bearing on forensic anthropological research, and concomitantly, forensic anthropological casework. Improperly pooled groups (e.g., Asian = Hispanic = Amerindian = Colombian) wrongly assume a level of homogeneity within often geographically, temporally, and genetically very different populations. As researchers continue to develop and refine methodological approaches to the assessment of ancestry, multiple groups should be included to fully account for human variation. Finally, regardless of the outwardly narrow needs of law enforcement or the medical examiner, who seek the peer-perceived ancestry of an unknown individual classified according to social race categories that have little biological meaning, forensic anthropologists—as observers of human variation—need to account for population-level variability.

Acknowledgments

This project was supported by Award No. 2015-DN-BX-K012 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice. The authors would like to thank the following organizations for access to data: Cementerio Museo San Pedro and Cementerio Universal Medellín for the Colombia data; the Macromorphoscopic Databank for the US reference data, which includes data from the following institutions and/or individuals: Dawnie Steadman and Lee Meadows Jantz (Department of Anthropology, University of Tennessee, Knoxville, TN), David Hunt (National Museum of Natural History, Smithsonian Institution, Washington, DC), Melody Ratliff (Pacific Islands data), and Bruce Anderson (Pima County

Medical Examiner's Office, Tucson, AZ) for providing access to skeletal collections. For assistance in data collection we would like to acknowledge the tireless efforts of the anthropology osteology laboratory assistants Olga Cecilia Londoño and Daniel Sampedro. The authors also wish to acknowledge the useful and insightful comments of the anonymous reviewer. Finally, we would like to acknowledge Kandus C. Linde for editorial comments on draft versions of this manuscript.

<REF>References

<BIBL>

- [1] Spradley MK, Jantz RL.;1; Ancestry Estimation in Forensic Anthropology: Geometric Morphometric versus Standard and Nonstandard Interlandmark Distances. J Forensic Sci 2016; <DOI> 10.1111/1556-4029.13081.</DOI>
- [2] Hefner JT, Spradley MK, Anderson BE.;1; Ancestry Assessment using Random Forest Modeling. J Forensic Sci 2014; 59(3): 583-89.
- [3] Hefner JT.;1; Cranial Variation nonmetric Ancestry and Estimating. J Forensic Sci 2009; 54(5): 985-995.
- [4] Hefner JT.;1; Cranial morphoscopic traits and the assessment of American Black, American White, and Hispanic Ancestry. In: Ta'ala S, Berg G, editors. Biological affinity in forensic identification of human skeletal remains: beyond black and white. <PL>Boca Raton, FL: </PL> <PN>Taylor &</PN> Francis Group, LLC. 2016: 27-42, [5] Hefner JT, Ousley SD.;1; Statistical Classification Methods for Estimating Ancestry using Morphoscopic Traits. J Forensic Sci 2014; 59(4): 883-890.
- [6] Hurst.;1; CV. Morphoscopic Trait Expressions Used to Identify Southwest Hispanics. J Forensic Sci 2012; 57(4):859-865.

