LJMU Research Online

Scott, GR, Navega, DS, Coelho, J, Vlemincq-Mendieta, T, Kenessey, D and Irish, JD

 rASUDAS2 A New Iteration of the Application for Assessing the Population Affinity of Individuals by Tooth Crown and Root Morphology

http://researchonline.ljmu.ac.uk/id/eprint/25587/

Article

rASUDAS2: a new iteration of the application for assessing the ancestry of individuals by tooth crown and root morphology

G. Richard Scott
University of Nevada Reno

David S. Navega
University of Coimbra

Joao Coelho
Oxford University

Tatiana Vlemincq-Mendieta
University of Nevada Reno

Dori Kenessey
University of Nevada Reno

Joel D. Irish
Liverpool John Moores University

Corresponding author:

G. Richard Scott
Department of Anthropology/MS 0096
University of Nevada Reno
Reno NV 89557
grscott@unr.edu
775 750 3091

rASUDAS2: a new iteration of the application for assessing the
population affinity of individuals by tooth crown and root morphology

SCHOLARONE™
Manuscripts

ABSTRACT

The web-based application rASUDAS, originally based on data from "The Anthropology of Modern Human Teeth," has been refined since its inception in 2015, undergoing two significant updates. The initial alpha version included 12 crown and five root traits. The beta version expanded the list to 15 crown traits and six root traits. The latest iteration, rASUDAS2, has expanded to include four additional traits along with seven backup traits, allowing for flexibility in cases where certain primary traits, like shoveling on UI1, cannot be scored so shoveling expression on UI2 is utilized. The Bayesian algorithm powering rASUDAS2 employs trait frequencies derived from archaeological finds. To evaluate its effectiveness with contemporary samples, rASUDAS2 was tasked with calculating posterior probabilities for samples of African and European descent. Utilizing between 12 to 25 traits, it assigns individuals to one of seven major biogeographic groups: Western Eurasia (WE), East Asia (EA), American Arctic (AA), non-Arctic America (n-AA), Southeast Asia (SEA), Australo-Melanesia (AM), and Sub-Saharan Africa (SSA). In a modern African sample (n = 159), the probability of classification into the Sub-Saharan African group was the highest at 68.6%, followed by Western Eurasian at 22.0%, with probabilities ranging from 0.6% to 4.4% for the other groups. Similarly, in a European-derived sample (n = 161), 75.2% were assigned to Western Eurasia and 13.0% to Sub-Saharan Africa, with the remaining 12% distributed among other groups. Approximately three out of four individuals from these regions could be accurately placed within their respective biogeographic groups. However, the likelihood of these individuals being assigned to any Asian-related group is low, at less than 10%. When assessing two mixed African and European samples, results indicated nearly equal percentages of affinity assignment, with around 40% for Sub-Saharan African and 30% for Western Eurasian.

In "The Anthropology of Modern Human Teeth" (AMHT), Scott and Turner (1997) describe global variation in 23 crown and root traits and note that dental morphology has the potential for assessing the population affinity of *individual* dentitions. This vision came to fruition in 2015 when David Navega and João Coelho of the University of Coimbra developed rASUDAS, a web-based application leveraging the trait frequencies outlined in Appendix 2 of AMHT. The initial alpha version, employing a naïve Bayes algorithm, encompassed the 21 regional samples listed in the appendix.

Building on this foundation, Navega and Coelho collaborated with Scott and utilized individual data sheets from the C.G. Turner II database to refine rASUDAS. This led to a more streamlined version accommodating 15 crown traits and six root traits across seven major biogeographic groups: Western Eurasia (WE), East Asia (EA), American Arctic (AA), non-Arctic America (n-AA; formerly American Indian), Southeast Asia (SEA), Australo-Melanesia (AM), and Sub-Saharan Africa (SSA). These enhancements were detailed in Scott et al. (2018a).

The current article has two primary objectives: first, to introduce rASUDAS2, an updated version of the application that includes additional traits and backup traits for certain variables; and second, to evaluate the performance of rASUDAS2 using data from recent modern samples of Western Eurasians, Sub-Saharan Africans, and populations with genetic contributions from both groups.

1. Changes to rASUDAS

Assuming sufficient geographic variation exists in a dental trait, the underlying principle is that including more traits in an analysis typically yields better results. There are two main reasons for expanding the trait list: (1) since traits vary in their capacity to discriminate among two or more

biogeographic groups, adding more traits enhances the precision of the Bayesian analysis; and

(2) in certain studies, particularly those with a bioarchaeological focus, a minimum of 12 traits is

often required for inclusion in an analysis. Therefore, the addition of traits enables researchers to

meet this criterion more consistently. In our effort to refine rASUDAS from its beta version, we

have incorporated four new traits and seven backup traits.

1.1 New traits

The four new traits and the specific teeth on which they are observed in the updated version of

rASUDAS include: (1) UI1 double-shoveling; (2) UC Bushman canine, also known as the mesial

canine ridge; (3) UC *tuberculum dentale*, or canine tubercle; and (4) LC distal accessory ridge.

