JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, September 2011, Vol. 56, No. 5
doi: 10.1111/j.1556-4029.2011.01875.x
Available online at: onlinelibrary.wiley.com

PAPER

ANTHROPOLOGY

Cris E. Hughes,[1,†] *Ph.D.; Chelsey A. Juarez,*[1] *Ph.D.; Taylor L. Hughes,*[2] *Ph.D.;*
Alison Galloway,[1] *Ph.D.; Gillian Fowler,*[3] *M.S.; and Shirley Chacon,*[4] *B.A.*

A Simulation for Exploring the Effects of the "Trait List" Method's Subjectivity on Consistency and Accuracy of Ancestry Estimations*

ABSTRACT: The nonmetric "trait list" methodology is widely used for estimating ancestry of skeletal remains. However, the effects of the method's embedded subjectivity on subsequent accuracy and consistency are largely unknown. We develop a mathematical simulation to test whether variation in the application of the "trait list" method alters the ancestry estimation for a given case. Our simulation explores how variations in (i) trait selection, (ii) number of traits employed, and (iii) ancestry choice thresholds affect the ancestry estimation of an unidentified skeleton. Using two temporally and geographically diverse samples, the simulation demonstrates that trait selection, trait quantity, threshold choices, and the exclusion of high-frequency traits had minimal effect on estimation of general ancestry. For all data sets and Runs, $\text{Accuracy}_{\text{AS}}$ was maintained above 90%. The authors close with a discussion on the logistical issues present when choosing traits, and how to avoid ancestry bias.

KEYWORDS: forensic science, forensic anthropology, inter-observer method bias, heritable quantitative traits, mathematical model, Asian continental ancestry group

The 2009 National Academy of Sciences report (*Strengthening Forensic Science in the United States: A Path Forward*) calls for a reassessment of the principles and assumptions upon which the various forensic disciplines are based. In this paper, we answer that call in the area of nonmetric ancestry assessment in forensic anthropology. The nonmetric trait list methodology, in addition to metric analyses, is widely used for assessing ancestry of skeletal remains (1–4). We first discuss how nonmetric traits are applied to estimate ancestry and the current limitations with this application. Next, we develop a mathematical simulation to test whether variation in the application of the trait list method alters the ancestry description for a given case. We discuss how to (i) move toward standardization of the trait list method, (ii) work with the method's embedded nuances and biases, and (iii) apply the trait list method responsibly.

Forensic anthropologists employ two types of nonmetric traits for qualitative ancestry estimation, which we define as "binary" and "morphoscopic." Binary nonmetric traits, such as sutural ossicals or postbregmatic depressions, are expressed as one of two states: "present" or "absent." Morphoscopic nonmetric traits have multiple states, described by the degree of expression, or the shape of the trait. Examples of morphoscopic trait expression include the degree of shoveling on shovel-shaped incisors (4–7) or the shape of the eye orbit as round, rectangular, or sloped (8–10). When forensic anthropologists analyze skeletal material, the presence (and/or degrees of presence) of a nonmetric trait is associated with a particular ancestry, while the absence of a trait generally carries no estimative weight. Similarly, morphoscopic trait shapes are associated with a particular ancestry. In the example cited above on orbit shape, round, rectangular, and sloped shapes are more frequent in individuals of Asian, African, and European ancestries, respectively (8,10). For morphoscopic traits with only two states, such as a round or oval external auditory meatus, two ancestries are associated with one state while the other ancestry is associated with the alternate state (10).

When assessing ancestry from nonmetric traits, forensic anthropologists employ what the authors call the "trait list method." A checklist of traits is completed, noting each trait's state expression. Next, the forensic anthropologist studies the distribution of the trait states among the three ancestries (Asian, African, and European). Trait states are not exclusive to a single ancestry; instead, the trait list method is grounded in the belief that individuals of a specific ancestry more often express a particular trait state than other ancestries. Because nonmetric traits are considered heritable, albeit to

---

[1] Department of Anthropology, University of California at Santa Cruz, 1156 High Street, Santa Cruz, CA 95060.
[2] Department of Physics, University of Illinois at Urbana-Champaign, 110 W. Green Street, Urbana, IL 61801.
[3] School of Natural and Applied Sciences, University of Lincoln, Brayford Pool, Lincoln LN6 7TS, UK.
[4] Forensic Anthropology Foundation of Guatemala, Avenida Siemon Cañas 10-64 Zona 2, Guatemala City, Guatemala.

*Funding support received by Dr. Taylor Hughes from the NSF under Grant DMR 0758462 at the University of Illinois, from the ICMT. Funding support received by Cris Hughes from the NSF Graduate Research Fellowship Program.

[†] Current address: University of Illinois at Urbana-Champaign, Department of Anthropology, 607 Matthews Ave., Davenport Hall, Urbana, IL 61801.

Received 17 April 2010; and in revised form 6 July 2010; accepted 8 Aug. 2010.

various degrees (11–18), genetic drift and gene flow must be considered when accounting for shifts in distributions of trait state expressions. The distribution of nonmetric traits is complex, and the relationship of genotype to phenotype is not fully understood. The polygenic nature of these traits maintains a complex path for variation in expression (19,20). When using the trait list method, the premise of trait state distribution and the influential genetic nuances can often be lost in the application. For example, a forensic anthropologist employs 20 traits to estimate ancestry of a given case; half of these trait states are consistent with Asian ancestry, and half are consistent with European ancestry. The forensic anthropologist may designate the ancestry of this individual as "admixed European and Asian." This designation implicitly relies on the concept that trait states are unique to a given ancestry. Combinations of trait state ancestries are interpreted to mean that an individual is of "mixed" ancestry, ignoring the possibility that this individual represents a nonconforming phenotype. An individual with these 20 trait states could have ancestral ties to any global location. All gene pools have the genetic potential to create individuals with any combination of trait states, with the exception of extreme cases of isolated island or indigenous populations (21). Samples studied to demonstrate trait frequency differences among ancestries cannot rely on pristine populations because pure "Asian," "African," or "European" ancestries never existed. As Weiss and Long (22, pp. 702–707) note:

> Even if one were to grant that contemporary data only provide estimates of, rather than actual, ancestral parental genotype frequencies, there is no reason to think that there ever were isolated, homogeneous parental populations at any point in our human past. Why do we so uncritically accept admixture-based analyses of global samples that give the appearance that human variation is clustered into a few major populations, portrayed in much the same way as classical races?

Once one understands the genetic and environmental factors influencing trait expression and distribution in a population, it seems unlikely that forensic anthropologists could successfully designate ancestry using nonmetric traits. The relative accuracy of the trait list method is unknown because there is no requirement or community space for forensic anthropologists to document the comparison of racial identities of adjudicated cases to the ancestry estimations generated by nonmetric methods. As a result of this deficiency, studies have attempted to refine the method by testing the reliability of particular traits in estimating ancestry. Hefner et al. (1) empirically demonstrated that frequencies of particular nonmetric traits are inconsistent with those listed in the foundational forensic literature as being associated with specific ancestries. Some of these core nonmetric studies used small samples ($n = 10$) with minimal sex diversity to establish associations between trait states and ancestries (10). This lack of adequate sampling cannot support an empirical discussion on the universal variation of trait expression. In addition to such inconsistencies, there is no standardization within forensic anthropology in applying the trait list method. Variations in which traits are used, the number of traits used, and the cut-off choices between one ancestry assessment and another must be studied.

Recent work attempts to avoid trait list methodology altogether (1,2,23,24). Instead, objective statistical methods such as discriminant analysis and neural networking are applied to nonmetric data. These methods negate the association of trait states to ancestries and provide a level of statistical significance through which to assess the result's accuracy (2,24). However, many of the newly proposed statistical methods are still based on subjective assessment of morphoscopic trait states at the initial data collection stage. The ambiguity of assigning qualitative descriptions to morphoscopic traits may be a serious source for subjectivity and nonstandardization, in that "where in the context of any particular systematic comparison does 'round' stop and 'oval' begin?" (25, p. 1). It is the opinion of the authors that the complexity and embedded subjectivity of choosing one trait state over another cannot be entirely eliminated by the creation of detailed descriptions and intense peer training. Statistically based methods are a necessary direction for nonmetric trait methodology in forensic anthropology, and the recent work cited above is progressing nonmetric analysis toward this goal. As a discipline, forensic anthropology must challenge itself to explore the subjectivity of these and future methodologies.

Mathematical Simulation of Trait List Analysis

Ancestry assessment from skeletal remains involves a series of decisions. There is no protocol for how to interpret a suite of traits into an ancestry estimation (2). For example, if 10 of 10 observed traits express the Asian state, the associated skeleton would typically be classified as having Asian ancestry. But what if only nine or eight or seven of 10 are associated with Asian ancestry? What is the threshold for considering the conventional admixture estimations (regardless of their validity) when using the "trait list" methodology?

In this paper, we use a mathematical simulation (developed by author TLH) that realistically represents the possible analytical variations of trait list ancestry estimation. Our simulation explores how (i) trait selection, (ii) number of traits employed, and (iii) ancestry choice thresholds affect the ancestry estimation of a skeleton. The simulation refines the widely used trait list method for ancestry estimation, which is known to suffer from empirical weakness and lack of standardization. The relative accuracy of the trait list method in actual casework has not been comprehensively examined. While valid critiques have been made (26–29), the method has endured because of its ease of application and the familiarity of certain traits to the anthropologist. Here, we are able to simulate the decision-making process of the anthropologist, taking into account possible variations. We assess how that variation in the application of the trait list method influences the method's outcome, that is the ancestry estimation.

Our simulation does not test how well an individual's or sample's nonmetric trait proportions conform to the paradigms of the trait list method. Instead, we want to know how variations in the decision-making process itself influence the ancestry estimation of any given individual. The broader application of our simulation examines the trends in ancestry estimations based on different populations and samples. In turn, this discussion assesses the use and limitations of the trait list method, of ancestry categories, and of specific nonmetric traits.