- [7] Hefner JT, Pilloud MA, Black CJ, Anderson BE.;1; Morphoscopic trait expression in ``Hispanic" populations. J Forensic Sci 2015; 60(5):1135-1139.
- [8] Spradley MK, Jantz RL, Robinson A, Peccerelli F.;1; Demographic change and forensic identification: Problems in metric identification of Hispanic skeletons. J Forensic Sci 2008; 53: 21-28.
- [9] Maier CA, Zhang K, Manhein MH, Li X.;1; Palate shape and depth: A shape-matching and machine learning method for estimating ancestry from human skeletal remains. Journal Forensic Sci 2015; 59(6):1493-1501.
- [10] Tise ML, Spradley MK.;1; Anderson BE. Postcranial sex estimation of individuals considered Hispanic. J Forensic Sci 2013; 58(s1):S9-S14.
- [11] Birkby WH, Fenton TW, Anderson BE.;1; Identifying Southwest Hispanics using non-metric traits and the cultural profile. J Forensic Sci 2008;53(1):29–33.
- [12] Hefner, JT.;1; 2013. Morphoscopic Cranial traits and the assessment of American Black, White American, and Hispanic. In: Berg, G. E., Ta'ala, S.C. (Eds.), Biological Affinity in Forensic Identification of Human Skeletal Remains: Beyond Black and White. <PN>Taylor and</PN> Francis, <PL>New York</PL>.
- [13] Ross AH, Slice DE, Ubelaker DH, Falsetti AB.;1; Population affinities of 19th century Cuban crania: implications for identification criteria in South Florida Cuban Americans. J Forensic Sci 2004;49(1):1–6.
- [14] Dane.;1; LG. La población étnica y el Censo General. <PL>Colombia</PL>: una nación multicultural. Su diversidad étnica, 2005: 35.
- [15] Bushnell, D.;1; The making of modern Colombia: A nation in spite of itself. <PL>Berkeley: </PL> <PN>University of</PN> California <PN>Press</PN>, 1993.

- [16] Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, Woodward SR, Salas A, Torroni A, Bandelt HJ.;1; The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. PLoS ONE 2008; 3(3):e1764.
- [17] Bedoya G, Montoya P, Garcia J, Soto I, Bourgeois S, Carvajal L, Labuda D, Alvarez V, Ospina J, Hedrick PW, et al.;1; Admixture dynamics in Hispanics: a shift in the nuclear genetic ancestry of a South American population isolate. Proc Natl Acad Sci 2006; 103:7234–7239.
- [18] Rodas C, Gelvez N, Keyeux G.;1; Mitochondrial DNA studies show asymmetrical Amerindian admixture in Afro-Colombian and Mestizo populations. Hum Biol 2003; 75:13–30.
- [19] Salas A, Acosta A, Alvarez-Iglesias C, Cerezo M, Phillips C, Lareu MV, Carracedo A.;1; The mtDNA Ancestry of Admixed Colombian Populations. Hum Biol 2008; 20:584-591.
- [20] Morales V. Poblamiento y población en el valle de Aburrá y Medellín, 1541- 1951. En: Historia de Medellín, J. O. Melo, Editor. Compañía Suramericana de Seguros. Medellín. 1996: 57- 84.
- [21] Melo, JO. Medellín 1880 – 1930: los tres hilos de la modernización. En: Seminario: una mirada a Medellín y al valle de Aburra. Medellín. 1993: 13-20.
- [22] Carvajal-Carmona IG, d. Soto I, Pineda N, Ortiz-Barrientos O, Duque C, Ospina-Duque J, McCarthy M, Montoya P, m. Álvarez M, Bedoya G, Ruiz-Linares A.;1; Strong Amerind/White sex bias and a possible sephardic contribution among the founders of a population in northwest Colombia. Am J Human Genetics 2000; 67(5): 1287-1295.
- [23] Melton PE, Briceno I, Gomez A, Devor EJ, Bernal JE.;1; Crawford MH. Biological relationship between Central and South American Chibchan speaking populations: evidence from mtDNA. Am J Phys Anthropol 2007; 133:753–770.