The frequencies for double-shoveling and the Bushman canine were calculated using hundreds of

samples from the C.G. Turner II database, which primarily consists of archaeological samples

from the last two millennia (Table 1). Since both double-shoveling and the Bushman canine were

part of Turner's 29 key traits, the data were available in Excel spreadsheets.

INSERT TABLE 1

Although the UC distal accessory ridge is one of Turner's 29 key traits, the LC distal

accessory ridge was selected for inclusion in rASUDAS2 due to its greater geographic variation.

Tuberculum dentale, a cingular variant expressed on all three upper anterior teeth, shows

significant within and between field correlations (Scott 1977a). Turner specifically included UI2

expressions in his key trait list. The box and whisker plots in Scott et al. (2018b) show limited

geographic variation for this trait on UI2, with all populations displaying a frequency of around

25% when the breakpoint is set at 2+. In contrast, the senior author observed the canine tubercle

across all seven biogeographic groups, noting significant variation—from 16.4% in East Asians

to 55.1% in Sub-Saharan Africans. The frequencies for the canine tubercle were derived from databases maintained by both Scott and Turner, encompassing modern and archaeological samples.

1.2 Backup traits

When analyzing human dentitions of various ages, several factors can affect the number of variables available for input into rASUDAS2. These include antemortem and postmortem tooth loss, dental attrition, and dental work such as fillings and crowns. Given that one assumption of Bayes theorem is the independence of traits, consistent within-field correlations (cf. Scott et al., 2018b) limit the use of the same trait across different teeth within the same dental district. That is, the human dentition is divided into four tooth districts: incisors, canines, premolars, and molars, with each quadrant containing two, one, two, and three members respectively, as noted by Dahlberg (1945). For example, traits such as the hypocone can be expressed on all members of the upper molar field.

One key trait for assessing variation by individuals or by groups is shovel-shaped incisors. There is a general shoveling field that affects all upper and lower incisors. For example, the Kendall's rank order correlation statistic *tau* between UI1 and UI2 shoveling is 0.50 (Scott 1977b). In the beta version of rASUDAS, only UI1 observations were used; even when UI2 shoveling was observable, it was not utilized for analysis despite its relevance in assessing population affinity. The population frequencies for shoveling on the two upper incisors are not the same, so using UI2 shoveling as a backup necessitated finding different point estimates. In the current framework, if UI1 cannot be scored for shoveling, UI2 can now serve as a backup, providing flexibility in data use. Details on point estimates for the original rASUDAS variables are documented in Scott et al. (2018a).

For some crown traits, the frequencies observed on different members within the same dental field may not be substantial enough to significantly impact the analysis. Notable examples include Carabelli's trait of the upper molar field and cusp 7 of the lower molar field. Typically, these traits are prominently expressed on the first molar, with expressions on the second and third molars being inconsistent and much less frequent. However, there are traits that are expressed frequently enough across multiple members of a field to provide valuable backups when the primary tooth for a trait cannot be scored. In the updated rASUDAS2, seven backup traits have been incorporated to enhance the robustness of the analysis. These include:

1. UI2 shoveling as a backup for UI1 shoveling.

2. UI1 *tuberculum dentale* as a backup for UC *tuberculum dentale*.

3. UM2 cusp 5 as a backup for UM1 cusp 5.

4. LM1 enamel extensions as a backup for UM1 enamel extensions.

5. LP1 multiple lingual cusps as a backup for LP2 multiple lingual cusps.

6. 4-cusped LM1 as a backup for LM2 with four cusps.

7. LM2 cusp 6 as a backup for LM1 cusp 6.

These backups ensure that even if the key tooth cannot be assessed due to damage or other factors, other teeth within the same field can still provide crucial information in an analysis. Using the current version of rASUDAS2 (available as freeware through osteomics.com), it would be possible to score an individual for all primary and backup traits. However, given the assumption of trait independence, researchers should only enter an observation on a backup trait when the primary tooth for that trait is unobservable.

2. Methods and Materials

2.1 Naïve Bayes classification algorithm

rASUDAS2, a web-based application developed on the R statistical platform, enables researchers

to estimate the likelihood of an individual belonging to one of seven biogeographic groups using

dental scores for up to 25 morphological traits. Utilizing a naïve Bayes classification algorithm,

rASUDAS2 assumes strong independence among traits, simplifying computation for practicality

(Hastie, Tibshirani & Friedman 2009). Its efficiency lies in ranking class probabilities, with

accurate predictions based on the probability of the true class exceeding others. For its

mathematical foundation, Zhang (2004) and Stephens, Huerta, & Linares (2018) provide

theoretical insights into the efficiency of this family of generative classifiers. The naïve Bayes

classifier for population affinity estimation from dental traits is expressed as

$$P(G_k \mid X) = \frac{P(G_k)\prod_{i=1}^{p} P(X_i \mid G_k)}{P(X_i)} = \frac{P(G_k)\prod_{i=1}^{p} P(X_i \mid G_k)}{\sum_{i=1}^{r} P(G_k)\prod_{i=1}^{p} P(X_i \mid G_k)}$$

where G represents ancestral groups, X represents the observed dental traits, *r* indicates the

number of groups in the analysis, and *p* denotes the number of traits considered in the

computation.