We begin our presentation of the simulation with the Materials and Methods section of this paper, in which we explain the groundwork behind the simulation we have developed. In the following section, we demonstrate how the simulation assesses the influence of methodological variation by testing it on 175 individuals from two different skeletal samples. Finally, we discuss the forensic implications of this simulation on the use of the trait list method.

Materials and Methods

To simulate the process of ancestry estimation in case analysis, we must have skeletal remains to analyze. It is important to emphasize that *any* sample could have been used for this, as the goal of this paper is not to see how the samples conform to the embedded paradigms of the trait list method but instead how the embedded subjectivity in the act of applying the trait list method alters the ancestry estimation outcome for a given skeleton. Obviously, the consistency of an ancestry estimation is dependent upon how well or poor a skeleton conforms to the paradigms of its particular ancestry. In the samples provided, we include individuals typically included under Asian ancestry: Native Americans and indigenous Guatemalans. However, there is a large range of diversity within these two samples in regard to the number of traits conforming to the Asian state. For some individuals, all traits conform to the Asian state while in others, only a few traits conform to the Asian state. Therefore, these two samples allow for an excellent platform for testing the consistency of the trait list method's ancestry estimations; if the consistency is low, this cannot be attributed to nonconforming skeletal remains as a range of conformity is present in our samples. On the other hand, if the ancestry estimations remain consistent throughout the simulations, this result cannot be attributed to a skeletal sample that wholly conforms to the paradigms for Asian ancestry.

*Skeletal Series: Historic and Prehistoric Northern California Native Americans (*n *= 107)*

Native American samples of adult males (*n* = 87) and females (*n* = 20) are included in this project. Authors CEH and CAJ collected these data. A group association of "NCNA" (Northern California Native Americans) was assigned to the samples that represent diverse temporal and geographic ranges within California's prehistory and more recent history. Alameda county samples consist of skeletal remains from sites ALA 309, ALA 307, ALA 328, ALA 329, and ALA 13. Calibrated radiocarbon assessments estimate varied prehistoric occupation dates for these sites from 3030 BC to AD 1395 (30–33). Contra Costa County samples come from sites CCO-300 and CCO-138 with uncalibrated radiocarbon assessments ranging from AD 150 to AD 1425 (30,31,33). San Joaquin county samples derive from site SJO-68 whose uncalibrated radiocarbon dates are estimated at 4350 BC to 2980 BC.

*Skeletal Series: Modern Guatemalan Maya (*n *= 221)*

The modern sample consists of adult male (*n* = 196) and female (*n* = 25) indigenous Guatemalan Maya (GUA). All skeletal remains used for this sample were victims of the Guatemalan genocide (1970s–1990s) and are in the custody of the Forensic Anthropology Foundation of Guatemala (FAFG). The FAFG gave permission for the collection of all data used in this study, collected by author CEH. The Guatemalan sample includes individuals from multiple indigenous Maya communities from the Guatemalan states Quiche, Chimaltenango, and Baja Verapaz. During the time range of the sample, the majority of Guatemalan Maya had limited contact with urban areas, and community endogamy practices were common (34). Marriage of Maya with Ladino and other non-Mayan individuals, even beyond their communities or linguistic groups is documented as rare (35,36). Anthropological interviews with Marco Perez (Director of Cultural Anthropology Branch, FAFG) have concluded that the majority of the cases included in the sample are indigenous Guatemalans. However, because two of the sites were military camps, we cannot currently confirm whether other populations (such as Ladinos from Guatemala City) are represented in this sample. DNA identification is under way (personal communication with John Crews, director of FAFG DNA laboratory), which will elucidate the assumptions about ethnic identity of the skeletal remains.

For both NCNA and GUA samples, crania with missing data are included, as it was difficult to find large undamaged samples. Crania exhibiting morphological deformation were not included in the samples.

Data Collection and Preparation

A combination of 25 binary and morphoscopic traits were collected for the two samples. The traits, their corresponding expressions, and associated ancestry were derived from Ossenberg (37) and Rhine (10) and are provided in the Appendix. While this study uses subjective morphometric traits as a point of analysis, the focus of this research is not in the subjectivity of the traits, but rather on other inconsistencies embedded in the trait list method. Only traits whose different states have been published and directly linked to a specific ancestry are included. Because of a significant lack of complete skulls in the NCNA samples, the authors chose not to collect mandibular data. Several postcranial traits were collected for both GUA and NCNA samples; however, the present analysis focuses only on cranial traits. The number of traits collected for each skull is contingent upon the conditions of the crania. The greatest number of traits collected from an individual was 19 and the least was 1.

We used the software JMP 7.0 (SAS Institute Inc., Cary, NC) and SYSTAT 12.0 (Systat Software, Chicago, IL) for statistical analyses. Tests for inter- and intra-observer error were conducted during data collection resulting in no significant differences at the alpha 0.05 level. Data for both sides were collected. To determine whether to use individual or side count methods, we tested for statistical significance of associations between trait expression and right and left sides using chi-square analysis. No significant associations were found at the alpha = 0.05 level. Therefore, we chose the side count method, using the left sides' expression for all bilateral traits. Chi-square tests for sex effects were conducted for the NCNA sample, and no significant associations between sex and trait expression at the alpha = 0.05 level were present.

Trait Weighting

If we simplify the trait list method, it is merely a tallying system of trait states to represent the proportion of each ancestry. To simulate realistic variations in ancestry assessments as described above, we must place assumptions on weighting traits. In this simulation for traits whose states are exclusively associated to one of the three ancestries, each trait state is weighted equally. For example, round, rectangular, or slanted orbit shape are each exclusively associated with individuals of Asian, African, and European ancestry; all trait states contribute the same amount to each ancestry. For traits whose states are shared between two ancestries the trait state's contribution is equally divided among both ancestries. For example, Rhine (10) suggests that "American Caucasoid" and "American Blacks" both share a round external auditory meatus, while "Southwestern Mongoloids" have an elliptical external auditory meatus. In this example, if a cranium exhibited a round external auditory meatus, then half the weight would be assigned to "American Blacks" and the other half assigned to "American Whites." In contrast, if the external auditory meatus was elliptical, the entire weight of the state would go to "Southwestern Mongoloids." This allows us to put individuals with the same number of present/known traits on equal footing when the specific subsets of traits collected differ for each individual in our samples.

Selecting a Data Set to Test Variability

The crania included in both the GUA and NCNA samples exhibited varying amounts of the 25 traits on which we could collect data. To accurately test variability of ancestry assessment, it is important to use as many traits *and* as large a sample as possible. However, a balance is required, as greatly increasing the number of traits used in a trait list analysis can severely limit the number of individuals within our sample meeting this requirement. We therefore restrict our sample to a reduced data set comprised of individuals with a minimum number of traits, designated as N_{trait}. After reducing the sample to a data set of individuals with at least N_{trait} traits, the *i*th individual in the sample has a number of traits N^{i}_{trait} that is greater than or equal to N_{trait}. For example, if $N_{\text{trait}} = 14$, only individuals from our sample that have data for *at least* 14 traits will be included in the reduced data set. To simulate variation in the ancestry estimation process using our data set, we assume that an observer only has access to or chooses to use a set of traits numbering less than or equal to N_{trait}. We designated N_{c} as the number of traits used to conduct the trait list method. N_{c} is fixed for the whole sample and is less than or equal to N_{trait}. Once N_{c} is fixed, the trait list method is performed on the data set using the following algorithm:

- Begin with individual 1 with N^{i}_{trait} traits.
- Take a subset of N_{c} traits out of the total N^{1}_{trait} traits and perform an ancestry estimation as detailed below.
- Repeat for all possible subsets of N_{c} traits for individual 1.
- Repeat for every individual in the data set.

As an explicit example, let $N_{\text{trait}} = 14$, $N_{\text{c}} = 10$. If individual 1 has $N^{1}_{\text{trait}} = 16$ traits, then the total number of subsets from which ancestry estimations are carried out for individual 1 is $C^{16}_{10} = 8008$, that is the number of combinations (C) of 10 traits out of a possible 16. This process mimics the selection of a certain subset of traits ($N_{\text{c}} = 10$) to apply in the trait list method. In actual casework, the selection of the subset of traits may be based on an array of circumstances, such as the analyst's preference/confidence in particular traits, or the absence of/damage to portions of the cranium. Simulating variation in the choice of traits used out of those possible for a given cranium helps us assess the effects of this choice on the derived ancestry estimation. In our analyses below, we also examine the effects of changing N_{c}.

Describing Trait Frequencies in terms of Ancestry

Three ancestries typically used in the trait list method are the basis of this simulation. A variety of descriptive labels are used in forensic literature (38,39); we chose to use the labels European (EU), Asian (AS), and African (AF). For each subset of N_{c} traits for an individual, the simulation calculates the percentage composition of EU, AS, and AF trait states. Based on our assumption that all traits are weighted equally in the application of the trait list method, the percentages are used to describe the amount of EU, AS, and AF contribution to an ancestry estimation. Our simulation describes the contribution of each ancestry to an individual's total ancestry estimation using one of the four "Proportion Descriptors" described below (for X = EU, AF, or AS):

- Full Ancestry X: Ancestry X's percentage is the only ancestry significantly reflected in the trait list analysis.
- Majority Ancestry X: Ancestry X's percentage comprises the major component of the trait list analysis, but not at a high or low enough percentage to be considered "Full Ancestry X" or "Minority Ancestry X," respectively.
- Minority Ancestry X: Ancestry X's percentage comprises the minor component of the trait list analysis, but not at a high or low enough percentage to be labeled as "Majority Ancestry X" or "Not Ancestry X."
- Not Ancestry X: Ancestry X's percentage does not comprise a large enough component to be significantly regarded.