- [24] Instituto Nacional de Medicina Legal y Ciencias Forenses en Colombia. <http://www.medicinalegal.gov.co/>. Accessed: 12 April 2016.
- [25] Ministerio del Interior. Dirección de Derechos Humanos. <http://www.mininterior.gov.co/sala-de-prensa/noticias/mininterior-encontro-28195-personas-no-identificadas-en-cementerios-del-pais> <C>Accessed 12 April 2016.</C>
- [26] Isaza J, Monsalve Vargas, T.;1; Biological characteristics of the osteological reference collection of the University of Antioquia: Preliminary Report. Bulletin of Anthropology, University of Antioquia 2011; 25 (42): 287-302
- [27] Wood C.;1; The age-related emergence of cranial morphological variation. Forensic Sci International 2015; 251:e1-e20.
- [28] Hefner JT, Dirkmaat DC, Ousley SD.;1; Morphoscopic traits and the assessment of ancestry. In: Dirkmaat DC, editor. A Companion to Forensic Anthropology. <PL>Malden, MA: </PL> <PN>Wiley-Blackwell</PN>, 2012: 287-310.
- [29] Ratliff M.;1; Evaluating morphoscopic trait frequencies of Southeast Asians and Pacific Islander [thesis]. Missoula (MT): The University of Montana, 2014.
- [30] R Development Core Team. R: a language and environment for statistical computing. <PL>Vienna, Austria</PL>: R Foundation for Statistical Computing, 2011.
- [31] SAS Institute Inc., SAS 9.1.3 Help and documentation. Cary, NC: SAS Institute Inc., 2000-2004.
- [32] Ruiz-Linares A, Adhikari K, Acuna-Alonzo V, Quinto-Sanchez M, Jaramillo C, Arias W, et al.;1; Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS Genet. 2014; 10(9):e1004572. <DOI>10.1371/journal.pgen.1004572.</DOI>

[33] Bravo ML, Moreno MA, Builes JJ, Salas A, Lareu MV, Carracedo A.;1; 2001. Autosomal STR genetic variation in Negroid Choco and Bogota populations. Int J Legal Med 2001; 115:102–104.

[34] Hudson R, Bushnell D.;1; The society and its environment. In, Colombia: A Country Study. Editor: R Hudson. Area Research Series, Federal Research Division: Library of Congress, 2010: 63-140.

[35] Hudson R.;1; Colombia: A Country Study. Area Research Series, Federal Research Division: Library of Congress, 2010.

[36] Spradley MK, Hefner JT.;1; Using non-metric traits to estimate ancestry in the 21st century. Proceedings of the American Association of Physical Anthropologists, 2016: 49.

</BIBL>

<Figure>Figure 1. Region/municipality map of Antioquia, Colombia.

<Figure>Figure 2. Joint display (biplot) of the row and column profiles derived from the contingency table data from Table 3 showing distribution of macromorphoscopic character states (columns) and birthplace cohort (rows). Profiles are plotted with respect to the existing principal axes.

<Figure>Figure 3. Diagrammatic representation of aNN algorithm using multilayer perceptron with 10 hidden layers.

<Figure>Figure 4. Scatterplot showing the first two principal coordinate axes of macromorphoscopic trait frequency profiles for several populations, including the pooled Colombian sample. See text for a description of the samples.

<Table>Table 1. Macromorphoscopic traits used in the current study

Trait	Abbreviation
Anterior Nasal Spine	ANS
Inferior Nasal Aperture	INA
Interorbit Breadth	IOB
Nasal Aperture Width	NAW
Nasal Bone Contour	NBC
Nasal Overgrowth	NO
Postbregmatic Depression	PBD
Posterior Zygomatic Tubercle	PZT
Zygomaxillary Tubercle	ZT
See Hefner (2009) for illustrations and definitions.	

<Table>Table 2. Groups and sub-groups used to identify Antioquia region.

Region/Municipality	Birthplace Cohort	Social group	Female	Male
Nordeste	Nordeste	C	1	3
Norte	Norte	D	4	13
Occidente	Occidente	A	9	7
Oriente	Oriente	B	5	21
Surdeste	Surdeste	B	12	23
Uraba	Uraba	A	3	2
Valle De Aburra	ValleAburra	E	12	59
Otre	Other	F	8	18
Pooled Sample*	--	--	67	177

*Note: Pooled sample includes individuals identified only as Colombian.

<Table>Table 3. Frequency distribution of morphoscopic traits in crania from Colombia.