To deduce population affinity using this method, the posterior probability of an

individual belonging to a specific biogeographic group is computed based on observed traits.

This calculation is done for each of the seven groups, and the one with the highest posterior

probability is considered the most likely 'ancestor.' The strength of the naïve Bayes classifier

lies in its reliance on conditional probabilities for affinity determination. It calculates the

probabilities of observing a morphological trait given a known population affinity and, using

Bayes theorem, inverts these probabilities to determine an individual's likelihood of belonging to a certain biogeographic group based on observed dental traits. Coelho et al. (2017) demonstrate how the naïve Bayes classification can be used with cranial morphoscopic traits to estimate an individual's population affinity.

To construct a naïve Bayes classifier, conditional frequency estimates are required for each trait. These frequencies are estimated from datasets or sourced from population studies. In this instance, frequencies were derived by analyzing multiple archaeological samples (dating mostly from the last two millennia) from each major geographic region. The algorithm was trained using trait frequencies from 166 populations, representing over 25,000 individuals globally. For rASUDAS2, posterior probabilities are calculated using Bayes theorem for up to 25 traits. For a detailed description of rASUDAS, the composition of the samples, and assumptions in the analysis, we direct the reader to Scott et al. (2018a).

With rASUDAS2, you can compute posterior probabilities for as few as two geographic groups and one trait, up to seven geographic groups and 25 traits. The ideal is to include as many traits as possible but given the limitations imposed by missing teeth, crown wear, and dental work/appliances, it is rarely possible to achieve the maximum of 25. As a middle ground to maximize sample size, we have adopted 12 as the minimum number of scorable traits for an individual to be included in an analysis when multiple individuals are evaluated in batch files (cf. Scott et al. 2021, 2023).

2.2 Ranked scores vs. template scores and how to deal with antimere asymmetry

When scoring crown and root morphology, the methodologies adhere to those specified in Turner et al. (1991) and Scott and Irish (2017). Crown traits are typically scored on ranked scales

ranging from absence (0) to varying degrees of presence (1...n). When an individual is evaluated,

there are drop-down boxes for each trait that include the grades of expression to be entered. For

running batch files with multiple individuals, ranked values are translated into the template

scores of 1 or 2, and sometimes 3. For instance, shoveling is graded on a scale from 0 to 7, but

with the template, these are consolidated into: score 1 for grades 0 and 1 (absence or minimal

expression); score 2 for grades 2 and 3 (moderate expression); and score 3 for grade 4 and above

(strong expression), as illustrated in Figure 1. Similarly, for Carabelli's trait: score 1 corresponds

to grades 0 and 1; score 2 encompasses grades 2, 3, and 4; and score 3 applies to grade 5 and

above.

INSERT FIGURE 1

On the application interface, when analyzing a single individual, two or three scores are

shown. The first typically indicates absence (either grade 0 or grades 0-1), and the second shows

presence (grades 1-n, or grades 2-n and above). Some traits have unique scoring methods that

deviate from this pattern. For LM2 groove pattern, "X" and "+" constitute the first grade

(template score 1), and "Y" is the second grade (template score 2). In the case of 4-cusped lower

molars, grades 1-5 (presence of the hypoconulid) are grouped under template score 1, while

grade 0 (absence of the hypoconulid) is template score 2.

For root traits, the highest root number typically corresponds to template score 2. An

exception is found with LM2, where 2-roots (grade 2) is template score 1, and 1-root (grade 1) is

template score 2. These modifications in template scoring are designed to accommodate the

geographic variability and complexity of dental morphology, making it easier to standardize and

compare data across different samples.

When asymmetry occurs between antimeres (corresponding parts on opposite sides of a bilateral structure), the longstanding practice is to use the higher grade of expression, following Turner and Scott (1977). This approach ensures that the most pronounced expression of a trait is considered in the analysis. For example, for LM2 groove pattern, if one side displays an X-+ pattern (grade 1) and the other side a Y pattern (grade 2), the higher grade, Y (template score 2), would be entered in the application for evaluating single individuals or running batch files for multiple individuals.

Similarly, for LM2 root number, if asymmetry presents with one antimere having 2 roots (graded as 1) and the other having 1 root (graded as 2), the higher grade, 2, is entered on the Excel template. In essence, whenever there is a discrepancy in grades or template scores between antimeres, the protocol dictates selecting the higher of the two scores to ensure that the most significant morphological expression is recorded and analyzed.