The final step of the simulation is generating an actual ancestry estimation by representing the combination of the EU, AF, and AS Proportion Descriptors into a cohesive ancestry estimation description. We generate 26 possible ancestry estimations listed in Table 1. These ancestry estimations do not represent a standardized labeling system used by forensic anthropologists—such standardization does not presently exist. The labeling system developed here was chosen based on the need to provide an ancestry estimation description that can represent all of the possible scenarios generated by the interaction of the Proportion Descriptors and the threshold variations (discussed below).

With the exception of a few, these descriptions are straightforward. We define the ancestry estimation of "Equal Ancestry X and Y" to be the case when exactly 50% of the trait states are consistent with X and 50% are consistent with Y. The ancestry estimation "Primarily X" is best explained by the following example. "Primarily AS" occurs when ancestry AS is represented by a percentage of traits consistent with the Proportion Descriptor "Majority Ancestry AS," but ancestries EU and AF are represented by a percentage consistent with the Proportion Descriptor "Not Ancestry EU" and "Not Ancestry AF." Because the percentage of AS traits is not great enough to be described at "Full AS," we explain such a circumstance as "Primarily AS."

Studying Trait Quantity and Trait Selection Variation

One of the uses for our simulation is to study how variability of trait selection and number of traits used in an ancestry analysis influences the ancestry estimation. For example, a skull has 30 possible traits for trait list method analysis; forensic anthropologist X uses 25 traits, forensic anthropologist Y also uses 25 traits, but some traits differ from those used by anthropologist X, and

TABLE 1—*The 26 possible ancestry estimations generated from the model.*

Full AS (AS)*	EU maj, AF min (EU_af)
Full EU (EU)	EU maj, AS min (EU_as)
Full AF (AF)	EU maj, AS and AF min (EU_as_af)
Equal AS, EU, AF (AS_EU_AF)	EU, AF, and AS min (eu_af_as)
Equal AS EU (AS_EU)	EU and AS maj, AF min (EU_AS_af)
Equal AS AF (AS_AF)	EU and AF maj, AS min (EU_AF_as)
Equal AF EU (AF_EU)	AS and AF maj, EU min (AS_AF_eu)
AS maj, EU min (AS_eu)	Primarily AS (P_AS)
AS maj, EU and AF min (AS_eu_af)	Primarily AF (P_AF)
AS maj, AF min (AS_af)	Primarily EU (P_EU)
AF maj, AS min (AF_as)	Primarily AS and AF (P_AS_AF)
AF maj, EU min (AF_eu)	Primarily AS and EU (P_AS_EU)
AF maj, EU and AS min (AF_eu_as)	Primarily AF and EU (P_AF_EU)

*Parentheses indicate the abbreviations used by the simulation and throughout this paper.

anthropologist Z uses 29 traits. Will the ancestry estimations be similar? The simulation answers such a question by following these steps:

- Choose the number of traits to use for the analysis N_{trait}. This can be however small or large the simulation's user chooses and is often determined by how many traits one was generally able to collect from the crania for a given sample. Selecting too large a number of traits can result in a diminished sample size. Here let us use 14 traits. Crania in the sample that have *at least* 14 traits will comprise the N_{trait} data set and are used for analyses.
- Next, choose the N_c to be applied to the data set. This is the actual number of traits that will be used to conduct the trait list method ancestry estimation. Recall that N_c must be less than or equal to N_{trait}. Here we choose 13.
- To test whether analyzing an equal number of, but different traits from a single cranium can result in differing ancestry estimations, we generate all possible 13-trait combinations per individual in the data set. For example, if Individual A has 15 traits, the program then generates all possible combinations of 13 traits from the 15 available traits, yielding 105 combinations for Individual A. The program repeats this step for all individuals in the $N_{\text{trait}} = 14$ data set.
- The program has now generated all possible trait combinations for every individual in the data set. The simulation then describes each of these combinations as one of the 26 possible ancestry estimation descriptions.
- The program displays the individual and sample distributions for ancestry assessments over the 26 categories. For example, of the 105 13-trait combinations for Individual A, the ancestry estimations were distributed between two ancestry categories; "Maj AS, Min EU" occurred 40 times and "Maj AS, Min AF" occurred 65 times.

This process shows that the assessments of anthropologists X and Y, who use the same number, but different subsets of traits, may yield different results if the traits they select are associated with EU or AF ancestry. The simulation allows us to quantify the variation in the results based on the selection of traits for assessment. The program can generate ancestry assessments for any number of traits (N_{trait}). To test whether the discrepancy between Anthropologists X and Z would result in different ancestry estimations, we simply change the number of traits (N_c) used in the analysis.

Studying Threshold Variations

In a single trait list method analysis, the forensic anthropologist must make a series of decisions for choosing how to describe the contribution of AF, AS, and EU ancestries. Variation in this decision-making process is a potential site for variation in the trait list method. For example, a forensic anthropologist may believe that, to describe an individual as "Full Ancestry X," 90% of their traits must express states consistent with Ancestry X. In contrast, another forensic anthropologist may use a lower threshold of 80%. We illustrate these thresholds schematically in Fig. 1 for Models A and B, which exhibit two different thresholds. The blue arrows indicate the thresholds for defining the four Proportion Descriptors: Not, Minority, Majority, and Full. Both models use the same trait list method results from a 20-trait analysis. Of the 20 traits used in the analysis, 85% conformed to the Asian ancestry, 15% conformed to the European ancestry, and 5% conformed to the African ancestry. On both models, the black arrows indicate where each ancestry's percentage would fall within the set thresholds. Based on these thresholds, Simulation A interprets the contribution of each ancestry as follows: Asian majority, European minority, and No African ancestry, conforming to the ancestry estimation label "AS maj, EU min." In contrast, Simulation B interprets the ancestries' contributions as Asian Full, No European, and No African ancestry, resulting in the ancestry estimation label "Full AS." As seen in Fig. 1, when using a fixed set of data (20 traits), but slightly altering the thresholds defining each Proportion Descriptor can influence the ancestry estimation generated by the simulation. To explore how threshold variations can influence ancestry estimations, our mathematical simulation provides flexibility for adjusting thresholds. However, a series of rules pertaining to the selection of thresholds are required for the simulation to ensure that only realistic and consistent ancestry estimations are provided.

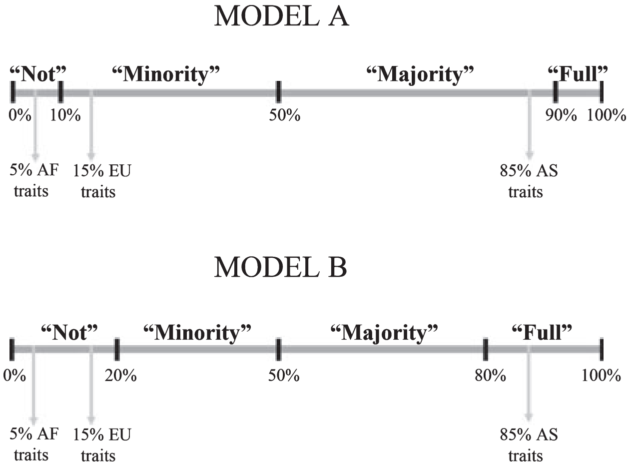

FIG. 1—*Models A and B demonstrate how differences in threshold cutoffs influence final ancestry estimation.*

In Fig. 1, each Proportion Descriptor (Full, Majority, etc.) has its own thresholds. Clearly, there is a lower limit at 0% and an upper limit at 100%. The boundary between Majority Ancestry X and Minority Ancestry X is fixed at 50%. In Fig. 1, Simulation A's threshold for Full Ancestry X is 90%, and the threshold for Not Ancestry X is 10%. Because 0, 50, and 100% are fixed boundaries, the only thresholds fluctuating are the upper threshold separating Majority and Full, and the lower threshold separating Minority and Not. Therefore, when describing the chosen thresholds, only these two need to be noted. We use the abbreviation 90/10 to represent when the upper threshold is set to 90%, and the lower threshold is set to 10%. This kind of abbreviation will be used throughout the rest of the article. To be explicit, for 90/10, we have:

- Full Ancestry X = 90–100% of traits express state consistent with Ancestry X.
- Majority Ancestry X = 50–89.9% of traits express state consistent with Ancestry X.
- Minority Ancestry X = 10.01–49.9% of traits express state consistent with Ancestry X.
- Not Ancestry X = 0–10% of traits express state consistent with Ancestry X.

The following restrictions pertain to the adjustment of thresholds:

- There is a fixed boundary at 50% trait composition separating "Majority" and "Minority" Proportion Descriptors. In other words, for an individual to have a "Minority Ancestry X," no *more* than 49.9% of the traits express states consistent with Ancestry X. Conversely, for an individual to have a "Majority

Ancestry X,” no less than 50% of the traits can express states consistent with Ancestry X.

- The upper threshold may be set to any percentage between 75% and 100%.
- The lower threshold may be set to any percentage between 25% and 0%.
- Full Ancestry is an exclusive ancestry estimation, so the upper and lower thresholds must sum together to be >100%. Without such a restriction, an individual could be labeled as “Full X with Y minority.” This ancestry description violates our definition of Full Ancestry X, which does not allow for the inclusion of other ancestries.

The thresholds can be adjusted by any numerical increment within the parameters outlined above. We ran our samples through the simulation to generate ancestry assessments using varying thresholds per Run. This allowed us to study how ancestry assessments are affected by this variation. However, not all incremental changes in thresholds will change an ancestry assessment, as this is dictated by the number of traits used in the analysis. For example, in the case where nine of 10 traits states are consistent with Ancestry X and one of these is shared, the total resulting weight is 8.5. In other words, 8.5 of 10 (85%) trait states are consistent with Ancestry X. If the threshold for “Full Ancestry X” is 84% (84/16), then this collection of traits results in the ancestry assessment of “Full Ancestry X” because 85% is included above the threshold of 84%. The next possible number of traits consistent with Ancestry X can only be a 0.5 increment, such as eight; when eight of 10 trait states are consistent with Ancestry X, this results in 80%. Because there are no possible outcomes for trait weighting between eight and 8.5, any threshold placed between 80% and 85% (such as 82/18, 84/16) will yield identical results. Thus, we must be careful to take the effects of the discrete nature of the data into account before drawing conclusions.