ANS			1	2	3				NO		0	1				
<i>Pooled Sample</i>			12	99	127				<i>Pooled Sample</i>		32	194				
<i>Nordeste</i>			-	2	2				<i>Nordeste</i>		1	3				
<i>Norte</i>			-	7	9				<i>Norte</i>		3	13				
<i>Occidente</i>			1	7	6				<i>Occidente</i>		2	13				
<i>Oriente</i>			2	10	13				<i>Oriente</i>		1	22				
<i>Surdeste</i>			1	11	22				<i>Surdeste</i>		6	27				
<i>Uraba</i>			1	1	3				<i>Uraba</i>		3	2				
<i>Valle</i>			2	26	43				<i>Valle</i>		6	59				
<i>Otre</i>			1	11	14				<i>Otre</i>		2	22				
INA			1	2	3	4	5		PBD		0	1				
<i>Pooled Sample</i>			11	29	73	56	71		<i>Pooled Sample</i>		146	97				
<i>Nordeste</i>					2	1	1		<i>Nordeste</i>		2	2				
<i>Norte</i>			1	2	6	5	3		<i>Norte</i>		10	7				
<i>Occidente</i>			1	1	5	3	4		<i>Occidente</i>		9	7				
<i>Oriente</i>			1	2	5	8	9		<i>Oriente</i>		12	14				
<i>Surdeste</i>			3	5	7	10	9		<i>Surdeste</i>		24	11				
<i>Uraba</i>			1	2	-	-	2		<i>Uraba</i>		1	4				
<i>Valle</i>			3	7	22	16	23		<i>Valle</i>		42	28				
<i>Otre</i>			-	3	6	6	11		<i>Otre</i>		16	10				
IOB			1	2	3				PZT		0	1	2	3		
<i>Pooled Sample</i>			55	98	91				<i>Pooled Sample</i>		1	114	104	25		
<i>Nordeste</i>			-	2	2				<i>Nordeste</i>		-	3	-	1		
<i>Norte</i>			9	1	7				<i>Norte</i>		-	8	8	1		
<i>Occidente</i>			5	4	7				<i>Occidente</i>		-	7	8	1		
<i>Oriente</i>			6	8	12				<i>Oriente</i>		-	13	13	-		

<i>Surdeste</i>			7	15	13				<i>Surdeste</i>		-	16	15	4		
<i>Uraba</i>			1	1	3				<i>Uraba</i>		-	5	-	-		
<i>Valle</i>			15	36	20				<i>Valle</i>		1	33	24	13		
<i>Otre</i>			5	11	10				<i>Otre</i>		-	9	16	1		

<Table>Table 3 (cont). Frequency distribution of morphoscopic traits in crania from Colombia.

NAW			1	2	3				ZS		0	1	2			
<i>Pooled Sample</i>			71	108	62				<i>Pooled Sample</i>		32	168	42			
<i>Nordeste</i>			1	3	-				<i>Nordeste</i>			4				
<i>Norte</i>			7	7	3				<i>Norte</i>		3	13	1			
<i>Occidente</i>			2	4	8				<i>Occidente</i>		5	9	2			
<i>Oriente</i>			8	11	7				<i>Oriente</i>		5	17	4			
<i>Surdeste</i>			10	17	8				<i>Surdeste</i>		2	25	8			
<i>Uraba</i>			2	-	3				<i>Uraba</i>		2	1	2			
<i>Valle</i>			23	34	13				<i>Valle</i>		6	53	11			
<i>Otre</i>			6	13	7				<i>Otre</i>		5	16	5			
NBC		0	1	2	3	4										
<i>Pooled Sample</i>		44	124	3	30	41										
<i>Nordeste</i>		-	1	-	3	-										
<i>Norte</i>		1	9	-	1	6										
<i>Occidente</i>		5	8	1	-	2										
<i>Oriente</i>		1	16	-	3	6										
<i>Surdeste</i>		9	17	-	6	3										
<i>Uraba</i>		3	1	-	-	1										
<i>Valle</i>		8	44	1	9	9										
<i>Otre</i>		7	9		4	6										