2.3 Samples

In 2018, dental morphological observations were made on European-derived, Sub-Saharan African, and samples with both European and African genetic components. The first samples included American white and American black skeletal series curated at the University of Tennessee Knoxville and Texas State University. Information regarding population affinity was either self-reported or reported by family members. Among a total of 355 individuals studied, 186 (52.4%) could be scored for 12 or more crown and root traits. For the Sub-Saharan African samples, observations were carried out on African blacks at Pretoria University and the University of Witwatersrand. Out of 378 individuals examined, 159 (42.0%) met the criteria for scoring 12 or more traits. Additionally, data were collected on an African sample with genetic contributions from both African and European gene pools (referred to as Cape Coloured in the

literature). This sample, comprising 99 individuals studied at Stellenbosch University, saw 58

(58.5%) individuals meeting the threshold for scoring 12 or more traits (see Table 2 for details).

INSERT TABLE 2

3. Results: testing rASUDAS2 on Sub-Saharan African, Western Eurasian, and samples of

diverse affinity

Table 3 presents the percentage and count distribution of each of the four samples relative to the

seven major biogeographic groups. American whites predominantly fall under the classification

of Western Eurasian, accounting for 75.2% of cases. The Sub-Saharan African group represents

the second highest percentage assignment, at 13.0%. Conversely, for African blacks, a reverse

pattern is observed. Here, 68.6% of individuals are assigned to Sub-Saharan Africa, with 22.0%

assigned to Western Eurasia. Notably, less than 10% of all individuals across the samples are

assigned to Asian or Asian-derived populations.

INSERT TABLE 3

Although genetic data that could help determine the extent of gene flow between Western

Eurasians and Sub-Saharan Africans are not available to us, it is noteworthy that the African

sample from Cape Town exhibits a frequency distribution almost identical to that of African

Americans. In both cases, approximately 40% of individuals are assigned to Sub-Saharan Africa,

with 30% assigned to Western Eurasia. Moreover, in these samples, over 10% of individuals are

classified as Australo-Melanesian, while assignments to East Asia and Native Americans are

minimal or absent. Figure 2 illustrates the percentage of individuals assigned to each of the major

biogeographic groups.

INSERT FIGURE 2

4.   Discussion and future directions

In their analysis of over 1400 individuals from Southeast Asia, East Asia, American Arctic,

Northwest Coast, North America, Mesoamerica, and South America using rASUDAS2, Scott et

al. (2023) observed very few individuals that classified as Western Eurasian or Sub-Saharan

African. In this specific test case, the results are reversed, with very few Western Eurasian or

Sub-Saharan African individuals classified within any Asian, Pacific, or Native American group.

While numerous crown and root traits differentiate Asian and Asian-derived populations

from Western Eurasians and Sub-Saharan Africans, our focus now shifts to identifying

morphological traits that more effectively distinguish between Western Eurasians and Sub-

Saharan Africans. Currently, we have identified three candidate traits that were not previously

observed by Christy G. Turner II, whose observations were utilized to estimate frequency set

points for the Bayes algorithm.

Firstly, the midline diastema (the space between left and right UI1) is more prevalent in

Sub-Saharan Africans than in Western Eurasians (Irish 1997; Scott and Irish 2017). Secondly,

crenulations of the upper and lower second and third molars are more common in Sub-Saharan

Africans compared to Western Eurasians (Pilloud et al. 2018). Crenulations are described as

"closely spaced enamel ridges of various lengths and complexities" (Scott and Fox, 2005:636).

The third trait, to our knowledge, has not been systematically classified or studied to date, and

that is a reduced entoconid (distolingual cusp) of the lower second molars. When making

morphological observations on a large sample of Hungarians, the senior author and his graduate

students noted a few instances where LM2 exhibited three cusps, with the entoconid entirely

absent (Dern, 2022). Preliminary assessments of this variable across diverse groups, including

Asiatic Indian, Bantu, native Australian, Chinese, and Malay samples, suggest that reduced

entoconids are more prevalent in Western Eurasians than in any other world population, including Sub-Saharan Africans.

Our next objective is to examine dentitions from the seven major biogeographic groups represented in the rASUDAS2 application and generate point estimates for these traits. These estimates will be incorporated into the rASUDAS2 algorithm to enhance its accuracy in distinguishing between Western Eurasians and Sub-Saharan Africans (see Figure 3 for trait illustrations).

INSERT FIGURE 3

While we have established 12 as the minimum number of traits for assessing population affinity, using as many traits as possible is always preferable. However, in certain forensic cases where only a few traits can be scored, the method remains powerful if key traits are observable. For instance, if an individual exhibits UI1 winging and grade 4 shoveling, the highest probability of affinity would be non-Arctic American (62.1%) and East Asian (22.2%), with extremely low posterior probabilities for Western Eurasia (0.3%) and Sub-Saharan Africa (0.1%). Conversely, if an individual shows no winging or shoveling, the highest probabilities would be Western Eurasia (34.8%) and Sub-Saharan Africa (37.0%), with the remaining 30% primarily assigned to Australo-Melanesia (14.5%) and Southeast Asia (11.8%). East Asia and the two Native American groups would each receive less than 1% probability.