Results

In this section, we review the simulation’s results on how variation in thresholds, trait quantity, and specific traits selected for use affect the ancestry estimation of a skeleton. We will present our results separately for each of these three factors.

Threshold Variation

To test the effect of threshold variation on ancestry estimations, we ran several scenarios on subsets of the GUA and NCNA samples. To isolate any effects of threshold variation, trait number and sample size were held constant. A sample size of 96 GUA individuals was derived from the cut-off $N_{\text{trait}} \geq 14$, that is, 96 individuals in the GUA sample have at least 14 traits (Table 2). Runs “A” test threshold variation while holding sample size and N_c (13 traits) constant. Runs “B” differ only in that N_c is decreased to and held constant at 12. The NCNA data set included 78 individuals for $N_{\text{trait}} \geq 11$. Runs “G” represent NCNA’s ancestry estimation distributions for when thresholds vary. For both data sets and all Runs, thresholds were initiated at 76/24 and were varied by 2% increments for each Run with the final Run at 90/10. Some threshold Runs exhibited identical results, an artifact of the number of traits used in the analysis. In these instances, only one of the identical Runs is included in the results. Because the number of N_c trait combinations per individual is dependent on and will vary according to the unique number of traits present per individual, we use percentages to avoid over- and under-emphasizing the distributions of ancestry estimations. From the individual percentages, we then calculated the average percentage distribution

TABLE 2—*Ancestry estimation distributions for runs A, B, and G.*

GUA: Threshold Variations					Average % Distribution of Averaged N_c Ancestry Estimations																							Variations in Distribution of Ancestry Estimations		
Run	Threshold	N_{trait}	N_c	n	E	AF	AS	E_af	E_as	AF_e	AF_as	AS_e	AS_af	E_af_as	AF_e_as	AS_e_af	E_eq_AF	E_eq_AS	AS_eq_AF	E_AF_AS	P_E	P_AF	P_AS	P_E_AF	P_E_AS	P_AF_AS	Row Sum	∑Top 10	∑Top 5	∑Top 10 − ∑Top 5
A	76/24	14	13	96	–	–	8.58	–	3.32	0.06	0.16	4.6	31.34	–	–	–	–	–	–	5.39	0.11	–	31.15	0.04	9.61	5.63	100	99.90	86.31	13.59
A	80/20	14	13	96	–	–	1.63	–	3.34	2.55E−03	0.16	11.22	48.25	0.1	0.05	5.12	–	–	–	13.89	–	–	9.45	–	4.79	2	100	99.84	87.93	11.91
A	84/16	14	13	96	–	–	1.63	–	3.19	2.55E−03	0.05	11.19	46.84	0.25	0.16	7.93	–	–	–	16.28	–	–	8.08	–	3.16	1.23	100	99.79	90.33	9.46
A	88/12	14	13	96	–	–	0.16	–	3.04	–	0.05	5.75	36.26	0.4	0.17	33.08	–	–	–	19.4	–	–	0.42	–	1.2	0.08	100	99.80	97.53	2.26
A	90/10	14	13	96	–	–	0.16	–	1.19	–	2.71E−03	4.75	33.86	2.25	0.21	36.52	–	–	–	20.32	–	–	0.38	–	0.34	0.001	100	99.79	97.71	2.08
B	90/10	14	12	96	–	–	0.38	–	0.88	3.93E−04	4.61E−03	6.24	33.65	1.11	0.09	22.64	–	–	–	29.46	–	–	0.89	–	4.27	0.38	100	99.90	93.10	6.80
B	88/12	14	12	96	–	–	0.38	–	0.88	3.93E−04	4.61E−03	6.24	33.65	1.11	0.09	22.64	–	–	–	29.46	–	–	0.89	–	4.27	0.38	100	99.90	93.10	6.80
B	86/14	14	12	96	–	–	0.38	–	1.85	3.93E−04	0.04	7.17	35.52	0.14	0.06	19.7	–	–	–	27.15	–	–	1.03	–	6.35	0.6	100	99.90	95.90	4.00
CNA: Threshold Variations					% Distribution of Averaged N_c Ancestry Estimations																									
Run	Threshold	N_{trait}	N_c	n	E	AF	AS	E_af	E_as	AF_e	AF_as	AS_e	AS_af	E_af_as	AF_e_as	AS_e_af	E_eq_AF	E_eq_AS	AS_eq_AF	E_AF_AS	P_E	P_AF	P_AS	P_E_AF	P_E_AS	P_AF_AS	Row Sum	∑Top 10	∑Top 5	∑Top 10 − ∑Top 5
G	90/10	11	9	78	–	–	0.18	0.03	0.9	0.1	0.51	4.82	11.17	1.15	4.72	42.6	–	–	–	33.57	–	–	0.17	–	0.08	–	100	99.79	96.89	2.91
G	86/14	11	9	78	–	–	2.26	0.28	1.23	0.98	1.25	11.94	23.98	0.53	3.07	17.05	–	–	–	30.97	0.03	0.03	3.71	0.51	1.03	1.14	100	96.60	87.65	8.96
G	82/16	11	9	78	0.02	0.02	2.26	0.15	1.69	0.81	2.65	14.14	26.18	0.08	1.67	8.56	–	–	–	26.99	0.14	0.19	7.81	0.51	2.91	3.24	100	96.40	83.67	12.73

The “–” indicates for the given run, no ancestry estimations were predicted to what is represented under the column containing the “–”. For example, none of the 96 individuals for Runs A were ever classified as a European (E) Ancestry.

for the 26 ancestry estimations for the entire data set (Table 2). Finally, for each threshold variation per Run, we summed the percentage contribution for the Top 10 and Top 5 most common ancestry estimations. For Runs A, B, and G, the Top 10 ancestry estimations held 96–99.9% of all the ancestry estimations generated by the simulation. Furthermore, 84–98% of all ancestry estimations were included in only five ancestry estimation categories. However, the most common ancestry estimations did vary slightly throughout the threshold adjustments; the five gray boxes per row in Table 2 indicate the five most common ancestry estimations for that threshold. While the large majority of results are focused within a few ancestry estimations, the distribution *within* these ancestry categories did vary with threshold changes (Fig. 2). The distribution can also be viewed numerically in Table 2 in the column labeled ΣTop 10-ΣTop 5. This column demonstrates that as the threshold ranges increase in size, the distribution of the ancestry assessments among the possible 26 also increases.

The distribution of ancestry estimations exhibited in Table 2 implies that threshold choice can alter the ancestry estimation. However, recall that the 26 ancestry estimation categories were generated specifically for our simulation as there is no standardization among practicing forensic anthropologists for how to describe an ancestry (or ancestries) on a case report. Table 2 shows that several of the most common ancestry estimation outcomes of Runs A, B, and G indicate something very similar—most of the Top 5 ancestry estimation labels insinuate a large Asian component. Does an interpretive difference between "Primarily AS" and "Full_AS" exist in actual casework? Does an "AS_eu" versus "AS_af" ancestry estimation influence the decisions of the law enforcement official reading a case report? Currently, mixed or multiple ancestries is not a searchable option when using the National Crime Information Center (NCIC) or the National Missing and Unidentified and Persons System (NamUs). Consequently, when selecting a single ancestry for searching the missing person's database, both AS_eu and AS_af ancestry estimations will most likely yield the same search for a person of Asian ancestry. While the results in Table 2 support that ancestry estimation can vary per threshold, it is more important to analyze this variation situated within the context and nuances of actual casework. Therefore we develop an alternative approach for conceptualizing ancestry estimation that is pertinent to our specific samples and the questions raised above. We collapse the 26 ancestry estimations into four classifications representing generalized categories of inclusion and exclusion. The four classifications are based on the conventional premise in forensic anthropology that both the GUA and NCNA samples' biohistories link them to Asian ancestry (8,40,41). Therefore, the proportions of Asian ancestry included in the simulation's 26 generated ancestry assessments are the foundation for the four ancestry classifications (Table 3).

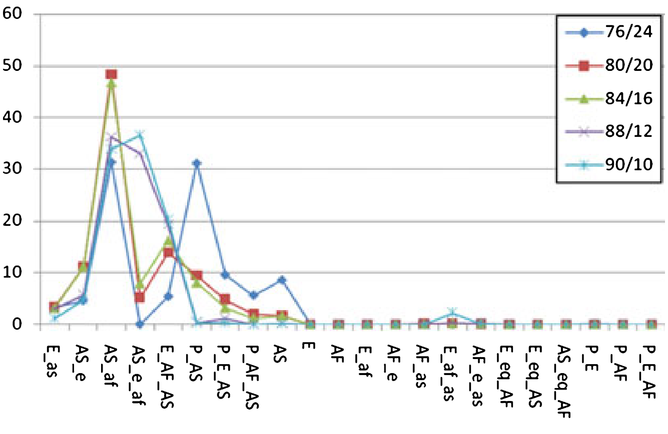

FIG. 2—*Average percentage distributions of ancestry assessments for Runs A.*

Based on the classification criteria outlined in Table 3, we define a new variable to describe our simulation's estimate of consistency among ancestry outcomes for the trait list method. Accuracy$_{AS}$ is the sum of the averaged percents of ancestry estimations consistent with classifications "Asian" or "Equal Asian." The implication of Accuracy$_{AS}$ pertains specifically to our samples and how law enforcement can interpret ancestry estimations on case reports to "fit" the searchable missing persons databases. An ancestry estimation that includes Asian ancestry as the prominent or equally dominant ancestry on a case report would match the assumed Asian ancestral link in our samples and most likely lead to an Asian ancestry-based search via NamUs or NCIC by law enforcement reading the case report. By nature of its definition, Accuracy$_{AS}$ will eliminate some of the variability presented in Runs A, B, and G. However, recall that this variation in ancestry estimation outcomes is an artifact of not only threshold changes, but also the availability of 26 possible ancestry estimations. Accuracy$_{AS}$ still considers ancestry estimation variation as a result of threshold changes, but is contextualized within more realistic applications of the ancestry estimations made by forensic anthropologists.