<Table>Table 4. Kruskal-Wallis ANOVA, by birth cohort

Dependent Variable	<i>H</i>	<i>N</i>	<i>p</i>
Anterior Nasal Spine	3.07	195	0.879
Inferior Nasal Aperture	4.01	196	0.778
Interorbital Breadth	4.17	200	0.761
Nasal Aperture Width	8.98	197	0.254
Nasal Bone Contour	12.42	200	0.088
Nasal Overgrowth	14.28	185	0.047

Postbregmatic Depression	6.42	199	0.492
Posterior Zygomatic Tubercle	6.87	200	0.442
Zygomaxillary Tubercle	6.78	200	0.453

<Table>Table 5. Polychoric correlation matrix.

	ANS	INA	IOB	NAW	NBC	NO	PBD	PZT
INA	0.200879	-						
IOB	-0.099880	-0.017931	-					
NAW	-0.214315	-0.209984	0.283458	-				
NBC	0.010970	0.030284	-0.233207	-0.188365	-			
NO	0.158827	0.127569	-0.169311	-0.156838	0.078421	-		
PBD	-0.076703	0.063029	-0.037244	0.002218	0.031882	0.000686	-	
PZT	-0.005957	0.029251	0.150030	0.005352	0.009346	-0.012384	-0.043472	-
ZS	0.130447	0.036462	0.008823	-0.213875	-0.019760	-0.078980	-0.076140	-0.049593
Bold values significant at the $p < 0.001$.								

<Table>Table 6. Artificial neural network model summaries.

Profile	Training	Tuning	Validation	Inputs	Hidden(1)	Hidden(2)
MLP 6:20-10-8:1	82.8%	11.6%	18.6%	9	10	0
Linear 1:3-8:1	44.1%	23.3%	30.2%	9	0	0
Linear 3:7-8:1	44.1%	20.9%	27.9%	9	0	0
RBF 6:20-1-8:1	44.1%	23.3%	30.2%	9	1	0
RBF 6:20-2-8:1	43.0%	23.3%	30.2%	9	2	0

<Table>Table 7. Artificial neural network model summaries comparing a pooled Colombian sample to US reference groups

Profile	Training	Tuning	Validation	Inputs	Hidden(1)	Hidden(2)
MLP 8:25-10-6:1	71.4%	49.6%	48.4%	8	10	0
Linear 5:15-6:1	55.3%	54.8%	46.0%	5	0	0
Linear 6:18-6:1	55.5%	49.2%	48.4%	6	0	0
RBF 8:25-71-6:1	66.8%	50.0%	48.0%	8	71	0
RBF 8:25-72-6:1	67.2%	51.2%	50.4%	8	72	0

<Table>Table 8. Classification matrix derived from the MLP 8:25 artificial neural network

Group	Colombian	Asian	Black	Pacific Island	Hispanic	White
Colombian	60.00	1.67	10.00	1.67	10.00	16.67
Asian	0.00	57.14	0.00	0.00	28.57	14.29
Black	9.52	6.35	50.79	7.94	6.35	19.05
Pacific Island	0.00	6.67	33.33	53.33	6.67	0.00
Hispanic	16.13	11.29	16.13	3.23	29.03	24.19
White	9.30	4.65	2.33	6.98	23.26	53.49

<Table>Table 9. Distance matrix derived from the principal coordinate axes analysis.

Group	Colombian	Asian	Black	Pacific Island	Hispanic
Asian	42.79991	-			
Black	92.86565	19.01395	-		
Pacific Island	52.00145	9.00602	11.62072	-	
Hispanic	26.11233	18.88864	60.63530	26.54032	-
White	40.25702	24.96489	70.49120	27.68553	7.51913

TDENDOFOCTD