In forensic cases, the significance of dental traits can be assessed using the rASUDAS2 application. By clicking on the "dental trait importance" button, users are directed to a screen displaying biogeographic groups and dental traits. To compute the importance of specific traits, select at least two groups and two traits. For instance, choosing Western Eurasia and Sub-

Saharan Africa along with all 32 traits reveals that the five most critical traits are 4-cusped LM2,

the UC Bushman canine, LM1 cusp 7, UM2 cusp 5, and UI2 interruption grooves. Notably,

incisor shoveling ranks 30th in importance for distinguishing between these two groups but

becomes the top trait when all seven groups are considered. In forensic contexts, it is advisable

for researchers to evaluate the relevance of dental traits observable in an individual, with

particular attention to geographic context. For example, in a U.S. city, it might be prudent to

exclude Australo-Melanesia and Arctic America from the analysis.

4.1 A forensic test case

In 2004, the senior author conducted a forensic examination focused solely on dental

morphology. The case originated from an archaeological site near Sutter's Fort in Sacramento,

CA, where a burial was being excavated due to impending construction. During the excavation, a

Native American monitor noted the individual had shovel-shaped incisors, even though the body

was dressed in a soldier's uniform. To adhere to the Native American Graves Protection and

Repatriation Act (NAGPRA) regulations, dental casts of the teeth were made. Given the

skeleton's fragile state, the casts were primarily of individual teeth.

Upon receiving the casts, the senior author arranged the teeth according to their positions

in the jaw and embedded them in modeling clay for photographic documentation (see Figure 4).

The dental traits that were evaluated, along with their scores, included UI1 shoveling (0-1),

double shoveling (0-1), UI2 interruption grooves (1-4), canine tubercle (2-7), Bushman canine

(0-1), Carabelli's trait (2-4), UM2 hypocone (4-6), UM1 cusp 5 (0), pegged-missing-reduced

UM3 (0), LP2 lingual cusp number (2-3), LM1 cusp 6 (0), LM1 cusp 7 (0), LM1 protostylid (0),

LM2 groove pattern (x-+), and 4-cusped LM2 (0).

INSERT FIGURE 4

The reader can input these traits and their grades into the rASUDAS2 application to obtain an estimate of the individual's population affinity. Our analysis revealed that when considering all seven biogeographic groups, Western Eurasia (WE) had the highest posterior probability of 0.710, followed by Southeast Asia (SEA) at 0.151 with all other groups less than 0.10. When SEA, AM, and American Arctic (AA) are excluded from the analysis, the results shift to WE at 0.933 and SSA at 0.062, with East Asia and Native Americans below 0.001. Considering only Western Eurasia and Native Americans (n-AA), the results show WE at 0.999 and n-AA at 0.0001.

When the initial analysis was conducted in 2004, there was no available method to estimate the likelihood of an individual belonging to one biogeographic group over another. With the advent of the rASUDAS2 tool, our analysis indicates that this individual most likely had Western Eurasian origins. The probability that this individual had Asian or Asian-derived ancestry is low, and the chance of finding these specific traits within a Native American population is nearly zero.

4.2 Potential of nonmetric dental traits in forensic analyses

Hefner et al. (2012:297) discussed the potential of using nonmetric traits, including dental morphology, for affinity assessment. They note "By studying the frequency of expression of individual traits, combining them into suites of significant traits, and then analyzing them within a statistical framework, we might begin to see patterns emerge that will allow us to make scientific and statistically valid assessments of ancestry based on nonmetric traits. These assessments, however, require (i) significant reference data (large sample sizes from multiple and

diverse skeletal populations), (ii) better and standardized protocols for trait and character state

recordation and coding, and (iii) rigorous classification statistics appropriate for categorical data

analysis."

We feel rASUDAS2 meets the requirements outlined above. Our reference dataset is

enormous and global in scope, involving the 20,000+ individuals scored by C.G. Turner II plus

data on an additional 10,000+ individuals from the work of J.D. Irish and G.R. Scott. The

protocols for scoring traits have been established for over 30 years (cf. Turner, et al. 1991) with

reference plaques available through Bone Clones as the Turner-Scott system (formerly

ASUDAS). As for classification statistics, rASUDAS2 is based on a naïve Bayes algorithm

which is rigorous and user-friendly. As points (i) and (ii) are well grounded, we remain open to

the possibility of using other classification statistics for affinity assessment using crown and root

morphology. We are optimistic that advances in the field (adding more traits and more samples

from diverse regions and temporal periods) will lead to further refinements of rASUDAS that

can be fruitfully employed when addressing either forensic and/or bioarchaeological problems.

References

Coelho C, Navega D, Cunha E, Ferreira MT & Wasterlain SN. Ancestry estimation based on morphoscopic traits in a sample of african slaves from Lagos, Portugal (15th–17th Centuries). *International Journal of Osteoarchaeology* 2017;27:320–326.

Dahlberg AA. The changing dentition of man. *Journal of the American Dental Association* 1945;32:676-690

Dern LL. *The Impact of Medieval and Early Modern Migrations on Dental Nonmetric Variation in Hungary*. Unpublished PhD dissertation, University of Nevada, Reno; 2023.