Table 4 displays the GUA and NCNA data sets' average percentage contributions to each of the four classifications in Runs A, B, and G. When the threshold boundaries change for each Run, the distribution of the Ancestry assessments among the four classifications remains highly stable, with associated Accuracy$_{AS}$ values at 96%, 98%, and 92% for Runs A, B, and G, respectively. In contrast, ancestry estimations that minimize or exclude Asian ancestry do occur, but at very low rates, around 3.6% for GUA and 7% for NCNA. These results indicate that for the GUA and NCNA samples, ancestry estimations cluster around those classifications that include Asian ancestry as a prominent figure, regardless of the threshold. While the expanded results using the 26 ancestry estimation labels exhibit more variation as a result of threshold changes, our collapsed results using the four Asian classifications demonstrate a strong stability and high Accuracy$_{AS}$ regardless of threshold variation.

Trait Quantity (N$_c$) *Variation*

Next, we utilized our simulation to explore how altering the number of traits used in the analysis (N_c) influences the ancestry estimation. Runs C, D, E, and F employ the same GUA data set of 96 individuals as Runs A and B. Runs C hold sample size and thresholds constant at 96 and 90/10, respectively, while decreasing the number of traits (N_c) used by increments of one (Table 5). Runs D, E, and F explore patterns of trait quantity variation by holding the thresholds constant at different levels. Runs H convey a similar pattern of variation for the NCNA data set using an 86/14 threshold only.

Runs C–F and H result in distribution patterns similar to Runs A and B; regardless of the number of traits used for analysis, 97–99% of the ancestry estimations were concentrated in the Top 10 categories. In contrast, 79–98% of all ancestry estimations were clustered in the Top 5 per N_c variation. Again, the percentage distribution among these five categories did vary as N_c changed. However, when collapsing the 26 categories into the four classifications, this variation collapses as well, but is still present (Table 6). For Runs C–F and H, Accuracy$_{AS}$ decreases as trait quantity decreases. Accuracy$_{AS}$ for the GUA data set decreases from 97 to 91% for 12 to seven traits, respectively. Ancestry assessments that minimize (minor Asian classification) or exclude Asian (no Asian classification) ancestry occur at average percentages ranging from 3.6 to 9.4% (Table 6). Accuracy$_{AS}$

TABLE 3—*Classification labels and descriptions.*

Label Number	Label Name	Label Description
Classification 1	Asian	Asian ancestry is prominent over other ancestries in the model's generated ancestry estimation (e.g., AS_af).
Classification 2	Equal Asian	Asian ancestry's dominance is shared by one or more ancestry in the model's generated ancestry estimation (eg., Equal_E_AF_AS).
Classification 3	Minor Asian	Asian ancestry is minimized in comparison to other ancestries in the model's generated ancestry estimation, but still included (eg., E_as)
Classification 4	No Asian	Asian ancestry is not included in the model's generated ancestry estimation (eg., Primarily_E).

TABLE 4—*Analysis of distributions using four classifications.*

					Average % Distribution of N_c Combinations				
Run	Threshold	N_{trait}	N_c	n	Asian	Equal Asian	Minor Asian	No Asian	$Accuracy_{AS}$
GUA: Threshold Variations									
A	76/24	14	13	96	75.67	20.63	3.48	0.21	96.3
A	80/20	14	13	96	75.67	20.68	3.65	0	96.35
A	84/16	14	13	96	75.67	20.68	3.65	0	96.35
A	88/12	14	13	96	75.67	20.68	3.65	0	96.35
A	90/10	14	13	96	75.67	20.68	3.65	0	96.35
B	90/10	14	12	96	63.81	34.1	2.09	0	97.91
B	88/12	14	12	96	63.81	34.1	2.09	0	97.91
B	86/14	14	12	96	63.81	34.1	2.09	0	97.91
NCNA: Threshold Variation									
G	90/10	11	9	78	58.94	33.65	7.28	0.13	92.59
G	86/14	11	9	78	58.94	33.14	6.08	1.84	92.08
G	82/16	11	9	78	58.94	33.14	6.08	1.84	92.08

for the NCNA data set decreased from 93 to 89% for 10 to seven traits, respectively. These results indicate that for the GUA and NCNA data sets and corresponding trait quantities tested here, ancestry assessments cluster around those classifications that include Asian ancestry as a prominent figure, regardless of the number of traits used in the trait list method.

Marked differences in the stability of ancestry estimations are present when comparing the results generated from alterations in (i) thresholds and (ii) N_c. For both the GUA and NCNA data sets, $Accuracy_{AS}$ is more stable when altering thresholds versus N_c. This feature suggests that N_c variation among practicing forensic anthropologists can result in contrasting ancestry estimations more often than thresholds variation. Overall, the simulation's results for $Accuracy_{AS}$ suggest that variations in applying the trait list method (thresholds or N_c) yield similar ancestry estimations at least 89% of the time for the GUA and NCNA data sets.

Trait Selection Variation

For every Run, the simulation generates all possible combinations of traits for each individual included in the data set. This process is performed to simulate variability in selecting different traits to use in a trait list method analysis among practicing forensic anthropologists. This variation is embedded within all Runs and is therefore taken into account in the previous discussion of our results. Next, we explore how including or excluding specific traits in the trait list analysis influence the ancestry estimations. For example, does excluding Asian-associated traits that are found in high frequencies in our samples from the simulation's Runs influence the ancestry estimations? From basic trait frequency analyses, we were able to discern that particular traits were "big hitters" for Asian ancestry in our two samples. For example, in the GUA sample traits such as shovel-shaped incisors, enamel extensions, buccal pits, canine fossa, and incisal rotation were found on 84–97% of individuals. How would the ancestry assessments vary if such influential traits were excluded? We tested this question by excluding the two "biggest hitters" from the raw data set: incisal rotation and shovel-shaped incisors. Because of the exclusion of these traits, N_{trait} was set to 12. From the GUA $N_{trait} = 14$ data set ($n = 96$), 73 individuals had data for at least 12 traits that did not include data for incisal rotation and shovel-shaped incisors. For Runs I, the threshold was held constant at 86/14 while N_c varied. For the four N_cs in Runs I, $Accuracy_{AS}$ ranged from *c.* 97 to 91% (Table 7). This range for $Accuracy_{AS}$ when excluding the two "big hitters" appears highly similar to $Accuracy_{AS}$ of Runs D, which is identical to Runs I except that it excludes the two "big hitters." To statistically discern whether excluding the two traits influenced the ancestry estimations, the percentage distribution over 26 ancestry estimations per individual was used to compare Runs I and Runs D. Each Run *excluding* the traits was paired with its identical Run (set at 86/14, N_c = 12, 11, 10, or 9) *including* the traits. We tested for significant differences in the "excluding" and "including" means for each of the 26 ancestries at each of the four N_cs using Tukey's HSD test at alpha = 0.05. For all comparisons, only three differences in means were statistically significant (Table 8). Next, we tested for significant differences of $Accuracy_{AS}$ for all N_c pairs of Runs D and I pairs using the same statistical parameters, and no significant differences were found. Therefore, we conclude that excluding two high-frequency traits from our GUA sample does influence the distribution of ancestry estimations when utilizing the 26 ancestries, but maintains high and stable estimations of $Accuracy_{AS}$.

Discussion

Outliers and Their Importance to Forensic Casework

The trends presented in Tables 2–7 demonstrate that ancestry assessments remain highly stable regardless of traits selection, trait number, and threshold variations. Additionally, the stable ancestry estimation distributions are largely consistent with an emphasis on

TABLE 5—*Ancestry estimation distribution for Runs C–F, H.*

Trait Quantity (N_c) Variation for GUA					Average % Distribution of Ancestry Estimations																							Variation in Distribution of Ancestry Estimations		
Run	Threshold	N_{trait}	N_c	n	E	AF	AS	E_af	E_as	AF_e	AF_as	AS_e	AS_af	E_af_as	AF_e_as	AS_e_af	E_eq_AF	E_eq_AS	AS_eq_AF	E_AF_AS	P_E	P_AF	P_AS	P_E_AF	P_E_AS	P_AF_AS	Row Sum	∑Top 10	∑Top 5	∑Top 10 − ∑Top 5
C	90/10	14	13	96	–	–	0.16	–	1.19	–	0.00	4.75	33.86	2.25	0.21	36.52	–	–	–	20.32	–	–	0.38	–	0.35	0.01	100.00	99.99	97.71	2.28
C	90/10	14	12	96	–	–	0.38	–	0.88	0.00	0.00	6.24	33.65	1.11	0.09	22.64	–	–	–	29.46	–	–	0.89	–	4.27	0.38	100.00	99.62	96.26	3.36
C	90/10	14	11	96	–	–	0.86	0.00	2.41	0.00	0.08	8.34	38.27	1.71	0.46	24.09	–	–	–	20.30	–	–	1.76	–	1.41	0.31	100.00	99.59	93.41	6.19
C	90/10	14	10	96	–	–	0.12	0.03	2.66	0.02	0.36	10.56	40.31	1.39	0.62	24.31	0.00	1.93	0.06	12.85	0.00	–	4.77	0.01	–	–	100.00	99.76	90.69	9.07
C	90/10	14	9	96	0.00	–	0.29	0.00	2.58	0.00	0.14	3.72	23.33	1.86	0.86	43.45	0.00	–	–	23.19	–	–	0.03	–	0.55	–	100.00	99.96	96.27	3.69
C	90/10	14	8	96	0.00	–	0.64	0.02	2.04	0.01	0.13	5.21	23.97	4.22	1.84	47.23	0.00	3.73	0.32	10.56	–	–	0.07	–	–	–	100.00	99.57	91.20	8.37
C	90/10	14	7	96	0.01	0.00	1.36	0.06	4.56	0.05	0.54	7.46	28.23	2.41	1.27	30.96	0.00	–	–	22.43	–	–	0.15	–	0.51	–	100.00	99.73	93.64	6.09
D	86/14	14	13	96	–	–	0.16	–	3.04	–	0.05	5.75	36.26	0.40	0.17	33.08	–	–	–	19.40	–	–	0.42	–	1.20	0.08	100.00	99.88	97.53	2.34
D	86/14	14	11	96	0.00	–	0.86	0.00	3.73	0.00	0.20	9.31	40.43	0.39	0.34	20.62	–	–	–	19.00	0.00	–	2.10	–	2.51	0.50	100.00	99.45	93.09	6.36
D	86/14	14	9	96	0.00	–	3.47	0.10	3.42	0.09	0.38	11.73	39.56	0.90	0.54	11.02	0.00	–	–	18.87	0.01	–	5.03	0.04	3.83	1.02	100.00	98.48	86.21	12.28
D	86/14	14	7	96	0.01	0.00	1.36	0.06	4.56	0.05	0.54	7.46	28.23	2.41	1.27	30.96	0.00	–	–	22.43	–	–	0.15	–	0.51	–	100.00	99.73	93.64	6.09
E	84/16	14	13	96	–	–	1.63	–	3.19	0.00	0.05	11.19	46.84	0.25	0.16	7.93	–	–	–	16.28	–	–	8.08	–	3.16	1.23	100.00	99.79	90.33	9.46
E	84/16	14	12	96	–	–	0.38	–	1.85	0.00	0.04	7.17	35.52	0.14	0.06	19.70	–	–	–	27.15	–	–	1.03	–	6.35	0.60	100.00	99.90	95.90	4.00
E	84/16	14	11	96	–	–	0.86	0.00	3.73	0.00	0.20	9.31	40.43	0.39	0.34	20.62	–	–	–	19.00	0.00	–	2.10	–	2.51	0.50	100.00	99.45	93.09	6.36
E	84/16	14	10	96	0.00	–	1.78	0.02	3.34	0.02	0.44	11.60	42.67	0.71	0.54	20.26	–	1.93	0.06	11.61	0.01	–	3.76	0.01	1.02	0.22	100.00	98.69	89.90	8.79
E	84/16	14	9	96	0.00	–	3.47	0.10	3.42	0.09	0.38	11.73	39.56	0.90	0.54	11.02	0.00	–	–	18.87	0.01	–	5.03	0.04	3.83	1.02	100.00	98.85	86.21	12.65
E	84/16	14	8	96	0.01	–	6.38	0.32	4.69	0.23	1.07	13.48	38.92	1.19	0.67	10.87	0.00	3.73	0.32	10.50	0.07	0.01	7.47	0.07	–	–	100.00	98.30	81.25	17.06
E	84/16	14	7	96	0.06	0.00	11.09	0.56	5.79	0.44	1.07	12.06	33.59	0.39	0.29	1.12	0.00	–	–	13.33	0.24	0.06	10.29	0.43	6.49	2.69	100.00	97.01	80.36	16.65
F	76/24	14	13	96	–	–	8.58	–	3.32	0.06	0.16	4.60	31.34	–	–	–	–	–	–	5.39	0.11	–	31.15	0.04	9.61	5.63	100.00	99.90	86.08	13.82
F	76/24	14	11	96	0.00	–	5.94	0.06	3.98	0.08	0.45	9.78	38.01	–	–	–	–	–	–	8.97	0.08	0.01	19.58	0.05	8.93	4.07	100.00	99.78	85.27	14.52
F	76/24	14	10	96	0.00	–	10.23	0.17	3.64	0.18	0.78	10.82	36.89	–	–	1.15	–	1.93	0.06	7.34	0.27	0.05	20.99	0.27	3.53	1.71	100.00	98.22	86.27	11.95
F	76/24	14	9	96	0.04	–	16.55	0.08	3.58	0.23	0.54	5.49	24.12	–	–	–	0.00	–	–	4.17	0.74	0.24	24.65	1.01	12.40	6.17	100.00	98.88	83.21	15.67
F	76/24	14	7	96	0.16	0.01	11.09	0.38	6.18	0.39	1.36	11.72	32.81	–	–	–	0.00	–	–	10.91	0.33	0.10	12.53	0.43	8.13	3.47	100.00	98.63	79.07	19.56

Trait Quantity (N_c) Variation for NCNA					Average % Distribution of Ancestry Estimations																							Variation in Distribution of Ancestry Estimations		
Run	Threshold	N_{trait}	N_c	n	E	AF	AS	E_af	E_as	AF_e	AF_as	AS_e	AS_af	E_af_as	AF_e_as	AS_e_af	E_eq_AF	E_eq_AS	AS_eq_AF	E_AF_AS	P_E	P_AF	P_AS	P_E_AF	P_E_AS	P_AF_AS	Row Sum	∑Top 10	∑Top 5	∑Top 10 − ∑Top 5
H	86/14	11	9	78	0.00	0.00	2.26	0.28	1.23	0.98	1.25	11.94	23.98	0.53	3.07	17.05	0.00	0.00	0.00	30.97	0.03	0.03	3.71	0.51	1.03	1.14	100.00	96.60	87.00	9.60
H	86/14	11	7	78	0.01	0.01	0.90	0.20	1.69	0.45	1.68	9.10	16.12	1.75	5.21	30.59	0.01	0.00	0.00	31.33	0.00	0.00	0.55	0.00	0.38	0.00	100.00	98.93	92.35	6.58

The "–" indicates that the given Run no ancestry estimations were predicted to what is represented under the column containing the "–".

TABLE 6—*Analysis of distributions using four classifications.*

					Average % Distribution of N_c Combinations				
Run	Threshold	N_{trait}	N_c	n	Asian	Equal Asian	Minor Asian	No Asian	$Accuracy_{AS}$
GUA: Trait Quantity Variation									
C	90/10	14	13	96	75.67	20.68	3.65	0	96.35
C	90/10	14	12	96	63.81	34.1	2.09	0	97.91
C	90/10	14	11	96	73.32	22.02	4.66	0.01	95.34
C	90/10	14	10	96	80.07	14.83	5.03	0.06	94.9
C	90/10	14	9	96	70.81	23.75	5.43	0.01	94.56
C	90/10	14	8	96	77.13	14.61	8.22	0.03	91.74
C	90/10	14	7	96	68.16	22.93	8.79	0.12	91.09
D	86/14	14	13	96	75.67	20.68	3.65	0	96.35
D	86/14	14	11	96	73.32	22.02	4.66	0.01	95.34
D	86/14	14	9	96	70.81	23.71	5.24	0.24	94.52
D	86/14	14	7	96	68.16	22.93	8.79	0.12	91.09
E	84/16	14	13	96	75.67	20.68	3.65	0	96.35
E	84/16	14	12	96	63.81	34.1	2.09	0	97.91
E	84/16	14	11	96	73.32	22.02	4.66	0.01	95.34
E	84/16	14	10	96	80.07	14.83	5.03	0.06	94.9
E	84/16	14	9	96	70.81	23.71	5.24	0.24	94.52
E	84/16	14	8	96	77.13	14.55	7.62	0.7	91.68
E	84/16	14	7	96	68.16	22.51	7.54	1.79	90.67
F	76/24	14	13	96	75.67	20.63	3.48	0.21	96.3
F	76/24	14	11	96	73.32	21.96	4.43	0.29	95.28
F	76/24	14	10	96	80.07	14.57	4.41	0.94	94.64
F	76/24	14	9	96	70.81	22.74	4.12	2.34	93.55
F	76/24	14	7	96	68.16	22.51	7.54	1.79	90.67
NCNA: Trait Quantity Variation									
H	86/14	11	10	78	69.14	23.87	6.05	0.94	93.01
H	86/14	11	9	78	58.94	33.14	6.08	1.84	92.08
H	86/14	11	8	78	68.36	20.83	7.54	3.27	89.19
H	86/14	11	7	78	57.27	31.71	10.33	0.69	88.98

TABLE 7—*Ancestry estimation % distribution with GUA trait exclusion.*

GUA Trait Quantity Variation					Average % Distribution of N_c Combinations				
Run	Threshold	N_{trait}	N_c	n	Asian	Equal Asian	Minor Asian	No Asian	$Accuracy_{AS}$
I	86/14	12	12	73	47.24	49.76	2.99	0.00	97
I	86/14	12	11	73	61.08	30.63	8.29	0.01	91.71
I	86/14	12	10	73	69.39	22.42	8.11	0.08	91.81
I	86/14	12	9	73	60.17	30.73	8.72	0.38	90.9

TABLE 8—*Statistically significant differences in the means of “including” and “excluding” traits.*

N_c	Ancestry Estimation	p-Value (Separate Variance)	Alpha
11	P_E_AF	0.046	0.05
12	P_AS	0.011	0.05
12	AS_E_AF	0.032	0.05

Thresholds held constant at 86/14 for all Runs.

Asian ancestry. These results focus on the average for the samples, but what about individual exceptions? Forensic anthropologists work at the individual case level, and we must also explore this level in our research. We reviewed the raw data for all individuals included in the $N_{trait} = 14$ data set ($n = 96$) to find those who had resulting ancestry descriptions that demonstrated deviations from the general trends. A few individuals’ N_c combinations resulted in ancestry descriptions indicating low or no levels of Asian ancestry such as “Full EU” or “EU maj_AS min.” We then selected a handful of these individuals to study (i) how their ancestry assessments were influenced by trait quantity and threshold variations and (ii) how coupling metric analysis with the trait list method analysis would affect the ancestry estimation that ends up on the case report.