Hastie T, Tibshirani R, & Friedman J. *The Elements of Statistical Learning*. New York, NY:Springer; 2009.

Hefner JT, Ousley S, & Dirkmaat DC. Morphoscopic traits and the assessment of ancestry. In: Dirkmaat DC, ed., *A Companion to Forensic Anthropology*, Chichester, West Suffix, UK:Wiley-Blackwell; 2012:287-310.

Irish JD. Characteristic high- and low-frequency dental traits in sub-Saharan African populations. *American Journal of Physical Anthropology* 1997:102:455-467.

Pilloud MA, Maier C, Scott GR, Edgar HJ. Molar crenulation trait definition and variation in modern human populations. *Homo* 2018;69:77-85.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2013.

Scott CS & Fox RC. Windows on the evolution of Picrodus (Plesiadapiformes: Primates):

morphology and relationships of a species complex from the Paleocene of Alberta.

Journal of Paleontology 2005;79:635–657.

Scott GR. Lingual tubercles and the maxillary incisor-canine field. *Journal of Dental*

Research, 1977a;56:1192-1192.

Scott GR. Interaction between shoveling of the maxillary and mandibular incisors. *Journal of*

Dental Research 1977b;56:1423.

Scott GR, Turner CG, II. *The Anthropology of Modern Human Teeth: Dental Morphology and Its*

Variation in Recent Human Populations. Cambridge: Cambridge University Press; 1997.

Scott GR & Irish JD. *The Arizona State University Dental Anthropology System: Tooth Crown and*

Root Morphology. Cambridge: Cambridge University Press; 2017.

Scott GR, Navega D, Coelho J, Pilloud MA, Cunha E, Irish JD. rASUDAS: a new web-based

application for estimating ancestry from tooth morphology. *Forensic Anthropology*

2018a;1:18–31.

Scott GR, Turner CG, II, Townsend GC, Martinón-Torres M. *The Anthropology of Modern Human*

Teeth: Dental Morphology and Its Variation in Recent and Fossil Homo sapiens, 2nd ed.

Cambridge: Cambridge University Press; 2018b

Scott GR, O'Rourke DH, Raff JA, Tackney JC, Hlusko LJ, Elias SA, et al. Peopling the Americas:

Not "out of Japan." *PaleoAmerica* 2021;7:309–332.

Scott GR, Navega D, Vlemincq-Mendieta T, Dern LL, O'Rourke DH, Hlusko LJ, Hoffecker JF.

Peopling of the Americas: a new approach to assessing dental morphological variation in

Asian and Native American populations. *American Journal of Biological Anthropology*

2023; DOI: 10.1002/ajpa.24878, early view.

Stephens C R, Huerta HF, Linares AR. When is the Naive Bayes approximation not so naive?

   Machine Learning 2018;107:397–441.

Turner CG, II, Scott GR. Dentition of Easter Islanders. In: Dahlberg AA, Graber T., eds.,

   Orofacial Growth and Development. The Hague: Mouton; 1977:229-249.

Turner CG, II, Nichol CR, Scott GR. The Arizona State University dental anthropology system:

   scoring procedures for key morphological traits of the permanent dentition. In: Kelley MA,

   Larson CS, eds., *Advances in Dental Anthropology*. New York: Wiley-Liss; 1991:13-31.

Zhang H. The optimality of Naive Bayes. *Proceedings of the Seventeenth International Florida

   Artificial Intelligence Research Society Conference, FLAIRS* 2004:2:562–567.

Figure captions

Figure 1. The ranked scores for UI1 and UI2 shovel-shaped incisors. Each incisor can be scored as absent (grade 0) or, when present, from slight to pronounced (grades 1-7). For processing multiple individuals in a batch file, grades are not used directly. All template scores are either 1 or 2. Where the range of expression is pronounced, as in shoveling and Carabelli's trait, there is a template score of 3. The correspondence between expression grades and template scores are shown in the figure. To assess the population affinity of an individual in a forensic context, the researcher should use the physical plaques produced by Bone Clones (Turner-Scott system) or the photos of the traits and plaques in Scott and Irish (2017).

Figure 2. Bar chart showing percent of highest assignments to each of the seven biogeographic groups for the four recent samples of American whites, African blacks, and two samples with genetic contributions from each of those groups.

Figure 3. Traits that may help better distinguish Western Eurasians and Sub-Saharan Africans: (A) three examples of reduced entoconids (distolingual cusp) that are 50% or less the size of the adjacent distobuccal cusp (hypoconid); (B) three examples of midline diastemas showing a distinct gap between the two upper central incisors; and (C) two examples of crenulated molars with extensive wrinkling on the crown surfaces of primarily LM2 and LM3 (see Pilloud et al. 2018 for a description of molar crenulations).