At a threshold of 86/14, Case A (GUA male) had 36 of 3432 possible eight-trait combinations that resulted in the exclusion of Asian ancestry in the final ancestry estimation. When more than eight traits were used ($N_c > 8$), the 36 non-Asian ancestry estimations were no longer present. Case A’s cranium underwent craniometric analysis by submitting 20 cranial measurements into Fordisc 3.0. Fordisc is a computer program that assists in the estimation of the biological profile for unidentified skeletal remains and is a common tool used in forensic anthropologists’ case analyses (42). The program uses a series of statistical analyses to determine to which sample population the skeleton in question is most similar. Fordisc classified Case A as Japanese male with a posterior probability of 0.839 and a Typicality Rank as 31/101. How the forensic anthropologist chooses to couple the metric and nonmetric assessments to conclude a final ancestry estimation on the case report is an interpretive and subjective procedure. If metric and trait list method analyses are given equal weight, there is a good chance that the case report for Case A will include the possibility of Asian ancestry in the final ancestry estimation. However, if more weight is place on the nonmetric assessment, then the final ancestry estimation may deviate from including an Asian ancestry component.

For Case B (GUA male), the Fordisc results provide little help with errant nonmetric ancestry assessments. The ancestry estimations for 2675 of Case B's 19,448 seven-trait combinations that underwent trait method analyses either minimized or excluded Asian ancestry. Because sex has been established as male, only male samples are used for the Fordisc analysis. As with Case A, *c*. 20 metric traits were input into Fordisc, which considered the metric data too dissimilar to make a classification. Here, the anthropologist would only have the nonmetric trait analysis on which to base the case report's estimation. Having to rely solely upon nonmetric analysis for ancestry assessment is a reality for the forensic anthropologist; oftentimes, cases may be deformed or fragmented. If circumstances (trait choice, fragmentation) allowed for one of the 2675 seven-trait combinations to be used for the trait list method analysis for Case B, then the resulting ancestry assessment that ended up on the case report would exclude or minimize the component of Asian ancestry when in fact Case B is known to have a large Asian ancestry component.

These examples provide just two scenarios to demonstrate how the coupling of metric and nonmetric analyses may aid or not aid in accurately assessing ancestry. Keep in mind that these cases represent the minority within our sample. Most individuals ancestries' generated by the simulation estimated a large proportion of Asian ancestry, consistent with the assumed ancestry for our samples. Additionally, for both of our samples, the majority of ancestry estimations that excluded or minimized the Asian component were products of using very few traits, such as seven- and eight-trait combinations. These nonconforming ancestry assessments completely disappear when the trait number or threshold is increased at or above 10-trait combinations or 88/12.

Trait List Method Biases

When only a limited number of traits can be included in a trait list method analysis, several biases need to be kept in mind. We will discuss these biases in light of the 25 nonmetric traits recognized in our study, but the results of this discussion should be applied to the use of any number of traits. Recall that the trait list method assumes an association between a particular ancestry and the presence of a trait or trait state. This association creates a *contribution potential* (CP) for each ancestry to the final ancestry estimation of an individual. The CP is dependent on the specific traits used in the analysis. For example, the trait "orbit shape" has trait states that are associated with each of the three ancestries, allowing for a CP for all three ancestries. In contrast, the trait "inca bone" has only one trait state that carries estimative weight: present. Because the presence of the inca bone is only associated with Asian ancestry, only this ancestry can glean a CP from this trait. Of the 25 traits on which we collected data for our two samples, the CPs for Asian, European, and African ancestries are as follows. Twenty of the 25 traits have trait states that are specifically associated with Asian ancestry, yielding a CP of 20. Eleven out of the 25 traits have trait states that are specific to European ancestry, and one trait with a trait state associated with *both* European and African ancestry, yielding a CP of 12.5 for European ancestry. Finally, the CP for African ancestry is 8.5. Because the CP varies for the three ancestries, so too does the possibility for having a "Full Ancestry X" Proportion Descriptor. For example, if we were able to collect 25 of 25 traits from a cranium, the absolute maximum number of traits that can exhibit a European state is 12.5. Let us assume that a cranium does indeed express all potential European states. Now let us set $N_c = 15$. For all possible combinations of 15 of the 25 traits, the maximum amount that can ever be contributed by European ancestry is 12.5 of 15 traits, or 83%. Therefore, if the threshold for the Proportion Descriptor "Full Ancestry X" is set above 83%, the European component will not be enough to ever be considered "Full Ancestry European." In contrast, if $N_c = 12$ or less, then the CP exceeds the number of traits used in the combinations and therefore has the potential to be classified as "Full Ancestry European."

For the 25 traits used in this paper's analysis, $N_c = 8$ or less would allow for the unbiased potential for *all* ancestries to be classified as "Full Ancestry X." We provided results for variations of N_c in Table 6. For $N_c < CP_{AF}$ and $N_c < CP_{EU}$, only extremely minimal shifts in the ancestry assessment distributions toward "Full European" or "Full African" occurred. The lack of shifting indicates that for our NCNA and GUA samples, the CP is an insignificant bias because rarely did an individual express all traits or trait states associated with a particular ancestry. Regardless of the CP's bias in our data sets, this is clearly a factor that has the potential to bias an ancestry estimation when employing the trait list method. Forensic anthropologists should be cognizant of the CPs they create when selecting the nonmetric traits to use in a trait list method ancestry analysis.

Conclusions

We developed a simulation to demonstrate how variations in the application of the trait list method influence ancestry estimations. Using two temporally and geographically diverse samples, we found that the trait list method provides ancestry estimation stability as thresholds, trait quantity, and specific traits are varied. For all data sets and Runs, Accuracy_{AS} was maintained above 90%. The simulation, its generated results, and the variable Accuracy_{AS} simply provide estimations of how ancestry estimation outcomes can differ when the trait list method is varied. True accuracy is a dynamic process, dependent on how the actual words used to describe the ancestry on a case report are interpreted by the law enforcement officials to successfully find a match in a missing person's database. However, in this paper, Accuracy_{AS} does serve to illuminate how well the trait list method results in a consistent ancestry estimation regardless of the variations placed on the method via the simulation. Such high levels of Accuracy_{AS} are counterintuitive when studying the diversity present in the raw frequencies of trait states for both the GUA and NCNA samples. For many of the morphoscopic traits, raw frequencies were roughly proportionate in each of the three ancestries. By using at least 10 traits and an 88/12 threshold, this raw diversity is minimized to result in similar ancestry estimations.

The simulations of accuracy and stability of ancestry estimations using the trait list method are only appropriate within the context of the two samples presented here. While we did choose samples that exhibited trait state diversity, verifying the results with additional samples can serve to verify the conclusions presented here. To generate a broader exploration on the effects of the lack of standardization of the trait list method, more samples need to be included. While we feel it is responsible to draw conclusions for only these two samples, we can make generalized recommendations based on our results. We recommend that N_c be less than or equal to the smallest CP for a given set of nonmetric traits being applied to a case analysis. In addition, keep in mind that if the lowest CP is small, then it may be best to increase N_{trait}. If it is impossible to increase N_{trait} because of missing portions of the cranium, we recommend that the forensic anthropologist be highly conservative with ancestry estimations solely based on a trait list method ancestry estimation, as the effects of the CP bias have not been generically tested yet. We also recommend that the forensic community move toward standardizing the minimum number of traits appropriate for conducting a trait list method analysis, as our results exhibit the greatest variation when N_c is changed. For the samples presented in this paper, using at least 10 traits in the trait list method analysis ($N_c \geq 10$) generates a limited number of ancestry estimations that minimize or exclude Asian ancestry. Additional samples must be tested in our simulation to confirm that this N_c is appropriate in the broad application of the trait list method.

In regard to trait selection, we have demonstrated that even when particular traits are excluded, the $\text{Accuracy}_{\text{AS}}$ remains high. For a generalized approach, the simulation employed here takes into account all possible combinations of traits, thereby excluding and including different traits at different times. Therefore, each of the trends presented here considers variations in the selection of traits used in a trait list method ancestry estimation. At this point, we make no recommendations for the necessary inclusion or exclusion of particular traits when applying the trait list method, as it is clear that even traits seen in the highest frequencies of our data sets (such as shovel-shaped incisors) only minimally alter the ancestry estimations when excluded.

Last, it is important to contextualize this paper within the realities of forensic anthropological casework. In order for the authors to explore their specific research questions, we made assumptions about the conventional employment of the trait list method, as well as the acceptance of associating specific trait states to specific ancestries. While the trait list method's discreteness clearly lends itself to cookie-cutter ancestry classifications, we also recognize that most likely the large majority of practicing anthropologists (i) understand the nuances of trait distributions throughout global populations, (ii) are aware of empirical critiques of the association of trait states to specific ancestries, and (iii) apply such insight when analyzing the presence or absence of nonmetric cranial traits. Because the trait list method is not grounded in statistical analyses, it is imperative to continue to explore how this analytical tool actually functions. Nuances, biases, and caveats can be embedded in every method of analysis and must be rigorously refined and validated to ensure that such methods can stand up to the requirements of empirical analysis. Here, we have elucidated several dark corners of the trait list method, but more analyses of this widely applied method are necessary. Aspects such as trait weighting, subjectivity of classifying trait states, sample variation, and the weighting of multiple ancestry estimation methods must also be explored to understand how the trait list method "works." As long as the trait list method and its variations continue to be applied to casework, forensic anthropologists bear the responsibility to know how the method's nuances ultimately influence ancestry estimation.

Acknowledgments

We thank Lauren Zephro for her challenging and supporting advice throughout this project. We also thank Dr. Linda Klepinger for her comments on various drafts. We thank the staff of the Pheobe A. Hurst Museum at Berkeley, who worked with us for 2 years. We thank the employees at the FAFG who supported Hughes' collection of data from 2007 to 2009.

References

1. Hefner J, Emanovsky P, Byrd J, Ousley SD. Morphoscopic traits the statistical determination of ancestry using cranial nonmetric traits. Proceedings of the 58th Annual Meeting of the American Academy of Forensic Sciences; 2006 Feb 20–25; Seattle, WA. Denver, CO: Publication Printers, 2006.
2. Hefner J, Emanovsky P, Byrd J, Ousley SD. The value of experience, education, and methods in ancestry prediction. Proceedings of the 59th Annual Meeting of the American Academy of Forensic Sciences; 2007 Feb 19–23; San Antonio, TX. Denver, CO: Publication Printers, 2007.
3. Spradley M, Ousley S, Jantz R. Evaluating cranial morphometric relationships using discriminate function analysis. Am J Phys Anthropol Suppl 2008;46:199.
4. Spradley K. Hispanic affiliations: definitions assumptions and biological reality. Proceedings of the 61st Annual Meeting of the American Academy of Forensic Sciences; 2009 Feb 16–21; Denver, CO. Denver, CO: Publication Printers, 2009.
5. Carbonell VM. Variations in the frequency of shovel-shaped incisors in different populations. In: Brothwell D, editor. Dental anthropology. New York, NY: Pergamon Press, 1963;211–34.
6. Crummett T. The evolution of shovel shaping: regional and temporal variation in human incisor morphology [Dissertation]. Ann Arbor (MI): University of Michigan, 1994.
7. Hinkes M. Shovel shaped incisors in human identification. In: Gill GW, Rhine S, editors. Skeletal attribution of race. Maxwell Museum of Anthropology Anthropological Papers No. 4. Albuquerque, NM: University of New Mexico, 1990;21–5.
8. Bruces A. People and races. New York, NY: Macmillan, 1977.
9. Krogman W. The human skeleton in forensic medicine. Springfield, IL: Charles C Thomas, 1962.
10. Rhine S. Nonmetric skull racing. In: Gill GW, Rhine S, editors. Skeletal attribution of race. Maxwell Museum of Anthropology Anthropological Papers No. 4. Albuquerque, NM: University of New Mexico, 1990;9–20.
11. Bamshad M, Wooding S, Watkins W, Ostler C, Batzer M, Jorde L. Human population genetic structure and inference of group membership. Am J Hum Genet 2003;72:578–89.
12. Cheverud J. A comparison of genetic and phenotypic correlations. Evolution 1988;42:958–68.
13. Cheverud J, Buikstra J. Quantitative genetics of skeletal nonmetric traits in the rhesus macaques of Cayo Santiago. III. Relative heritability of skeletal nonmetric and metric traits. Am J Phys Anthropol 1981;59(2):151–5.
14. Cheverud J. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 1982;36(3):499–516.
15. Cheverud J, Buikstra J. Quantitative genetics of skeletal nonmetric traits in rhesus macaques on Cayo Santiago. I. Singe trait heritabilities. Am J Phys Anthropol 2005;54(1):43–9.
16. Ishida H, Dodo Y. Nonmetric cranial variation and the population affinities of the Pacific peoples. Am J Phys Anthropol 1993;90(1):49–57.
17. Ishida H, Dodo Y. Cranial variation in prehistoric human skeletal remains from the Marianas. Am J Phys Anthropol 1997;104(3):399–410.
18. McGrath J, Cheverud J, Buikstra J. Genetic correlations between sides and heritability of asymmetry for nonmetric traits in rhesus macaques on Cayo Santiago. Am J Phys Anthropol 1984;64(4):401–11.
19. Cavalli-Sforza L, Bodmer W. The genetics of human populations. New York, NY: Dover and Mineola, 1999.
20. Jorde L, Wooding S. Genetic variation, classification, and 'race.' Nat Genet Suppl 2004;36:S28–33.
21. Long J, Kittles R. Human genetic diversity and the nonexistence of biological races. Hum Biol 2003;75(4):449–71.
22. Weiss K, Long J. Non-Darwinian estimation: my ancestors, my genes' ancestors. Genome Res 2009;19(5):703–10.
23. Hefner J. Morphoscopic traits: mixed ancestry, Hispanics, and biological variation. Proceedings of the 61st Annual Meeting of the American Academy of Forensic Sciences; 2009 Feb 16–21; Denver, CO. Denver, CO: Publication Printers, 2009.
24. Parr N, Hefner J. A statistical assessment of cranial and mandibular morphoscopic traits used in the determination of ancestry. Proceedings of the 61st Annual Meeting of the American Academy of Forensic Sciences; 2009 Feb 16–21; Denver, CO. Denver, CO: Publication Printers, 2009.
25. MacLeod N, Forey P. Introduction: morphology, shape and phylogenetics. In: Macleod N, Forey P, editors. Morphology, shape and phylogeny. London, UK: CRC Press, 2002;1.
26. Hefner J. An assessment of craniofacial nonmetric traits currently used in the forensic determination of ancestry. Proceedings of the 54th Annual Meeting of the American Academy of Forensic Sciences; 2002 Feb 10–16; Atlanta, GA. Denver, CO: Publication Printers, 2002.
27. Wheat A. Estimating ancestry through nonmetric traits of the skull: a test of education and experience. Proceedings of the 61st Annual Meeting of the American Academy of Forensic Sciences; 2009 Feb 16–21; Denver, CO. Denver, CO: Publication Printers, 2009.
28. Williams S. A new method for evaluating orbit shape. Proceedings of the 59th Annual Meeting of the American Academy of Forensic Sciences; 2007 Feb 19–24; San Antonio, TX. Denver, CO: Publication Printers, 2007.
29. Ousley S. Anthropology and its relationship to human variation. Public presentation at the Science and Justice Panel: race work, the promise and peril of using racial categories in forensic, genetic, and biomedical

research.Santa Cruz, CA: University of California at Santa Cruz, 2008;May 15.
30. Bartelink E. Resource intensification in pre-contact Central California: a bioarchaeological perspective on diet and health patterns among hunter–gathers from the lower Sacramento Valley and San Francisco Bay [Dissertation]. College Station (TX): Texas A&M University, 2006.
31. Bean J. The Ohlone past and present: native Americans of the San Francisco bay region. Journal of California and Great Basin Anthropology. Banning: Ballena Press, 1995.
32. Lightfoot K, Luby E. Late Holocene in the San Francisco bay area: temporal trends in the use and abandonment of shell mounds in the East Bay. In: Erlandson J, Jones T, editors. Catalysts to complexity: late Holocene Societies of the California Coast. Los Angeles, CA: Cotsen Institute of Archaeology, University of California, 2002:121–39.
33. Pheobe A. Hearst Museum: records. Department of Special Collections and Archives, University of California at Berkeley, http://troy.oditech.com/fmi/iwp/cgi?-db=CaliforniaRadiocarbonDates&-loadframes (accessed February 10, 2009).
34. Holian J. The fertility of Maya and ladino women. Lat Am Res Rev 1985;20(2):87–103.
35. Lovell G, Lutz C. Conquest and population: Maya demography in historical perspective. Lat Am Res Rev 1994;29(2):133–40.
36. Montejo V, editor. Maya intellectual renaissance. Austin, TX: University of Texas Press, 2005.
37. Ossenberg NS. Discontinuous morphological variation in the cranium [Dissertation]. Toronto (Canada): Toronto University, 1969.
38. Stewart T. Essentials of forensic anthropology: especially as developed in the United States. Springfield, IL: Charles C Thomas, 1979.
39. Cartmill M. The status of the race concept in physical anthropology. Am Anthropol 1998;100(33):651–60.
40. Dillehay T. Probing deeper into first American studies. Proc Natl Acad Sci USA 2009;106(4):971–8.
41. Powell J, Neves W. Cranio-facial morphology of the first Americans: pattern and process in the peopling of the New World. Am J Phys Anthropol 1999;110(S29):153–88.
42. Ousley S, Jantz R. Fordisc 3.0: personal computer forensic discriminant functions. Computer program. Knoxville, TN: University of Tennessee, 2005.

Additional information and reprint requests:
Cris Hughes, Ph.D.
University of Illinois at Urbana-Champaign
Department of Anthropology
607 Matthews Ave., Davenport Hall
Urbana, IL 61801
E-mail: crisehughes@gmail.com

Appendix: Traits, Trait States, Associated Ancestries and References

The traits used in this research are binary and morphoscopic in nature. When a trait state could not be confidently declared by the authors, an "unknown" was given for the trait state and was treated the same as a missing trait state in the analysis. The authors conformed to the descriptions of the states given in the references when descriptions were given. The association "n/a" indicates that no ancestry is weighted for that state.

Trait	Reference	Trait States	Ancestry Association
Metopic suture	10	Present	European
		Obliterated	n/a
		Trace	Asian
Infraorbital suture	37	Present	Asian
		Absent	n/a
		Trace	Asian
Inca bone	10	Present	Asian
		Absent	n/a
		Trace	Asian
Postbregmatic depression	10	Present	African
		Absent	n/a
Inion hook	10	Present	European
		Absent	n/a
Nasal overgrowth	10	Present	Asian
		Absent	n/a
Marginal process	10	Present	Asian
		Absent	n/a
Canine fossa	10	Present	European
		Absent	n/a
Buccal pits	10	Present	Asian
		Absent	n/a
Enamel extensions	10	Present	Asian
		Absent	n/a
Carabelli's cusp	10	Present	European
		Absent	n/a
Zygomatic hook	10	Present	Asian
		Absent	n/a
Zygomaticamaxillary suture	10	Curved	African or European
		Angled	Asian
Incisal rotation	10	Present	Asian
		Absent	n/a
Peg tooth	10	Present	Asian
		Absent	n/a
Sutural ossicles	10	Present	Asian
		Absent	n/a
Keeling	10	Present	European
		Absent	n/a
Shovel shaped incisors	10	Present	Asian
		Absent	n/a
Prognathism	10	Large	African
		Medium	Asian
		None	European
Dental arcade shape	10	Elliptic	Asian
		Hyperbolic	African
		Parabolic	European
Palatine suture	10	Straight	Asian
		Bulging	African
		Jagged	European
Nasal spine	10	Large	European
		Intermediate	Asian
		Small	African
Nasal sill	10	Deep	European
		Shallow	Asian
		Blurred	African
Orbital shape	10	Rounded	Asian
		Rectangular	African
		Sloping	European
Nasal form	10	High and angled	European
		Wide and slightly concave	Asian
		Low and rounded	African