Figure 4. Dental casts from teeth found at Sutter's Fort in Sacramento, CA, along with grades of expression in parentheses. A. Lingual view of upper anterior teeth: UI1 shoveling (1:0-1), UI2 interruption grooves (2:1-4), Bushman canine (0-1), and canine tubercle (3:2-7). B. Labial view of upper anterior teeth: double shoveling (0-1). C. Left and right upper molar tooth districts: Carabelli's trait (1:2-4), UM2 hypocone (2:4-6), UM1 cusp 5 (0) and PRM UM3 (3:0). D. Left and right lower molars: LM1 cusp 6 (0), LM1 cusp 7 (0), LM1 protostylid (0), LM1 deflecting wrinkle (1:3), LM2 groove pattern (2:X-+), and 4-cusped LM2 (0). E. UM1: grade 4 Carabelli's trait. F. LP2: lingual cusp number (2-3). Note: only traits that are present are noted by numbers and arrows.

Table 1. Point estimates (trait frequencies) for new traits and backup traits used in rASUDAS2

Trait and tooth	Rank	Template score	WE	EA	AA	n-AA	SEA	AM
BU: Shoveling UI2	0+1	1	0.891	0.062	0.093	0.025	0.458	0.694
	2+3	2	0.109	0.473	0.646	0.448	0.477	0.277
	4+	3	0.000	0.465	0.261	0.527	0.065	0.029
NT: Double-shov UI1	0-1	1	0.860	0.576	0.449	0.227	0.746	0.896
	2+	2	0.140	0.424	0.551	0.773	0.254	0.104
NT: Bushman Canine	0-1	1	0.972	0.992	1.000	0.994	0.980	0.972
	2+	2	0.028	0.008	0.000	0.006	0.020	0.028
NT: Canine tubercle	0-1	1	0.690	0.836	0.691	0.699	0.522	0.617
	2+	2	0.310	0.164	0.309	0.301	0.468	0.383
BU: UI1 TD	0-1	1	0.563	0.651	0.649	0.551	0.674	0.559
	2+	2	0.437	0.349	0.551	0.449	0.326	0.441
BU: Cusp 5 UM2	0	1	0.892	0.916	0.802	0.864	0.878	0.656
	1+	2	0.108	0.089	0.138	0.136	0.122	0.344
BU: Enamel extensions LM1	0-1	1	0.959	0.458	0.448	0.630	0.669	0.881
	2+	2	0.041	0.542	0.552	0.370	0.231	0.119
NT: Lower canine distal accessory ridge	0-1	1	0.889	0.602	0.482	0.364	0.863	0.853
	2+	2	0.111	0.398	0.518	0.636	0.137	0.147
BU: multiple lingual cusps LP1	0-1	1	0.867	0.654	0.803	0.675	0.592	0.604
	2+	2	0.132	0.346	0.197	0.325	0.408	0.396
BU: 4-cusped LM1	5	1	0.863	1.000	1.000	0.996	0.990	0.945
	4	2	0.137	0.000	0.000	0.004	0.010	0.055
BU: Cusp 6 LM2	0	1	0.983	0.858	0.782	0.746	0.915	0.864
	1+	2	0.017	0.142	0.218	0.254	0.095	0.136

*see text for group abbreviations; BU = backup trait, *NT = new trait*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59

SSA
0.845
0.155
0.000
0.981
0.019
0.771
0.229
0.449
0.551
0.533
0.467
0.686
0.314
0.997
0.003
0.619
0.381
0.752
0.248
0.975
0.025
0.902
0.098

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 2. Samples used to test rASUDAS2

Sample	n	Sample source
American white	161	University of Tennessee Knoxville &
American black	25	Texas State University
African black	159	Pretoria University & University of Witwatersrand
African mixed (Cape Coloured)	58	Stellenbosch University
Total	403	

Table 3. Percent of highest assignments to each of seven biogeographic groups

Sample*	n		WE	EA	AA	n-AA	SEA	AM	SSA
American white	161	Percent	0.752	0.000	0.019	0.000	0.025	0.075	0.130
		Count	121	0	3	0	4	12	21
African black	159	Percent	0.220	0.006	0.044	0.006	0.006	0.031	0.686
		Count	35	1	7	1	1	5	109
African (mixed)	58	Percent	0.293	0.034	0.086	0.000	0.069	0.103	0.414
		Count	17	2	5	0	4	6	24
African American	25	Percent	0.320	0.000	0.000	0.000	0.120	0.150	0.400
		Count	8	0	0	0	3	4	10

*Abbreviations are noted in text

Figure 1. The ranked scores for UI1 and UI2 shovel-shaped incisors. Each incisor can be scored as absent (grade 0) or, when present, from slight to pronounced (grades 1-7). For processing multiple individuals in a batch file, grades are not used directly. All template scores are either 1 or 2. Where the range of expression is pronounced, as in shoveling and Carabelli's trait, there is a template score of 3. The correspondence between expression grades and template scores are shown in the figure. To assess the population affinity of an individual in a forensic context, the researcher should use the physical plaques produced by Bone Clones (Turner-Scott system) or the photos of the traits and plaques in Scott and Irish (2017).
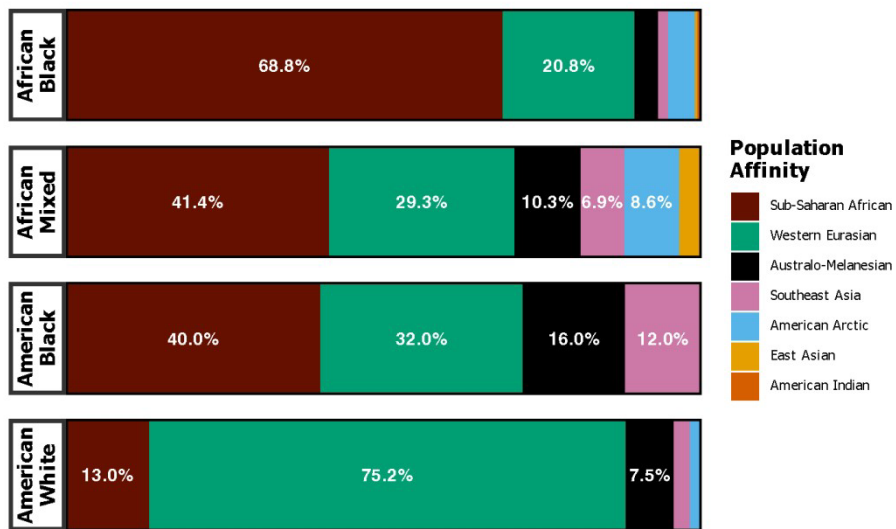
297x209mm (300 x 300 DPI)

Figure 2. Bar chart showing percent of highest assignments to each of the seven biogeographic groups for the four recent samples of American whites, African blacks, and two samples with genetic contributions from each of those groups.

203x127mm (200 x 200 DPI)

Figure 3. Traits that may help better distinguish Western Eurasians and Sub-Saharan Africans: (A) three examples of reduced entoconids (distolingual cusp) that are 50% or less the size of the adjacent distobuccal cusp (hypoconid); (B) three examples of midline diastemas showing a distinct gap between the two upper central incisors; and (C) two examples of crenulated molars with extensive wrinkling on the crown surfaces of primarily LM2 and LM3 (see Pilloud et al. 2018 for a description of molar crenulations).
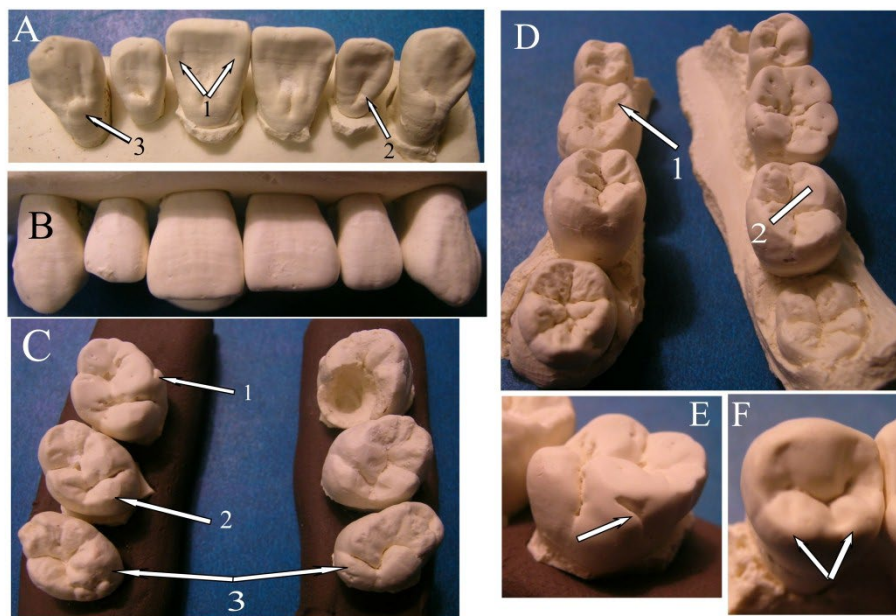
297x209mm (300 x 300 DPI)

Figure 4. Dental casts from teeth found at Sutter's Fort in Sacramento, CA, along with grades of expression in parentheses. A. Lingual view of upper anterior teeth: UI1 shoveling (1:0-1), UI2 interruption grooves (2:1-4), Bushman canine (0-1), and canine tubercle (3:2-7). B. Labial view of upper anterior teeth: double shoveling (0-1). C. Left and right upper molar tooth districts: Carabelli's trait (1:2-4), UM2 hypocone (2:4-6), UM1 cusp 5 (0) and PRM UM3 (3:0). D. Left and right lower molars: LM1 cusp 6 (0), LM1 cusp 7 (0), LM1 protostylid (0), LM1 deflecting wrinkle (1:3), LM2 groove pattern (2:X-+), and 4-cusped LM2 (0). E. UM1: grade 4 Carabelli's trait. F. LP2: lingual cusp number (2-3). Note: only traits that are present are noted by numbers and arrows.

297x209mm (300 x 300 DPI)